Cross-Language Influences in the Processing of Multi-Word Expressions: From a First Language to Second and Back

By

Lingli Du

A thesis submitted to Victoria University of Wellington

in fulfilment of the requirements for

the degree of Doctor of Philosophy in Applied Linguistics

Victoria University of Wellington, New Zealand

2022

Abstract

Cross-language similarity in form and meaning between the first language (L1) and the second language (L2) benefits bilingual language processing. The facilitative role of congruency is well-established in bilingual lexical (single word) processing, such as cognate facilitation effects. Recent research has suggested that the processing advantage afforded by congruency also extends to units beyond the word level, to multi-word expressions (MWEs). Congruent L2 MWEs that have an equivalent form in the L1 are likely to be processed faster and more accurately than incongruent L2-only MWEs that have no equivalent in the L1. However, it is still unclear what mechanisms underpin the congruency effect in the processing of L2 MWEs, and whether congruency can also affect L1 MWE processing. To further understand cross-language influences in the processing of MWEs, the present study investigated cross-language influences in the processing of binomial expressions (*knife and fork*) in the L1→L2 and L2→L1 directions. Two groups of unbalanced bilinguals (Chinese-English and English-Chinese) and a control group of English monolinguals performed a visual lexical decision task that incorporated unmasked priming. To assess cross-language influences, I used three types of expressions: congruent binomials (English binomials that have translation equivalents in Chinese), English-only binomials, and Chinese-only binomials translated into English. Lexical decision latencies to the last word (*fork*) in a binomial (*knife and fork*) were compared with response latencies to the same word in a matched control phrase (*spoon and fork*). I found that (1) Chinese-English bilinguals in Experiment 1a showed a significant priming effect for congruent binomials but no facilitation for English-only binomials, (2) English monolingual controls in Experiment 1b showed comparable priming for congruent and English-only binomials, (3) English-Chinese bilinguals in Experiment 2 showed a trend toward priming for congruent binomials, which did not reach statistical significance, and no priming for English-only binomials. None of the three participant groups showed priming for translated Chinese-only binomial over controls. These findings suggest that L1 influences the processing of L2 binomials, and that there may be some cross-linguistic influence in the opposite direction, i.e., from L2 to L1, although to a lesser extent. For the Chinese-English bilinguals, the facilitation for congruent binomials is probably not due to the automatic activation of L1 translation equivalents of the L2 binomials. I conclude that exposure to binomial phrases in the L2 is needed for the congruency effect to occur. For the English-Chinese bilinguals, they showed no priming for English-only binomials. I propose that their L1 may be inhibited in L2 learning and immersion contexts and switching back to L1 may come at a cost. However, the English-Chinese bilinguals showed a weak priming in the processing of congruent binomials, providing initial evidence that crosslinguistic influence can occur from the non-dominant L2 to the dominant L1, even in an entirely within-L1 task.

Acknowledgements

The journey to the completion of this thesis has by no means been smooth! It would not have been possible without the help I received from the following people:

First and foremost, I would like to express my deepest gratitude to my supervisors, Anna Siyanova-Chanturia and Irina Elgort, for their expert advice, constructive criticism, and professional guidance. Many thanks to Anna for suggesting me to enrol in Statistics even before I started my PhD study, which has laid a solid foundation for my research. A very special thanks to Irina for supervising and helping me complete my research proposal while Anna was on leave. I would also like to thank them for their hard but thought-provoking questions, which help me to think independently and critically. I am especially grateful to them for their generous compliments on any little progress I have made, which greatly boosts my confidence and helps me to reach my full potential. They encourage and guide me to attend conferences and present seminars. They assist and collaborate with me in publication of work arising from this thesis. Their supervision and support make this PhD journey such a pleasant and rewarding experience. Importantly, I owe a great deal to my supervisors for always being empathetic, giving prompt feedback, and supporting me, even when they have many responsibilities to take care of themselves. Finally, their passion for research, positive attitudes and resilience will keep inspiring me to learn and take on new challenges.

Many thanks to Lisa Woods for her expertise and incredible patience with my statistical analyses. Without her help, the data analysis would have taken ages to be done. Thank you also to learning advisers, Deborah Laurs and TJ Boutorwick, at Student Learning, for their help with my academic writing.

I would also like to express my sincerest gratitude to staff at School of Linguistics and Applied Language Studies at Victoria University of Wellington. Many thanks to Professor

Paul Warren for answering my questions about study design with great patience. Thanks also to Janet Attrill for paying a visit to my office when she heard my eyes suffered from strong sunshine. A very special thanks to Bernie Hambleton for letting me have the luxury of staying in my old office alone for six months. It was during this period that most hard thinking was done.

A very special thanks to my officemates, HaThi Phuong Pham, Taha Omidian, and Ken Smith, for their company and being great inspirations to me. Many thanks to Oliver Ballance and Ha for helping me prepare my presentation.

I gratefully acknowledge the support provided to me by Victoria University of Wellington in the form of a Faculty Research Grant and a Wellington Doctoral Submission Scholarship. I am also grateful to the individuals who participated in my experiments.

My most special thanks I have saved till the end. I do not think I can possibly express my gratitude to my parents, my brother and sisters for their unconditional love and support. Many thanks to my elder son, Zhenyuan Liu, for being so understanding and tolerant with my busy research. I owe so much to my younger son, Zhaosong Liu. This thesis basically stripped him of his mum. A very special thanks to my parents-in-law for taking care of him. Finally, this study would not have been possible without the unconditional support from my husband, Qi Liu. Thank you so much for supporting the whole family. This thesis is for you!

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

AoA	Age-of-acquisition
BIA	Bilingual Interactive Activation
CCL	Center for Chinese Linguistics
CLI	Cross-language influence
COCA	Corpus of Contemporary American English
EFL	English as a foreign language
EPP	English Proficiency Program
ERP	Event-related brain potential
ESL	English as a second language
HSK	Hanyu Shuiping Kaoshi
IELTS	International English Language Testing System
ISI	Inter-stimulus interval
L1	First language
L2	Second language
LDT	Lexical decision task
MS	Milliseconds
MWEs	Multi-word expressions
Q-Q plot	Quantile-quantile plot
RT	Reaction time
SD	Standard deviation
SOA	Stimulus onset asynchrony
SOPHIA	Semantic, orthographic, and phonological interactive activation
TOEFL	Test of English as a Foreign Language
USF	University of South Florida
VUW	Victoria University of Wellington

Chapter 1 Introduction

1.1 Setting the Scene

Cross-language influence (CLI) refers to the phenomenon that a person's knowledge of one language can affect that person's knowledge or use of another language (Jarvis & Pavlenko, 2008, p. 1). CLI abounds in various areas of language use by bilinguals, such as word order, relativization, negation, vocabulary, phonology, narrative structure, discourse style, conversational strategies, etc. It can also be identified in a variety of psycholinguistic processes, such as lexical and syntactic processing (Cook et al., 2003; Dijkstra, 2003; Kroll & Dussias, 2012), in listening and reading comprehension (Su, 2001; Upton & Lee-Thompson, 2001), in tip-of-the-tongue states (Ecke, 2001; Gollan & Silverberg, 2001), and in conceptual representation (Jarvis, 1998; Kroll & de Groot, 1997). Crucially, research on bilingual language processing has found that cross-language overlap or congruency (i.e., similarity in form and meaning between the first language and the second language) benefits bilingual language processing. When there is an overlap between the first language (L1) and the second language (L2), L1 plays an important facilitative role in L2 processing, and vice versa. A classic example of congruency advantage is cognate facilitation effect, wherein cognates (words that have a similar form and meaning across languages, such as *film, taxi*, and *restaurant* in French and English) are processed faster and more accurately than matched control words (de Groot et al., 2000, Experiment 2; Dijkstra, Van Heuven, et al., 1998; Libben & Titone, 2009). Even for bilinguals whose languages differ in script and share no orthographic similarity (e.g., Japanese-English, Korean-English, Hebrew-English, Greek-French), shared phonological and semantic similarity can also facilitate processing of cognates (Gollan et al., 1997; Hoshino & Kroll, 2008; Kim & Davis, 2003; Taft, 2002; Voga

& Grainger, 2007). Further, for words which share no formal features but only meaning (e.g., non-cognate translation equivalents), cross-linguistic semantic overlap appears to provide processing advantages for them in a variety of priming tasks, at least when primes are in the L1 (de Groot & Nas, 1991; Gollan et al., 1997; Jiang, 1999). A number of studies have shown that when bilinguals process language in their L2, they obligatorily activate the L1 translation equivalents (i.e., cross-language translation priming: e.g., the L2 word *horse* primes its L1 translation equivalent *ma*/*马*) (Wu et al., 2013; Wu & Thierry, 2010; Zhang et al., 2011). These robust facilitation effects underpinned by overlap in semantics, phonology and/or orthography are widely interpreted as evidence for language co-activation. That is, the two languages of a bilingual can be activated simultaneously or non-selectively, and this cross-language activation influences processing (for a review, see Van Hell & Tanner, 2012).

A number of models have been proposed to account for cross-language activation in bilingual lexical processing, such as the Bilingual Interactive Activation (BIA) model (Dijkstra & Van Heuven, 1998; Dijkstra, Van Heuven, et al., 1998), its successor, BIA+ (Dijkstra & Van Heuven, 2002), and the Multilink model (Dijkstra et al., 2019). The BIA model assumes four hierarchical levels of representation units or “nodes”, starting from visual letter features, letters, orthographic forms of words to language information. The activation is assumed to start from letter feature nodes, leading to the activation of word nodes from both languages of the bilingual. To illustrate, when a Dutch-English bilingual sees an English word *land*, orthographic neighbours in both languages become activated (such as the English neighbour *lend* and the Dutch neighbour *mand*). However, the BIA model restricts cross-language activation/interaction to shared orthographic information (Kroll & Dussias, 2004). To account for the effects of phonological and semantic similarities in bilingual language processing, Dijkstra and Van Heuven (2002) extended the BIA model by adding phonological and semantic representations, proposing a new model, SOPHIA (semantic, orthographic, and phonological interactive activation). Instead of the inhibitory connections from the language nodes proposed in the BIA, Dijkstra and Van Heuven (2002) proposed a task/decision system (a control system) to account for the extra-linguistic influences on language processing. The task/decision system and SOPHIA's word identification system together form the new model, BIA+ (Dijkstra & Van Heuven, 2002). Moreover, built on the BIA+ model and the Revised Hierarchical Model (RHM; Kroll & Stewart, 1994), Dijkstra et al. (2019) proposed the Multilink model to provide a unified account of bilingual word comprehension and word production.

Importantly, recent research has suggested that the processing advantage afforded by cross-linguistic similarity or congruency may also extend to units beyond the word level, to multi-word expressions (henceforth, MWEs), such as idioms (*spill the beans*) and collocations (*spread news*). A number of studies have shown that congruent L2 MWEs that have an equivalent form in the L1 are likely to be processed faster and more accurately than incongruent L2-only MWEs that have no equivalent in the L1 (Carrol & Conklin, 2014, 2017; Carrol et al., 2016; Wolter & Gyllstad, 2011, 2013; Yamashita & Jiang, 2010). For example, Wolter and colleagues (Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2018) observed that high-proficiency L2 speakers showed a robust processing advantage in response times for congruent versus incongruent collocations. They observed that congruent collocations (verb-noun and adjective-noun) were processed significantly faster and more accurately than incongruent (English-only) collocations by advanced Swedish learners of English, in lexical decision (Wolter & Gyllstad, 2011) and acceptability judgement experiments (Wolter & Gyllstad, 2013). Comparable results were reported with Japanese-English bilinguals (Wolter & Yamashita, 2018). Similarly, L2 idiom processing studies also found a facilitative effect of congruency (Carrol et al., 2016; Titone et al., 2015). For example, in an eye-tracking study, Carrol et al. (2016) found that advanced Swedish learners of English processed congruent idioms faster than literal controls, whereas they processed incongruent (English-only) idioms and literal controls in a similar way, as indexed by the likelihood of skipping of the final word. Together, these studies suggest that congruency between languages can facilitate L2 MWE processing.

The congruency effect in MWE processing is interpreted by some researchers as evidence that cross-language activation extends to units beyond the word level. It is argued that known L1 MWEs are automatically activated in L2 processing, leading to their faster processing (e.g., Carrol & Conklin, 2014; Carrol et al., 2016; Wolter & Gyllstad, 2011). For example, Carrol and Conklin (2014, concerning idioms) proposed that L2 words automatically activate L1 equivalents in bilinguals, which, in turn, trigger a known L1 sequence via direct retrieval of a unitary form. Likewise, concerning collocations, Wolter and Gyllstad (2011) proposed that an L2 word activates not only its L2 collocates (e.g., *strong* activates its collocate *tea*), but also its L1 translation equivalent (*strong – nong*/*浓*), which in turn activates its L1 collocates via collocational priming (*cha*/*茶 – tea*). Thus, they argued that due to cross-language activation in bilinguals, congruent L2 MWEs are processed faster than incongruent L2-only MWEs.

If the above account holds, some facilitatory L1 influence should also be observed for the processing of translation equivalents of L1-only MWEs that are first encountered in an L2 (Carrol et al., 2016). If bilinguals process translated L1-only MWEs faster than L2 controls, without ever experiencing those exact L2 word combinations previously, then we could argue that this facilitation is due to the online activation of known L1 MWEs; that is, known MWEs are simultaneously activated in L2 processing (see also, Carrol & Conklin, 2017; Yamashita, 2018; Zeng et al., 2020). Several empirical studies did find that translated L1-only idioms are processed faster than literal controls by bilinguals, in spite of being encountered in the L2 for the first time (Carrol & Conklin, 2014, 2017; Carrol et al., 2016). Carrol and Conklin (2014), for example, found that high-proficiency Chinese-English bilinguals showed priming in a lexical decision task for translated Chinese-only idioms (e.g., *draw a snake and add … feet*) relative to matched controls (e.g., *draw a snake and add … hair*), whereas a control group of English monolinguals showed no priming. Similar findings were reported in a follow-up eye-tracking study with a similar population (Carrol & Conklin, 2017). However, contrasting results are reported in empirical studies with other types of MWEs, such as collocations. Studies on L2 *collocational* processing did not report a processing advantage for translated L1-only collocations compared to matched controls. Wolter and Yamashita (2015), for example, found that translated Japanese-only collocations (*high effect*) and non-collocational controls (*bad gift*) are processed in a similar way by intermediate and advanced Japanese-English bilinguals. The same results were replicated in their follow-up study that employed a phrase-acceptability judgment task (Wolter & Yamashita, 2018). Thus, studies with translated L1-only collocations provides no empirical support for the account that known MWEs are simultaneously activated in L2 processing (Wolter & Yamashita, 2015, 2018).

What lies behind the conflicting findings? If the underlying mechanism behind the congruency effect in L2 MWE processing is not the on-line activation of known L1 MWEs in L2 processing, then what underpins the robust facilitation for congruent L2 MWEs over incongruent L2-only MWEs? The aim of the present thesis is to probe into the CLI in MWE processing and the underlying mechanisms behind it. Further, little research, if any, has been conducted on the issue of whether the weaker L2 can affect the processing of MWEs in the L1. Thus, the present thesis also aims to extend previous studies by investigating whether CLI occurs in the processing of L1 MWEs. I am interested in exploring how directionality modulates CLI in MWE processing. Therefore, the focus in the thesis is on the role of cross-linguistic influence in the processing of MWEs, from L1 to L2 as well as from L2 to L1.

1.2 Aims of the Thesis

The present study is an empirical investigation of CLI in the on-line processing of one type of MWEs: binomial expressions (e.g., *knife and fork*). My research investigates CLI as a psycholinguistic phenomenon. The first aim is to investigate whether a bilingual's L1 influences the processing of MWEs in the L2, i.e., CLI in MWE processing in the direction of L1→L2. Specifically, it will investigate whether a bilingual's L1 influences the processing of congruent L2 MWEs, and whether this influence extends to the processing of translated L1-only MWEs (i.e., MWEs that do not exist in the L2). The effect of the L1 in the processing of congruent L2 MWEs will be investigated in two behavioural experiments. Experiment 1a will be with bilinguals and Experiment 1b with monolingual controls. Thus, the three key variables manipulated in the experiments are phrase type (binomials versus matched control phrases), congruency (i.e., whether the English binomial has a binomial equivalent in Chinese), and participant type (Chinese-English bilinguals and English monolinguals).

The second aim is to test whether a bilingual's L2 influences the processing of MWEs in the L1, that is, whether CLI in MWE processing occurs in the reverse direction, from the weaker L2 to the stronger L1. CLI in the direction of L2→L1 has received much less attention in the literature than in the L1→L2 direction. By investigating cross-language influences in the processing of MWEs in both directions, the present thesis attempts to address an important gap in the literature. It raises an important question regarding CLI, namely, the influence of directionality. It will be investigated (1) whether bilinguals show a greater processing advantage for congruent L1 MWEs relative to incongruent L1-only MWEs, (2) and if so, whether L2 influence can extend to the processing of translated L2-only MWEs (i.e., MWEs that do not exist in the L1). Employing the same paradigm and testing materials as above, Experiment 2 will investigate how L2 influences L1 MWE processing with a different group of bilinguals.

Finally, by investigating CLIs in MWE processing in both directions, I will be able to directly compare CLI in the direction of L1→L2 and CLI in the reverse direction, L2→L1, and to test whether CLIs in the different directions are equally strong. This will advance our understanding of the contexts where CLI is enhanced or decreased. It has been found that L1 typically has a higher impact on L2 processing than vice versa (Jiang, 1999; Keatley et al., 1994; Schoonbaert et al., 2009). It remains to be investigated whether the CLI asymmetry extends to units above the word level. In other words, the present thesis aims to investigate how directionality modulates CLI in MWE processing. I am interested in investigating whether CLI in the L2→L1 direction, if any, is weaker than that in the L1→L2 direction.

1.3 Structure of the Thesis

Chapter 2 introduces MWEs, their types, and key properties, such as frequency, familiarity, predictability, fixedness, and phonological properties. It also introduces major approaches to identifying MWEs, such as phraseological and frequency-based approaches.

Chapter 3 reports on previous research on MWE processing in an L1 and L2, focusing on the role of the following factors: frequency, familiarity, and predictability. After broadly classifying MWEs into figurative non-compositional and literal compositional, I first review studies with non-compositional MWEs in an L1 as well as L2, and then I review studies on compositional MWE processing by L1 and L2 speakers.

This is followed by a detailed discussion of the factor that has been found to affect MWE processing in the L2 – cross-language overlap or congruency. Chapter 4 reviews the existing studies on L1 influence in L2 MWE processing, pointing out the common themes and discrepancies among these studies.

Chapter 5 presents an empirical investigation of CLI in L2 MWE processing. It focuses on the comprehension of binomial expressions by Chinese-English bilinguals (Experiment 1a) and English monolingual controls (Experiment 1b). Using a primed lexical decision task, I look at the processing of three types of binomial expressions versus matched control phrases, including congruent (*sun and moon*), incongruent L2-only (*bread and butter*), and translated L1-only binomials (*wisdom and strength*). Chapter 6 looks at CLI in L1 MWE processing. In Experiment 2, using the same experimental design, I test the involvement of L2 in the processing of L1 binomials.

In Chapter 7, I consider the findings of the three experiments and the theoretical implications of my findings with respect to the literature. The general discussion focuses on two themes, CLI in the direction of L1→L2 and CLI in the reverse direction, L2→L1. The underlying mechanisms behind CLIs in MWE processing are discussed. Finally, Chapter 8 provides a summary of the thesis, focusing on the findings of the present study, limitations, and directions for future work.

Chapter 2 Multi-Word Expressions (MWEs)

2.1 Terminology and Definitions

Natural language has plenty of (semi-) fixed expressions which are made up of at least two words (collocations as, for instance, *strong tea*, binomials as *bride and groom*, idioms as *kick the bucket*, etc.). These word clusters or word combinations, often termed multi-word expressions (MWEs), are "over-learned, literal and non-literal sequences of words whose representations are stored in semantic memory" (Cacciari, 2014, p. 267). MWEs have been defined as "fixed and recurrent pattern of lexical material sanctioned by usage" (Grant & Bauer, 2004, p. 38), and as "familiar phrases that exhibit a certain degree of fixedness and are recognized as conventional by a native speaker" (Siyanova-Chanturia, 2013, p. 246). Namely, high levels of frequency, familiarity and fixedness characterise MWEs. From a probabilistic perspective, they are combinations of words that co-occur more often than would be expected by chance alone (Manning & Schutze, 1999; Siyanova-Chanturia & Martinez, 2015). Despite the potentially infinitive creativity of language, a large amount of natural language is formulaic, automatic and rehearsed, rather than completely novel and newly assembled on each utterance.

MWEs are pervasive in natural language. Biber et al. (1999) reported that multi-word speech constituted 28% of the spoken and 20% of the written discourse they analysed. Erman and Warren (2000) and Howarth (1998b) estimated that multi-word speech of various types amounted to 52.3% and 40%, respectively, of the written discourse they looked at. Such pervasiveness of MWEs lends itself well to an extensive study of MWEs from various perspectives such as theoretical, applied and corpus linguistic, psycholinguistic, neurolinguistic perspective, and so on.

MWEs have traditionally been studied in a wide range of linguistic subdisciplines, such as linguistics, applied linguistics, and others. Each field has its own terminologies and key issues and topics in the examination of MWEs (Siyanova-Chanturia & Omidian, 2020; Wray, 2002). Thus, MWEs have also been known under different names, such as "automatic language", "chunks", "collocations", "conventionalized forms", "fixed expressions", "formulae", "idioms", "idiomatic forms of expressions", "lexicalised sentence stems", "multiword units", "prefabricated word sequences", "phraseme", "phrasal expressions", "lexical phrases", "word combinations", and so on (for a summary of terms used to describe MWEs, see Wray, 2002).

In the present study, I chose one of the most inclusive terms, MWEs, as an umbrella term to refer to strings of language beyond the word level. Another commonly used umbrella term is *formulaic language* (and by extension, *formulaic sequences* referring to individual instances of formulaic language). Although both terms refer to strings of language beyond the word level, formulaic language is more inclusive than MWEs in that it also comprises single-word items, such as expletives and exclamations (*damn*, *hurrah*), speech formulas (*yeah*, *bye*), and so on (Siyanova-Chanturia & Pellicer-Sánchez, 2019; Van Lancker Sidtis & Rallon, 2004). In the present study, I chose MWEs as an umbrella term to refer to the general linguistic phenomenon and individual instances of strings of language beyond the word level.

2.2 Properties of MWEs

MWEs are heterogeneous, consisting of a large set of expression types, such as idioms (*kick the bucket*), lexical bundles (*in the middle of*), binomials (*bride and groom*), collocations (*strong tea*), and other phrasal elements (Siyanova-Chanturia & Van Lancker Sidtis, 2019). They differ in a lot of aspects such as compositionality, figurativeness, syntactic fixedness, etc. However, they share several important features when compared to novel, creative language, such as high degrees of frequency, familiarity, predictability, fixedness, and so on.

2.2.1 Frequency

Frequency is one of the most prominent *statistical* properties of MWEs (Siyanova-Chanturia & Omidian, 2020). It plays an important role in how MWEs are defined and identified. As opposed to novel language, speakers tend to use the same MWEs over and over again in natural language (Cowie, 1998; Sinclair, 1991). Thus, this repetition in usage is a key characteristic that distinguishes MWEs from novel language – which has low frequency. Some researchers even go as far as to define some categories of MWEs (such as formulas) purely based on this property. For example, De Cock (1998) and Ellis (2012) defined formulas as *recurrent* word combinations. Moreover, frequency is often used as an important criterion to identify MWEs from corpora. For example, Biber et al. (2004) took a frequency-driven approach to the identification of lexical bundles (*do you want to*) by setting the frequency cut-off of 40 times per million words. This solely frequency-based approach, or "lexical bundle approach" (Ellis, 2012, p. 27), provides a straightforward measure to identify recurrent sequences.

However, not all MWEs are of particularly high frequency (Moon, 1998). Some types of MWEs such as proverbs (*in for a penny, in for a pound*) and idioms (*kick the bucket*) have low frequency. They are highly salient and are thus easily recognized and retrieved, due to uniqueness of their function and/or non-compositionality of meaning (Huang, 2009; Wray, 2012). In addition, not all high-frequency recurrent word sequences are MWEs, because they do not have communicative functions or distinctive meanings, such as *and of the,* or *but it is* (examples taken from Ellis, 2012). Thus a "function-first" approach is proposed to identify MWEs as the recurrent linguistic units associated with various communicative functions (e.g., Durrant & Mathews-Aydınlı, 2011). In summary, frequency in itself is insufficient to indicate the formulaic status of all MWEs.

Nevertheless, although frequency does not equate with formulaicity, it contributes greatly to the formulaic status of a MWE. First, frequency plays a central role in MWE processing (Conklin, 2020). Over and above other factors, frequency of occurrence accounts for much of the processing advantage for MWEs over novel language (Carrol & Conklin, 2020). For example, Carrol and Conklin (2020) compared the processing of three subtypes of MWEs – idioms, binomials, and collocations, and they found that overall frequency can explain the advantage for all three types of MWEs, although other factors also affect the processing of each specific type: familiarity and decomposability for idioms; predictability and semantic association for binomials; and mutual information for collocations. Indeed, there is a large amount of evidence attesting the importance of frequency effects in MWE processing (Arnon & Cohen Priva, 2013; Arnon & Snider, 2010; Bannard & Matthews, 2008; Carrol & Conklin, 2017; Durrant & Doherty, 2010; Jacobs et al., 2016; Siyanova-Chanturia, Conklin, & Van Heuven, 2011; Tremblay et al., 2011; Vilkaitė, 2016). A detailed review of research on frequency effects in the processing of larger units (collocations, binomials, lexical bundles, phrases, etc.) is provided in Chapter 3.

Second, frequency conditions chunking processes where the frequently co-occurring individual parts fuse into larger units (Bybee & Scheibman, 1999). As Bybee (2006) put it, "words used together fuse together". Formulaic word strings arise through fusion caused by frequent use (Bybee, 2010). In other words, frequency can affect the relationship between the parts and the whole for MWEs. Sosa and MacFarlane (2002) found that reaction times to the target word *of* in high frequency collocations (*kind of*, *sort of*) were significantly slower than in lower frequency ones, which indicates that highly frequent collocations may be chunked into single units, and thus access to the constituent *of* may be impeded. In addition, Kapatsinski and Radicke (2009) found that the detection of the particle *up* in extremely high and low frequency *verb* + *up* collocations (*sign up*, *run up*) was slower than in medium-frequency ones. This finding confirms the proposal that highly frequent co-occurring words can fuse into larger units in the lexicon, and the fusion makes it harder to detect the constituent. It needs to be noted that it does not mean that MWEs become unanalyzed, holistic units where the parts are no longer available. Instead, it means that the effect of the parts on MWEs processing gradually weakens, while the effect of the whole increases (Arnon & Cohen Priva, 2014; Bybee, 2002).

Overall, phrase frequency is one of the most salient and determining characteristics of MWEs. It plays a key role in defining and identifying MWEs. It is a key driver of the processing advantage for MWEs over novel language. It can affect the relationship between the parts and the whole and render difficult the detection of the constituent word. With increase of frequency in use, the whole grows more prominent than the parts, and eventually the components get chunked together and fuse into larger units.

2.2.2 Familiarity

In addition to frequency, familiarity is another important property of MWEs. It is highly correlated with, but is different from, frequency. A MWE can become familiar to language users as a function of frequency. The more frequent a MWE is, the more likely it is to become familiar to language users, and subsequently become represented in their mental lexicon (Siyanova-Chanturia, 2010). However, frequency and familiarity reflect different properties of MWEs, and differ in their link with formulaicity. While frequency indicates the *statistical* distribution of MWEs, familiarity reflects the *psychological* reality in language users of MWEs, that is, the state of being stored in memory and personally known to speakers (Van Lancker Sidtis, 2004). In addition, frequency can be measured objectively by referring to large reference corpora, while familiarity is usually assessed with subjective norm ratings.

More importantly, familiarity is a property of all MWEs (Hallin & Van Lancker Sidtis, 2017). Although not all MWEs are frequent, they are all familiar to competent language speakers. For example, the proverb '*never too late to mend…*' has frequency of 1 in the Corpus of Contemporary American English (COCA), but it is no doubt well known to any L1 speaker of English. As pointed out by Siyanova-Chanturia and Van Lancker Sidtis (2019, p. 38), "Formulaic sequences are by definition familiar phrases; many – although not all – also enjoy high frequencies of occurrence."

Familiarity plays a particularly important role in the processing of MWEs which are formulaic and familiar to the linguistic community but are, nevertheless, rather infrequent, such as idioms and proverbs. In comparison with other types of MWEs, such as lexical bundles (*in the middle of the*) and conversational speech formulas (*How are you?*), idioms (*kick the bucket*) and proverbs (*a rolling stone gathers no moss*) are relatively infrequent in natural language (Hallin & Van Lancker Sidtis, 2017). Nevertheless, they still exhibit a processing advantage relative to novel, newly computed language due to subjective familiarity (Carrol et al., 2016). For example, Hallin and Van Lancker Sidtis (2017) investigated the production of low frequency Swedish proverbs in L1 speakers, whose frequencies were between 0 and 0.3 occurrences per million words in a reference corpus. Despite low frequency, all participants reported that they had heard and could recognize them. They found that adult L1 speakers produced proverbs faster and with less stressed tonal patterns (i.e., uniform pattern) than novel matched control sentences. This finding was taken as the evidence for the dual-process model (Van Lancker Sidtis, 2012), whereby MWEs are processed holistically while novel language is processed analytically.

2.2.3 Predictability

As mentioned above, MWEs are familiar to a language community, and many of them are also highly frequent. Familiarity with MWEs and/or high frequency can make later parts of MWEs become highly predictable given earlier parts of MWEs (Siyanova-Chanturia & Omidian, 2020). That is, the reader can predict the upcoming word(s) by reading the initial constituent(s) of a MWE. For instance, once reading the initial constituent(s) of an idiom (*spill the* …), a proverb (*an apple a day keeps the doctor* …), or a binomial (*bride and* …), a proficient language user is very likely to come up with the correct completion (i.e., *bean*, *away*, *groom*, respectively).

The predictability of a word driven by it occurring in a highly conventional string of language (the word *chips* following *fish and* …) is generally measured in two ways. One can be established subjectively by asking participants to supply the word that best completes each phrase via a cloze task (e.g., Siyanova-Chanturia, Conklin, & Van Heuven, 2011). The percentage of participants who provide the correct completion is taken as the probability of the word (i.e. close probability/predictability). Alternatively, predictability can be measured objectively using a corpus. Predictability in this case is generally conceived as contingencies between words, or word-to-word contingency, often referred to as transitional probability (Carrol & Conklin, 2020; Gregory et al., 1999; McDonald & Shillcock, 2003b). For example, Carrol and Conklin (2020) determined the predictability of "*beans*" in the idiomatic structure "*spill the* …" by comparing the number of times the sequence "*spill the beans*" occurs in the corpus to the number of times that "*spill the*" is followed by a different word. Thus, the transitional probability of "beans" is calculated as: *spill the beans* (overall phrase frequency: 39) ÷ *spill the* (frequency of the initial constituents: 93) × 100 = 42%.

The predictability of MWEs, in addition to their frequency, defines how MWEs are processed in the brain (Siyanova-Chanturia et al., 2017). In general, predictability facilitates comprehension and production. More predictable multiword sequences have decreased processing times, and in reading tasks, the final constituent(s) of MWEs are more likely to be skipped altogether (Conklin, 2020). According to probabilistic models of language, the statistical information implicit in language input is represented in a speaker's mind through language exposure, that is, frequency (Gregory et al., 1999; McDonald & Shillcock, 2003b). This information includes not only a word's frequency of occurrence itself, but also the statistical likelihoods of it occurring in a specific context (i.e., the probability of the occurrence of *Word n* following *Word n-1*, or preceding *Word n+1*). The language processor is able to draw on the statistical information to estimate the lexical probabilities of upcoming words (McDonald & Shillcock, 2003b; Siyanova-Chanturia, Conklin, & Van Heuven, 2011; Underwood et al., 2004). For example, using eye-movements for phrases in identical neutral sentences, McDonald and Shillcock (2003a) compared the reading of verb-noun sequences which differed in their transitional probability (e.g., high probability – *avoid confusion*; low probability – *avoid discovery*), but were matched in the frequency of the (constituent) nouns (e.g., *confusion* versus *discovery*). They found that verb-noun combinations with a high transitional probability were read faster than pairs with a low transitional probability, indicating the facilitatory effect of probability in language comprehension (McDonald & Shillcock, 2003a).

Not only are highly predictable MWEs processed faster than matched novel controls or less predictable MWEs, but also distinct mechanisms may be involved in the processing of highly (even uniquely) predictable MWEs, where *Word n* is uniquely predicted given the occurrence of *Word n-1*, versus novel sequences (Molinaro & Carreiras, 2010; Siyanova-Chanturia et al., 2017; Vespignani et al., 2010). For example, in a neurolinguistic study investigating the comprehension of Italian idioms, Vespignani et al. (2010) found that the processing of highly expected words in idioms, where the prediction is based on the knowledge of idioms, is different from the processing of highly expected words in literal compositional sentences, where the prediction is based on context- and sentence-level information. They proposed that a categorical template matching mechanism specifically underpins the processing of idioms, whereby the beginning fragment of an idiom is matched to the stored idiom template (i.e., idiom configuration) in the mental lexicon. If the expected idiom template mismatches the actual unfolding constituent, the reader needs to revise the interpretation and use more time to read it.

Similarly, in another electrophysiological study looking at the comprehension of English binomials (*knife and fork*), Siyanova-Chanturia and colleagues (Siyanova-Chanturia et al., 2017) compared the ERPs elicited by the final word "*Word n+1*" when it was preceded by "*Word n + and*" and when it was preceded just by "*Word n*" without "*and*" (e.g., *knife and fork* versus *knife-fork*). They found the processing advantage (shown by larger P300s and smaller N400s) for the final word in the binomial condition with "and" but not in the violation condition without "and", suggesting the activation of a "template" in the mental lexicon.

Together, the finding of the electrophysiological differences in the processing of frequent, predictable versus novel phrases indicates that MWEs are stored in the memory of the reader as exemplars, irrespective of compositionality (e.g., figurative, non-compositional idioms and literal, compositional binomial expressions). As a result, the human processor can draw on these stored mental templates (i.e., exemplars) to predict the upcoming information, resulting in a reduced processing load (Siyanova-Chanturia et al., 2017).

2.2.4 Fixedness and Insertion

In the above, I have reviewed the statistical (i.e., frequency) and psycholinguistic (i.e., familiarity and predictability) properties of MWEs, which are, by and large, experience-based. In this section, I will describe *formal and structural* properties of MWEs.

One of the formal and structural properties of MWEs is fixedness. It is often assumed that MWEs show a degree of fixedness in word order and specific lexical composition (Moon, 1998; Schmitt, 2004; Siyanova-Chanturia & Omidian, 2020; Wray, 2002). Fixedness is an important feature of MWEs, such that some researchers even use "fixed expressions" as an umbrella term to refer to all MWEs (Sprenger et al., 2006). As compared to novel, newly computed language which is characterised by full syntactic and lexical flexibility, MWEs are typically rather fixed and constrained in their syntactic and lexical flexibility (Van Lancker Sidtis, 2012). For example, no changes are permitted in fully fixed (i.e., frozen) MWEs without the phrase losing its original meaning (*kick the bucket* vs. *the bucket was kicked*) or its MWE status (*bride and groom* vs. *groom and bride*). However, other semi-fixed MWEs allow some degree of variation or syntactic and lexical flexibility, such as insertions (*provide some of the information*), adjectival modification (*make a good impression*), passivization (*spilled the beans → the beans were spilled*), pluralization (*red herrings*[s]), change in word order ([*you can't] teach an old dog new tricks → teach new tricks to an old dog*), and so on. Therefore, MWEs are often characterized as lying along a continuum of fixedness (Molinaro et al., 2013; Wray, 2002), which shows the heterogeneity of MWEs (Howarth, 1998a).

Fixedness contributes greatly to the formulaic status of MWEs. In the case of idioms, it is often argued that it is the structural or lexical invariance that is more associated with the formulaicity of idioms than the frequency with which such expressions are encountered in language (Cowie, 1998). The lexical invariance or structural properties of idioms, rather than frequency, are assumed to determine a speaker's judgement of how familiar they are with these expressions (Libben & Titone, 2008). For example, Reuterskiöld and Van Lancker Sidtis (2013) showed that children were significantly better at recognizing (low-frequency) idioms than matched novel (non-figurative) phrases after only one exposure, suggesting that there must be something inherently salient or noticeable in idioms (e.g., inherent holistic characteristic). Additionally, fixedness also plays an important role in the formulaic status of other subsets of MWEs. For instance, binomials (A and Y, e.g., *bride and groom*) are sequences where a specific word order is highly preferred to the reversed form (Carrol & Conklin, 2020).

The issue of how the (relative) fixedness of MWEs affects their processing has been of interest to researchers. Is it possible to argue that, due to their fixedness, MWEs are processed as a holistic unit and not subject to syntactic processes? In other words, do regular decompositional analyses take place in the processing of MWEs? As the most prototypical of all fixed expressions, idioms have, in particular, received attention in the literature (Konopka & Bock, 2009; Snider & Arnon, 2012). One proposition is that idioms are stored and processed as unanalysed wholes, or as single morphemes (e.g., Underwood et al., 2004). However, current empirical evidence is against this proposition, and supports the idea that idioms have internal structure and are governed by regular decompositional analyses (Konopka & Bock, 2009; Snider & Arnon, 2012; Sprenger et al., 2006). For example, idioms can be successfully primed by one of the constituent words, and their literal word meanings are activated during the production of idioms (e.g., Sprenger et al., 2006).

In sum, fixedness is an important aspect of MWEs. Fixedness plays a crucial role in the formulaic status of MWEs. Nevertheless, it is important to note that MWEs vary along the continuum of fixedness, and modifications to the internal structure of some MWEs is acceptable. Despite their relative fixedness, MWEs still undergo analytical processes and maintain their internal structure during the on-line processing.

2.2.5 Phonology

Apart from the above differences between MWEs and novel (newly computed) language in terms of frequency, familiarity, predictability, and (syntactic and lexical) fixedness, MWEs also differ from novel language in articulatory patterns, such as articulatory duration, (sound) reduction, stress, pauses, etc. L1 listeners can draw on the prosodic contrasts alone to identify sentences as either literal or idiomatic, which are observed in different languages, such as English, French, Korean, and so on (Van Lancker & Canter, 1981; Yang et al., 2015).

The difference in articulatory characteristics between MWEs and novel expressions was noted a long time ago (see Lieberman, 1963). Van Lancker and Canter (1981) showed that L1 listeners of American English could correctly discriminate whether the intended meaning of a spoken sentence containing a ditropic idiom (which may carry either a literal or an idiomatic meaning, e.g., *at the end of your rope*) is literal or idiomatic. Acoustic analyses in a later study found that literal utterances have longer word and phrase durations, more inter-word pauses, more pitch contours (discernible rise-fall excursions of fundamental frequency; i.e., more changes in pitch) and open junctures than idiomatic utterances (Van Lancker et al., 1981). In addition, they also found that lexical and phrase boundaries were more strongly marked, and that the constituent words are more salient in literal than idiomatic sentences, reflecting the holistic/analytic distinction between idiomatic and literal sentences. Recently, in a reading-aloud experiment with L1 speakers, Siyanova‐Chanturia and Lin (2018) found that English idioms were articulated faster than controls, and figurative meanings were articulated faster than their literal counterparts.

Similar studies have been conducted in other languages. L1 listeners of other languages were also able to discriminate between idiomatic and literal exemplars of ambiguous sentences containing ditropic idioms. The prosodic contrasts between idiomatic and novel expressions seem to be universal, although they are different for different languages. For example, Yang et al. (2015) found that in Korean, idiomatic sentences were characterized by greater variation in intensity, shorter duration, and greater variation in syllable duration compared to literal sentences. The phonological prosodic contrasts between idiomatic and novel expressions may thus be universal, although they are realised differently for different languages.

While the above studies focus on prosodic contrasts between idiomatic and literal expressions, the production of compositional MWEs has also been found to be different from novel language. Bybee and Scheibman (1999) found that the word *don't* was more likely to be reduced when used in a frequent three-unit phrase (e.g., *I don't know*, *I don't think*) than in an infrequent phrase (e.g., *we don't see*). This finding shows that individual elements that are frequently used together show a tighter constituent structure (i.e., become chunked) than those that co-occur less frequently. Bybee (2001, 2002) found that high-frequency phrases are more likely to undergo phonological reduction. Bell et al. (2003) showed that frequent function words (*the*, *that*, *and*, and *of*) are more likely to be shorter and reduced when they are more predictable given the surrounding words or in predictable multiword collocations. For example, they found that *and* was significantly shorter when it was in frequent phrases (e.g., *trucks and stuff*; *lockers and everything*) than in other occurrences (e.g., *And, and I get mail*). Recent production studies have showed similar findings. Frequent n-grams are produced (articulated) faster than infrequent controls by adult speakers, such as trigrams (Arnon & Cohen Priva, 2013; Arnon & Cohen Priva, 2014), binomials (Siyanova-Chanturia & Janssen, 2018), four-word sequences (Ellis et al., 2008, Experiment 2; Tremblay & Tucker, 2011).

In sum, MWEs have been found to have articulatory patterns distinct from novel sequences. In general, MWEs, at least in English, are articulated faster, have more reduction and less inter-word pauses than novel controls. These findings provide further support for the psycholinguistic distinction between MWEs and novel expressions.

2.3 Identification of MWEs

The multifarious nature of MWEs is also reflected in the different procedures used to identify MWEs in language (Boers, 2020; Wood, 2020). There are two major approaches to identifying MWEs: phraseological and frequency-based approaches. The phraseological approach existed long before the advent of corpus linguistics, whereas the frequency-based approach has become increasingly informed by corpus data (Boers, 2020).

In the phraseological approach, MWEs are identified more on typological grounds, with degree of semantic compositionality and structural fixedness as guiding principles (Gyllstad & Wolter, 2016). On the semantic grounds, if the meaning of a given word string can be inferred straightforwardly by adding up the meanings of the constituent words, then the word string is considered to be "compositional". If not, then it is considered to be "non-compositional" (Boers, 2020, p. 144). Idioms are traditionally identified using this approach. Within the class of idioms we find gradation in transparency, in that some idioms may be interpretable thanks to cultural background knowledge (e.g., *break the ice*), while others (e.g., *by and large*) are truly opaque (Grant & Bauer, 2004). One of the attempts to identify MWEs in the phraseological tradition is Howarth's Continuum Model (Howarth, 1998a), which classified MWEs into four categories: free combinations, restricted collocations, figurative idioms, and pure idioms. Free combinations, such as *thank you*, are word combinations in which the lexical elements are used in the literal sense. Restricted collocations, such as *fast food*, are word combinations in which one of the component words is used in a figurative sense, whereas the other constituent(s) are used literally. In addition, substituting one constitute with another word carrying the same meaning (*quick food*) will render the expression unnatural or simply wrong. Figurative idioms (or ambiguous idioms) can appear both with a holistic metaphorical meaning as well as a literal meaning, such as *at the end of day* (literal: in the evening; figurative: eventually), but pure idioms are truly non-compositional and opaque (*by and large*).

With the understanding that formulaic sequences such as *fast food* and *thank you* can also be regular in their form and meaning and that any quantity of our language could be formulaic, we need a procedure to tell apart formulaic and newly generated language that will look identical (Wray, 2000). The frequency-based approach to identifying MWEs is built upon the assumption that MWEs are more frequent than other word strings and that frequency is a central definitional criterion of MWEs (Wray, 2000). One procedure is to screen a corpus for highly frequent *uninterrupted* word strings, the so-called n-grams such as ‘in the middle of’ (Boers, 2020, p. 144). Word strings that meet a certain frequency criterion stipulated by researchers are customarily labelled “lexical bundles” such as *for instance, as soon as*, and *one of the* (Biber & Barbieri, 2007; Biber et al., 2004). For example, Biber et al. (2004) took a frequency-driven approach to the identification of lexical bundles (*do you want to*) by setting the frequency cut-off of 40 times per million words.

However, frequency counts may not be a reliable means of differentiating formulaic and non-formulaic language (or semantically transparent and grammatically regular strings versus novel strings), in that some MWEs are actually not very frequent in discourse (e.g., *long live the King*; *All for one and one for all*), and some frequent sequences are not formulaic, such as *that there is a* (Moon, 1998; Siyanova-Chanturia & Omidian, 2020). Frequency aside, another corpus-based procedure is to look for strong word partnerships, i.e., frequent co-occurrences of content words regardless of whether they are immediately adjacent (Boers, 2020). Word strings that occur together more frequently than would be expected by chance (i.e., the above-chance co-occurrence) such as *strong coffee* are commonly called “collocations” (Sinclair, 1991). Collocations identified in this approach refer to a very general category of MWEs that comprise many linguistic phenomena, such as idioms, binomials, cliches, and so on (Fellbaum, 2007). The strength of co-occurrence between two words that form a collocation is often determined on the basis of mutual information (MI), which reflects the extent to which two words seek each other’s company rather than the company of other words (Boers, 2020). Typically, a MI score of 3 is taken as the threshold above which a word pair can be considered as a potential collocation (Hunston, 2002). For example, the combination *tell* + *joke* should have high frequency of occurrence as well as a high MI score in a representative corpus. However, word strings consisting of words with low frequency (e.g., *wreak havoc*) can also yield very high MI score. Therefore, when identifying collocations using corpus-based procedures, frequency thresholds as well as collocational strength should both be considered (Martinez & Schmitt, 2012; Simpson-Vlach & Ellis, 2010).

2.4 Conclusion

This chapter provides a brief overview of what MWEs are. Although MWEs are extremely heterogeneous as reflected by various terms used, there are common properties shared by idioms, collocations, binomials, phrasal verbs, and other types of MWEs, such as frequency, familiarity, predictability, (semi-)fixedness, etc. A key characteristic of MWEs, however, is the processing advantage they have as compared to novel (newly created) language. I will now look in detail at how MWEs are processed relative to novel language. The following chapter will review literature on MWEs processing in L1 and L2, focusing on the mechanisms associated with MWE processing in L1 and L2.

Chapter 3 The On-Line Processing of MWEs

3.1 Introduction

As mentioned in the above sections, MWEs are extremely heterogeneous, varying in size and complexity, compositionality, fixedness, etc. They can be short (*strong tea*) or long (*you can't judge a book by its cover*); literal (*bride and groom*) or figurative (*break the ice*); fixed (*kick the bucket*) or relatively flexible (*spilled the beans → the beans were spilled*), etc. Nevertheless, what MWEs have in common is that: (1) they are recurrent, in the sense that they occur in natural language more frequently than novel (i.e., newly created) phrases; (2) proficient speakers are familiar with them (Carrol & Conklin, 2020; Siyanova-Chanturia & Van Lancker Sidtis, 2019). Frequency and familiarity have far-reaching implications for how MWEs are processed relative to novel language, with MWEs showing a processing advantage relative to matched control novel phrases. In addition, in line with usage-based accounts of how language develops and is organized (Bybee, 2006; Goldberg, 2006; Tomasello, 2003), recurrent MWEs should be processed faster compared to control novel phrases. In the past decades, there have been numerous studies that have shown that MWEs, compositional (literal) or non-compositional (figurative), are indeed processed differently from novel strings of language (Siyanova-Chanturia & Van Lancker Sidtis, 2019). These studies generally compare participants' decision/response to, reading or production of MWEs and matched novel language in various comprehension and production tasks. Overall, they have demonstrated that L1 speakers, as well as proficient L2 speakers, are sensitive to phrase frequency and formulaicity of MWEs. In other words, MWEs enjoy a processing advantage over novel language.

To explore the processing of MWEs in L1 and L2 speakers, the *online* production and comprehension of MWEs have been investigated employing a wealth of methodologies, tasks, and paradigms. "On-line" processing means that language processing happens under time pressure with no advance preparation, contrasting with "off-line" processing where participants have time to reflect on it. In what follows, I first give a very brief review of the *production* of MWEs and then focus on the *comprehension* of MWEs, not least because the present study focuses on the comprehension, and not production.

3.1.1 On-Line Production of MWEs

From a production (articulation) perspective, compositional MWEs have been found to have distinct articulatory patterns from infrequent sequences, such as shorter production duration, in spontaneous and elicited speech. For example, words in frequent and thus predictable phrasal contexts such as *I <u>don't</u> know* are more likely to be phonetically reduced (e.g., Bell et al., 2003; Bybee & Scheibman, 1999); MWEs are more likely to be produced as a single intonation unit (e.g., Lin, 2010); frequent n-grams are produced (articulated) faster than infrequent controls by adult speakers, such as trigrams (Arnon & Cohen Priva, 2013; Arnon & Cohen Priva, 2014), binomials (Siyanova-Chanturia & Janssen, 2018), and four-word sequences (Ellis et al., 2008, Experiment 2; Tremblay & Tucker, 2011); frequent phrases are repeated faster and more accurately than less frequent controls by young children (Bannard & Matthews, 2008). Further, these phrase frequency effects in production have also been found in languages other than English (Janssen & Barber, 2012). However, unlike L1 speakers, L2 speakers have not always been reported to have shorter articulatory durations for MWEs than infrequent control phrases (Siyanova-Chanturia & Janssen, 2018). There have been a very small number of production studies with L2 speakers, which makes it difficult to come to any meaningful conclusions.

Akin to literal, compositional MWEs, figurative MWEs – idioms and proverbs – have also been found to show shorter articulatory durations and distinctive prosodic characteristics versus novel language (for a review, see Siyanova-Chanturia & Van Lancker Sidtis, 2019). For example, idioms are articulated faster than controls (Siyanova-Chanturia & Lin, 2018), the utterances conveying the literal meaning have longer durations, more pausing, and greater numbers of pitch contours (i.e., more changes in pitches) versus the identical utterances conveying the figurative meaning in English language (Van Lancker et al., 1981) and in languages other than English such as Korean (Yang et al., 2015). L1 speakers can draw on these prosodic cues to judge whether the target utterances are intended to be figurative or literal better than proficient L2 speakers, while English as a second language (ESL) learners may not be able to do so (Van Lancker Sidtis, 2003). In addition to idioms, proverbs have also been shown to have shorter production duration and more tonal patterns than matched novel sequences in L1 adult and child speakers (Hallin & Van Lancker Sidtis, 2017).

3.1.2 On-Line Comprehension of MWEs

As can be seen from the above, the majority of the studies on the production of MWEs have been conducted with L1 speakers, with little research done with L2 speakers. Most of the current research on L2 speakers comes from language comprehension. Crucial evidence from a comprehension perspective suggests that L1 speakers, as well as proficient L2 speakers, process MWEs differently from novel language. Therefore, in what follows, I cover the studies within the *comprehension* domain in a first and second language. Following the literature (e.g., Siyanova-Chanturia & Van Lancker Sidtis, 2019), MWEs may be broadly classified into figurative non-compositional and literal compositional. Figurative, non-compositional MWEs such as idioms possess highly idiosyncratic semantics that literal, compositional language does not have (Siyanova-Chanturia & Van Lancker Sidtis, 2019).

Thus, I first review studies with respect to non-compositional MWEs and then compositional ones.

It is noteworthy that the broad division of MWEs into compositional and non-compositional does not mean that these two categories of MWEs are necessarily processed differently. On the contrary, they have in common the fact that they, as recurrent sequences, are processed faster compared to novel control phrases (Carrol & Conklin, 2020). In addition, in line with a dual-process processing model (Van Lancker Sidtis, 2012), all MWEs are processed in the same way, i.e., direct access/retrieval, compared to novel language which is computed at the time of usage. Familiarity – whether an item was known or unknown – is regarded as the key driver of MWE processing (Tabossi et al., 2009; Van Lancker Sidtis, 2012). It overshadows other factors such as transparency, compositionality, frozenness and so on, leading to faster processing of recurrent sequences compared to novel control phrases (Van Lancker Sidtis, 2012). Furthermore, as we will see below, a similar pattern of results was observed with respect to the brain's electrophysiological response to compositional and non-compositional MWEs (Siyanova-Chanturia et al., 2017; Vespignani et al., 2010), providing more evidence that highly conventional language, both compositional and non-compositional, may be processed in a comparable way.

3.2 Processing of Non-compositional MWEs

As the most representative members of MWEs, idiomatic expressions such as *break the ice*, *bark up the wrong tree*, have long been studied in psycholinguistics and neurolinguistics (Cacciari, 2014). Idioms have conventional figurative meanings which may not derive from the meanings of the component words such as *kick the bucket* 'die' (Swinney & Cutler, 1979). Therefore, most idioms have the characteristic of ambiguity (Swinney & Cutler, 1979). That is, they convey an acceptable literal as well as figurative meaning, such as *at the end of the day* (literal: 'in the evening'; figurative: 'eventually'). How idioms are processed and stored in the mental lexicon has drawn a fair amount of attention. Research into idiom processing typically compares how L1 speakers, as well as L2 speakers, process the figurative versus literal meaning of an ambiguous idiom, and idiomatic expressions versus novel phrases (Siyanova-Chanturia, Conklin, & Schmitt, 2011; Siyanova-Chanturia & Van Lancker Sidtis, 2019). In what follows, I first cover studies on idiom processing in the L1 and then in the L2.

3.2.1 Processing of Non-compositional MWEs in L1 Speakers

A range of theories has been put forth to account for idiom comprehension, such as the idiom list hypothesis (Bobrow & Bell, 1973), the lexical representation hypothesis (Swinney & Cutler, 1979), the direct access hypothesis (Gibbs, 1980, 1986), and the configuration hypothesis (Cacciari & Tabossi, 1988). One of the prominent theories of idiom comprehension is the lexical representation hypothesis by Swinney and Cutler (1979). In contrast to the idiom list hypothesis (Bobrow & Bell, 1973), this hypothesis holds that idioms are stored and retrieved from the lexicon in the same manner as any other word (e.g., compounds, *hotdog*), and that there is no special idiom list (which is not part of the normal lexicon) or any special idiom processing mode when literal analysis fails. According to the lexical representation hypothesis, the retrieval of the figurative meaning takes place in parallel with computation of the literal meaning, as soon as the first word of an idiom string is encountered. Swinney and Cutler (1979) argued that the figurative meaning should become activated first because the retrieval of the figurative meaning is less time-consuming than computation of the literal meaning. In contrast, the direct access hypothesis by Gibbs (1980, 1986) claims that only the idiomatic interpretation is initiated upon presentation of an idiom string, and that the literal interpretation starts only if the idiomatic interpretation does not fit the context.

To test the lexical representation hypothesis and the direct access hypothesis, Cacciari and Tabossi (1988) investigated the processing of idioms with three cross-modal priming experiments, where participants listened to a (non-biasing) sentence containing an idiom and then performed lexical decisions to visually presented targets at the end of the idiomatic string. In Experiment 1, they found that L1 speakers were faster at performing a lexical decision to idiom targets (a word related to the idiomatic meaning of an idiom, e.g., *in seventh heaven* → *happy*) than to literal targets (a word associated with the meaning of the last word in the string, e.g., *in seventh* <u>*heaven*</u> → *saint*). This result seemed to be consistent with the lexical representation theory and the direct access theory, both of which predicted that idiomatic meaning was activated faster than literal meaning. However, when idioms were not predictable (i.e., an idiom was not recognizable as idiomatic until its completion, e.g., *go to the devil*), participants were faster on literally related targets (i.e., *horns*) than on idiomatically related targets (*away*) (Experiment 2). Interestingly, when targets words were presented 300 ms after the end of idioms was heard (rather than right after the end of idioms as in Experiment 1 and 2), participants were faster both for literally related targets (*horns*) and for idiomatically related targets (*away*), relative to control targets (*trout*) (Experiment 3). They concluded that idioms are initially processed *only* literally (Experiment 2), and some time is required before the retrieval of the idiomatic meaning (Experiment 3); but, after idioms are recognized, any remaining lexical items in the string may not be literally processed (Experiment 1). Based on these findings, Cacciari and Tabossi (1988) proposed the Configuration Hypothesis. Under this hypothesis, idioms are not stored as separate entries in the mental lexicon; instead, they are represented as configurations of lexical items. The individual words that participate in a configuration (i.e., an idiom) are accessed until the configuration can be recognized. In this account, the computation of the literal meaning goes on until the 'idiomatic key' – a place where an idiom can be identified – has been reached, and then the idiomatic configuration becomes recognized and the idiomatic meaning is directly retrieved. At this point, the literal interpretation is rejected as no longer viable.

Empirical evidence for the Configuration Hypothesis has been reported in a number of studies (e.g., Fanari et al., 2010; Tabossi et al., 2005; Titone & Connine, 1994). For example, using a cross-modal priming paradigm, Titone and Connine (1994) investigated the influence of predictability on idiom processing. Participants heard neutral sentences containing idioms (*George wanted to* <u>*bury the hatchet*</u> *soon after Susan left*), and then made lexical decisions in response to visually presented targets (idiom target: *forgive*; control target: *gesture*). Targets were presented at two positions, at the idiom offset (Experiment 1) and at the offset of the second to last word of the idiom (penultimate position, Experiment 2). In Experiment 1, idioms with both high predictability (*bury the hatchet*) and low predictability (*hit the sack*) showed significant priming effects, suggesting no effect of predictability. However, because low-predictable idioms used in the experiment were also highly familiar, high familiarity may have contributed to the facilitation for them. To partial out the effect of familiarity, they presented visual targets at the offset of the second to last word of the idiom in Experiment 2. The results showed that high-predictable idioms showed greater priming than low-predictable idioms, suggesting that predictability does facilitate the figurative interpretation of idioms.

There is also evidence showing that different cognitive processes are involved before and after the idiom's recognition point. Using event-related potentials (ERPs), Vespignani et al. (2010) investigated the electrophysiological correlates of highly expected words in Italian idioms. To test the prediction by the Configuration Hypothesis that a qualitative change (i.e., from computation to retrieval) occurs after the idiom's recognition point (RP), they compared three types of neutral sentences: one contains a predictable Italian idiom (idiomatic condition: *Giorgio aveva un* ***buco*****(**RP**)** ***nello*****(**RP+1**)** *stomaco quella mattina*; English translation:

George had a hole in the stomach *that morning*, meaning 'George was hungry that morning', whose RP is considered to be 'hole'), one contains the same idiomatic expression but with the RP replaced (substitution condition: *Giorgio aveva un* ***delore nello*** *stomaco quella mattina*; *George had a* pain *in the stomach that morning*), and one contains the same expression but with the word after the RP replaced (violation condition: *Giorgio aveva un* ***buco sulla camicia*** *quella mattina*; *George had a hole on the* shirt *that morning*). The ERPs elicited by the idiomatic and substitution conditions at the recognition point (*hole* versus *pain*) are comparable, suggesting that at this point the fragment of an idiom is still perceived as literal. In contrast, the ERPs elicited by the idiomatic and violation conditions after the recognition point were different, with larger P300 amplitudes (an event-related potential component elicited in the process of decision making) on the word following the recognition point in the idiomatic condition (*stomach*) compared to the violation condition (*shirt*). They concluded that after recognition of an idiomatic configuration, a categorical (rather than probabilistic) prediction mechanism operates, confirmed by a P300 in idiomatic condition.

In addition, models which view idioms as both individual words and whole units, similar to the Configuration Hypothesis, are broadly referred to as hybrid models (Cacciari & Glucksberg, 1991; Cacciari & Tabossi, 1988; Cutting & Bock, 1997; Sprenger et al., 2006; Tabossi et al., 2005; Titone & Connine, 1999). In these models, idioms are "not semantically empty long words" and the meaning of individual components of idioms are accessed as well (Cacciari, 2014, p. 276). That is, idioms are described as compositional, i.e., composed of constituent parts, as well as non-compositional (or unitary), i.e., processed as whole units (Van Lancker Sidtis, 2012). The whole units, as a separate representation of the whole idiomatic phrase, are variously described as configurations (Cacciari & Tabossi, 1988), superlemmas (Sprenger et al., 2006) or formulemes (Van Lancker Sidtis, 2012). Due to specific idiomatic meaning which is usually different from the literal meaning of the component words, idioms need their own semantic entry (Bybee, 2006; Wray, 2012). Nevertheless, they are still related to the lexemes and construction from which they arise (Nunberg et al., 1994). A number of studies suggest that individual words of an idiom are activated in its processing (Konopka & Bock, 2009; Snider & Arnon, 2012; Sprenger et al., 2006). For example, idioms can be successfully primed by one of the constituent words, and their literal word meanings are activated during the production of idioms (e.g., Sprenger et al., 2006), suggesting the compositional nature of idioms.

In addition to the research on access of figurative versus literal meanings, researchers have also been interested in the processing of figurative expressions versus matched novel (i.e., literal propositional) language. Idioms (*break the ice*) have long been found to be processed more quickly than matched novel phrases (*break the cup*) by L1 speakers (Gibbs, 1980; Swinney & Cutler, 1979). However, opinions differ on whether or not decomposability of idioms contributes to the processing advantage of idioms over novel language. According to the idiom decomposition hypothesis (Gibbs et al., 1989), decomposability of idioms – the degree to which the literal meanings of individual components independently contribute to the overall figurative meaning of the idioms – plays an important role on how an idiom is processed relative to novel control phrases. Using a phrase judgement task, where participants decided if a (visually presented) word string formed a meaningful phrase, Gibbs et al. (1989) found that participants were faster at responding to decomposable idioms (*pop the question*) than to their control phrases (*ask the question*), but they were slower at responding to non-decomposable idioms (*kick the bucket*) than to their matched control strings (*fill the bucket*). They argued that the speed with which idioms are processed depends on the degree of semantic decomposition, suggesting that compositional analysis is undertaken when understanding idioms. However, empirical evidence on the role of decomposability remains mixed, with some studies reporting a processing advantage for decomposable over non-decomposable idioms (Caillies & Butcher, 2007), and other studies not (Titone & Connine, 1999). In an eye-movement by Titone and Connine (1999), participants read sentences containing decomposable (*save your skin*) and non-decomposable (*kick the bucket*) idioms. Decomposable idioms were read faster than non-decomposable idioms when disambiguating contexts preceded the idiom (e.g., <u>*Because she never took care of herself*</u>, *Carolina suddenly kicked the bucket*), whereas the two types of idioms were read at a similar speed when disambiguating contexts followed the idiom(e.g., *Carolina suddenly kicked the bucket*, <u>*because she never took care of herself*</u>), evidenced in the analysis of first pass fixation duration. This suggests that decomposability may only affect later stages of idiom processing (e.g., non-decomposable idioms take longer for readers to integrate a contextually appropriate meaning than decomposable idioms), but not initial stage of processing. According to the authors, the activation of idiomatic and literal meanings of idiomatic phrases (decomposable or non-decomposable) is mandatory, thus decomposable and non-decomposable idioms were read at a similar speed when there was no preceding context.

In contrast, the dual route model holds that all idioms are processed via the same route, i.e., direct retrieval route, as long as they are familiar expressions to speakers (Van Lancker Sidtis, 2012; Wray, 2002; Wray & Perkins, 2000). In this model, two different routes are employed in idiomatic and novel language processing: direct retrieval route for frequent, familiar phrases which are stored in long-term memory, and computation route for novel phrases using a words-and-rules approach. In this model, idioms are processed via the retrieval route, because they are all previously encountered and, thus, familiar phrases. There is no distinction in the route of processing between decomposable and non-decomposable idioms. It is subjective familiarity that ultimately determines whether or not the direct route can be available. For example, using a semantic judgement task, Tabossi et al. (2009) found that L1 Italian speakers judge both decomposable and non-decomposable idioms faster than novel control phrases, suggesting that the link between the meaning of the idiom's components and the idiom's overall figurative meaning, or lack thereof, does not affect its processing. The dual route model is consistent with the configuration hypothesis (Cacciari & Tabossi, 1988). They both hold that after an idiom becomes recognized, its idiomatic meaning is directly retrieved, making the idiom processed faster than its matched novel control, or the idiomatic meaning activated faster than the literal meaning of its components.

To sum up, idioms are both compositional and non-compositional. Compositional analysis and retrieval are both involved in idiom processing. Idioms are not directly retrieved from semantic memory without any linguistic processing (as claimed by the lexical representation theory or the direct access theory). Rather, they are processed word by word until enough information has accumulated to trigger recognition of the idiom. Only then can the idiomatic meaning be retrieved from semantic memory. Predictability, familiarity and, possibly, decomposability, together affect idiom processing.

3.2.2 Processing of Non-compositional MWEs in L2 Speakers

Similar to L1 research on idiom processing, L2 research also addresses the following issues: (1) whether there is a difference in the processing of literal versus figurative meaning of ambiguous idioms, and (2) whether there is a difference in the processing of idioms versus matched novel strings (Beck & Weber, 2016; Siyanova-Chanturia, Conklin, & Schmitt, 2011). Research with L2 speakers has found that while idiom processing is relatively easy for L1 speakers, it is not the case for L2 speakers (Conklin, 2020). With respect to the issue of the access to figurative vs. literal meaning, it has been found that the literal meaning enjoys a processing advantage over the figurative meaning (Carrol & Conklin, 2017; Cieślicka, 2006; Matlock & Heredia, 2002; Siyanova-Chanturia, Conklin, & Schmitt, 2011). However, research on the comparison of idiom processing to novel language is mixed, with the processing advantage for idioms over matched novel strings reported in some studies (Beck & Weber, 2016; Carrol et al., 2016; Conklin & Schmitt, 2008; Underwood et al., 2004), but not in others (Carrol & Conklin, 2014, 2017; Cieślicka, 2006; Cieślicka & Heredia, 2011; Siyanova-Chanturia, Conklin, & Schmitt, 2011).

Concerning activation of idioms' figurative vs. literal meanings, one view holds that L2 speakers have a fundamentally different approach to processing idioms compared to L1 speakers. While L1 speakers tend to directly retrieve the figurative meaning of idioms (Titone & Connine, 1999), L2 speakers are likely to process idioms in a fully compositional manner, whereby they analyse the individual constituents of idioms (Cieślicka, 2006, 2013; Liontas, 2002; Matlock & Heredia, 2002; Siyanova-Chanturia, Conklin, & Schmitt, 2011). For example, using the cross-modal lexical priming paradigm, Cieślicka (2006) had late L2 learners listen to non-biasing sentences that contained familiar idioms ('George wanted to *bury the hatchet* soon after Susan left.') and then perform a lexical decision task on a word associated with the idiom's figurative meaning ('forgive'), or its control ('gesture'); a word associated with the idiom's literal meaning ('axe'), or its control ('ace'). She found that target words related literally to the auditorily presented idiomatic sentences (i.e., literal targets) showed a priming advantage over target words related to the figurative meaning of these sentences (i.e., idiomatic targets), indicating that idioms' literal meanings are activated first and contribute to the processing of the figurative meanings of idioms. Following the graded salience hypothesis proposed by Giora (1997, 2003), Cieślicka (2006) put forth the Literal Salience Model which maintains that literal meanings are always most salient for L2 users, regardless of their frequency and familiarity. On this account, literal meanings of individual components are acquired earlier, better established in the mental lexicon (in terms of length of storage and completeness of representation) and are likely to remain more frequently used than idiomatic ones in an L2. Thus, literal meanings of L2 idiomatic items are always more salient, and are thus activated faster, than their figurative meanings, regardless of contextual bias or degree of familiarity. That is, the compositional analysis of idioms in L2 speakers is obligatory and prioritized.

Further evidence for the priority of literal meaning in L2 idiom processing is presented by Siyanova-Chanturia, Conklin and Schmitt (2011). Using eye-tracking, the authors found that after a preceding disambiguating story context, the figurative meanings (*at the end of day* – 'eventually') were read with the same speed as the literal ones (*at the end of day* – 'in the evening') by the L1 speakers. However, they did not find any facilitation in the reading of the figurative meanings of ambiguous idioms in L2 speakers: figurative meanings still required more reading time than literal ones, even in the presence of a preceding disambiguating context. The authors argued that when the form-meaning connection between an idiom and its figurative meaning is not as strong as the connection between the form and the meaning of the individual lexical items, the figurative meaning (*at the end of day* – 'eventually') would not be activated as quickly as the literal *(at the end of day* – 'in the evening').

Similarly, in another eye-tracking study, Carrol and Conklin (2017) found that L2 speakers (Chinese native speakers learning English) showed significant facilitation for the final words of L1-only idioms translated into the L2 (*draw a snake and add feet*) compared to a control phrase (*draw a snake and add hair*), as indexed by first fixation duration and total reading time. However, they found that figurative meanings (*add oil and vinegar* – 'to embellish a story') were still read more slowly than literal meanings of translated L1-only idioms (*add oil and vinegar* – 'to add some dressing'), just as what has been found for L2-only idioms. This suggests that the non-compositional nature of idioms makes their processing problematic in an L2, and this also extends to L1 idioms translated into the L2.

Another issue of research on idiom processing with L2 speakers concerns the comprehension of idioms vs. matched novel control strings. In one of the earliest such studies employing eye-tracking, Underwood et al. (2004) investigated whether the initial components of idioms can provide sufficient context to facilitate the processing of their terminating word in L1 and L2 speakers of English. They compared how the terminal words in idioms (*hit the nail on the head*) and the same words in nonformulaic contexts (*pick up a terrible cold in his head*) are processed when they both were embedded in a story context. They found that L1 speakers showed fewer and shorter fixations on the terminal words when they were part of an idiom than when they were embedded in non-formulaic text, indicating an effect of predictability on the number of fixations and fixation duration. Most importantly, it extended to L2 speakers of English too. L2 speakers of English also had fewer fixations on the terminal words when they were in an idiom than in the middle of a nonformulaic text. However, unlike L1 speakers, they did not show processing advantage in terms of the duration of fixations. That is, there was no significant difference in the duration of fixations when the terminal words were in an idiom than when they were not. Thus, L2 speakers only showed a partial processing advantage for idioms vs. control phrases. The authors attributed the finding that L2 speakers made fewer, but not shorter, fixations on the terminal words to the possibility that these L2 participants only partial, mastery of MWEs.

In addition, in a self-paced, line-by-line reading experiment, Conklin and Schmitt (2008) compared the processing of the figurative meanings of ambiguous idioms (e.g., *a breath of air* meaning ‘a new approach’), their literal meanings (‘breathing clean air outside’), and control phrases (*a fresh breath of some air*). They found that like L1 speakers, L2 speakers read idioms faster than controls, regardless of whether they were used figuratively or literally. However, this result was not replicated by their follow-up study employing eye-tracking (Siyanova-Chanturia, Conklin, & Schmitt, 2011). In this study, the authors found that L1 speakers showed a processing advantage for idioms (*at the end of the day*) over novel phrases (*at the end of the war*), as evidenced by fewer and shorter fixations. However, they did not find any processing advantage for idioms over novel phrases in L2 speakers. The lack of processing advantage for idioms over matched novel control phrases in L2 speakers may be due to the fact that idioms are generally of lower frequency (compared to other instances of MWEs) and are thus less likely to be encountered and acquired by L2 learners (Siyanova-Chanturia et al., 2019).

Indeed, it has been shown that if L2 speakers know the idioms, they can also demonstrate the processing advantage over matched novel control strings like L1 speakers (Conklin, 2020). A set of studies by Carrol and Conklin (2014, 2017) investigated this question. They presented English monolinguals and L1 Chinese-L2 English bilingual speakers with English-only idioms and control (*spill the beans*/*chips*) and translated Chinese-only idioms and controls (*draw a snake and add feet*/*hair*). Although bilingual speakers did not demonstrate faster processing for English-only idioms than for their controls as monolinguals did, they did demonstrate a processing advantage for translated Chinese-only idioms compared to their controls. This result suggests that (1) when L2 speakers are familiar with idioms, they do demonstrate a processing advantage over matched novel control phrases – even when the idioms are presented in the L2 (Conklin, 2020), and (2) there is no difference in the way L1 and L2 speakers process idioms.

At the electrophysiological level, the reduction of the N400 component (i.e., the amplitude of a negativity peaking around 400 ms post-stimulus onset which is highly associated with processing at the level of meaning) in the comprehension of idioms in L2 speakers has been reported. In an ERP study with highly proficient English-Spanish bilingual speakers, Moreno et al. (2002) found that expected completions (*mind*) to English idiomatic expressions (*Out of sight, out of* …) elicited reduced N400s relative to unexpected but plausible completions (*brain*). This finding indicates that high-probability completions to idioms are easier to integrate semantically. This shows that L2 speakers can demonstrate a processing advantage for idioms, similar to L1 speakers.

Based on the above findings, we can arrive at the following conclusions. First, while L1 speakers read figurative and literal meanings in a comparable way, literal meanings of ambiguous idioms are easier and faster for L2 speakers to process (Cieślicka, 2006; Siyanova-Chanturia, Conklin, & Schmitt, 2011; for a review, see Siyanova-Chanturia & Van Lancker Sidtis, 2019; Siyanova-Chanturia et al., 2019). Second, unlike L1 speakers who read idioms consistently faster than novel controls, L2 speakers may not always be able to process idioms faster than novel propositional language (Siyanova-Chanturia, Conklin, & Schmitt, 2011). However, with ample exposure to linguistic input and high level of proficiency, L2 speakers may show a native-like pattern in the processing of idioms (Beck & Weber, 2016; Carrol & Conklin, 2014, 2017; Moreno et al., 2002; Paulmann et al., 2015).

3.3 Processing of Compositional MWEs

Whereas the processing of figurative MWEs like idioms have long been of interest for researchers (Kuiper & Haggo, 1984; Pawley & Syder, 1983), not least because of their idiosyncrasy and salience (Siyanova-Chanturia & Martinez, 2015), the investigation of the processing of *compositional* MWEs, such as collocations, binomials, lexical bundles, is a relatively recent phenomenon. This is possibly due to the traditional view that formulaic language is just a peripheral phenomenon, restricted to idiomatic expressions like *kick the bucket* (Ellis et al., 2008). With the understanding that formulaic language is fundamental to the way language is used, processed, and acquired (Biber et al., 1999; Ellis et al., 2008; Martinez & Schmitt, 2012; Meunier & Granger, 2008; Nattinger & DeCarrico, 1992; Schmitt, 2004, 2010; Sinclair, 1991; Wray, 2002, 2008), the focus has gradually extended to on-line processing of literal, compositional MWEs. Research into the comprehension of compositional MWEs has by and large addressed the following issues: (1) the role of phrase frequency and predictability in the processing of frequent MWEs versus novel strings of language, (2) the role of age of acquisition in the processing of MWEs, and (3) the impact of language proficiency and L2 exposure in MWEs processing in an L2. Below, I review major findings of current research addressing the above issues in the L1 and L2.

3.3.1 Processing of Compositional MWEs in L1 Speakers

There is considerable evidence showing that L1 speakers can recognize, read and respond to MWEs significantly faster than matched novel strings of language. Behavioural studies have showed that compositional MWEs are processed faster than novel strings of language in L1 speakers, employing lexical decision task (Ellis et al., 2009) and phrasal-decision task (e.g., Arnon & Snider, 2010), self-paced reading (Tremblay & Baayen, 2010, Experiment 1; Tremblay et al., 2011), eye-tracking (Vilkaitė, 2016), and priming paradigm (Durrant & Doherty, 2010). Electrophysiological evidence further shows that MWEs are not only processed faster but also associated with easier semantic integration and template matching mechanisms (Siyanova-Chanturia et al., 2017).

3.2.1.1 Phrase Frequency Effect. The processing advantage for MWEs is predominantly attributed to their statistical or distributional properties, such as frequency of occurrence. Frequency is one of the most prominent statistical properties of multiword items (Siyanova-Chanturia & Omidian, 2020, p. 531). It is often regarded as a primary characteristic of MWEs (Siyanova-Chanturia & Martinez, 2015). It has been shown that, due to their frequency, MWEs are processed faster than matched novel phrases. For example, using a (double) lexical decision task, Ellis et al. (2009, Experiment 1) investigated whether or not L1 speakers were sensitive to collocational frequency. Participants had to judge whether a pair of letter strings, shown simultaneously on the screen, were words (*cause problem*) or not (*phrup problems*). They found that frequent adverb-adjective collocations (*entirely blameless*) and verb-object collocations (*end war*) were responded to faster compared to infrequent control items (*entirely fledged* or *end weight*). Furthermore, the processing advantage was restricted to *actual* collocations which occurred in language use and did not extend to word pairs which were consistent in semantic prosody (i.e., the general tendency of certain words to co-occur with either negative or positive expressions, such as *cause* has a negative prosody [cause an accident /damage/harm], while *attain* has a positive prosody [attain goals/benefit/maturity]) but were not attested in the corpus (zero frequency). This finding suggests that it is the memory for *particular* word combinations that affords faster lexical access.

The frequency effect in MWE processing has not only been observed for bigrams, but also for longer units. For instance, using a phrasal-decision task, in one of the earliest studies, Arnon and Snider (2010) investigated the role of phrase-frequency in the processing of compositional 4-word phrases (*don't have to worry*) versus the control phrase (*don't have to wait*). The target phrases and controls differed in whole-phrase frequency but were matched in the frequencies of component words. They divided these phrases into several frequency bins and compared higher phrases with lower ones within each frequency bin. They found that L1 speakers responded to more frequent 4-word phrases faster than to less frequent 4-word phrases. This finding highlights the graded nature of phrase frequency effect (akin to word frequency effect) and the need to treat phrase frequency as a continuous variable. It further suggests that speakers can notice, learn and store frequency information of larger units across the frequency continuum (not just for the very frequent ones), as they do for words.

Further evidence for the role of whole-form frequency in MWE processing can be found in the processing of lexical bundles, which usually span phrasal boundaries (Biber et al., 1999). For example, in a self-paced reading study, Tremblay et al. (2011) investigated the role of whole-string frequency in the processing of lexical bundles (*in the middle of the*) versus lower frequency control sequences (*in the front of the*), which were embedded in sentence context. They found that lexical bundles were processed more quickly than matched controls, regardless of how the items were presented: word-by-word, portion-by-portion, and sentence-by-sentence. In two additional sentence and word recall experiments, they found that sentences containing lexical bundles were recalled better than those containing comparable control sequences, irrespective of how the sentences were presented, auditorily or visually. These authors concluded that "regular multiword sequences leave memory traces in the brain." (p. 595). A similar result was confirmed in Tremblay and Baayen (2010, Experiment 1). Using an immediate free recall task, participants were presented with six 4-word sequences at a time and then asked to recall as many sequences as possible. They found that higher frequency four-word sequences were recalled better than lower frequency ones. These findings suggest that the frequency with which MWEs occur facilitates the subsequent processing and recall of these MWEs.

In subsequent studies, the frequency effect was also observed in modified MWEs with intervening elements (Vilkaitė, 2016). For example, using eye-tracking, Vilkaitė (2016) compared the processing of adjacent (*provide information*) and nonadjacent collocations (*provide* some of the *information*) with the processing of their matched infrequent control phrases (*compare information*; *compare* some of the *information*), which were embedded in sentences. The author found that not only adjacent *but also* non-adjacent collocations showed processing advantage over their corresponding non-collocating control items, although adjacent collocations showed larger facilitation than nonadjacent ones. This finding shows that the elements of a MWE do not have to co-occur adjacently to be processed faster than infrequent control phrases, and thus supports Wray's (2000) definition of formulaic sequences which include both continuous and discontinuous sequences. On the other hand, this finding suggests that the facilitation effect may be attributed to links between words rather than holistic storage and access, which is widely assumed (but controversial) in MWE research (for a deeper discussion of holistic storage, see Siyanova-Chanturia, 2015; Wray, 2012). That is, the elements of a collocation should prime each other and be activated together to facilitate the processing, even when there are intervening words between the components.

The link or association between individual parts of MWEs has also been investigated using the priming paradigm. For example, using a primed lexical decision task with visually presented stimuli, Durrant and Doherty (2010) investigated whether the priming effect exists between constituents of two-word collocations (*foreign-debt*) versus low-frequency two-word combinations (*direct-danger*). To disentangle collocational priming from associative priming, they identified collocations which were strong associates (*card-game*) and ones which were not (*foreign-debt*). They found that high-frequency collocations, associated or non-associated, demonstrated significant priming relative to non-collocating combinations, supporting the existence of collocational priming due to frequency of co-occurrence rather than semantic association. This study provides empirical evidence for Hoey's lexical priming theory (Hoey, 2003, 2004, 2005), which holds that collocation is the product of drawing upon the mental store of lexical combinations (Hoey, 2013). According to Hoey (2013), the lexical priming (e.g., the collocational priming observed in Durrant and Doherty, 2010) is in essence repetition priming (i.e., frequency effect), where word *a* and word *b* co-occur often as a word combination, so that later exposure to word *a* can accelerate recognition of word *b*. The lexical priming theory indicates that, every time language users encounter a word, they note not only the word itself but also the other words with which it occurs together (Hoey, 2013).

3.2.1.2 Predictability Effect. Another body of research shows that processing time is not only affected by the frequency of a MWE, but also by its predictability (Conklin, 2020; McDonald & Shillcock, 2003b). Although MWEs vary greatly in frequency of occurrence

(Moon, 1998), what they have in common is that they are highly familiar to a language community. One consequence of being highly familiar is that the sequence becomes highly predictable (Siyanova-Chanturia & Omidian, 2020). That is, the reader can predict the upcoming word(s) by reading the initial constituent(s) of a MWE. For instance, upon reading the initial constituent(s) of an idiom (*spill the* …), a proverb (*an apple a day keeps the doctor* …), a binomial (*bride and* …), or a collocation (*fast* …), a proficient language user is very likely to complete the phrase with the expected completion (i.e., *beans*, *away*, *groom*, *food*, respectively). Thus, the final constituent(s) of a MWE decrease processing time, or can be skipped altogether, because they are highly, or even uniquely, anticipated prior to the reader reaching them (Conklin, 2020; Siyanova-Chanturia & Omidian, 2020).

Indeed, a multitude of studies employing a range of paradigms and tasks have confirmed that predictability plays an important role in the processing of MWEs. Psycholinguistic (McDonald & Shillcock, 2003a, 2003b) and neurolinguistic studies (Molinaro & Carreiras, 2010; Siyanova-Chanturia et al., 2017; Vespignani et al., 2010) have investigated the effect of predictability in the processing of MWEs. For instance, using eye-tracking, McDonald and Shillcock (2003a) compared L1 speakers' reading of verb-noun sequences embedded in identical neutral sentence context. The word sequences differed in their transitional probability, i.e., the statistical likelihood that a word precedes or follows another word (e.g., high probability – *avoid confusion*; low probability – *avoid discovery*), when they were matched in the frequency of the (constituent) nouns (e.g., *confusion* versus *discovery*). They found that verb-noun combinations with a high transitional probability were read faster than pairs with a low transitional probability, as evidenced by shorter initial fixation duration on the target nouns, which is a measure of processing effort that is sensitive to variables such as a word's frequency of occurrence (Rayner, 1998; Rayner et al., 1996) and its predictability from context (Rayner & Well, 1996). They took the results as evidence for the contribution to reading behaviour of the low-level predictability (in the form of corpus-derived transitional probabilities), rather than the high-level conceptual knowledge about word and context meaning (but see Frisson et al., 2005, for a different discussion which argues that transitional probability effects could be traced back to contextual predictability effects).

Electrophysiological research into MWEs also supports the argument that the expectations driven by a highly conventional string of language are different from the expectations based on the more general discourse-based constraints (Molinaro & Carreiras, 2010; Siyanova-Chanturia et al., 2017; Vespignani et al., 2010). For example, in an ERP study with L1 speakers, Siyanova-Chanturia et al. (2017) reported both N400 and P300 effects in L1 readers' comprehension of English literal compositional phrases, such as binomial expressions (*knife and fork*) versus associates (*spoon and fork*) versus semantic violations (*theme and fork*). They found that the second content word in the binomial condition elicited a larger P300 and a smaller N400 compared to the other two conditions (Experiment 1a). However, when phrases were presented without the conjunction "and" (Experiment 1b), no differences were observed between binomials and associates. The elicitation of smaller N400s and larger P300s when the binomial expressions were presented with the conjunction "and", but not when the stimuli presented without the conjunction "and", suggests that it is the phrasal, prefabricated, conventional status of binomial expressions that lead to the processing differences between binomials and associates in Experiment 1a. The result also suggests that frequent multi-word expressions are not only characterized by a reduced processing load and easier semantic integration (eliciting reduced N400s), but also by pre-activation of the mental template that uniquely matches the unfolding configuration (leading to increased P300s). The findings suggest that literal compositional phrases such as binomial expressions can be stored and represented in the mental lexicon similarly to single words, which supports frequency-based accounts of language acquisition, processing, and use.

3.2.1.3 Age-of-Acquisition Effect. Another issue that has received some attention in the literature is whether age-of-acquisition (AoA) has an effect on the processing of sequences above the word level (Arnon et al., 2017). Research into lexical (single word) AoA has shown that early-acquired words are processed faster than late-acquired words, i.e., AoA effect (Ghyselinck et al., 2004), when other factors affecting the speed of processing are controlled for. More recently, AoA has been found to play a role in the processing of sequences beyond single word level. Using a phrasal decision task on three-word sequences (i.e. lexical bundles), Arnon et al. (2017) examined whether or not AoA effect holds true for sequences beyond word level. They found that adults responded faster to early acquired phrases (*a good girl*) compared to later acquired ones (*a good dad*), suggesting the effect of AoA for units beyond single word level. The finding confirms the proposition that the building blocks of language vary in size and complexity and are not limited to single words (Siyanova-Chanturia & Van Lancker Sidtis, 2019). Similar issue has also been addressed in research on compounds. Using the eye movement paradigm, Juhasz (2018) investigated the role of AoA in the processing of compound words in sentence context in L1 speakers of English. Juhasz (2018) found that the AoA of the compound word (e.g., *airport, bodyguard*), rather than that of the individual morphemes, affected their processing, as shown by the gaze durations and total fixation durations.

Taken together, frequency, predictability and age-of-acquisition play an important role in the processing of MWEs in L1 speakers. Frequent and predictable MWEs are not only processed faster than infrequent novel sequences, but are also processed qualitatively differently from infrequent novel language in that different ERP components are involved in their processing. These phrase frequency effects show that language users notice, learn, store and represent the distribution of larger units of language, and that this information facilitates the processing of this linguistic material in its further usage (Siyanova-Chanturia & Martinez, 2015). Crucially, the processing of uniquely predictable MWEs is characterized by easier semantic integration and the activation of template matching mechanisms.

3.3.2 Processing of Compositional MWEs in L2 Speakers

In line with studies with L1 speakers, the issue of whether there is difference in the processing of compositional MWEs versus novel language controls has also been addressed in studies with L2 speakers. Most of the literature has shown that, similar to L1 speakers, L2 speakers demonstrate a processing advantage for frequent MWEs versus infrequent novel controls (Hernández et al., 2016; Jiang & Nekrasova, 2007; Kim & Kim, 2012; Siyanova-Chanturia, Conklin, & Van Heuven, 2011; Siyanova-Chanturia & Schmitt, 2008, Study 3; Sonbul, 2015; Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2018), whereas only a handful of studies have failed to report a processing advantage (Babaei et al., 2015; Siyanova-Chanturia & Janssen, 2018; Valsecchi et al., 2013; Vilkaitė & Schmitt, 2019). For example, using a phrase grammaticality judgement task, in which participants had to judge whether formulaic sequences (*to tell the truth*) and related nonformulaic control phrases (*to tell the price*) formed a grammatical phrase, Jiang and Nekrasova (2007) looked at the online processing of compositional MWEs by L1 and L2 speakers. They found that both L1 and highly proficient L2 speakers showed faster reaction times and lower error rates on formulas than controls. Based on this, they claimed that formulas were represented and processed holistically as single lexicalized units. Although this conclusion has been contested, not least because this study showed only processing speed advantage rather than the activation of the parts within the whole (Edmonds, 2014; Siyanova-Chanturia, 2015), this study was important as one of the first to look at the phrase frequency effects in MWEs processing by L2 users (Siyanova-Chanturia & Van Lancker Sidtis, 2019). Such findings highlight the parallels in the processing of MWEs between L1 and L2 speakers and suggest that, similar to L1 speakers, L2 speakers develop sensitivity to distributional properties of linguistic units.

L2 speakers' sensitivity to phrase frequency effects in the processing of MWEs is particularly related to two factors – language exposure and proficiency. These two factors reflect L2 speakers' experience with language in addition to frequency obtained from corpora. Corpora only provide an approximation of actual exposure, and do not necessarily reflect actual experience with a language (Conklin, 2020). To better understand the impact of language experience on the processing advantage for MWEs versus novel language in L2 speakers, researchers further investigate the role of language exposure and proficiency on MWE processing in an L2. With respect to language exposure, recent studies have shown that L2 speakers, on a par with L1 speakers, are also sensitive to multiword frequency. More importantly, this effect is not restricted to higher frequency phrases (but see Kim & Kim, 2012), but occurs across the frequency continuum (Hernández et al., 2016). In addition, English as a foreign language (EFL) learners are more sensitive to MWEs which are frequent in the input they receive (e.g. textbooks) than those which are frequent in large reference corpora but not present in the textbooks (Northbrook & Conklin, 2019). Further, L2 speakers can even develop and demonstrate sensitivity to novel multiword items after encountering them several times (Toomer & Elgort, 2019). Nevertheless, the type of L2 exposure – immersion-based versus classroom-based – has not been found to be a significant determinant of L2 phrase frequency effect (Hernández et al., 2016). These findings highlight the importance of the amount of exposure – in other words frequency of occurrence – in eliciting a processing advantage for multiword items.

Proficiency has also been found to be an important factor which affects L2 users' sensitivity to phrase frequency. To understand how proficiency plays a role in the processing of MWEs by L2 speakers, researchers typically compare the processing of MWEs versus novel controls between L2 speakers with varying levels of L2 proficiency. For example, using eye-tracking, Siyanova-Chanturia, Conklin and Van Heuven (2011) examined the processing of binomials (*bride and groom*) and their reversed forms (*groom and bride*) by L1 and proficient L2 English speakers. They found that both L1 and L2 speakers, higher and lower proficiency, were sensitive to phrase frequency. However, only higher proficiency L2 speakers showed sensitivity to the word order of the phrases (binomial vs. reversed), whereas lower proficiency L2 speakers did not, indicating that proficiency plays a crucial role in binomial versus reversed form processing and that frequency of exposure determines what is represented in the mental lexicon.

Similar results were reported in other studies with L2 speakers (Wolter & Yamashita, 2018; Yamashita & Jiang, 2010). For example, using an online acceptability judgment task, Wolter and Yamashita (2018) investigated how L2-English proficiency affected L1-Japanese speakers' sensitivity to adjective-noun collocations (*thick fog*) versus non-collocational controls (*hot parents*). They found that both L1 speakers and L2 speakers – intermediate and advanced – demonstrated sensitivity to collocational frequency. Moreover, with gain in L2 proficiency, L2 speakers tend to rely more on collocational-level frequency than on word-level frequency, trending toward a system that is more similar to that of L1 speakers.

Overall, the above studies showed that similar to L1 speakers, L2 speakers are attuned to phrase frequency distributions of MWEs in an L2. The results indicate that all language users, L1 or L2, attend to frequency information at multiple levels of representation – single word level as well as units above single word level (for a review, see Siyanova-Chanturia, 2015). They further indicate that experience with language plays an important role in the processing of linguistic materials, as proposed by emergentist approaches (Bybee, 2006; Goldberg, 2006; Tomasello, 2003).

3.4 Conclusion

This chapter reviewed the online processing (mainly comprehension) of MWEs, focusing on the role of various factors on MWE processing in an L1 and L2, including phrase frequency, predictability, age-of-acquisition, figurativeness, L2 proficiency, etc. The central finding reported in the studies on the processing of figurative non-compositional MWEs is an important role of familiarity and predictability. Familiar idioms have been found to be processed quantitatively (processing speed) and qualitatively (different ERP amplitudes) differently from novel control strings (Moreno et al., 2002; Vespignani et al., 2010). Concerning L2 speakers, the literal interpretation of an ambiguous idiom always precedes the figurative interpretation, which is different from that observed for L1 speakers (Cieślicka, 2006; Siyanova-Chanturia, Conklin, & Schmitt, 2011). With respect to the processing of compositional MWEs, phrase frequency plays a key role. Both L1 and L2 speakers have been shown to be sensitive to phrase frequency distributions (Arnon & Snider, 2010; Siyanova-Chanturia, Conklin, & Van Heuven, 2011; Tremblay et al., 2011). Predictability or probability also plays a key role in MWE processing, with more predictable words having decreased processing times (McDonald & Shillcock, 2003b).

Chapter 4 Cross-Language Influences in the Processing of MWEs

The present chapter[1] provides a state-of-the-art review of what is currently known about cross-language influences in the processing of MWEs in an L2. Two lines of research are considered: first, how L2 speakers process congruent MWEs versus L2-only MWEs; second, how L2 speakers process L1-only MWEs translated into the L2 compared with control novel phrases (Conklin & Carrol, 2018; Du, et al., under review). Studies have shown that congruent MWEs generally have a processing advantage over L2-only MWEs in L2 speakers. In contrast, evidence is mixed with regard to whether or not translated L1-only MWEs exhibit a processing advantage over matched controls in L2 speakers, with facilitation so far observed for idioms, but not for other types of MWEs (e.g., collocations). I consider possible reasons for these mixed findings.

Cross-language overlap or congruency benefits bilingual language processing. L1 plays an important facilitative role in L2 processing, if there is an overlap between the L1 and L2. Importantly, recent research has suggested that the processing advantage afforded by congruency may also extend to units beyond the word level, i.e., MWEs. The aim of the present chapter is to provide an up-to-date account of what is currently known about cross-language influences in the processing of MWEs in an L2. I will focus on empirical evidence for, and against, cross-language influences in L2 MWE processing, the mechanisms behind L1 influence on the processing of MWEs in an L2, and present gaps in this line of research.

Due to their frequency, familiarity, and predictability, MWEs have been found to be processed faster than matched novel strings of language by L1 speakers (Arnon & Snider,

[1] Parts of this chapter have been included in a chapter for an edited volume (*Cross-Language Influences in Second Language Acquisition and Processing*: *Interdisciplinary Insights and Perspectives*), as part of the *Bilingual Processing and Acquisition* book series published by John Benjamins.

2010; Tremblay et al., 2011; Vilkaite, 2016) and L2 speakers (Hernández et al., 2016; Jiang & Nekrasova, 2007; Siyanova-Chanturia, Conklin, & Van Heuven, 2011). While much of the research centred on the factors such as frequency, familiarity, and predictability in the processing of MWEs by L1 and L2 speakers, how *cross-language congruency* affects the processing of MWEs in L2 speakers[2] has so far received limited attention, with only a handful of studies investigating this issue (Carrol & Conklin, 2014, 2017; Carrol et al., 2016; Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2015, 2018; Yamashita & Jiang, 2010). These studies have mainly focused on how L2 speakers process congruent MWEs (i.e., L2 expressions that have a translation equivalent in the L1) versus incongruent L2-only MWEs (i.e., L2 expressions that do not have a translation equivalent in the L1), or with translated L1-only MWEs (i.e., L1 expressions that do not have a translation equivalent in the L2 and are translated into the L2) versus matched control (novel) phrases. While studies agree that congruent MWEs are processed faster and more accurately than incongruent L2-only MWEs, and that L1 influences L2 MWE processing, research evidence is mixed with regard to whether translated L1-only MWEs are processed faster than matched controls. In what follows below, I first review the studies looking at the processing of congruent versus incongruent L2-only MWEs (collocations and idioms), followed by the research on the processing of translated L1-only MWEs versus matched controls.

4.1 Processing of Congruent and L2-only MWEs

Emerging evidence indicates that congruency plays an important role in L2 MWE processing. L2 speakers can process congruent MWEs more rapidly and accurately than incongruent L2-only MWEs that are matched for phrase frequency (Carrol et al., 2016; Wolter & Gyllstad, 2011, 2013; Yamashita & Jiang, 2010). The relative processing advantage

[2] In this review, I use the terms second language (L2) speakers and bilinguals interchangeably. Both early and late learners of an additional language are deemed bilinguals, following Siyanova-Chanturia, Canal, & Heredia (2019). When reviewing individual studies, I use the terms adopted in the studies concerned.

in speed and accuracy for congruent over incongruent L2-only MWEs has been interpreted as cross-language influences in the processing of MWEs in an L2. Yamashita and Jiang (2010) were among the first to investigate the role of congruency in the processing of MWEs in an L2. Using a phrase-acceptability judgment task (i.e., whether an item is acceptable in English, YES/NO), this study investigated the processing of congruent verb-noun and adjective-noun collocations versus incongruent combinations. Three groups of participants were recruited: English as a *foreign* language (EFL) Japanese learners, English as a *second* language (ESL) Japanese users, and L1 speakers of English. Congruent (*make lunch*) and incongruent collocations (*kill time*) were matched in terms of length, word frequency, and phrase frequency in English. They found that lower proficiency EFL learners made more errors with and responded more slowly to incongruent collocations (*kill time*, whose Japanese equivalent literally translates as ‘crush/break time’) than to congruent collocations (*kill animals*). In contrast, although higher proficiency ESL users also made more errors on incongruent collocations than on congruent ones, they responded equally fast to the two types of collocations. The control group of L1 speakers, however, processed congruent and incongruent collocations with no difference in speed or accuracy. Based on these results, Yamashita and Jiang (2010) concluded that L1 influences L2 collocation processing only in early stages of acquisition, but not in late stages of acquisition. However, a limitation of this study is that individual words in test items were repeated, such that a total of 39 of the 48 test items contained repeated words (Wolter & Yamashita, 2018). For example, *tea* appears once in a congruent item (*cold tea*) and twice in incongruent items (*make tea* and *weak tea*). This factor may have influenced reaction times (RTs) to congruent and incongruent items.

More recent evidence suggests that the L1 continues to affect the processing of MWEs in an L2 even when L2 speakers are highly proficient and in advanced stages of L2 acquisition (Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2018). Using a primed visual lexical decision task, Wolter and Gyllstad (2011) investigated the processing of verb-noun collocations with advanced Swedish learners of English and L1 speakers of English. The participants were first shown the prime (the verb in a verb-noun collocation, e.g., *give*) and then asked to decide whether or not the target (the noun which collocates with the verb, e.g., *answer*) was a legal word in English. They observed that in L2 speakers, congruent collocations (e.g., *give an answer*) were processed significantly faster and more accurately than L2-only collocations (e.g., *pay a visit*, whose Swedish equivalent *betala ett besök* literally translates as 'make a visit/do a visit'). The effect was found despite the fact that congruent and L2-only collocations were matched in terms of the collocational strength and phrase frequency in the L2. In contrast, L1 speakers showed no difference in terms of speed or accuracy for congruent and English-only collocations. Interestingly, L2 speakers demonstrated an even stronger priming effect for congruent collocations than L1 speakers, suggesting there was a 'doubling up' of activation (i.e., simultaneous activation) for congruent collocations in both the L1 and L2 which contributed to the stronger facilitation in L2 speakers (Wolter & Gyllstad, 2011: 443). Note that what Yamashita and Jiang (2010) call 'incongruent collocations', Wolter and Gyllstad (2011) label 'L2-only collocations'. Wolter and Gyllstad (2011) argued that corresponding L1 collocations are likely to be activated in the processing of collocations in an L2, even when L2 speakers are highly proficient and have established direct links between L2 collocations and the concepts, which contradicts Yamashita and Jiang (2010).

Similar results were reported in a follow-up study (Wolter & Gyllstad, 2013). Using the same task as Yamashita and Jiang (2010) – a phrase-acceptability judgment task – Wolter and Gyllstad (2013) investigated how the L1 influences the processing of adjective-noun collocations in the L2 with advanced Swedish learners of English and L1 speakers of English. Three types of items were employed: L1-L2 congruent collocations (*handsome man*), incongruent L2-only collocations (*identical twins*), and non-collocational novel items (*angry use*). Consistent with Wolter and Gyllstad (2011), Wolter and Gyllstad (2013) found that congruent collocations were judged more quickly and accurately than incongruent collocations by the L2 speakers, but were judged as quickly and accurately as incongruent collocations by the L1 speakers. Notably, the L2 speakers responded to the congruent collocations as fast as the L1 speakers, suggesting they were highly proficient. The L1 speakers responded faster than the L2 speakers only on the incongruent items. They further found that collocational frequency in the L2 (English), rather than that in the L1 (Swedish), affected variation in RTs between the congruent and incongruent collocations for the L2 speakers. This suggests that, in addition to congruency, frequency in an L2, but *not* frequency in an L1, is an important factor for explaining RTs in making judgements about the acceptability of collocations. A limitation of the study is that most non-collocational control items were not plausible (e.g., **angry use*, **legal plant*, **final others*), which is likely to have caused a significant delay in the processing of such phrases (Sonbul & Siyanova-Chanturia, 2021). Of note is that this also applies to Wolter and Gyllstad (2011).

In addition, the effect of congruency on collocation processing in an L2 is reported in a recent study with a different group of L2 speakers – L1 Japanese speakers of English. Employing the same task and collocation type as Wolter and Gyllstad (2013), namely, phrase-acceptability judgment task and adjective-noun collocations, Wolter and Yamashita (2018) investigated how intermediate and advanced Japanese speakers of English, as well as L1 speakers of English, processed congruent (*strong wind*) and incongruent L2-only (*busy road*) collocations. Both groups of L2 speakers (intermediate and advanced) responded significantly faster to congruent than to incongruent collocations. Crucially, L1 speakers showed no differences in RTs for these two types of collocations. In line with Wolter and Gyllstad (2011, 2013), this study further indicates that congruency with the L1 influences collocation processing in the L2. However, in this study congruent collocations were significantly more frequent than English-only items, which may have contributed to their faster processing compared with English-only collocations.

Yamashita (2018) argues that the above studies on collocations have a possible confounding effect of semantic transparency – the extent to which constituents contribute straightforwardly to the meaning of a MWE. Yamashita (2018) pointed out that there were more transparent and fewer opaque collocations in the congruent than incongruent collocations in the studies reviewed above (i.e., Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2018; Yamashita & Jiang, 2010), which could have affected response times. Semantic transparency has been previously found to affect the processing of collocations. For example, using a visual semantic judgment task (i.e., Is the item meaningful and natural in English? YES/NO), Gyllstad and Wolter (2016) investigated how high-proficiency Swedish learners of English and L1 English-speaking controls processed (1) verb-noun collocations (*draw a conclusion*, *run a risk*) with the verb appearing in a specialized sense and the noun used literally, and (2) verb-noun free combinations (*write a letter*, *kick a ball*) with both constituent words used literally. Collocations were less transparent than free combinations, but the two types of combinations were matched for phrase frequency in English. Collocations and free combinations were congruent, in the sense that they could be translated word for word into L1 Swedish. Gyllstad and Wolter (2016) found that not only L2 learners, but also L1 speakers, showed slower RTs and higher error rates for collocations than free combinations, indicating that the lower level of transparency in collocations may have led to an increased processing cost.

Similarly, the facilitative effect of congruency has also been observed in the processing of idioms in an L2. Using the eye movement paradigm, Carrol et al. (2016) investigated how high-proficiency Swedish learners of English and L1 speakers of English process congruent (*break the ice*) and L2-only (*kick the bucket*) idioms, compared to their corresponding matched literal controls (*crack the ice*, *drop the bucket*). Carrol et al. (2016) found that L2 speakers processed congruent idioms (*break the ice*) faster than literal controls (*crack the ice*), as suggested by early measures such as likelihood of skipping for the final words, and late measures such as total reading time and regression path durations. However, for L2-only idioms, L2 speakers only showed facilitation for the meaning integration (via late measures), but no facilitation for the form (via early measures). In contrast, L1 speakers showed comparable advantage in both early and late measures when reading congruent and incongruent, English-only idioms compared to their matched controls.

Taken together, the findings presented above suggest that congruency does facilitate MWE processing in an L2, in that congruent MWEs show a greater processing advantage over incongruent L2-only MWEs. The congruency effect has been found robust and across different types of MWEs, collocations and idioms, and with different paradigms employed (RTs[3] and eye movements).

4.2 Mechanisms Underpinning the Congruency Effect in L2 MWE Processing

To understand what underpins the congruency effect in the processing of idioms and collocations in an L2, research has investigated whether translated L1-only MWEs (e.g., Chinese-only idiom *画蛇添足 = draw a snake and add feet*, meaning “to ruin with unnecessary detail”) can also show a processing advantage over matched controls. If bilinguals process translated L1-only MWEs faster than L2 controls, without ever experiencing these L2 word combinations previously, we could argue that this facilitation is due to the online activation of known L1 MWEs; that is, known MWEs are simultaneously activated in L2 processing (see also, Carrol & Conklin, 2017; Yamashita, 2018; Zeng et al.,

[3] Some of the RT measures reported above are not measures of processing the MWEs but of making judgments about the MWEs.

2020). This account is referred to as the L1 multi-word expression activation account (Pulido & Dussias, 2020; Yamashita, 2018). Alternatively, if bilinguals do not process translated L1-only MWEs faster than L2 controls, we could argue that prior exposure to these L2 word combinations is necessary for them to show any processing advantage. The congruency effect may be due to the acquisition order of congruent L2 MWEs, i.e., congruent MWEs may be acquired earlier than L2-only MWEs, as a result of cross-language transfer. This account is referred to as the L2 multi-word expression experience account.

4.2.1 L1 MWE Activation Account

In the L1 MWE activation account, known L1 MWEs are assumed to be automatically activated in L2 processing, leading to a greater processing advantage for congruent MWEs when compared to incongruent L2-only ones (e.g., Carrol & Conklin, 2014; Carrol et al., 2016; Wolter & Gyllstad, 2011). According to this account, the initial words of an L2 MWE automatically activate their L1 translation equivalents via cross-language translation priming (Thierry & Wu, 2007; Wu & Thierry, 2010; Zhang et al., 2011), which, in turn, triggers the activation of a known L1 MWE via within-language lexical priming between the constituents of MWEs (e.g., idiom priming, collocational priming). This account is not unlike the lexical-translation mechanism proposed by Carrol and Conklin (2014). For example, when L2 speakers of English (whose L1 is Chinese) read *draw a snake and add* …, the Chinese translation of *hua she tian* /*画蛇添* … might be automatically activated as each word is encountered. The activated Chinese words then trigger a known idiomatic sequence *hua she tian zu* /*画蛇添足* via direct retrieval, making the final character zu /*足* available, which in turn primes its English translation *feet* (Carrol & Conklin, 2014, 2017). Likewise, for collocations, Wolter and Gyllstad (2011) hypothesized that when an L2 word is presented, it activates not only its L2 collocates (e.g., *strong* activates its collocate *tea*), but also its L1 translation equivalent (*strong – nong* /*浓*), which in turn activates its collocates in the L1 via collocational priming (*cha* /*茶* – *tea*). The activated L1 collocate (*cha* /*茶* – *tea*), in turn, primes its L2 translation (*tea*). That is, the target word (*tea*) is primed as both an L1 and L2 collocate (i.e., doubled activation), leading to a stronger priming effect than that for L2-only collocations. As Carrol and Conklin (2014) argue, MWEs may show cross-language priming in the same way as single words.

This account is built on the literature on cross-language activation in bilingual processing and on within-language lexical priming between the constituents of MWEs. The literature on bilingual language processing has found that the two languages of a bilingual can be activated simultaneously, even when processing happens entirely in one language (for a review, see Van Hell & Tanner, 2012). For example, using a within-L2 semantic relatedness task and event-related brain potentials, Thierry and Wu (2007) investigated how Chinese-English bilinguals and English monolinguals process semantically related (e.g., *post-mail*) and unrelated (e.g., *train-ham*) English word pairs. Half of the word pairs shared a character when translated into Chinese. For example, *train-ham* are not related in meaning but their Chinese translations *huo che* (*火车*) and *huo tui* (*火腿*) share one Chinese character *huo* (*火*); while *apple-table* are neither related in meaning nor overlap in their Chinese translations (i.e., *ping guo* [*苹果*] - *zuo zi* [*桌子*]). They found that L2 speakers showed reduced N400 amplitudes for English word pairs that shared a character in their Chinese translations (such as *train-ham*), but not for those that did not (such as *apple-table*). This suggests that L1 translation equivalents are automatically activated in L2 processing, and they in turn activate associative links within the L1. MWE studies have found lexical priming between the constituents of a MWE, such that the beginning of a MWE can facilitate the recognition of its terminal word (Carrol & Conklin, 2017; Durrant & Doherty, 2010;

Siyanova-Chanturia, Conklin, & Schmitt, 2011; Underwood et al., 2004; Wolter & Gyllstad, 2011).

4.2.2 L2 MWE Experience Account

In the L2 MWE experience account, congruent MWEs are assumed to be acquired before incongruent MWEs, because acquisition is more straightforward when there is correspondence between the L1 and L2 (due to positive cross-language transfer). Thus, congruent MWEs should be processed faster than incongruent MWEs due to the age/order-of-acquisition effect (e.g., Wolter & Gyllstad, 2013; Wolter & Yamashita, 2015, 2018). Rather than assuming that L1 MWE translation equivalents are activated simultaneously in L2 processing due to cross-language activation and that this activation leads to faster processing, the experience account holds that prior exposure to a word sequence in an L2 is essential for it to be processed faster than its control, whether or not it has an equivalent form in the L1. Having an equivalent form in the L1, however, does facilitate the acquisition of a (congruent) L2 MWE, in that, acquiring congruent MWEs may take less time and require less exposure to the L2 than acquiring incongruent L2-only MWEs. Congruent MWEs may also be acquired faster because they are more likely to be noticed in the L2 input, since they share form (translation equivalents), structure (fixed word order), and referential meaning (same construct, e.g., *sun and moon*) in the L1 and L2 (Yamashita & Jiang, 2010). Earlier acquisition of congruent MWEs results in more experience with the expression in the L2 and leads to greater processing advantage for congruent MWEs compared to incongruent MWEs. According to this account, when encountered in an L2 for the first time, translated L1-only MWEs will not be processed faster than matched L2 control phrases, because L2 speakers are not familiar with them. This account is in line with usage- and exemplar-based acquisition and processing accounts which assume that frequency of encounters with and use of a lexical item (word, MWE) determines the strength of its mental representations and ease of processing (Bybee, 2006; Langacker, 2000).

The L2 MWE experience account is based on the literature on crosslinguistic influences in MWE acquisition and, to a lesser extent, on the age/order-of-acquisition effect in MWE processing. A number of studies have shown that L1 influences the *acquisition* of L2 MWEs (Nesselhauf, 2005; Römer et al., 2014; Sonbul et al., 2020; Yamashita & Jiang, 2010). For instance, using a phrase-acceptability judgment task, Yamashita and Jiang (2010) found that both lower-proficiency EFL and higher-proficiency ESL users of English, but not L1 English-speaking controls, made fewer errors on congruent collocations than incongruent L2-only collocations. They concluded that congruent L2 collocations may be acquired faster than incongruent L2-only collocations. Because a congruent L2 MWE and its L1 counterpart share the identical or very similar concept, it can be easily accepted and stored in memory by simply resorting to its L1 equivalent (Yamashita & Jiang, 2010, p. 662). In addition, there is also evidence showing that age-of-acquisition affects the processing of units longer than a word (e.g., phrases: Arnon et al., 2017; compound words: Juhasz, 2018).

Taken together, the L1 MWE activation account and L2 MWE experience account make differential predictions about the processing of translated L1-only MWEs when encountered in the L2 for the first time. In what follows below, empirical evidence on the processing of translated L1-only MWEs is reviewed in detail.

4.3 Processing of Translated L1-only MWEs

The processing of translated L1-only MWEs has so far been investigated with two types of MWEs, idioms and collocations. However, conflicting results have been reported. Studies with idioms have found a processing advantage for translated L1-only idioms over control novel phrases, suggesting direct L1 involvement in L2 processing, even when presented and processed entirely in the L2 (Carrol & Conklin, 2014, 2017; Carrol et al., 2016). In contrast, studies with collocations have not found a robust processing advantage for such items over control phrases (Wolter & Yamashita, 2015, 2018). In what follows, these studies and the possible reasons for the pattern of findings are discussed in some detail.

4.3.1 Research on Idioms

Carrol and Conklin (2014, 2017) were the first to investigate the processing of translated L1-only MWEs. These authors focused on the processing by intermediate proficiency Chinese speakers of English and L1 speakers of English of Chinese-only idioms translated into English. Carrol and Conklin (2014) investigated whether the initial words of English-only idioms (*to spill the* …) and translated Chinese-only idioms (*draw a snake and add* …) primed the final words (*beans*, *feet*) in a lexical decision task. L2 speakers responded faster to targets that formed idioms in their L1 (*draw a snake and add feet*) than to targets in corresponding matched control items (*draw a snake and add hair*). That is, they showed priming for translated L1-only idioms. However, they showed no advantage for English-only idioms (*spill the beans*) over matched novel language (*spill the chips*). In contrast, L1 speakers showed a different pattern of results. Targets that formed English idioms were reliably faster than controls; targets that formed Chinese idioms were processed with the same speed as controls. Both L1 and L2 speakers showed priming only for idioms taken from their respective L1s. The results suggest L1 influence extends to the processing of translated L1-only idioms. However, because the task was not performed under time pressure and participants could view the primes for as long as they wanted, it may have allowed participants to actively anticipate the completion of a phrase.

The processing advantage for translated L1-only idioms relative to control novel phrases was also reported in a follow-up study that used eye movements (Carrol & Conklin, 2017). This study compared reading times for idioms versus control novel phrases (*draw a snake and add feet* vs. *draw a snake and add hair*) in Experiment 1, and figurative versus literal uses of idioms (*add oil and vinegar* – figurative meaning ‘to embellish a story’ vs. literal meaning ‘to add some dressing’) in Experiment 2. Target items were embedded in short sentence context biasing the figurative meaning. Consistent with earlier behavioural evidence (Carrol & Conklin, 2014), Chinese speakers of English, but not L1 speakers of English, showed significant facilitation for the final word of translated L1-only idioms compared to control phrases, evidenced in the analysis of first fixation durations and total reading time. In contrast, L1 speakers, but not L2 speakers, showed significant priming for English-only idioms relative to matched controls, as suggested by the likelihood of skipping and total reading time. Similar to Carrol and Conklin (2014), both L1 and L2 speakers only showed facilitation for idioms taken from their L1. However, in Experiment 2, L2 speakers read expressions used figuratively more slowly than those used literally, regardless of whether the idioms were English-only or Chinese-only, evidenced in the analysis of total reading times. This indicates that either the figurative meanings of idioms were unknown to L2 speakers, or the figurative meanings were only accessed after the literal meaning had been rejected. In contrast, L1 speakers showed no difference in reading times for literal or figurative uses of English-only idioms, whereas they read the figurative uses of translated Chinese-only idioms more slowly than the literal uses. Carrol and Conklin (2017) took these results as evidence that the recognition of the form of a translated L1-only idiom (as suggested in Experiment 1) did not automatically lead to the access of the figurative meaning of the idiom. According to the authors, the processes underlying recognition of form, and access to the phrase-level figurative meaning of an idiom, may not be the same. Recognition of form may be affected by strong intra-lexical links among the individual words of an idiom, while access to phrasal meaning may be affected by familiarity and (language specific) frequency of encounters with a whole form structure and its associated figurative meaning. In addition, the recognition of form may be realized via a lexical/translation route whereby L2 words automatically activate L1 equivalents (i.e., cross-language translation priming), which in turn trigger a known L1 idiom, facilitating the L2 processing. Access to meaning, however, may be realized via a conceptual route, whereby L2 words directly trigger their underlying concepts (e.g., DRAW, SNAKE, ADD), the association of which in turn triggers the underlying idiom concept (‘run with unnecessary detail’). These two routes are referred to as the lexical-translation mechanism and the conceptual priming mechanism, respectively. As Carrol and Conklin (2017) argue, for Chinese speakers of English, translated L1-only idioms had never been encountered in English and, thus, L2 representations of whole forms and their associated figurative meanings were likely weak. Therefore, idioms were more difficult to process when used figuratively than literally. However, intra-lexical links among the individual constituents of an idiom may be triggered by fast, automatic cross-language translation priming, whereby the initial words of an English expression (*draw a snake and add* …) automatically activate their translation equivalents in the L1 (here, Chinese), which in turn trigger a known idiomatic expression in the L1 (e.g., *hua she tian zu* /*画蛇添足*). The triggered L1 idiom facilitates form recognition, making the final word (*zu* /*足*) available and in turn priming its translation equivalent in English (*feet*).

Interestingly, translated L1-only idioms have shown a similar level of facilitation as congruent idioms relative to their corresponding matched controls. Carrol et al. (2016) investigated native L1 English and advanced L2 English (L1 Swedish) speakers’ reading patterns of idioms versus matched control phrases (*spill/drop the beans*) in three conditions: congruent, English-only, and translated Swedish-only. They used eye-tracking to compare reading patterns for the whole idiom (looking at *phrase-level measures*, such as total reading time, etc.) and its final word (looking at *word-level measures*, such as likelihood of skipping, etc.). The L2 speakers showed a processing advantage for translated Swedish-only (*play monkey* vs. *taste monkey*) and congruent idioms (*lose your head* vs. *hurt your head*) over controls in the word-level analysis, as suggested by the likelihood of skipping. Further, they observed a processing advantage for translated Swedish-only idioms over controls extended to the phrase-level, as evidenced by shorter total reading times and fewer fixations for the former over the latter. The results suggest that there was form activation and meaning activation in the processing of translated L1-only idioms by L2 speakers. The results, however, are inconsistent with Carrol and Conklin (2017), who only found evidence for form recognition of translated L1-only idioms, but no evidence for access to their figurative meanings. The results from Carrol et al. (2016) provide evidence against the argument proposed by Carrol and Conklin (2017) that access to meanings of idioms may be due to familiarity and (language specific) frequency of encounters of the structure and its associated figurative meaning. Instead, the results suggest that the figurative meanings of translated L1-only idioms can also be activated with no prior L2 exposure. More importantly, in Carrol et al. (2016), the processing advantage for Swedish-only and congruent idioms over controls was found to be comparable, suggesting that congruent idioms may not show more facilitation than translated L1-only idioms, despite (additional) L2 exposure. The authors argue that it is L1 (rather than L2) knowledge that determines the ease of processing of idioms in L2 speakers. This conclusion, however, does not seem to align with the finding of Wolter and Gyllstad (2013) that collocational frequency in the L2, rather than that in the L1, affected the difference in RTs between the congruent and incongruent L2-only collocations for the L2 speakers.

4.3.2 Research on Collocations

Unlike the facilitation effect observed for translated L1-only idioms relative to control phrases, studies with collocations paint a different picture. For example, using a double lexical decision task that presented both words on the screen simultaneously, Wolter and Yamashita (2015) investigated how the L1 influences collocation processing in the L2 with intermediate and advanced Japanese speakers of English, as well as L1 speakers of English. Three types of items were used: translated Japanese-only verb-noun and adjective-noun collocations (*buy anger*), English-only collocations (*catch breath*), and non-collocational baseline items (*bad gift*). The results showed that there was no significant difference for either RTs or error rates between translated Japanese-only items and non-collocational baseline items, suggesting that there was no activation of L1 collocations when processed entirely in the L2. However, of note is that the non-collocational baseline items used in this study (e.g., *careful branch*, *dark guess*, *false months*) were not semantically plausible, as in Wolter and Gyllstad (2011, 2013).

Similar results were observed in a follow-up study with intermediate and advanced Japanese speakers of English that used a phrase acceptability task (Wolter & Yamashita, 2018). In this study, four types of adjective-noun constructions were employed: congruent, English-only, Japanese-only collocations, and baseline items which were not semantically meaningful. Note that the English-only and Japanese-only collocations were adopted from Wolter and Yamashita (2015), while congruent collocations were developed for this study. Similar to Wolter and Yamashita (2015), the two groups of L2 speakers, intermediate and advanced, judged translated L1-only collocations and matched baseline items with a similar speed. Congruent collocations, however, were processed significantly faster than English-only collocations by both groups of L2 speakers. In contrast, L1 speakers of English processed both types of items (congruent and English-only) in a similar way. This once again suggests that there is a processing advantage for congruent over incongruent L2-only collocations, but there is no clear processing advantage for translated L1-only collocations over matched controls. However, the findings of this study may be affected by the previously identified issue of low semantic plausibility of the non-collocational controls (*proud idea*, *open teeth*, *lucky tea*), which may have contributed to their slower processing relative to the other conditions. As noted by Sonbul and Siyanova-Chanturia (2021), control items should be infrequent and not strongly associated phrases that are nonetheless semantically plausible.

In sum, studies on idioms show priming for translated L1-only idioms over novel L2 phrase controls, suggesting that there might be automatic activation of known L1 MWEs (Carrol & Conklin, 2014, 2017; Carrol et al., 2016). Conversely, studies on collocations have not found facilitation in the processing of translated L1-only collocations (Wolter & Yamashita, 2015, 2018). Therefore, it is still insufficient to unequivocally support one of the two accounts, and pin down the origin of the congruency effect in bilingual processing of MWEs. Further research is needed to investigate the processing of translated L1-only MWEs in bilinguals and probe into the underlying mechanisms of the congruency effect in MWE processing.

4.4 Conclusion

This chapter has reviewed research on cross-language influences in the processing of L2 MWEs. The congruency effect was found in the literature that compared the processing of congruent and incongruent L2-only collocations and idioms. However, the findings are mixed with regard to the processing advantage of translated L1-only MWEs over matched controls, with a processing advantage reported in studies with idioms *but not* collocations. Thus, further research is needed to investigate the processing of translated L1-only MWEs in bilinguals. In addition, little is known about whether a nontarget and weaker L2 can influence the processing of MWEs in the bilinguals' L1. In my thesis, I aim to address these gaps.

Chapter 5 Experiment 1: Probing L1 Influence in L2 MWE Processing

As reviewed in Chapter 4, studies on crosslinguistic influences in the processing of L2 MWEs converge on the finding that congruent MWEs show a processing advantage over incongruent L2-only MWEs (i.e., the congruency effect), whereas studies on translated L1-only MWEs, encountered in the L2 for the first time, report mixed results. Namely, translated L1-only idioms are processed faster than matched controls, while translated L1-only collocations show no such processing advantage. It is thus unclear what the origin of the congruency effect is in the processing of L2 MWEs, that is, whether the congruency effect is due to the automatic activation of known L1 MWEs. In addition, the issue of whether the weaker L2 can influence the MWE processing in the L1 has not been sufficiently addressed in the literature, so far. In order to better understand these issues, the present research focused on one particular type of MWEs, binomials (*bride and groom*) – for which this issue has not yet been addressed. Binomials are three-word phrases that are realised in English in the form of *A and B*, where a specific-word order is preferred in that *A and B* is always more frequent than *B and A* such as *knife and fork* vs. *fork and knife* (Benor & Levy, 2006; Carrol & Conklin, 2020). This chapter[4] covers the issue of whether L1 influences binomial processing in the L2, and the following chapter covers the issue of whether L2 affects binomial processing in the L1.

In Experiment 1, I wanted to determine whether congruent English binomials (i.e., English binomials whose Chinese translation equivalents are also binomials in Chinese) are processed faster than English-only binomials (i.e., English binomials whose Chinese

[4] The studies reported in this chapter and next chapter have been published as Du, L., Elgort, I., & Siyanova-Chanturia, A. (2021). Cross-language influences in the processing of multiword expressions: From a first language to second and back. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2021.666520

translation equivalents are not binomials in Chinese) by Chinese-English bilinguals in their second language. If there was indeed a processing advantage for congruent English binomials over English-only binomials in Chinese-English bilinguals, but not in English monolinguals, it would be taken as evidence for the congruency effect, a marker of cross-linguistic influences. To further understand and interpret the mechanisms underpinning the congruency effect in L2 MWE processing, I sought to determine whether translated Chinese-only binomials (i.e., Chinese binomials whose English translation equivalents are not binomials) were processed faster than their control phrases.

Therefore, I sought to answer the following research questions:

1. Is CLI observed in the processing of congruent English binomials by Chinese-English bilinguals?
2. Is CLI observed in the processing of Chinese-only binomials translated into English by Chinese-English bilinguals?

To answer the two questions, the present investigation recruited two groups of participants, Chinese-English bilinguals and English monolinguals. All participants completed the same *English* lexical decision experiment with a binomial priming manipulation. That is, they were shown the first two words of a binomial or control phrase (*knife* + *and* OR *spoon* + *and*), one word at a time, and then were asked to make a lexical decision on the final word (*fork*). This enabled me to directly compare the RTs of target words (*fork*) that appear in binomials (*knife and <u>fork</u>*) versus controls (*spoon and <u>fork</u>*). In Experiment 1a, the processing of binomials is compared with that of controls for Chinese-English bilinguals. Experiment 1b looks at the same issue with English monolinguals, who served as a baseline group.

5.1 Experiment 1a

Experiment 1a had several goals. Firstly, I wished to determine if Chinese-English bilinguals exhibit a greater processing advantage for congruent binomials relative to their matched controls (*sun and moon* vs. *star and moon*), when compared with English-only binomials relative to their matched controls (*bread and butter* vs. *toast and butter*). Such a finding would further support L1 influence on L2 MWE processing and extend it beyond idioms and collocations, to a new type of MWEs – binomials. Secondly, I investigated if translated Chinese-only binomials show a processing advantage relative to their matched controls (*wisdom and strength* vs. *exercise and strength*) in Chinese-English bilinguals. If so, this would suggest that L1 influence extends to the processing of translated L1-only binomials, which has not been previously shown in L2 studies. This would support the claim that corresponding L1 binomials are automatically activated in L2 processing, as predicted by the L1 activation account of CLI in bilingual MWE processing. If not, this would suggest that there is no automatic activation of L1 binomials in L2 processing and that prior exposure to L2 word sequences is necessary for them to show significant facilitation compared to control novel phrases. This is predicted by the L2 experience account of CLI in bilingual MWE processing. I tested these predictions in a reaction time study with adult Chinese-English bilinguals using three types of binomials: congruent (*sun and moon*), incongruent English-only (*bread and butter*), and translated Chinese-only (*wisdom and strength*).

5.1.1 Participants

Fifty-two Chinese-English bilinguals participated in the experiment (40 females; age: 18-45; mean age = 28.46 years). The number of participants was estimated based on a repeated-measures design with the expected effect size being around d = 0.3 for the power of 80% (Brysbaert & Stevens, 2018). Participants were recruited through Victoria University of Wellington networks, through advertising posters, university Facebook pages and personal contacts. They were undergraduate, postgraduate international students and young professionals studying or working at Victoria University of Wellington (VUW). The study was conducted with the ethics approval from VUW (see Appendix 1). All participants were introduced to the research and shown an information sheet about the study (See Appendix 2). Each participant gave informed consent (see Appendix 3) and received $10 for their participation in the experiment. They completed a language background questionnaire (see Appendix 4) before the experiment, in which they reported their English proficiency test score (International English Language Testing System [IELTS] or Test of English as a Foreign Language [TOEFL]), the number of years in an English-speaking country (average = 3.8 years, range: 0.5 – 17 years), and an estimate of their daily usage of English (average = 48%, range: 10% – 90%). Their mean IELTS score was 6.67 (range: 6 – 8; roughly equivalent to the levels B2-C1 of the Common European Framework of Reference for Larson et al., 2014). They were thus regarded as advanced speakers of English as a second language. All participants had normal or corrected-to-normal vision. Table 5-1 summarizes Chinese-English participants' language proficiency characteristics.

Table 5-1 *Means (standard deviations) of self-reported age, English proficiency levels, daily usage of English, years of exposure to English in English-speaking countries for Chinese-English bilinguals*

	Chinese-English bilinguals (N=52)
Age	28.46 (6.16)
English proficiency	Advanced
Daily usage of L2	English: 48% (23%)
Years of exposure to L2	3.86 (3.88)

Note. Among the 52 Chinese-English bilinguals, there were 39 participants with IELTS scores, 4 participants with TOEFL scores, and 9 participants who had no English proficiency test record but took the University's English Proficiency Program (EPP) and met the University's English proficiency requirement. TOEFL scores were converted to IELTS band scores for the ease of comparison based on the Comparison Table provided on the website: https://www.ets.org/toefl/institutions/scores/compare/.

5.1.2 Materials

The critical items included 60 binomials and 60 controls, which were all presented in English. The binomials belonged to one of the following conditions: (1) congruent (e.g., *sun and moon* ↔ 太阳和月亮, taiyang he yueliang), (2) incongruent English-only (e.g., *bread and butter* ↔ 面包和黄油, mianbao he huangyou), and (3) translated Chinese-only (e.g., *wisdom and strength* ↔ 智慧和力量, zhihui he liliang), which were categorized based on their overlap with Chinese. Each binomial had a corresponding control item. Control items were created by replacing the first word (*knife*) of the binomial (*knife and fork*) with an alternative word (*spoon*). Thus, binomials and their corresponding controls differed in the first content word, but shared the other two words, that is, the conjunction "and" and the second content word (e.g., *knife* <u>*and fork*</u> vs. *spoon* <u>*and fork*</u>). Control items were grammatically plausible but infrequent novel three-word combinations (*spoon and fork*). In addition, the two content words of the binomial and the control conditions were equally strongly associated. This resulted in 120 experimental stimuli (60 binomials and 60 controls, see Appendix 5). Examples of the materials for each condition are presented in Table 5-2.

Table 5-2 *Example of stimulus materials for each condition*

Condition	Binomial	Control
Congruent	*sun and* <u>*moon*</u>	*star and* <u>*moon*</u>
English-only	*bread and* <u>*butter*</u>	*toast and* <u>*butter*</u>
Chinese-only	*wisdom and* <u>*strength*</u> *智慧和力量*	*exercise and* <u>*strength*</u> *锻炼和力量*

5.1.2.1 The Binomials and Their Phrase Frequency. As the relative frequency of "*A and B*" vis-à-vis the reversed form "*B and A*" is central to formulaic status of binomials, the frequency-based approach was adopted to identify binomials. First, a set of English binomial expressions was identified by using the 560-million-word version of the Corpus of Contemporary American English (COCA: Davies, 2008). Second, these English binomial expressions were categorized into congruent binomials that exist in both English and Chinese, or incongruent English-only binomials that exist solely in English, based on the relative frequency of their Chinese translations vis-à-vis the reversed form of these translations in the corpus of Centre for Chinese Linguistics Peking University[5] (CCL: 437.5 million words, Zhan et al., 2003). Third, the Chinese corpus CCL was used to find a set of binomial expressions that exists solely in Chinese (i.e., Chinese-only binomials). In what follows below, the item development is described in some detail.

English binomial expressions were selected according to the following criteria. First, the target binomials had to be frequent word combinations. Critically, they had to be

[5] The CCL corpus contains 700 million Chinese characters. However, the character number is not equivalent to word number (e.g., 东西 dong xi, meaning 'thing(s)' in English, is considered as a two-character word). Following the Lancaster Corpus of Mandarin Chinese (McEnery & Xiao, 2004), I adopted a ratio of 1:1.6 between words and characters to calculate the number of words in CCL, which amounts to 437.5 million.

numerically more frequent than the reversed form. That is, they had to be frequent word combinations with word-order preference. If there was no word-order preference (e.g., *green and yellow* versus *yellow and green*), they would not be considered to be binomial expressions. Using the LIST search in COCA, all three-word strings formed by two content words from the same lexical class connected by the conjunction *and* (e.g., 'N and N', 'V and V', 'Adj and Adj') were first extracted. The first 1000 most frequent three-word combinations were chosen, which would provide me a broad enough range to select potential suitable items while avoiding very low-frequency word combinations. Second, binomials which cannot be used literally but only figuratively were excluded. The majority of items used in the experiment are literal, transparent phrases. Finally, these binomials and their reversed forms should have the same meaning (e.g., *knife and fork* means the same as *fork and knife*). This resulted in a total of over 200 English binomials.

After selecting English binomials, determinations regarding congruency and incongruency with Chinese were made. First, these English binomials were translated into Chinese, word by word. The translations were then checked by three high-proficiency Chinese-English bilinguals. They agreed that they were correct literal translations which used the core meaning of each word. Second, the frequencies of the Chinese translations and their reversed forms were extracted from CCL. An English binomial expression was deemed congruent with Chinese if the frequency of its Chinese translation was also numerically greater than that of the reversed form in the Chinese corpus, CCL. That is, there is the same word preference for a phrase to its reversed form in Chinese as in English. If, however, there was no significant difference in their frequencies, the English binomial was not considered to be a binomial in Chinese and, thus, was classified as an English-only binomial**.** For example, the binomial, *sun and moon* (太阳和月亮, taiyang he yueliang), is much more frequent than the reversed form, *moon and sun* (月亮和太阳, yueliang he taiyang), in English and Chinese:

30.54 versus 6.25 occurrences (per 100 million words) in COCA (English), and 38.63 versus 7.77 occurrences (per 100 million words) in CCL (Chinese). It can thus be classified as a true binomial both in English and Chinese. In contrast, the English binomial *bread and butter* is more frequent than the reversed form *butter and bread* in COCA (71.79 vs. 2.32 occurrences), whereas the Chinese translation equivalent for the binomial *bread and butter*, 面包和黄油 (mianbao he huangyou), is almost as frequent as the reversed form *butter and bread*, 黄油和面包 (huangyou he mianbao): 3.89 versus 1.83 occurrences in CCL. It is therefore classified as an English-only binomial. Finally, to ensure the differences in the processing of congruent and English-only binomials could be attributed to the difference in congruency rather than phrase frequency, I also matched congruent and English-only binomials for phrase frequency in English (congruent binomials: mean = 69.54, SD = 90.89; English-only binomials: mean = 106.90, SD = 217.56; t = -0.08, p = .94). The phrase frequency values were log-transformed prior to running t-tests. This resulted in a list of 20 congruent and 20 English-only binomials. Congruent binomials and their reversed forms differed in phrase frequency in English (binomials: mean = 69.54, SD = 90.89; reversed forms: mean = 8.63, SD = 9.46; t = 6.95, p < .0001) and Chinese (binomials: mean = 66.21, SD = 75.41; reversed forms: mean = 6.57, SD = 9.38; t = 6.97, p < .0001). English-only binomials differed from their reversed forms in phrase frequency (binomials: mean = 106.90, SD = 217.56; reversed forms: mean = 6.28, SD = 10.10; t = 7.20, p < .0001), whereas their Chinese translation equivalents were as frequent as their reversed forms (binomials: mean = 2.07, SD = 2.11; reversed forms: mean = 1.12, SD = 1.10; t =1.66, p = .11). In addition, to ensure that these binomials were well known to L1 English speakers, a group of 20 L1 speakers of English rated their familiarity with the binomials on a scale from 1 to 5, with 1 being “not familiar at all” and 5 being “extremely familiar”. These binomials were judged to be highly familiar (mean = 4.65). (Sample of the norming study is supplied in Appendix 6).

Finally, I selected Chinese-only binomials in CCL, using the same criteria as described above for the selection of English binomials. That is, the binomial expressions had to be frequent and, critically, more frequent than their reversed forms. They also had to be transparent and were required to have the same meaning as their reversed forms. It is important to note that unlike English binomials which have a fixed structure of *A and B*, Chinese binomials are more flexible in form, in that they can take the following three forms: *A and B*, *AB*, and *A*、*B* (e.g., *knife and fork*: 刀和叉，刀叉，刀、叉). This is related to the characteristics of Chinese language, which is a paratactic language. In such languages, connective elements are often optional or unnecessary (Li & Ho, 2016). Therefore, for binomials in Chinese, the word order is the most important attribute (i.e., A precedes B, rather than B precedes A), while the coordinator is not necessary. Thus, when identifying the frequency of occurrence of a Chinese phrase in CCL, I extracted its frequency in the forms of *A and B*, *AB*, and *A*、*B*, and used the sum of their frequency as the frequency of occurrence of this phrase. After I identified a set of Chinese binomials, I translated them into English character by character, adding conjunction 'and' for those binomials which do not usually include the coordinator '和' (=and). The translations were then verified by three high-proficiency Chinese-English bilinguals as correct literal translation. The frequencies of the English translations were then checked against the online version of the COCA to ensure that they were low frequency phrases (mean frequency = 0.02 per million) and they were not more frequent than their reversed forms (t = 0.29; p = .77). Chinese-only binomials and their reversed forms differed in phrase frequency in Chinese (binomials: mean = 213.17, SD = 353.71; reversed forms: mean = 7.09, SD = 11.72; t = 8.98, p < .0001), but not in English (binomials: mean = 1.97, SD = 2.12; reversed forms: mean = 1.53, SD = 1.60; t = 0.50, p = .62). Additionally, to ensure that these binomials were well known to L1 Chinese speakers, a group of 20 Chinese L1 speakers rated their familiarity with them (the binomials were in

Chinese) on a scale from 1 to 5 as in the norming test with L1 speakers of English and the mean was 4.65. This produced the list of translated Chinese-only binomials. (Sample of the norming study is supplied in Appendix 7.)

With regard to literality, most of the binomials used in the present study are literal phrases. A few binomials, however, can be used both literally and figuratively (e.g., *bread and butter*). In the congruent category, two items have a figurative and a literal meaning (*song and dance*, *thick and thin*). In the English-only category, three items have both a figurative and a literal meaning (*bread and butter*, *sticks and stones*, and *bed and breakfast*). In the Chinese-only category, three items have both a figurative and a literal meaning (*flowers and applause* [meaning 'success and recognition'], *dragon and phoenix* [meaning 'excellence'], and *wine and meat* [meaning 'good food and drink']).

5.1.2.2 Association Strength. As mentioned above, each binomial was paired with a control item, which differed from the binomial only in the first content word (*knife and fork* vs. *spoon and fork*). Critically, the two content words in the binomial and control conditions were also matched in semantic association strength (the forward association was measured). This was needed to ensure that any processing advantage for binomials over their corresponding controls was not due to the first word in the binomials (*knife*) being a better prime than the first word in the control items (*spoon*) for the same target (*fork*) (e.g., Siyanova-Chanturia, Conklin, & Van Heuven, 2011). Following Siyanova-Chanturia, Conklin and Van Heuven (2011), the University of South Florida (USF) Free Association Norms database (http://w3.usf.edu/FreeAssociation/) was used to match the constituents (i.e., the first content word and the second content word) of the binomials (*knife and fork*) and the control items (*spoon and fork*) in forward association strength (*knife-fork* vs. *spoon-fork*: 0.33 vs. 0.61). The components of congruent and English-only binomials and their corresponding controls were equally strongly associated in forward association strength (congruent condition, *t* = 1.41, *p* = .17; English-only condition, *t* = 1.75, *p* = .15). However, since the above Free Association Norms database is based on English, the association strength between the constituents of translated Chinese-only binomials and their corresponding controls could not be determined. In addition, no comparable Chinese database exists for the Chinese language. Therefore, I could only obtain association strengths for items which existed in the USF database. For Chinese-only items whose association strengths were not available in the database, their association strengths were given the value of 0.

5.1.2.3 Word Length and Frequency of the First Content Word. Additional steps were taken to ensure comparability across the binomial and control items. First, the first words in the binomial and control conditions were matched for part of speech. Second, word length of the first word was matched between the binomial and control items. There was no significant difference between the first word in the binomial and the control conditions for word length (congruent condition, *t* = -1.19, *p* = .24; English-only condition, *t* = -1.04, *p* = .28; Chinese-only condition, *t* = -0.07, *p* = .94). Additionally, the first words in the binomial and control conditions were also matched for lexical frequency (where possible). The first words in congruent and Chinese-only binomials and their corresponding control phrases were matched in terms of lexical frequency (congruent condition, *t* = 1.24, *p* = .23; Chinese-only condition, *t* = 0.52, *p* = .61). However, the first words in English-only binomials were significantly more frequent than those in the matched controls (*t* = 3.01, *p* = .005). It was impossible to create plausible control items matched in frequency as well as forward association strength. To partial out any possible effect of the first word's lexical frequency, the frequency of the first word was added as a covariate in the initial statistical model (see section 5.1.5). The properties of the experimental items are presented in Table 5-3.

Table 5-3 *Means (standard deviations) of phrase frequency, word length and frequency of first word, and semantic association strength for the binomial and control items (counts based on occurrences per 100 million words)*

	Congruent		English-only		Chinese-only	
	Binomial	Control	Binomial	Control	Binomial	Control
Phrase frequency (English corpus)	69.54 (90.89)	0.95 (1.03)	106.90 (217.56)	0.76 (0.94)	1.97 (2.12)	0.52 (0.57)
(Chinese corpus)	66.21 (75.41)	0.55 (1.21)	2.07 (2.11)	0.13 (0.23)	213.17 (353.71)	0.56 (1.28)
First word length	5.45 (1.70)	6.05 (1.70)	4.8 (1.28)	5.3 (1.56)	6.4 (2.28)	6.55 (2.48)
First word frequency	7636.60 (7400.19)	15633.22 (46006.48)	8559.96 (8485.87)	4544.17 (7898.84)	5700.86 (5918.46)	5698.47 (9866.77)
Association strength	0.24 (0.22)	0.17 (0.15)	0.14 (0.17)	0.07 (0.09)	0.01 (0.01)	0.01 (0.03)[6]

5.1.2.4 Fillers and Non-Word Items. In addition to the two experimental conditions (binomial versus control), a set of filler phrases with the same syntactic structure (e.g., 'N and N', 'V and V', 'Adj and Adj') as binomials was constructed to reduce the proportion of related prime-target pairs, following 1/5 ratio (fillers/target items = 1/5) proposed by McNamara (2005). The filler phrases were grammatical but semantically implausible (e.g., *business and soul*), whereas the control items were grammatical and plausible (*spoon and fork*). Non-word items were created to make an equal number of word/non-word responses (i.e., 50/50 word/non-word ratio), with the syntactic structure of *word* + *and* + *non-word*. It was done to avoid a task-wide bias toward the word/nonword response (Perea & Rosa, 2002).

[6] For items in the Chinese-only category, I could only obtain the association strength between the constituents of 6 binomials and 4 control phrases. The values reported were based on the 10 items which existed in USF norm database.

All non-words came from the ARC nonword database (Rastle et al., 2002). They conformed to the phonotactic rules of English and were matched with the other items for length (mean = 5.88 letters). Primes for the non-word targets were words that were not used in other conditions. (See Appendix 8 for fillers and non-word items used in the experiment.)

5.1.3 Design

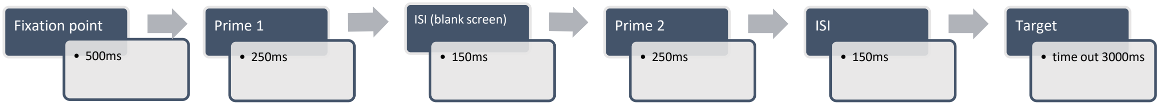
A within-participant design was used, with each participant exposed to the critical items in both conditions (binomial and control); this allowed for a within-participant comparison of response times in the two experimental conditions (binomial and control), providing better control for individual differences (Millar, 2011). Individuals differ considerably in reading speed, varying from 260 words per minute for average readers to 586 words per minute for fast readers (Jackson & McClelland, 1975). Many psycholinguistic experiments have adopted such a design (Finkbeiner et al., 2004; Siyanova-Chanturia et al., 2017; Wolter & Gyllstad, 2011).

Such a design would, however, be likely to lead to repetition priming due to prior exposure to the same target stimulus (Gabrieli, 1998). To control for the repetition effect, two counterbalanced presentation lists were constructed with a binomial and its control appearing in opposite lists. For example, when *knife and fork* appeared on List A, its corresponding control *spoon and fork* appeared on List B. Half of the critical targets per list were presented as binomials and half as control phrases. In addition, the numbers of stimuli of each congruency type in each list were also balanced, such that each list contained an equal number of congruent, English-only, and Chinese-only binomials. Participants were randomly assigned to one of the two groups in the order of their participation. Group 1 saw List A first and then List B, and for Group 2 the order was reversed. The same number of participants was assigned to each group.

A primed visual lexical decision task (LDT) was employed in the experiment whereby the initial two words of the binomial were the prime and the last word was the target. Participants had to decide whether the last string of letters was a real English word or not by pressing the appropriate key (YES or NO) on a response box. Research has shown that prime-target stimulus onset asynchrony (SOA) can affect the priming effects in a LDT (e.g., de Groot et al., 1986). Previous studies on the online processing of MWEs using priming paradigms varied considerably in their prime-target SOA. For example, a 150-ms SOA was adopted in Sonbul and Schmitt (2013), a 300-ms SOA in Wolter and Gyllstad (2011), and a 600-ms SOA in Durrant and Doherty (2010), which all used the first word of two-word collocations as the prime and the second word as the target and reported a significant collocational priming. In another study where a prime consisted of two words (a verb + noun phrase), a 1300-ms SOA was adopted for L1 speakers and a 2000-ms SOA for L2 speakers (Matsuno, 2017). Moreover, in cases where the prime was more than two words (e.g., 'open your books to', 'where are the': Ellis et al., 2008), a prime-target SOA as long as 3000 ms was used. Also, in an ERP study looking at the processing of binomials (e.g., *knife and fork* vs. *spoon and fork*: Siyanova-Chanturia et al., 2017), the authors presented the words using rapid serial visual representation, that is, one word at a time. They set the presentation duration of each word at 300 ms and inter-stimulus interval (ISI) at 200 ms in each trial. However, a short SOA (e.g., 200 ms or less) was advised by McNamara (2005) if the automatic component of semantic priming is investigated. Nevertheless, I was not concerned with the question whether automatic priming or strategic priming were involved (also see Carrol & Conklin, 2014; Wolter & Gyllstad, 2013). With these considerations in mind, I decided to present each prime word for 250 ms and set ISI at 150 ms. As a result, the prime-target SOA is as long as 800 ms.

5.1.4 Procedure

Each trial consisted of a fixation point for 500 ms, the first prime word for 250 ms, an inter-stimulus interval (black screen) for 150 ms, the second prime word for 250 ms, another inter-stimulus interval for 150 ms, and a target word or nonword for 3000 ms. All items were presented in the middle of the screen in light grey lowercase letters in Courier New font, size 24 pt, over a black background. At the start of each trial, a fixation point ("+++++", always five characters wide) was presented in the middle of the screen for 500 ms. It was replaced with the first word prime (*knife* in "*knife and fork*"), which was displayed for 250 ms. After that, a blank screen (i.e., inter-stimulus interval/ISI) was presented for 150 ms. Then the second word prime "*and*" was displayed for 250 ms, followed by the same ISI (150 ms). Finally, the target was shown for up to 3 seconds or until a response was registered, whichever came first. The procedure is summarized in the following diagram:



The experiment was conducted in a laboratory using DMDX software (Forster & Forster, 2003) on a HP Omen Laptop 15. Participants were assigned to one of the two presentation lists randomly. They first read instructions on the computer screen and then completed 20 practice trials before the experimental trials. Participants were instructed to decide whether the target was a word or not in English by pressing one of the two triggers on a Logitech cordless rumblepad 2 (right trigger for YES and left trigger for NO, and vice versa for left-handed participants), with the dominant hand for the YES button. They were asked to respond as quickly and accurately as possible. The items were presented in two counterbalanced blocks of 154 trials each, with a self-paced break after the first block. Within each block, the trial order was randomized for each participant. I also ensured that a binomial and its control were not close together (e.g., if a binomial came at the end of the first block, its control would not come at the beginning of the second block). Upon the completion of the experiment, an exit interview was given. The whole experiment took approximately 20 minutes to complete.

5.1.5 Analysis and Results

I analysed accuracy and response latencies (RT). In the accuracy analysis, all responses were included. The mean response accuracy to non-word items was 82.18% for the Chinese-English bilinguals. On word trials, the mean accuracy was 97.27% for the Chinese-English bilinguals. However, there was no significant difference in response accuracy between the binomial and control conditions for the three congruency types (i.e., congruent, English-only, and Chinese-only). That is, there was no response accuracy priming for any of the three types of binomials. The accuracy from Experiment 1a is presented in Table 5-4.

Table 5-4 *Descriptive statistics: Accuracy (%) for Chinese-English bilinguals*

Chinese-only		Congruent		English-only		Fillers	Nonwords
binomial	control	binomial	control	binomial	control		
96.54	95.38	98.65	98.46	98.46	98.65	95.76	82.18

For the RT analyses, the data for non-word and filler items and for incorrect responses were excluded from the analysis. I performed the analyses on RTs to 60 binomials (20 items for each congruency type: congruent, English-only, Chinese-only) and their corresponding controls (120 items in total). This is because my objective was to compare RTs to the terminal words of binomials (*knife and fork*) with RTs to the terminal words of control phrases (*spoon and fork*), in order to examine whether or not binomial priming effects would occur. Incorrect responses were removed from the RT analysis, resulting in the loss of 2.31% data. Extreme values (RTs longer than 2000 ms or shorter than 250 ms) were also excluded (e.g., Matsuno, 2017; Sprenger et al., 2006), which resulted in the loss of 0.56% data. Means of RTs (descriptive statistics) by condition for Chinese-English bilingual participants are shown in Table 5-5.

Table 5-5 *Descriptive statistics: Mean response times in ms (standard deviations) and difference between mean response times to the binomial and control phrases for Chinese-English bilinguals in each of the six experimental conditions*

	Binomial	Control	Difference
Congruent	637.84 (221.73)	657.86 (213.51)	20.02
English-only	641.41 (204.58)	645.09 (205.91)	3.68
Chinese-only	676.97 (241.64)	669.25 (222.02)	-7.72

Note. Difference is calculated with mean RT to controls minus mean RT to binomials.

RT data was analysed with linear mixed effects models (LMEM) using *lme4* package (Bates et al., 2015) and *lmerTest* (Kuznetsova et al., 2017) packages. I opted for LMEM because they can account for the variability of the fixed effects across experimental participants and linguistic items by including participants and items as random effects (Barr et al., 2013; Winter, 2013). In the experiment, multiple observations were taken for each participant and for each item, thus data points were not truly independent. RTs from the same participant and for the same item might be correlated. Thus, the variations coming from participants and items needed to be controlled for. LMEMs have the ability to simultaneously handle *crossed* random effects (i.e., all the participants have experienced all the levels of the fixed effects) (Barr et al., 2013). Therefore, it became the method of choice in this study.

Following Brysbaert and Stevens (2018), RTs were inverse transformed (i.e., -1000/RT) to bring the data closer to normal distribution. Inverse-transformed RTs were normally distributed, with skewness of -0.21 and kurtosis of 0.52. Inverse transformed RT to the final word of each phrase was used as the response variable.

The maximal model used inverse-transformed RTs as the response variable and incorporated items and participants as random intercepts. I did not include random slopes at this stage, because the maximal model with random slopes failed to converge. The following variables were included as fixed effects: (1) item type (binomial vs. control), (2) congruency (congruent vs. English-only vs. Chinese-only), (3) English phrase frequency (counts based on occurrences per 100 million words, log transformed), (4) the frequency of the first content word of a phrase (counts based on occurrences per 100 million words, log transformed), and (5) forward association strength between the first word and the last word of a phrase (based on USF database, log transformed). I log transformed all continuous covariates because log-transforming can help address the issue of skewness and make the regression model adhere to the assumption of normality (Winter, 2019). In addition, the value "1" was added to each association strength score before log transforming them, since many Chinese-only items had association strengths with the value of "0". It is important to note that phrase frequency and item type (binomial versus control) may be collinear (i.e., correlated). In order to address the issue of the collinearity between phrase frequency and item type, I orthogonalized phrase frequency by fitting a linear model in which phrase frequency was predicted by item type, following Siyanova-Chanturia, Conklin and Van Heuven (2011). The residuals of this model (EngPhrFreq.Residual) were then used as the predictor of phrase frequency, such that effects of item type were partialed out. Additionally, block order (order in which participants saw the two presentation lists: Order 1 vs. Order 2) and the trial number of the presentation of the phrase in the experiment (scaled) were considered as fixed effects to account for repetition priming and the longitudinal effect of the experimental task on the behaviour of the participants. This was done because order effects (e.g., participants usually respond faster as a result of practice or perform differently at the end of an experiment because of being bored or tired) are of special concern in within-participant designs (Shaughnessy et al., 2000). Trial number was scaled to normalize the data and to ensure that the estimated coefficients were all on the same scale, making it easier to compare effect sizes (Howell, 2012). Moreover, the maximal model also included the following interactions between the fixed effects: (1) item type and congruency, (2) item type and phrase frequency, (3) item type and the frequency of the first content word, (4) congruency and phrase frequency, and (5) association strength and congruency. The initial model was as below:

> RTinv ~ ItemType*Congruency + EngPhrFreq.Resid *ItemType + EngPhrFreq.Resid * Congruency + Wrd1Freq.log.c*ItemType + AssoStrength.log*Congruency + TrialNum.sc + BlockOrder + (1| Participant) + (1| Target)

Starting with the maximal model, I used the step() function in *lmerTest* to arrive at the best model fit. After fitting the best model, I conducted a forward stepwise model selection to identify the appropriate random effects structure with random slopes, using the ranova() function (Kuznetsova et al., 2017) and using Akaike information criterion (AIC) values. A design-driven approach was adopted to construct random effects structure (Barr et al., 2013). In this study, participants saw two English strings of letters (*knife – and*) and then decided whether or not the third letter string (*fork*) was a legal English word. Each participant was presented with the same target word (*fork*) twice, when it was preceded by the first two words in a binomial (<u>*knife and*</u> *fork*) and by the first two words in its control phrase (<u>*spoon and*</u> *fork*). This manipulation was *within-participants* as each participant was measured to both levels of item type (binomial and control). It was also a *within-item* manipulation because the targets (*fork*) were held constant across conditions (binomial and control), but the prime words varied (*knife and* vs. *spoon and*). Within-participants and within-items design would call for by-participant and by-item random-intercepts-and-slopes models (Barr et al., 2013). That is, LMEMs with maximal random effects structure should be used. Thus, I included by-item random slopes for predictors which vary within participants, including item type, congruency, word 1 frequency and trial number. With regards to random effects of item, I included by-participant random slopes for predictors which vary within items, including item type, word 1 frequency and trial number. For all analyses, the α level for testing the fixed-effect slope was set to .05. In this manner, I identified the following best-fit model for Chinese-English participants with random intercepts and random slopes:

RTinv ~ ItemType * Congruency + AssoStrength.log + TrialNum.sc + (1 + TrialNum.sc | Participant) + (1 + EngPhrFreq.Resid + TrialNum.sc | Target)

After identifying the best model with random slopes, I visually inspected a quantile-quantile plot of the model's residuals to assess whether the residuals were normally distributed. However, there were some deviations from the line and the residuals did not conform with normal distribution. Therefore, I removed 2.5 SD from the residuals to satisfy the assumption of homoscedasticity and normal distribution, which resulted in the loss of 1.85% data. I fitted the model to the new data again. The Q-Q plot (as shown in Figure 5-1) showed that the residuals conformed the assumption of normal distribution this time. To assess the constant variance assumption, the residuals were plotted with residuals (y-axis) against the fitted values (x-axis). The residual plot (as shown in Figure 5-2) showed that the spread of the residuals was approximately equal across the range of fitted values. Thus, the model satisfied the normality and constant variance assumptions.

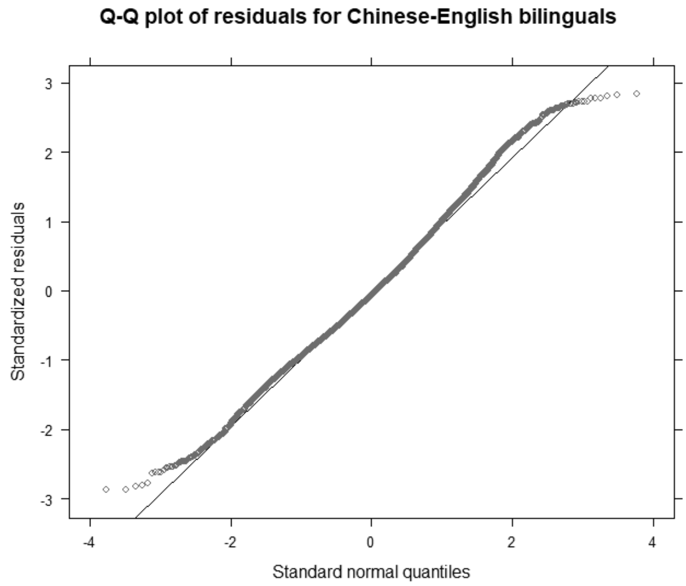


Figure 5-1 *Q-Q plot of residuals for Chinese-English bilinguals*

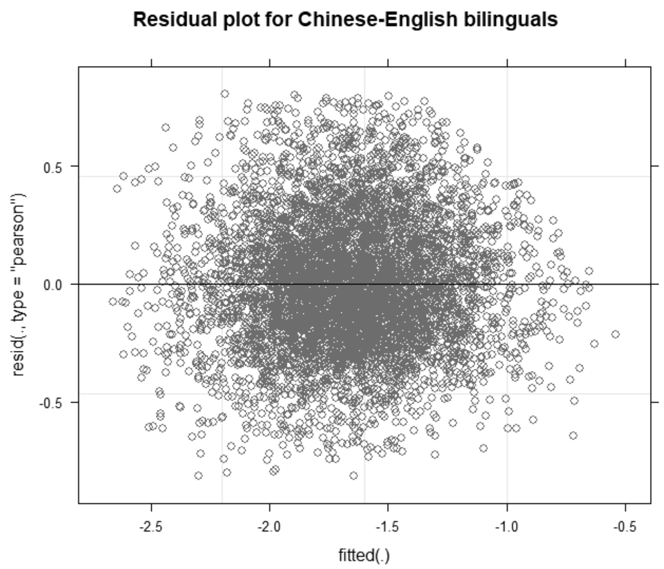


Figure 5-2 *Residual plot for Chinese-English bilinguals*

The results for the identified model for the Chinese-English bilinguals are shown in Table 5-6.

Table 5-6 *Results of mixed model for Chinese-English bilinguals*

Fixed effect	Estimate	Std.Error	df	*t*	*p*
Intercept	-1.61	0.05	118.70	-29.81	<.001
ItemType (binomial)	-0.05	0.02	159.50	-3.39	<.001
Congruency (Chinese-only)	0.01	0.06	61.70	0.09	.93
Congruency (English-only)	-0.06	0.06	58.76	-1.01	.32
AssoStrength.log	-0.18	0.08	139.03	-2.27	.02
TrialNum.sc	-0.08	0.01	52.99	-7.53	<.001
ItemType (binomial) * Congruency (Chinese-only)	0.06	0.02	58.92	2.38	.02
ItemType (binomial) * Congruency (English-only)	0.06	0.02	186.20	2.63	.009
Random effects	Variance	SD			
Target	0.03	0.18			
EngPhrFreq.Resid | Target	0.001	0.03			
TrialNum.sc | Target	0.0001	0.01			
Participant	0.06	0.24			
TrialNum.sc | Participant	0.004	0.07			
Residual	0.08	0.28			

Note. df = degrees of freedom; Intercept levels: ItemType = Control; Congruency = Congruent; Marginal R^2 = 0.04, Conditional R^2 = 0.57.

There was a significant two-way interaction between item type and congruency (χ^2 = 8.70, *p* = .012) from model comparison performed with the anova() function. There were also statistically significant main effects of association strength (χ^2= 5.21, *p* = .02) and trial number (χ^2 = 39.13, p < .0001). This means that words within more strongly associated phrases had overall shorter response latencies and that as the number of trials increased, the response time became faster. The significant interaction between item type and congruency was explored through post-hoc comparisons using the emmeans() function in the *emmeans* package (Lenth, 2019), with Bonferroni adjustments. The results revealed significant differences between binomials and controls in the congruent condition. The two-way interaction (Figure 5-3) showed that the Chinese-English bilinguals processed congruent binomials significantly faster than the controls, but there was no difference between their processing of English-only binomials versus controls, nor any difference between translated Chinese-only binomials versus controls. That is, only congruent binomials showed a priming effect. The results are shown in Table 5-7.

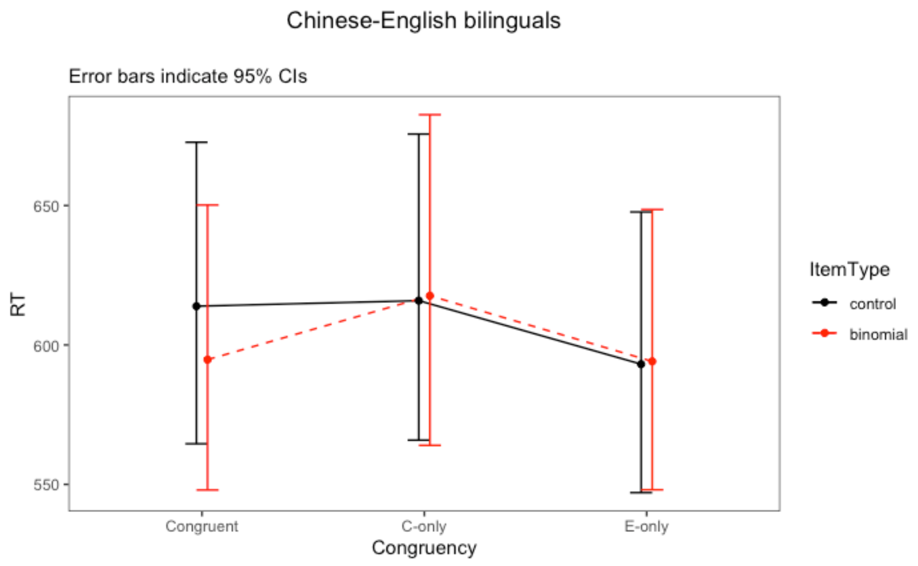


Figure 5-3 *Interaction plot of Item type * Congruency for Chinese-English bilinguals*

Note: RTs were back-transformed data (unit = ms).

Table 5-7 *Results of post-hoc test of RTs for congruent, English-only, and Chinese-only binomial items relative to the control items for Chinese-English bilinguals*

Contrast	Estimate	Std.Error	*t*	*p*	ED (ms)
Control - Congruent	0.053	0.016	3.388	.013	19
Control - English-only	-0.003	0.015	-0.187	.99	-1
Control - Chinese-only	-0.005	0.018	-0.253	.99	-2

Note. ED (estimated difference) is calculated with RTs to controls minus RTs to binomials.

For the Chinese-English bilinguals, the priming effect was only observed in the congruent condition (*t* = 3.39, *p* = .013), with RT[7] to the terminal word in the binomials 19 ms faster than RT to the control items (595 vs. 614 ms, respectively). No priming effect was present for the English-only binomials (*t* < 1, *p* = .99) nor translated Chinese-only (*t* < 1, *p* = .99) binomials: there was no difference between the binomials and the control items in the English-only condition (594 vs. 593 ms) nor in the translated Chinese-only condition (618 vs. 616 ms). This suggests that only congruent expressions were processed as binomials, whereas English-only and translated Chinese-only items were not.

In sum, the relative processing advantage for congruent over English-only binomials compared to their corresponding controls was found for the Chinese-English participants. In other words, the congruent binomials had a processing advantage over the English-only binomials for the Chinese-English bilinguals, even though the two types of binomials had been matched in phrase frequency in L2/English. Critically, no processing advantage for translated Chinese-only binomials over matched controls was found for the Chinese-English participants, suggesting that there was no automatic activation of the L1/Chinese.

5.1.6 Discussion

The aim of Experient la was to examine whether or not the Chinese-English bilinguals' L1 influenced the processing of frequent and familiar L2 word sequences. I focused on binomial expressions, i.e., literal and compositional MWEs, which have so far received little attention in bilingual processing research. To this end, I used a primed lexical decision task to examine the processing of congruent (*sun and moon*), English-only (*bread and butter*), and translated Chinese-only (*wisdom and strength*) binomial expressions versus

[7] Here and the following reported in-text are back-transformed model estimates. I reported descriptive statistics calculated prior to the data analyses in Table 5-5.

infrequent but equally strongly associated control phrases (*star and moon*, *toast and butter*, *exercise and strength*, respectively).

5.1.6.1 The Congruency Effect in the Processing of Binomials in the L2. The Chinese-English bilingual participants showed significant facilitation in the processing of the final word in congruent binomial phrases (*sun and moon*) compared to control phrases (*star and moon*), with a positive priming effect of 19 ms. However, they exhibited no priming effect in the processing of English-only binomials (*bread and butter*) (-1 ms), although the two types of binomials were matched for English phrase frequency (t = -0.08, p = .94). In other words, the Chinese-English bilinguals in this experiment showed a processing advantage for congruent binomials over controls, but not for English-only binomials over controls, reflecting the effect of congruency on the processing of congruent L2 binomials. This result is consistent with previous studies which found that bilinguals show an advantage in the processing of congruent over incongruent, L2-only idioms (Carrol et al., 2016) and collocations (Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2018). For example, in an eye-tracking study, Carrol et al. (2016) found that L2 speakers showed facilitation for the form and meaning integration for congruent idioms (*break the ice*) relative to matched controls (*crack the ice*), while they only showed facilitation for the meaning integration for L2-only idioms (*kick the bucket*) compared to their corresponding matched literal controls (*drop the bucket*). Similarly, Wolter and Gyllstad (2011, 2013) found that Swedish-English bilinguals showed a congruency advantage in the processing of congruent over L2-only collocations, whereas English monolinguals processed congruent and L2-only collocations in a similar way. Since congruent and English-only binomials were matched in English phrase frequency in the present study, the greater processing advantage for congruent binomials over English-only binomials (relative to their corresponding controls) by the Chinese-English bilinguals could be taken as evidence for the congruency effect. That is, cross-language influence was observed in the processing of congruent L2 binomials by the Chinese-English bilinguals.

5.1.6.2 Processing of Translated L1-only Binomials. With regard to the question whether or not the facilitatory effect of L1 extends to translated L1-only binomials, I compared the Chinese-English bilinguals' RTs to the terminal words in the translated Chinese-only binomials (*wisdom and strength*) versus controls (*exercise and strength*). The Chinese-English bilinguals showed no priming for the translated Chinese-only binomials (cf. RQ3), suggesting that known L1 MWEs are not automatically activated in L2 processing.

This finding is inconsistent with the L1 MWE activation account of L2 processing, used to explain the findings in previous studies on the processing of translated L1-only *idioms* (e.g., Carrol & Conklin, 2014, 2017; Carrol et al., 2016). Using priming (2014) and eye-movement (2017) paradigms, Carrol and Conklin reported a processing advantage for translated Chinese-only idioms over matched controls with Chinese-English bilinguals in the two studies. Moreover, in an eye-tacking study with a different group of bilinguals – Swedish-English bilinguals, Carrol et al. (2016) replicated and extended this finding. They found that translated Swedish-only idioms (*play monkey*) showed a processing advantage over matched controls (*taste monkey*). Crucially, the processing advantage was comparable with that for congruent idioms (*lose your head*) relative to matched controls (*hurt your head*). That is, there was no more facilitation for congruent idioms than for Swedish-only ones due to their additional experience in the L2. This led them to conclude that L1 MWE knowledge is the key determinant of how translation equivalents are processed in an L2, over and above direct experience in an L2.

However, the finding that translated Chinese-only binomials were not processed faster than matched controls is consistent with the results reported in Wolter and Yamashita (2015, 2018). Wolter and Yamashita did not observe a processing advantage for translated Japanese-only collocations (*buy anger*) compared to non-collocational matched controls (*bad gift*) with Japanese-English bilinguals in two response-based tasks: a double lexical decision task (2015) and an acceptability judgment task (2018). Both in the present study and in Wolter and Yamashita (2015, 2018), translated L1-only MWEs were processed as unknown word combinations since processing of the translated L1-only collocations was not different from controls, suggesting that there was no automatic activation of known L1 MWEs in L2 processing. The absence of priming for translated L1-only MWEs thus seems to support the L2 MWE experience account. In this account, although having an equivalent form in the L1 facilitates the acquisition of a (congruent) L2 MWE, prior exposure to L2 word sequences is necessary for them to show significant facilitation compared to control novel phrases. This view is in line with usage- and exemplar-based acquisition and processing accounts, which assume that frequency of encounters with and use of a lexical item (words, MWEs) determines quality of its mental representations and its ease of processing (Bybee, 2006; Langacker, 2000). A number of empirical studies have shown that frequency plays a key role in MWE processing (e.g., Arnon & Snider, 2010; Siyanova-Chanturia, Conklin, & Van Heuven, 2011). Due to their frequency, MWEs are processed faster than matched novel phrases by L1 as well as L2 speakers (for a review, see Siyanova-Chanturia & Van Lancker Sidtis, 2019). Since translated L1-only MWEs have very low frequency (close to zero) in an L2, they are unlikely to show a phrase frequency effect in an L2.

The discrepancy between the present results and those of Carrol and colleagues may be due to the type of MWEs (i.e., binomials versus idioms) and to the methodological differences between the studies (response time versus eye-movement). According to Carrol and Conklin (2020), different types of MWEs have different properties contributing to their processing advantage relative to control phrases. For example, familiarity and decomposability are relevant to the processing of idioms, predictability and semantic association for binomials, and mutual information for collocations (Carrol & Conklin, 2020, p. 95). Therefore, it is possible that the conflicting findings are related to the type of MWEs being processed, literal (collocations and binomials) versus figurative (idioms) expressions. That is, mostly figurative (e.g., idioms) and mostly literal (e.g., some collocations and binomials) language may be processed differently. Since idioms have a conventional figurative phrase meaning which is different from literal meaning of the individual words, figurative meaning of idioms has to be activated in their processing (i.e., meaning activation), in addition to the recognition of specific word combinations presented in a particular order or configuration (i.e., form activation) (Carrol et al., 2016). However, for literal MWEs (e.g., collocations and binomials), the source of the processing advantage is likely due to form activation. The locus of form activation is lexical, while the locus of meaning activation is semantic or conceptual. The nature of CLI may thus differ in accordance with the type(s) of knowledge involved. Thus, for L1-only literal MWEs translated into L2, they may not show any processing advantage relative to control phrases, because bilinguals are likely to be unfamiliar with the translated forms. However, when L1-only figurative MWEs are presented in their unfamiliar translated forms, figurative meaning of L1-only figurative MWEs may still be activated in their processing, although the form of figurative MWEs may not be recognized. In other words, it is possible that the facilitation for translated L1-only idioms is driven by the conceptual L1-L2 overlap, in which case, the figurative meaning of idioms is activated even if they have not been encountered in the L2 (Carrol et al., 2016, but see Carrol and Conklin, 2017).

However, the hypothesis that facilitation for translated L1-only idioms is driven by the conceptual L1-L2 overlap needs further verification. First, evidence is mixed with regard to whether or not the figurative meaning of translated L1-only idioms is activated in L2 processing. Carrol et al. (2016) found that advanced Swedish-English bilinguals showed a processing advantage for translated Swedish-only idioms (*play monkey* vs. *taste monkey*) in the phrase-level analysis, as evidenced by shorter total reading times and fewer fixations on the whole phrase. They took the results as evidence that the figurative meaning of translated L1-only idioms is activated in the processing in bilingual speakers. However, Carrol and Conklin (2017) found no evidence for access to the figurative meaning of translated L1-only idioms (Experiment 2). Chinese-English bilingual speakers read idiomatic expressions used figuratively more slowly than those used literally, suggesting that bilinguals might not know the figurative meanings of idioms, or they only access the figurative meanings after the literal meaning had been rejected. Second, there is also evidence for form activation in the processing of translated L1-only idioms (Carrol & Conklin, 2017; Carrol et al., 2016). For example, Carrol and Conklin (2017) found that Chinese-English bilinguals showed significant facilitation for the final word of translated L1-only idioms compared to control phrases, evidenced in the analysis of first fixation durations and total reading time. Carrol et al. (2016) found that Swedish-English bilinguals showed a processing advantage for translated Swedish-only idioms (*play monkey*) over controls (*taste monkey*) in the word-level analysis, as suggested by the likelihood of skipping on the final word. Therefore, it is unclear whether the processing advantage for translated L1-only idioms is driven by form activation or meaning activation, or even both. Importantly, if form activation occurs in the processing of translated L1-only idioms via the lexical-translation route, as suggested by Carrol and Conklin (2014), one may wonder why form activation does not occur in the processing of translated L1-only collocations or binomials. Therefore, to test the hypothesis that the facilitation for translated L1-only idioms is driven by meaning activation (because of conceptual L1-L2 overlap), rather than by form activation, future research should investigate the processing of L1-only idioms which are paraphrased into the L2, rather than translated word-by-word (e.g., instead of, *draw a snake and add feet*, researchers may use, *add feet to a* *snake*). If the modified translated L1-only idioms, which retain their figurative meanings, show a processing advantage over control novel phrases, the conceptual-overlap priming explanation would be confirmed.

Differences in methodology may also cause the discrepancy between studies with idioms and the present study. Most studies with idioms investigated the processing of idioms embedded in sentence contexts (Carrol & Conklin, 2017; Carrol et al., 2016), while the present study investigated the processing of binomials in isolation. A biasing context could have contributed to the facilitation for translated L1-only idioms over matched controls, by greatly increasing predictability in the processing of idioms (Cieślicka, 2013; Titone & Connine, 1999). (Note, however, that Carrol et al. (2016) found a processing advantage for translated L1-only idioms over control phrases for L2 speakers without using biasing contexts.) In addition, idioms used in the reviewed studies (Carrol & Conklin, 2014, 2017) are relatively long (e.g., *draw a snake and add … feet*), which may have allowed participants to actively anticipate the completion of a phrase (Carrol et al., 2016). Critically, most of the studies that found facilitation for translated L1-only idioms employed eye-tracking, while studies with collocations and binomials employed behavioural measures such as phrase-acceptability judgment tasks and primed lexical decision tasks. Eye-tracking allows researchers to analyse separately early and late stages of reading (Rayner, 2009; Siyanova-Chanturia, 2013), as well as look at different areas of interest. Early measures reflect how easily the expected word combinations are activated, and late measures indicate how easily the overall meaning is activated and integrated into the context (Carrol et al., 2016). Unlike eye movements, primed lexical decisions mainly assess automatic lexical level activation, and are less sensitive to conceptual meaning activation (Du et al., 2021).

To sum up, in Experiment 1a, congruent binomials showed a processing advantage over incongruent, English-only items. This result reaffirms the influence of congruency on the processing of L2 MWEs reported in previous studies (e.g., Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2018; Yamashita & Jiang, 2010). Importantly, translated L1-only binomials showed no processing advantage over matched controls, which supports the L2 MWE experience account, but goes against the L1 MWE activation account. This result thus highlights the role of phrase frequency in L2 processing and argues against the view that known L1 MWEs are automatically activated in L2 processing.

5.2 Experiment 1b

In Experiment 1a, the Chinese-English bilinguals showed significant priming for congruent binomials (*sun and moon*) relative to infrequent but equally strongly associated controls (*star and moon*), but no priming for English-only binomials (*bread and butter*) or translated Chinese-only binomials (*wisdom and strength*) versus matched controls (*toast and butter*, *exercise and strength*, respectively). Using the same stimuli and task, the aim of Experiment 1b was to determine if (a) English monolinguals exhibit significant priming effects for any of the three item types (i.e., congruent, English-only, and translated Chinese-only binomials), and (b) if there are significant differences in terms of priming effects among the three item types. For the first research question, I predicted that binomials that exist in English (i.e., congruent and English-only binomials) would produce a significant priming effect, and that binomials that exist solely in Chinese but not in English (i.e., translated Chinese-only binomials) would not produce any significant priming effect in English monolinguals. This was predicted by the usage-based models of language acquisition and processing, according to which language speakers are sensitive to MWE frequency and thus process frequently occurring sequences faster than less frequent ones as a result of repeated exposure (Arnon & Cohen Priva, 2014; Ellis, 2002). For the second research question, I predicted that no significant difference in priming effect would be found between congruent and English-only binomials, because English monolinguals would not be aware of or affected by the distinction between congruent and L2-only items, i.e., the two item types (congruent and English-only binomials) are equivalent from an L1 monolingual's perspective. If this is confirmed, the difference in processing between congruent and English-only binomials observed for Chinese-English bilinguals can be attributed to the different status of the items held in the Chinese-English bilinguals' L1.

5.2.1 Participants

Fifty-two English monolingual speakers participated in the experiment (33 females; age 18-43; mean age = 23.85 years). Participants were recruited through Victoria University of Wellington networks, through advertising posters, university Facebook pages and personal contacts. They were undergraduate and postgraduate university students, as well as young professionals at VUW. All participants were introduced to the research and shown an information sheet about the study (see Appendix 2). Each participant gave informed consent and received $10 for their participation in the experiment. They completed a language background questionnaire (see Appendix 4) before the experiment to make sure they had no knowledge of Chinese. English monolinguals identified themselves as not being proficient in any language other than English, although some had studied a language other than Chinese. All participants had normal or corrected-to-normal vision.

5.2.2 Materials and Design

The materials and design were the same as Experiment 1a.

5.2.3 Procedure

The procedure was the same as Experiment 1a.

5.2.4 Analysis and Results

Data analysis was performed in the same way as in Experiment 1a. Namely, accuracy and response latencies (RT) were analysed. For the English monolinguals, the mean response accuracy to non-word items was 95.94%, and the mean accuracy to word items was 98.39%. Similar to the results of Experiment 1a, there was no significant difference in response accuracy between the binomial and control conditions for any of the three congruency types (i.e., congruent, English-only, and translated Chinese-only). The accuracy from Experiment 1b is presented in Table 5-8.

Table 5-8 *Descriptive statistics: Accuracy (%) for English monolinguals*

Chinese-only		Congruent		English-only		Fillers	Nonwords
binomial	control	binomial	control	binomial	control		
99.04	98.94	99.04	98.94	98.95	98.46	96.66	95.94

For the RT analyses, the data for the non-word and filler items and for incorrect responses were excluded from the analysis as in Experiment 1a. In addition, incorrect responses were removed from the RT analysis, resulting in the loss of 1.12% data for English monolinguals. Extreme values (RTs longer than 2000 ms or shorter than 250 ms) were also excluded, which resulted in the loss of 0.14% data. Means of RTs (descriptive statistics) by condition for English monolingual participants are shown in Table 5-9.

Table 5-9 *Descriptive statistics: Mean response times in ms (standard deviations) and difference between mean response times to the binomial and control phrases for English monolinguals in each of the six experimental conditions*

	Binomial	Control	Difference
Congruent	479.42 (125.88)	505.11 (134.61)	25.69
English-only	471.07 (132.57)	492.68 (129.82)	21.63
Chinese-only	519.12 (138.14)	513.56 (128.80)	-5.56

Note. Difference is calculated with mean RT to controls minus mean RT to binomials.

In Experiment 1b, RTs were also inverse transformed to bring the data closer to normal distribution as in Experiment 1a. Inverse-transformed RTs were normally distributed, with skewness of 0.10 and kurtosis of 0.05. I started with the same maximal model as in

Experiment 1a (see section 5.1.5) and adopted the same procedure to identify the best-fit model for English monolinguals with random intercepts and random slopes, which is shown as follows:

RTinv ~ ItemType * Congruency + ItemType * Wrd1Freq.log.c + EngPhrFreq.Resid + AssoStrength.log + TrialNum.sc + BlockOrder + (1 + TrialNum.sc | Participant) + (1 | Target)

After identifying the best model with random slopes, I first visually inspected a quantile-quantile plot of the model's residuals to assess whether the residuals were normally distributed. Then I removed 2.5 SD from the residuals to satisfy the assumption of homoscedasticity and normal distribution, which resulted in the loss of 1.83% data for English monolinguals. The new Q-Q plot (Figure 5-4) and the residual plot (Figure 5-5) showed that the model satisfied the normality and constant variance assumptions.

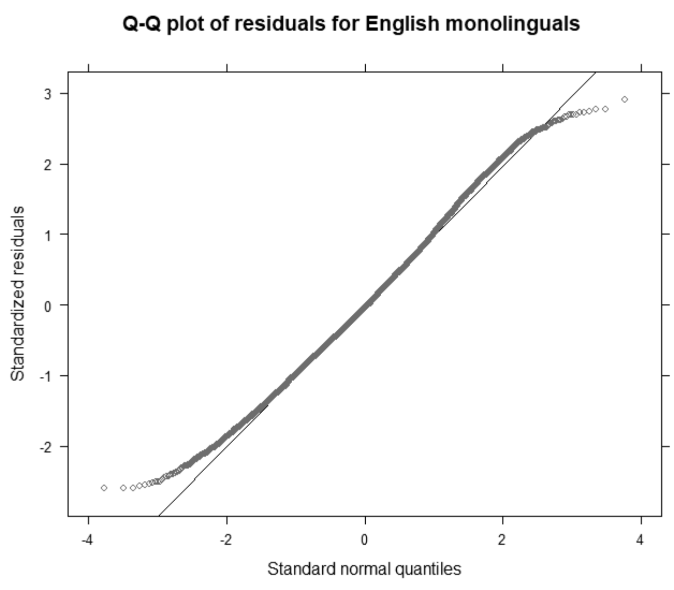


Figure 5-4 *Q-Q plot of residuals for English monolinguals*

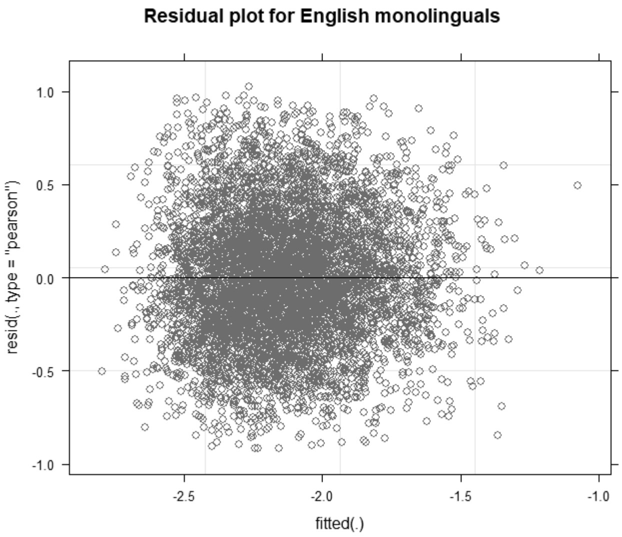


Figure 5-5 *Residual plot for English monolinguals*

The results for the identified model for English monolinguals are shown in Table 5-10.

Table 5-10 *Results of mixed model for English monolinguals*

Fixed effect	Estimate	Std.Error	df	*t*	*p*
Intercept	-1.98	0.04	97.86	-45.14	<.001
ItemType (binomial)	-0.10	0.02	5563	-5.58	<.001
Congruency (Chinese-only)	0.03	0.03	82.24	0.11	.92
Congruency (English-only)	-0.06	0.03	76.85	-1.93	.05
EngPhrFreq.Resid	-0.01	0.01	1308	-1.12	.26
AssoStrength.log	-0.37	0.07	628.9	-5.01	<.001
TrialNum.sc	-0.05	0.01	50.84	-4.41	<.001
BlockOrder (2)	-0.13	0.05	50.08	-2.66	.01
ItemType (binomial) * Congruency (Chinese-only)	0.08	0.03	2963	2.45	.01
ItemType (binomial) * Congruency (English-only)	-0.01	0.02	5921	-0.46	.65
ItemType (control) * Wrd1Freq.log.c	0.01	0.01	1386	2.25	.02
ItemType (binomial) * Wrd1Freq.log.c	0.04	0.01	1546	4.87	<.001
Random effects	Variance	SD			
Target	0.008	0.09			
Participant	0.03	0.18			
TrialNum.sc | Participant	0.005	0.07			
Residual	0.12	0.35			

Note. df = degrees of freedom; Intercept levels: ItemType = Control; Congruency = Congruent; Marginal R^2 = 0.08, Conditional R^2 = 0.33.

The final model for the English monolinguals included two significant two-way interactions (item type × congruency, item type × Word 1 frequency). There were also statistically significant main effects of association strength ($\chi2 = 25.07$, $p < .0001$), block order ($\chi2 = 6.68$, $p = .01$) and trial number ($\chi2 = 16.75$, $p < .0001$). The model suggested that association strength has an important effect in binomial processing such that more strongly associated phrases had overall shorter response latencies. Also, block order affects the overall response as well such that participants responded faster in Block 2 than in Block 1. English monolinguals also went faster as the number of trials increased. The significant interaction between item type and congruency was explored through post-hoc comparisons, which revealed significant differences between binomials and controls in the congruent and English-only conditions. The two-way interaction between item type and congruency ($\chi2 = 7.62$, $p = .02$) (Figure 5-6) showed that the English monolingual speakers processed congruent and English-only binomials significantly faster than their corresponding controls (i.e., priming effects are observed for congruent and English-only binomials), but there was no difference between Chinese-only binomials and their controls (i.e., no priming for Chinese-only). It also showed that priming effects for congruent and English-only binomials were comparable. The result of the test is shown in Table 5-11.

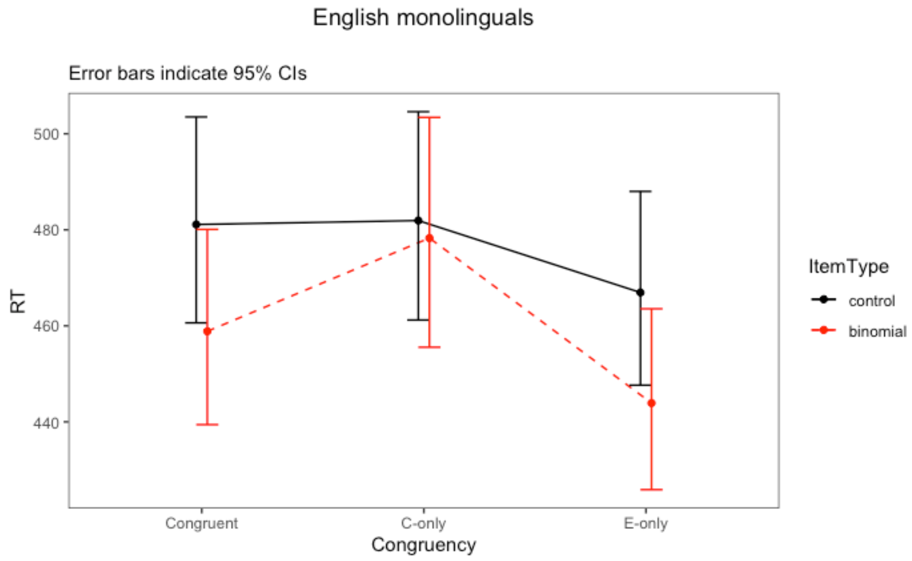


Figure 5-6 *Interaction plot of Item type * Congruency for English monolinguals*

Note: RTs were back-transformed data (unit = ms).

Table 5-11 *Results of post-hoc tests of RTs for congruent, English-only, and Chinese-only binomial items relative to the control items for English monolinguals*

Contrast	Estimate	Std.Error	*t*	*p*	ED (ms)
Control - Congruent	0.101	0.018	5.573	<.0001	22
Control - English-only	0.111	0.019	5.944	<.0001	23
Control - Chinese-only	0.016	0.025	0.645	.99	4

Note. ED (estimated difference) is calculated with RTs to controls minus RTs to binomials.

For the English monolinguals, priming effect was observed for the congruent (t = 5.73, $p < .0001$) and English-only conditions (t = 5.94, $p < .0001$). The mean RT (model estimates) to the binomials was 22 ms faster than RT to the control items (459 vs. 481 ms) in the congruent condition and 23 ms faster in the English-only condition (444 vs. 467 ms). I did not find priming for the Chinese-only condition ($t < 1$, $p = .99$), which confirmed that the translated Chinese-only items were not processed as binomials by the English monolinguals.

5.2.5 Discussion

Using a priming paradigm, I tested whether English monolinguals responded faster to words (*fork*) that appear in binomials (*knife and fork*) than to words that appear in infrequent but equally strongly associated controls (*spoon and fork*) – that is, if they showed binomial priming effects. I further tested whether priming effects were similar for both types of English binomials, congruent and English-only. If English monolinguals show a similar level of facilitation for the two types of binomials, then the difference in priming effects between congruent and English-only binomials examined in the Chinese-English bilinguals (Experiment 1a) should be attributed to their knowledge of the L1/Chinese.

As expected, the English monolingual participants showed significant facilitation in the processing of the final word in English binomials compared to control phrases. The magnitude of the priming effect for congruent binomials was 22 ms and that for English-only binomials was 23 ms. Also, as expected, the English monolingual speakers showed no priming for the translated Chinese-only binomials over controls, since these word sequences had very low frequency (close to zero) in English. The results offer further support to the tenet that frequency of encounters with MWEs determines their ease of processing. A plethora of empirical studies have shown that L1 speakers can recognize, read and respond to MWEs significantly faster than matched novel strings of language (Arnon & Snider, 2010; Durrant & Doherty, 2010; Ellis et al., 2009; Tremblay & Baayen, 2010; Tremblay et al., 2011; Vilkaitė, 2016). Concerning binomials, previous studies have found that binomials are processed differently from the matched novel controls employing paradigms such as eye movement (Siyanova-Chanturia, Conklin, & Van Heuven, 2011) and ERP (Siyanova-Chanturia et al., 2017). Using eye-tracking, for example, Siyanova-Chanturia, Conklin and Van Heuven (2011) found that L1 speakers are sensitive to the frequency with which three-word binomial phrases occur in a language, such that L1 speakers read binomials significantly faster than the reversed forms which are identical in syntax and meaning but differ in phrase frequency (*bride and groom* vs. *groom and bride*). They concluded that both the overall frequency of a phrase and whether the phrase is in the preferred (binomial) or nonpreferred (reversed) configuration affect the processing of binomial phrases. Siyanova-Chanturia and colleagues' follow-up ERP study (2017) has further found that when the conjunction "and" was removed from binomials, no differences were observed between the binomial (*knife-fork*) and the control conditions (*spoon-fork*). In contrast, when the conjunction "and" was present in the stimuli, binomials (*knife and fork*) elicited larger P300s (reflecting the activation of a template that matches the upcoming information) and smaller N400s (indexing easier semantic integration) compared to the matched controls (*spoon and fork*). This indicates that seeing *knife* (with "and") should activate the representation (i.e., mental template) of the configuration (*knife and fork*) in the brain, while seeing *knife* (without "and") should not. These studies suggest that for predictable sequences such as binomials, distinct mechanisms underlie their processing compared to novel sequences.

The results are also consistent with the results of previous studies on the processing of MWEs using priming paradigms, suggesting that there is link or association between the constituents of a MWE due to frequency of co-occurrence. For example, Durrant and Doherty (2010) found that adjective-noun collocations (*foreign debt*) produced significant priming effects when compared to infrequent two-word combinations (*direct danger*). Critically, collocations which are psychological associates (*card game*) exhibited similar statistical robustness of the priming effect as collocations which are not (*foreign debt*). This indicates that priming can exist between the constituents of collocations simply due to frequency of co-occurrence, irrespective of semantic association. Similarly, in a primed lexical decision task, Wolter and Gyllstad (2011) found that L1 speakers of English responded faster to the target in a verb-noun collocation (*find job*) than the word in an infrequent non-collocational item (*create issues*). Likewise, Carrol and Conklin (2014) found that the initial part of the idioms (*to spill the …*) can prime the last word in the idioms (*beans*), such that L1 speakers of English responded faster to targets in the idiom than control conditions (*to spill the chips*). In the present study, English monolinguals also responded faster to targets when they were in a binomial (*knife and fork*) than when not (*spoon and fork*), providing further support to the tenet that word *a* can accelerate recognition of word *b* if speakers have encountered the word combination *a* + *b* sufficiently frequently (i.e., lexical priming: Hoey, 2013). Importantly, in the present study, the two content words within the binomial (*knife and fork*) and the matched control (*spoon and fork*) were equally strongly associated. Similar to Durrant and Doherty (2010), this study reaffirms that priming can exist between constituents of frequent co-occurring multi-word sequences, regardless of whether the constituents are associated.

Crucially, in contrast to the Chinese-English bilinguals, English monolingual controls exhibited similar facilitation for congruent and English-only binomials. This indicates that from L1 monolinguals' perspective, the two types of binomials, congruent and English-only, are equivalent. This finding reaffirms the interpretation that the processing advantage for congruent over English-only binomials is attributed to the differential status (congruent vs. incongruent) the items held in the Chinese-English bilinguals' L1.

5.3 Conclusion

Using a priming paradigm, I set out to investigate the L1 influence on the processing of frequent familiar MWEs – binomials (*knife and fork*) – relative to infrequent novel phrases (*spoon and fork*) in a lexical decision task. Namely, the study aimed at investigating the role of congruency between the L1 and L2 in the processing of L2 by demonstrating that congruency facilitates MWE processing in bilinguals, but not in monolinguals. Crucially, one of the goals was to investigate the underlying mechanisms for the congruency effect, if any, in L2 MWE processing.

In Experiment la, the Chinese-English bilinguals showed a significant priming effect for congruent binomials (*sun and moon*) relative to infrequent novel phrases (*star and moon*), but no priming for English-only binomials (*bread and butter*) relative to the matched controls (*toast and butter*). However, in Experiment lb, the monolingual English controls showed no difference in priming effects between the two types of binomials, congruent and English-only, relative to their matched controls. This congruency advantage observed in the Chinese-English bilinguals, but not in monolingual English controls, was interpreted as a significant marker of cross-language influence in L2 MWE processing. In previous studies, this congruency advantage for congruent L2 MWEs over incongruent L2-only MWEs was reported for idioms (Carrol et al., 2016) and collocations (Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2018). The present study reaffirms and extends previous findings by showing that congruency also facilitates the processing of frequent *literal* MWEs – binomials – in the L2.

However, similar to the monolingual English controls in Experiment 1b, the Chinese-English bilinguals in Experiment 1a did not show any facilitation for translated Chinese-only binomials (*wisdom and strength*), which had very low frequency (close to zero) in English, when compared to infrequent novel phrases (*exercise and strength*). This result is inconsistent with studies with idioms that found a processing advantage for translated L1-only idioms relative to control novel phrases (Carrol & Conklin, 2014, 2017; Carrol et al., 2016), while it is consistent with studies with collocations that did not report such a processing advantage for translated L1-only collocations compared to matched controls (Wolter & Yamashita, 2015, 2018). Therefore, the finding of the present study does not provide support for the L1 MWE activation account, according to which MWEs across the L1 and L2 are both activated in L2 processing. Instead, the present study suggests that prior exposure to L2 word sequences is necessary for any significant facilitation to occur, compared to control novel phrases, as suggested by the L2 MWE experience account.

Chapter 6 Experiment 2: Probing L2 Influence in L1 MWE Processing

Although previous studies as well as Experiment 1 of the present study have established crosslinguistic influences in the processing of MWEs in the direction of L1→L2, the issue of whether crosslinguistic influences in the processing of MWEs exist in the reverse direction of L2→L1 is still an open question. The current literature on L2 influence in L1 lexical, single word, processing, however, has suggested that crosslinguistic influences are bidirectional. That is, the weak, non-dominant L2 may affect the processing of words in the dominant L1 (Degani et al., 2011; Schoonbaert et al., 2009; Van Hell & Dijkstra, 2002). One may wonder whether bidirectional crosslinguistic influences extend to units above the word level, namely, MWEs. Therefore, the purpose of Experiment 2 in the present study was to determine whether or not the L2 of bilinguals can in turn affect MWE processing in their L1. To answer these questions, I examined binomial priming in English-Chinese bilinguals with the same experimental design used in Experiment 1.

Experiment 2 had several goals. Firstly, I wished to determine if English-Chinese bilinguals exhibit greater facilitation for congruent binomials relative to their matched controls (*sun and moon* vs. *star and moon*), when compared with English-only binomials relative to their matched controls (*bread and butter* vs. *toast and butter*). If so, it would suggest that the weaker L2 can also influence L1 MWE processing in bilinguals, and that the cross-language influence in MWE processing is bidirectional (i.e., in the direction of L1→L2 and the reverse direction of L2→L1). Secondly, I tested if translated Chinese-only binomials show a processing advantage relative to their matched controls (*wisdom and strength* vs. *exercise and strength*). If so, this would suggest that L2/Chinese influence extends to the processing of translated Chinese-only MWEs which have not been previously seen in the

L1/English and support the claim that the corresponding L2 MWEs are automatically activated in L1 processing. If not, this would suggest that there is no automatic activation of L2 MWEs in L1 processing, and that prior exposure to the word sequences shown in the L1 is necessary for them to show significant facilitation compared to control novel phrases.

Based on previous findings that the L1 influences L2 MWE processing (Carrol & Conklin, 2014, 2017; Carrol et al., 2016; Wolter & Gyllstad, 2011, 2013) and that cross-language influences occur in both directions in bilingual lexical (single word) processing (Basnight-Brown & Altarriba, 2007; Schoonbaert et al., 2009; Van Hell & Dijkstra, 2002), I predicted that L2 could also influence the processing of L1 MWEs such that congruent binomials should show greater priming effects than English-only binomials in English-Chinese bilinguals. However, based on the findings that the dominant L1 typically has a higher impact on the non-dominant L2 processing than vice versa (Jiang, 1999; Keatley et al., 1994; Schoonbaert et al., 2009), it is likely that cross-language influence in the L1→L2 direction will be greater than that in the L2→L1 direction. In addition, on the basis of the results that the Chinese-English bilinguals in Experiment la showed no priming for translated Chinese-only binomials, it was expected that the English-Chinese bilinguals would behave in a similar manner by showing no priming for translated Chinese-only binomials. Lastly, it was expected that the English-Chinese bilinguals might process L1 binomials in a different way from English monolinguals in Experiment lb. This prediction was based on previous research that suggests that bilinguals immersed in an L2 environment (like the English-Chinese bilinguals tested here) show slower processing speed in L1 processing compared to their monolingual counterparts (Baus et al., 2013; Linck et al., 2009; Morales et al., 2014).

6.1 Participants

Fifty-one English-Chinese bilingual participants (29 females; age 18-31; mean age = 22.88 years) were undergraduate and postgraduate students from Peking University and

Tsinghua University, China. Participants were recruited through advertising posters and personal contacts. They were L1 English speakers who came to study Chinese or other subjects in Beijing as international students. They were dominant in their L1 English. They completed a language background questionnaire before the experiment (see Appendix 9), in which they reported their Chinese proficiency (self-reported), the number of years of exposure in China (average = 1.92 years, range: 0.60 – 8.00 years), and the estimation of their daily usage of Chinese (average = 37%, range: 5% – 90%). Only nine participants had taken the standardised Chinese proficiency test called Hanyu Shuiping Kaoshi (HSK), which was not compulsory for their programmes. Therefore, I used self-reported measures to assess their Chinese proficiency. Twenty-two participants reported themselves as intermediate speakers of Chinese as an L2, and twenty-nine participants as advanced speakers of Chinese as an L2. All participants had normal or corrected-to-normal vision. Each participant received $10 for their participation in the experiment. Table 6-1 summarizes English-Chinese participants' language proficiency characteristics.

Table 6-1 *Means (standard deviations) of self-reported age, Chinese proficiency levels, daily usage of Chinese, years of exposure to Chinese in L2-speaking countries for English-Chinese bilinguals*

	English-Chinese bilinguals (N = 51)
Age	22.88 (2.85)
Chinese proficiency	Intermediate+
Daily usage of L2	Chinese: 37% (22%)
Years of exposure to L2	1.92 (1.84)

6.2 Materials and Design

The materials and design were the same as Experiment 1.

6.3 Procedure

The procedure was the same as Experiment 1.

6.4 Analysis and Results

Data analysis was performed in the same way it was done in Experiment 1a and 1b. That is, both accuracy and RTs were analysed. For English-Chinese bilinguals, the mean response accuracy to non-word items was 96.54%, and the mean accuracy to word items was 97.68%. As in Experiment 1a and 1b, there was no significant difference in response accuracy between the binomial and control conditions for the three congruency types. The accuracy from Experiment 2 is presented in Table 6-2.

Table 6-2 *Descriptive statistics: Accuracy (%) for English-Chinese bilinguals*

Chinese-only		Congruent		English-only		Fillers	Nonwords
Binomial	control	binomial	control	binomial	Control		
97.94	97.75	98.73	97.65	98.24	97.84	96.48	96.54

For the RT analyses, only data for binomials and their corresponding controls was analysed (120 items in total). Incorrect responses were removed from the RT analysis, resulting in the loss of 1.98% data for English-Chinese bilinguals. Extreme values (RTs longer than 2000 ms or shorter than 250 ms) were also excluded, which resulted in the loss of 0.17% data. Means of RTs (descriptive statistics) by condition for English-Chinese bilingual participants are shown in Table 6-3.

Table 6-3 *Descriptive statistics: Mean response times in ms (standard deviations) and difference between mean response times to the binomial and control phrases for English-Chinese bilinguals in each of the six experimental conditions*

	Binomial	Control	Difference
Congruent	495.43(128.65)	516.26(14.013)	20.83
English-only	501.56(152.66)	504.96(141.42)	3.4
Chinese-only	528.79(158.83)	523.23(152.79)	-5.56

Note. Difference is calculated with mean RT to controls minus mean RT to binomials.

In Experiment 2, the response variable – inverse-transformed RTs – were also normally distributed, with skewness of 0.24 and kurtosis of 0.08. The same maximal model and procedure were used to identify the best-fit model for the English-Chinese bilinguals in Experiment 2 as it was in Experiment 1a and 1b. The best-fit model with random intercepts and random slopes was as follows:

RTinv ~ ItemType*Congruency + EngPhrFreq.Resid + Wrd1Freq.log.c*ItemType + AssoStrength.log + TrialNum.sc + (1 + TrialNum.sc | Participant) + (1 + TrialNum.sc + Wrd1Freq.log.c | Target)

A quantile-quantile plot of the model's residuals was inspected to assess whether the residuals were normally distributed. However, there were deviations from the line and the residuals did not conform to normal distribution. Thus, I removed 2.5 SD from the residuals to satisfy the assumption of homoscedasticity and normal distribution, which resulted in the loss of 1.87% data for the English-Chinese bilinguals. The new Q-Q plot (Figure 6-1) and the residual plot (Figure 6-2) showed that the model satisfied the normality and constant variance assumptions.

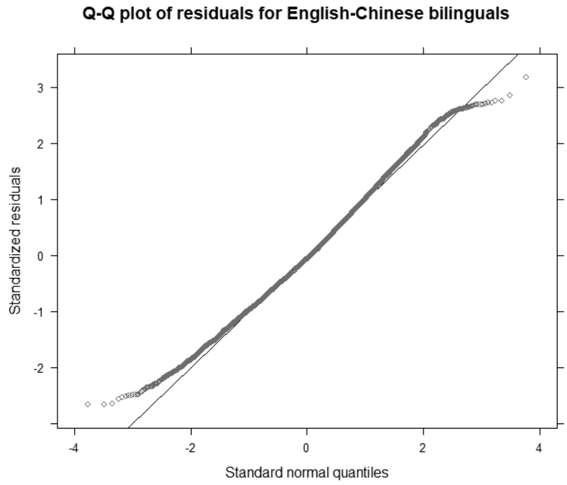


Figure 6-1 *Q-Q plot of residuals for English-Chinese bilinguals*

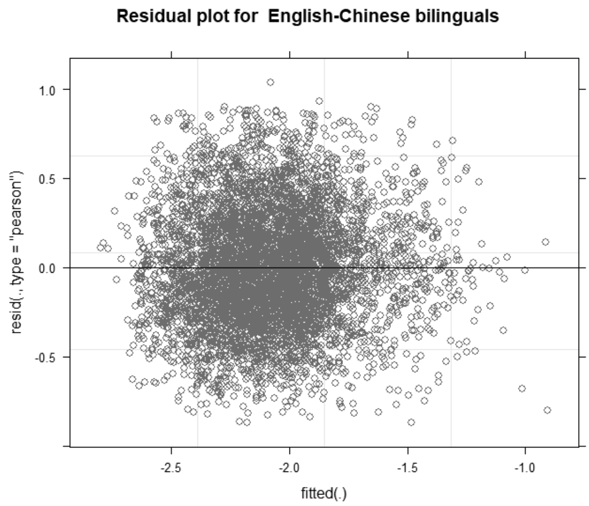


Figure 6-2 *Residual plot for English-Chinese bilinguals*

The results for the identified model for the English-Chinese bilinguals are shown in Table 6-4.

Table 6-4 *Results of mixed model for English-Chinese bilinguals*

Fixed effect	Estimate	Std.Error	df	*t*	*p*
Intercept	-2.01	0.04	99.99	-48.47	<.001
ItemType (binomial)	-0.05	0.02	130.49	-2.72	.007
Congruency (Chinese-only)	-0.04	0.03	84.37	-1.36	.18
Congruency (English-only)	-0.07	0.03	77.60	-2.59	.01
EngPhrFreq.Resid	-0.03	0.01	177.49	-2.74	.006
Wrd1Freq.log.c	0.01	0.01	40.61	0.77	.44
AssoStrength.log	-0.25	0.08	161.17	-3.21	.002
TrialNum.sc	-0.04	0.01	51.70	-3.51	<.001
ItemType (binomial) * Congruency (Chinese-only)	0.02	0.04	173.83	0.55	.59
ItemType (binomial) * Congruency (English-only)	0.06	0.02	179.42	2.32	.02
ItemType (binomial) * wrd1Freq.log.c	0.02	0.01	176.52	2.08	.04
Random effects	Variance	SD			
Target	0.007	0.08			
TrialNum.sc | Target	0.0003	0.02			
Wrd1Freq.log.c	0.0009	0.03			
Participant	0.06	0.24			
TrialNum.sc | Participant	0.004	0.07			
Residual	0.11	0.33			

Note. df = degrees of freedom; Intercept levels: ItemType = Control; Congruency = Congruent; Marginal R^2 = 0.02, Conditional R^2 = 0.41.

The final model for the English-Chinese bilinguals revealed a marginally significant (with alpha of 0.05) two-way interaction between item type and congruency (χ^2 = 5.55, *p* = .06). There were also statistically significant main effects of English phrase frequency (χ^2 = 7.53, *p* = .006), association strength (χ^2 = 9.89, *p* = .002) and trial number (χ^2 = 11.24, *p* = .0007). The model suggested that phrase frequency was always facilitative (led to lower overall RTs). Association strength was also facilitative whereby more strongly associated phrases led to lower overall RTs. Participants responded faster as the trial number increased. The interaction between item type and congruency was explored through post-hoc comparisons using the emmeans() function, with Bonferroni adjustments. The two-way interaction (Figure 6-3) showed that the English-Chinese bilinguals processed congruent binomials somewhat faster than the controls, but there were no processing differences either between the English-only binomials versus matched controls, or between Chinese-only binomials versus matched controls. The result is shown in Table 6-5.

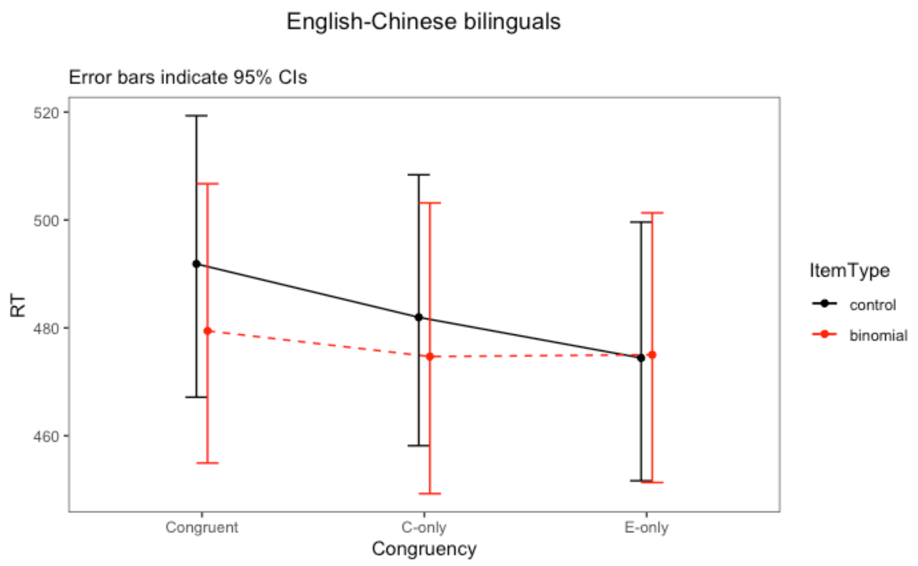


Figure 6-3 *Interaction plot of Item type * Congruency for English-Chinese bilinguals*

Note: RTs were back-transformed data (unit = ms).

Table 6-5 *Results of post-hoc tests of RTs for congruent, English-only, and Chinese-only binomial items relative to the control items for English-Chinese bilinguals*

Contrast	Estimate	Std.Error	*t*	*p*	ED (ms)
Control - Congruent	0.053	0.019	2.717	.11	12
Control - English-only	-0.003	0.019	-0.135	.99	-1
Control - Chinese-only	0.032	0.027	1.19	.99	7

Note. ED (estimated difference) is calculated with RTs to controls minus RTs to binomials.

For the English-Chinese bilinguals, there was a priming of 12 ms (model estimate) for congruent binomials. The priming did not reach statistical significance after applying a correction for multiple comparison (t = 2.72, p = .11). However, it can be viewed as a *weak* priming effect for congruent binomials. No priming effects was observed in English-only (t < 1) or Chinese-only (t = 1.19) conditions.

Nevertheless, it needs to be noted that there is variation in the responses to the control items between English-Chinese bilinguals and the other two groups of participants (i.e., Chinese-English bilinguals and English monolinguals). In both Experiment 1a and 1b, the control items for the English-only condition were responded to the fastest, and the RTs for the congruent and Chinese-only control items were about the same. In contrast, in Experiment 2, although the control items for the English-only condition were also responded to the fastest, as in Experiment 1a and 1b, RTs for the control items in the congruent condition were slower than those in the Chinese-only condition. Within group post-hoc comparisons showed that the differences within the control item category were not significant in Experiment 1a (congruent – Chinese-only: t = -0.09, p = .99; congruent – English-only: t = 1.01, p = .57; Chinese-only – English-only: t = 1.10, p = .52), nor in Experiment 1b (congruent – Chinese-only: t = -0.11, p = .99; congruent – English-only: t = 1.93, p = .13; Chinese-only – English-only: t = 2.05, p = .11). In Experiment 2, the differences in the RTs to the congruent and Chinese-only control items (t = 1.36, p = .36) and the RTs to the

Chinese-only and English-only control items (*t* = 1.17, *p* = .47) were not significantly different. However, control items in the congruent condition were responded significantly slower than those in the English-only condition by English-Chinese bilinguals (*t* = 2.59, *p* = .03). This leads to the possibility that the numerical facilitation for congruent binomials compared to control phrases by English-Chinese bilinguals might be attributed to the slow responses to the control items in the congruent condition, rather than the fast responses to the binomials in the congruent condition. Therefore, caution needs to be taken when interpreting the marginally significant priming effect for congruent binomials by English-Chinese bilinguals.

In addition, to investigate the effect of Chinese proficiency on the responses to the target words, a separate model was fitted which replaced the two-way interaction of Item Type and Congruency with a three-way interaction of Item Type* Congruency*Proficiency in the original maximal model (see section 5.1.5). Twenty-two participants reported themselves as intermediate speakers of Chinese as an L2, and twenty-nine participants as advanced speakers of Chinese as an L2. Thus, Chinese proficiency was treated as a categorial variable with two levels, intermediate and advanced. The new maximal model was as below:

> RTinv ~ ItemType*Congruency*Proficiency + EngPhrFreq.Resid *ItemType + EngPhrFreq.Resid * Congruency + Wrd1Freq.log.c*ItemType + AssoStrength.log*Congruency + TrialNum.sc + BlockOrder + (1| Participant) + (1| Target)

I used the step() function in *lmerTest* to arrive at the best model fit, which was the same as the best model fit that I got from the former model without the three-way interaction of ItemType*Congruency*Proficiency. Therefore, no effect of Chinese proficiency was found for the RTs to the target words.

6.5 Discussion

The aim of Experient 2 was to examine whether or not the L2 of English-Chinese bilinguals' influenced the processing of binomial expressions in their L1. The role of language direction in MWE processing has so far received little attention in bilingual processing research. To this end, I used the same primed lexical decision task as in Experiment 1a to examine how the English-Chinese bilinguals process the same items, that is, congruent (*sun and moon*), English-only (*bread and butter*), and translated Chinese-only (*wisdom and strength*) binomial expressions versus infrequent but equally strongly associated control phrases (*star and moon*, *toast and butter*, *exercise and strength*, respectively).

The key findings were as follows. The English-Chinese bilingual participants showed no facilitation in the processing of the final word in English-only binomial phrases (*bread and butter*) compared to control phrases (*toast and butter*), whereas their monolingual English counterparts in Experiment 1b showed a 23 ms priming effect for English-only binomials. However, the English-Chinese bilinguals exhibited a trend toward priming for the congruent binomials (*sun and moon*) compared to control phrases (*star and moon*). They processed congruent binomials quantitatively faster than their controls (mean difference = 20 ms, model estimate = 12 ms). However, after applying a correction for multiple comparison, the priming did not reach statistical significance (p = .11). Thus, the English-Chinese bilinguals showed a weak congruency effect, that is, a greater processing advantage for congruent than English-only binomials. Recall that the Chinese-English bilinguals in Experiment 1a showed significant priming for congruent binomials and a clear congruency advantage over English-only binomials. The English-Chinese bilinguals' performance was more similar to the Chinese-English bilinguals in Experiment 1a than to the English monolinguals in Experiment 1b. This is because the English-Chinese bilinguals showed no priming effect for English-only binomials but some congruency advantage in the processing of congruent over English-only binomials. Lastly, similar to the English monolinguals in Experiment 1b and the Chinese-English bilinguals in Experiment 1a, the English-Chinese bilinguals showed no processing advantage for the translated Chinese-only binomials (*wisdom and strength*) over controls (*exercise and strength*).

6.5.1 The Inhibition of the L1

The finding that the English-Chinese bilinguals showed no processing advantage for English-only binomials over controls seems inconsistent with the literature on MWE processing in L1 speakers. It has been established that L1 speakers can recognize, read, and respond to MWEs significantly faster than matched novel strings of language (Arnon & Snider, 2010; Durrant & Doherty, 2010; Siyanova-Chanturia et al., 2017; Vilkaite, 2016). In fact, in Experiment 1b, the English monolingual controls showed a significant priming effect for English-only binomials. What, then, might have contributed to the absence of priming for English-only binomials in the English-Chinese bilinguals, who performed the task in their first and dominant language?

One possibility is that the L1 of the English-Chinese bilinguals had to be inhibited in the L2 immersion environment (while studying Chinese in China). When they had to switch back to their strongly inhibited L1, for the purpose of completing the experiment, their L1 processing could have been impaired. The result that the mean RTs on L1 (English) lexical decisions were slower for the English-Chinese bilinguals than for the English monolinguals (512 ms vs. 497 ms; $p < .0001$) provides some evidence to support this conjecture. It has been shown that, after immersion in a foreign language, even just for a few months, bilinguals may experience delay when retrieving L1 words (Baus et al., 2013; Linck et al., 2009). Immersion is argued to enable bilinguals to attenuate the activity of the L1, thus enabling speakers to better control L1 lexical competition and facilitating L2 learning (Linck et al., 2009). For instance, in a comprehension task (translation recognition), Linck et al. (2009) found that the immersed English-Spanish bilinguals showed no sensitivity to English distractors which had form overlap with the presented Spanish words (e.g., *cara*-*card*). The results were interpreted as evidence that immersed bilinguals suppress the visually presented distractors from intruding on their judgments, and that L1 was inhibited frequently during immersion to facilitate L2 learning. Recent evidence from classroom learning also indicates that the inhibition of L1 equivalents improves learning and retrieval of L2 MWEs even in an L1-speaking environment (Pulido, 2021b; Pulido & Dussias, 2020).

According to the Inhibitory Control model (Green, 1998), the non-target language is inhibited, preventing it from disrupting the selection of target language words. The amount of inhibition applied to the non-target language is proportional to the baseline strength of its activation. The more dominant the language, the stronger inhibition is needed. Since the L1 of an unbalanced bilingual is dominant, it is strongly suppressed whenever bilinguals need to use the L2. As a result, the cost of reactivating the L1 after using an L2 is likely to be greater than a switch in the opposite direction (Mosca & de Bot, 2017; Wodneicka et al., 2020), having a greater effect on L1 performance. Numerous studies have shown that switching costs are larger for the stronger than for the weaker language (i.e., asymmetrical switching costs) (Jackson et al., 2001; Macizo et al., 2012; Meuter & Allport, 1999). For our unbalanced English-Chinese bilinguals, the L1 had to be strongly inhibited to enable them to use the L2 in the immersion context. Switching back to their strongly suppressed L1, in order to perform an L1 lexical decision task, likely came at a cost. The absence of priming for English-only binomials in the English-Chinese group may have thus been a result of their weakened L1 performance.

6.5.2 L2 influence in L1 MWE Processing

This difficulty of retrieving the dominant L1 in the L2 immersion context may also explain the findings for the processing of congruent L1 binomials by the English-Chinese bilinguals. The English-Chinese bilinguals showed some facilitation in the processing of the final word in congruent L1 binomials relative to control phrases, which, however, did not reach significance after applying a correction for multiple comparison (t = 2.72, p = .11). It is possible that due to L1 inhibition, L1 access was attenuated so that the priming effect for congruent L1 binomials was less robust than that observed for the Chinese-English bilinguals and the English monolinguals. This result is compatible with their performance on English-only binomials. However, I observed a weak priming for congruent L1 binomials that suggests possible activation of known, corresponding L2, binomials. Since English-only and congruent L1 binomials were matched in L1 phrase frequency and the English monolingual controls showed comparable facilitation towards them, activation of L2 binomial equivalents during the L1 task by the English-Chinese bilinguals seems to be the likely explanation of this weak priming for congruent L1 (but not English-only) binomials. This is evidence of cross-language influence in the L2→L1 direction in an entirely within-L1 task. This result is not unlike the findings of an automatic activation of single words in the weaker language in mixed stimulus lists (e.g., Dijkstra et al., 2000) and in L1-only lists (e.g., Van Hell & Dijkstra, 2002). This finding suggests that known L2 MWEs may be automatically activated in L1 processing, leading to the faster processing of MWEs that exist in both languages.

Finally, the English-Chinese bilinguals showed no facilitation for translated Chinese-only binomials over controls. The same pattern of results was observed in English monolinguals and Chinese-English bilinguals. This indicates that there was no activation of translated Chinese-only MWEs (i.e., L2-only in the case of English-Chinese bilinguals). It is not surprising given that there was no activation for translated Chinese-only binomials over controls in Chinese-English bilinguals. In other words, the effects in the L2→L1 direction were less likely to take place when no such effects were observed in the L1→L2 direction, because cross-language influence in the L1→L2 direction is normally stronger than in the opposite direction.

To sum up, the English-Chinese bilinguals showed no priming for English-only binomials, but weak priming for congruent binomials. These results support the view that L1 may be inhibited in L2 learning and immersion contexts and, thus, switching back to L1 may come at a cost. The results also support the view that crosslinguistic influence can occur from the non-dominant L2 to the dominant L1, even in an entirely within-L1 task. Thus, I conclude that crosslinguistic influences in the processing of binomials are bidirectional, although the influence in the direction of L1→L2 is stronger than in the reverse direction of L2→L1. This conclusion is in line with studies with bilingual word processing which suggest that crosslinguistic influences are bi-directional (Degani et al., 2011; Schoonbaert et al., 2009; Van Hell & Dijkstra, 2002). The present study is the first study, to my knowledge, that investigated bi-directional cross-language influences in the processing of binomials – a less commonly studied type of MWEs.

Chapter 7 General Discussion

In this chapter, I compare and discuss the findings of my three experiments altogether. I firstly briefly review the research questions the present study addressed and outline the key findings. Next, I discuss the theoretical implications of my findings with respect to the current literature.

7.1 Research Questions and Key Findings

While previous research converges on the finding that congruency across the L1 and L2 facilitates the processing of L2 MWEs such that congruent L2 MWEs show a greater processing advantage over incongruent L2-only MWEs, evidence is mixed with regard to whether L1 influence extends to the processing of translated L1-only MWEs that do not exist in the L2. The issue of whether or not bilinguals show an advantage in the processing of translated L1-only MWEs has important implications for understanding the mechanisms behind the congruency effect in the processing of L2 MWEs. If translated L1-only MWEs show a processing advantage over control novel phrases, it would suggest that L1 MWEs are automatically activated in L2 processing, as predicted by the L1 MWE activation account. If translated L1-only MWEs are processed in a similar way to control novel phrases, it would suggest that L1 MWEs are not activated in L2 processing and that without prior exposure to a word sequence in the L2, no processing advantage would be observed for it, as predicted by the L2 MWE experience account. If it is the case, then the congruency effect in L2 MWE processing may stem from the likely earlier acquisition (and, therefore, greater experience with their L2 form) of congruent MWEs (compared with incongruent, L2-only MWEs), rather than from the automatic activation of corresponding L1 MWEs. Therefore, one of the aims of the present study was to probe the underlying mechanisms behind the congruency effect in L2 MWE processing. In addition, the present study aimed to investigate whether cross-language influence in MWE processing would also occur in the reverse direction, from the non-dominant L2 to the dominant L1. This issue, to my knowledge, has not been sufficiently addressed in the literature. Research on lexical, single word, processing suggests that cross-language influence is bi-directional such that the non-dominant L2 may also affect the processing of words in the dominant L1 (Schoonbaert et al., 2009; Van Hell & Dijkstra, 2002). The present study thus tests whether the influence of the non-dominant L2 in L1 processing can extend to units beyond the word level (to MWEs) and, if so, whether or not this influence is equally strong as the cross-language influence in the L1→L2 direction.

To address these questions, the present study tested cross-linguistic influences in the processing of binomials, a type of MWEs for which this issue has not been investigated. I compared the processing of three types of binomials relative to matched controls, that is, congruent, English-only, and translated Chinese-only binomials (all presented in English), in three groups of participants, Chinese-English and English-Chinese bilinguals and English monolinguals. The three groups of participants completed the same English lexical decision experiment with a binomial priming manipulation. This design allowed me to test the involvement of L1 in L2 MWE processing in the case of Chinese-English bilinguals, and to test the involvement of L2 in L1 MWE processing in the case of English-Chinese bilinguals. The English monolingual group of participants served as a baseline group. Specifically, I asked the following two questions: (1) Do Chinese-English and English-Chinese bilinguals show a greater processing advantage for congruent binomials over English-only binomials? (2) Do Chinese-English and English-Chinese bilinguals show a significant processing advantage for translated Chinese-only binomials over control novel phrases? The following findings emerged.

In Experiment 1a, the Chinese-English bilinguals showed a significant priming effect (19 ms) for congruent binomials, but no priming effect for English-only binomials. In contrast, in Experiment 1b, the English monolinguals showed comparable priming effects for congruent (22 ms) and English-only (23 ms) binomials. Interestingly, in Experiment 2, the English-Chinese bilinguals did not show a statistically significant priming effects for either congruent or English-only binomials, but they did show a trend towards priming for congruent binomials (12 ms). With regard to the processing of translated Chinese-only binomials, none of the three groups of participants showed any significant processing advantage for these items relative to matched controls. I discuss each of these findings below.

7.2 Phrase Frequency Effect in the Processing of Binomials

Similar to the English monolinguals, the Chinese-English bilinguals showed no priming in the processing of the terminal words of the *translated Chinese-only* binomials (*wisdom and strength*) which had very low frequency (close to zero) in English. This result indicates that monolingual and bilingual speakers are sensitive to phrase frequency effects. Specifically, they do not show facilitation for word combinations that they have not encountered before. Prior exposure to an L2 word combination is needed for any facilitation for it to occur, regardless of whether this MWE does or does not have an equivalent form in the L1. This is in line with previous studies that found that it is the memory for particular word combinations that affords faster lexical access. For example, using a (double) lexical decision task, Ellis et al. (2009, Experiment 2) investigated whether or not L1 speakers of English showed any significant facilitation for word pairs that they had never seen before but were consistent in semantic prosody (i.e., the general tendency of certain words to co-occur with either negative or positive expressions). They compared equally *infrequent* two-word pairs which were matched in semantic prosody (*cause bad*, *cause harm*) and pairs which were not matched (*cause good*, *cause benefit*). They found that semantic prosody matching pairs (*cause bad*, *cause harm*) and mismatching pairs (*cause good*, *cause benefit*) were processed with a similar speed, suggesting it is the memory for *particular* word combinations that affords faster lexical access and there are no top-down semantic generalizations upon this level of processing. This is in line with the finding in single word processing, where *token* frequency determines the processing of individual word forms (Croft and Cruse, 2004).

However, unlike the English monolinguals, the Chinese-English bilinguals showed no priming for English-only binomials. This result is consistent with the results reported in Wolter and Yamashita (2015, 2018). They found Japanese-English bilinguals showed no processing advantage for English-only collocations (*catch breath*, *busy road*) over novel controls (*bad gift*, *active gift*). However, this result is inconsistent with Wolter and Gyllstad (2011, 2013), who reported a significant processing advantage for English-only collocations (*pay a visit*, *identical twins*) over control items (*defend sales*, *angry use*) in Swedish-English bilinguals. Note that these studies by Wolter and colleagues (2011, 2013) used semantically implausible items as controls, which probably increased the processing cost of control items and overestimated the processing advantage for English-only collocations. The present study, however, used semantically plausible phrases as controls (*spoon and fork*), giving a more accurate picture of the phrase frequency effect. The result of the present study suggests that although phrase frequency in the L2 is a prerequisite of the processing advantage for binomials over control novel phrases (as shown in the case of translated Chinese-only binomials), it is not the only factor that may affect their processing. Having an equivalent form in the L1 may be another important determinant of how L2 binomials are processed. For example, the Chinese-English binomials showed a significant priming effect (19 ms) in the processing of the terminal words of the congruent binomials (*sun and moon*) compared to the infrequent but equally strongly associated control phrases (*star and moon*). This suggests that phrase frequency in the L2 and having an equivalent form in the L1 (i.e., congruency) combined to affect the processing of L2 binomials by Chinese-English bilinguals.

When both driving forces for the processing of L2 binomials exist, phrase frequency in the L2 and having an equivalent form in the L1, L2 speakers can show a significant processing advantage for MWEs similar to L1 speakers. In the present study, the Chinese-English bilinguals showed a significant priming effect for congruent binomials. This suggests that similar to the English monolinguals, the Chinese-English bilinguals are also sensitive to phrase frequency as well as whether a phrase occurs in a particular configuration (binomial vs. control). For both groups of participants, the terminal word of *congruent* binomials was primed by the first two constituent words of binomials (*sun and → moon*), while the same target word of matched controls was not primed by the first two constituents of controls (*star and → moon*). Note that the two content words within the congruent binomial (*sun – moon*) and the matched control (*star – moon*) were equally strongly associated ($t = 1.41$, $p = .17$), suggesting that the priming effect was not due to the first word in the binomials (*sun*) being a better prime than the first word in the control items (*star*) for the same target (*moon*). Thus, the priming effect for congruent binomials suggests the contribution of entrenchment of a particular phrase of an *A and B* form in memory. This finding is consistent with an ERP study with binomials. Siyanova-Chanturia et al. (2017) reported both N400 and P300 effects in L1 readers' comprehension of English binomial expressions (*knife and fork*). They found that the second content word in binomial condition (*knife and fork*) elicited a larger P300 and a smaller N400 than the same word in novel but equally strongly associated phrases (*spoon and fork*) or non-associated, unattested semantic violations (*theme and fork*) in Experiment 1a. However, when phrases were presented without the conjunction "and" in Experiment 1b, no differences in waveforms in the P300 and N400 time windows were observed between binomials (*knife-fork*) and novel associates (*spoon-fork*). According to Siyanova-Chanturia et al. (2017), a mental template that uniquely matches the unfolding configuration is preactivated, leading to increased P300s. Seeing *knife and* should activate the template, while seeing *knife* (without *and*) should not activate the template. Similarly, in the present study, see *knife and* primed the word *fork,* while seeing *spoon and* did not. Both studies suggest that it is the phrasal, prefabricated, conventional status of binomial expressions that leads to the processing differences between binomials and controls. The findings also suggest that akin to single words, binomial expressions can be stored and represented in the mental lexicon, which supports frequency-based accounts of language acquisition, processing, and use. Further, according to Langacker (2000), when a structure no longer requires conscious attention to its parts through repeated use, it has a status of unit and is automatized (or routinized). Higher frequency of a unit results in a greater degree of entrenchment, i.e., cognitive routinization (Kemmer & Barlow, 2000, p. x). On the basis of frequency of occurrence, binomials (*knife and fork*) eventually become highly familiar and predictable to L1 speakers, and hence access to their final constituent (*fork*) given the initial two constituents is facilitated. For Chinese-English bilinguals, congruent binomials also become entrenched, and hence they showed a significant priming effect.

In addition, significant priming effects for congruent binomials by the English monolinguals and Chinese-English bilinguals in the present study also suggest that an intralexical link is formed between individual words which frequently co-occur in a lexical combination. Above and beyond links between semantically related words and between words and underlying concepts (see e.g., Kroll & Stewart, 1994), links can also be formed among the constituents of a MWE. This finding is in line with the results showing that the initial parts of a MWE can prime the terminal parts of it in priming research (Carrol & Conklin, 2014; Durrant & Doherty, 2010; Wolter & Gyllstad, 2011). For example, Durrant and Doherty (2010) investigated collocational priming by looking at whether a word can accelerate subsequent recognition of its collocate. The participants were first shown the prime (the adjective in an adjective-noun collocation, e.g., *music*) and then asked to decide whether or not the target (the noun which collocates with the adjective, e.g., *hall*) was a legal word in English. They found that L1 speakers of English responded more rapidly to the target in a collocation (*music* hall) than to the same word in a control phrase (*special* hall). Similarly, using a primed lexical decision task, Wolter and Gyllstad (2011) found significant collocational priming for verb-noun collocations (*give* [*an*] *answer*) by L1 speakers of English, as well as by advanced Swedish-English bilinguals. The present study provides further support for the formation of interlexical links between the constituents of MWEs (i.e., MWE priming) in monolingual and bilingual speakers. According to Hoey (2013, p. 2), priming between the individual words of a word combination is in essence repetition priming. When a listener or reader encounters *a* and *b* in combination, later exposure to *a* can accelerate the recognition of *b*. This indicates that speakers note subconsciously the linguistic context in which a word occurs, and they can identify other word/s that it often occurs with. Furthermore, the present study is in line with the finding of Wolter and Gyllstad (2011) that bilingual speakers are capable of developing links among the individual words of a word combination, especially when this word combination has an equivalent form in the L1.

The processing advantage observed for the congruent binomials (*sun and moon*) in the English monolinguals and Chinese-English bilinguals could also be attributed to the effect of predictability. As Siyanova-Chanturia, Conklin and Van Heuven (2011, p. 7) put it, being predictable is an intrinsic characteristic of a MWE. The reader can predict the upcoming word(s) by reading the initial constituent(s) of a MWE. Binomials are, by and large, fixed and highly predictable MWEs. Upon reading the initial constituent(s) of a binomial (*knife and* …), a proficient language user is very likely to come up with the correct completion (i.e., *fork*). Thus, the final constituent(s) of a binomial should have decreased processing times, or even be skipped altogether, because they are highly, or even uniquely, anticipated prior to the reader reaching them. Predictability has been found to play an important role in MWE processing. For example, in an eye-tracking study with L1 speakers of English, McDonald and Shillcock (2003a) found that sentences containing verb-noun combinations with a higher transitional probability (*avoid confusion*) were read faster than sentences containing combinations with a lower transitional probability (*avoid discovery*), evidenced by shorter initial fixation durations on the target nouns. Note that the length and frequency of the nouns (*confusion* vs. *discovery*) were matched, and the verb-noun combinations in different conditions (higher vs. lower transitional probability) were embedded in the identical sentence contexts (*One way to avoid <u>confusion/discovery</u> is to make the changes during vacation*.). At the electrophysiological level, familiar, predictable MWEs showed easier semantic integration of familiar information (leading to reduced N400s) and pre-activation of the mental template for uniquely predictable linguistic information (leading to larger P300s) (Molinaro & Carreiras, 2010; Siyanova-Chanturia et al., 2017; Vespignani et al., 2010).

In summary, the results reported in the present study confirm that phrase frequency affects the processing of binomials in L1 monolingual speakers and, possibly, in bilingual speakers. It suggests that language users, be they monolingual or bilingual speakers, are sensitive to the distribution of linguistic information at various grain sizes. In addition, an intralexical link can develop between individual words that frequently co-occur in a lexical combination such that the initial parts of a MWE can prime its terminal part(s). Consequently, language users may predict the upcoming word(s) by reading the initial constituent(s) of a MWE, greatly reducing the processing cost. However, phrase frequency in the L2 may not be the only factor that affects binomial processing in bilinguals, especially for unbalanced bilinguals whose L2 is distant from their L1, such as the Chinese-English bilinguals in the present study. In additional to phrase frequency in the L2, congruency across the L1 and L2 is another key factor that affects the processing of L2 binomials. A detailed discussion of the role of congruency is provided in Section 7.4.

7.3 L1 Inhibition

Surprisingly, unlike the English monolinguals, the English-Chinese bilinguals showed no significant facilitation for the congruent and English-only binomials, although English is also their native and dominant language. One may ask why the robust phrase frequency effect was not observed for the English-Chinese bilinguals. As discussed in Chapter 6, one possibility is that when the English-Chinese bilinguals were immersed in an L2-speaking country, their L1 had to be inhibited, in order to better control the competition from the L1 and facilitate L2 learning. Therefore, when the English-Chinese bilinguals had to switch back to their inhibited L1 to complete the experiment (which was conducted in their L1, English), a delay occurred in retrieving L1 words and, thus, their L1 processing was negatively affected.

This interpretation is supported both theoretically and empirically. During bilingual language processing, words from the non-target language are simultaneously activated (i.e., language non-selective access: e.g., Dijkstra et al., 2000; Dijkstra, Van Jaarsveld, et al., 1998; Van Heuven et al., 1998). To minimize cross-language interference and to confine processing to the relevant language, a language control process is implemented (Declerck, 2020). According to the Inhibitory Control Model (Green, 1998), when one language is used, the other language of a bilingual speaker is inhibited to prevent it from disrupting the selection of target language words. Such an inhibitory control process plays an important role in overcoming interference from the non-target language which is activated simultaneously (Levy et al., 2007). Specifically, the amount of inhibition applied to the non-target language depends on the baseline strength of its activation. The more dominant the non-target language, the stronger inhibition is needed. If bilinguals are unbalanced and dominant in the L1, their L1 will be strongly suppressed whenever they need to use the L2. However, the non-dominant L2 does not need to be suppressed as strongly when bilinguals use their dominant L1. As a result, the cost of switching back to using the L1 should be higher than switching back to using the L2, leading to a greater performance decrease in L1 than in L2 (e.g., asymmetric switching costs). Empirical studies have found that switching costs from the weaker language to the stronger language are larger than vice versa, in production (Jackson et al., 2001; Macizo et al., 2012; Meuter & Allport, 1999) and in comprehension (Jackson et al., 2004; Mosca & de Bot, 2017). For instance, Mosca and de Bot (2017) found that unbalanced Dutch-English bilinguals showed language switching cost only in the L1 but not in the L2 while performing a lexical decision task. Similarly, using a parity judgment task (i.e., classifying a digit as odd or even), Jackson et al. (2004) found that unbalanced bilinguals (L1 English-different L2s) showed language switching cost only in the L1 but not in the L2.

In addition, empirical studies have found that the L1 is inhibited while bilinguals are immersed in an L2 environment (Baus et al., 2013; Linck et al., 2009; Morales et al., 2014). For example, using a translation recognition task, Linck et al. (2009) compared two groups of unbalanced English-Spanish bilinguals: an immersed group in L2 environment and a classroom group in L1 environment. They found that the immersed group showed no sensitivity to English distractors that had form overlap with the presented Spanish words (correct translations: e.g., *cara-face*; English lexical-neighbor distractors: e.g., *cara-card*), compared to the classroom group who were matched in L2 proficiency. The results suggest that the L1 is less accessible for the immersed group than for the classroom group, and that L1 processing is impaired during L2 immersion. Therefore, when the L1 is inhibited during L2 use, it may result in a cost in overcoming the inhibition when switching back to the L1. It follows that the English-Chinese bilinguals may process L1 MWEs in a different way from their monolingual L1 counterparts, due to the need to inhibit interference from their L1 in an L2 immersion context, especially if their knowledge of L2 is comparatively weak. Future work would need to explore whether immersion affects L1 MWE processing by investigating the processing of MWEs by bilingual speakers in the L1 versus L2 immersion context.

Similar to the English-Chinese bilinguals in Experiment 2, the Chinese-English bilinguals in Experiment la were also immersed in an L2-speaking country. Thus, their L1 was likely to have been inhibited. The need to inhibit the non-target language (here, the participants’ L1/Chinese) may have delayed the potential activation of the original L1 MWEs. This might have caused the absence of priming for translated Chinese-only binomials for the Chinese-English bilinguals. Neurological studies have shown that competing information in the L1 needs to be suppressed to access information in an L2 (Abutalebi & Green, 2007; Pulido, 2021b). Inhibiting L1 interference can improve L2 performance, in both immersion and non-immersion context (i.e., the L1 Regulation Hypothesis: Bogulski et al., 2019). Therefore, when the Chinese-English bilinguals were in the L2 immersion context, their L1 might have been inhibited to improve L2 performance (Linck et al., 2009). However, different from the English-Chinese bilinguals, the Chinese-English bilinguals might have been more balanced in their knowledge of the two languages. For instance, they reported longer average years of L2 exposure than the English-Chinese bilinguals (3.9 vs. 1.9 years). Therefore, the amount of inhibition of the L1 (the non-target language) may decrease along with the relative balance between the two languages of a bilingual. This would explain why their L1 could have been more readily activated when compared to the L1 of English-Chinese bilinguals, facilitating the processing of congruent L2 binomials. This account is compatible with the extended Inhibitory Control model that is based on the language balance model (Casado, et al., 2021), which holds that the amount of inhibition applied to L1 during L2 use is related to the relative balance between the two languages. Studies have shown that when the two languages of a bilingual speaker are relatively balanced, the switching costs between languages becomes comparable, i.e., symmetrical switching costs (Christoffels et al., 2007;

Declerck et al., 2013; Schwieter & Sunderman, 2008). In order to test this account, future research would need to compare the processing of L1 MWEs by bilinguals in an immersion versus non-immersion context. This would allow us to examine how L2 immersion affects the L1 MWE processing (Linck et al., 2009; Morales et al., 2014).

7.4 Cross-language Influence in the L1→L2 Direction

While the Chinese-English bilinguals showed a significant priming effect for congruent binomials, they showed no priming effect (-1 ms) in the processing of the terminal words of the *English-only* binomials (*bread and butter*) compared to the control phrases (*toast and butter*), despite the fact that congruent and English-only binomials were matched in English phrase frequency (t = -0.08, p = .94) and showed no difference in monolingual processing (a 22 ms priming effect for congruent binomials and 23 ms priming effect for English-only binomials). Thus, congruent English binomials showed a processing advantage in the Chinese-English bilinguals but English-only binomials did not, suggesting cross-language influences in the processing of L2 binomials. This indicates that above and beyond phrase frequency in the L2, having or not having an equivalent form in the L1 is a key factor that affects the processing of L2 binomials in Chinese-English bilinguals. Specifically, having an equivalent form in the L1 (i.e., congruency) may greatly facilitate the processing of congruent English binomials for the Chinese-English bilinguals. Without having an equivalent form in the L1 (e.g., English-only binomials), however, it may be difficult for such an advantage to occur (e.g., it might need more L2 exposure).

This result is in line with the results of previous research showing that for bilinguals, congruent L2 MWEs enjoy a processing advantage over incongruent L2-only MWEs even when they are matched for phrase frequency in the L2 (i.e., Carrol et al., 2016; Wolter & Gyllstad, 2011, 2013; Yamashita & Jiang, 2010). For example, using a phrase-acceptability judgment task, Yamashita and Jiang (2010) found that Japanese-English bilinguals showed a processing advantage for congruent collocations (*kill animals*) than incongruent collocations (*kill time*, whose Japanese equivalent literally translates as ‘crush/break time’). This congruency effect has been reaffirmed in following studies (Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2018). Advanced L1 Swedish-L2 English bilinguals showed a significant processing advantage for congruent verb-noun or adjective-noun collocations (*give an answer, handsome man*) over incongruent, L2-only collocations (*pay a visit*, *identical twins*) in a primed lexical decision task (Wolter & Gyllstad, 2011) and in a phrase-acceptability judgment task (Wolter & Gyllstad, 2013). In contrast, in the two studies, monolingual English controls responded to congruent collocations as quickly and accurately as to incongruent English-only collocations. Similar results were reported with L1 Japanese-L2 English bilinguals (Wolter & Yamashita, 2018). Wolter and Yamashita (2018) investigated how intermediate and advanced Japanese-English bilinguals, as well as English monolinguals, processed congruent (*strong wind*) and incongruent English-only (*busy road*) adjective-noun collocations in a phrase-acceptability judgment task. Both groups of bilingual speakers responded significantly faster to congruent than to incongruent English-only collocations. Crucially, monolingual English speakers showed no differences in RTs for these two types of collocations. Further, using eye-tracking, Carrol et al. (2016) found that high-proficiency Swedish-English bilinguals processed congruent idioms (*break the ice*) faster than literal controls (*crack the ice*), evidenced in early measures and late measures. However, bilinguals processed English-only idioms (*kick the bucket*) faster than literal controls (*drop the bucket*) in the late measures, and not in the early measures. In contrast, monolingual English controls processed congruent and English-only idioms relative to their corresponding controls in a similar way, evidenced by the analysis of the early and late measures. Together, these studies suggest that congruency plays an important facilitative role in L2 MWE processing, a prominent marker of cross-language influences.

In addition, unlike the studies by Wolter and colleagues (Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2018) which used implausible, non-collocational items as controls (e.g., **angry use*, **legal plant*, **final others*), the present study used infrequent but semantically plausible phrases as controls (*spoon and fork, toast and butter, exercise and strength*). As pointed out by Sonbul and Siyanova-Chanturia (2021), implausibility is likely to cause a significant delay in the processing. Control items should thus be infrequent and not strongly associated phrases that are nonetheless semantically plausible (Sonbul & Siyanova-Chanturia, 2021). Previous studies that used implausible items as controls probably overestimated the processing advantage for both congruent and, *importantly*, incongruent (L2-only) MWEs. For example, Wolter and Gyllstad (2011, 2013) reported a significant processing advantage for English-only collocations (*pay a visit*, *identical twins*) over implausible controls (**defend sales*, **angry use*) in Swedish-English bilinguals. However, conflicting findings were reported in Wolter and Yamashita (2015, 2018), which found no processing advantage for English-only collocations (*catch breath*, *busy road*) over control items (**bad gift*, **active gift*) in Japanese-English bilinguals. Similarly, no processing advantage was found in Chinese-English bilinguals for English-only binomials over control phrases in the present study. Importantly, by using semantically plausible phrases as controls, the present study gives a more accurate picture of the phrase frequency effect, suggesting that L2-only MWEs is less likely to show a processing advantage over novel phrases when compared with congruent MWEs over controls.

In addition, in the present study, the two content words within the binomial (*knife and fork*) and the matched control (*spoon and fork*) were equally strongly associated in forward association strength (congruent condition, t = 1.41, p = .17; English-only condition, t = 1.75, p = .15). This was done to ensure that any processing advantage for binomials over their corresponding controls was not due to the first word in the binomials (*knife*) being a better prime than the first word in the control items (*spoon*) to the same target (*fork*) (e.g., Siyanova-Chanturia et al., 2017; Siyanova-Chanturia, Conklin, & Van Heuven, 2011).

7.4.1 L1 MWE Activation Account

Although I found a clear evidence of L1 influence on L2 binomial processing (i.e., congruent binomials showed a robust priming effect, while English-only binomials showed no priming in the Chinese-English bilinguals but did show priming in the English monolinguals), I did not find any evidence that L1 influence extends to the processing of translated Chinese-only binomials. Similar to the English monolinguals, the Chinese-English bilinguals did not show any significant priming effect for the translated Chinese-only binomials, suggesting that translated Chinese-only binomials were not processed as binomials in the L2. This finding indicates that the congruency effect – a greater processing advantage for congruent over English-only binomials – may not be attributed to the simultaneous activation of L1 MWE translation equivalents in L2 processing, as predicted by the L1 MWE activation account.

This finding is contrary to the findings of past studies with idioms that translated L1-only idioms were processed faster than matched controls (i.e., Carrol & Conklin, 2014, 2017; Carrol et al., 2016). For example, Carrol and Conklin (2014) found that intermediate proficiency Chinese-English bilinguals responded faster to the final word of translated Chinese-only idioms (*draw a snake and add* <u>*feet*</u>) than to the final word of its corresponding matched control items (*draw a snake and add* <u>*hair*</u>) in a lexical decision task, despite the fact that the target words were matched in lexical frequency and word length, and that English monolingual controls responded similarly to the target words in the different conditions (idiom vs. control). The processing advantage for translated L1-only idioms relative to control novel phrases by bilinguals was also reported in a follow-up study that used eye movements (Carrol & Conklin, 2017). Idioms (*draw a snake and add* <u>*feet*</u>) which were embedded in a short sentence context biasing the figurative meaning were read faster than control novel phrases (*draw a snake and add hair*) in Experiment 1 by Chinese-English bilinguals, but not by English monolinguals, evidenced in the analysis of first fixation durations and total reading time. This result was taken as evidence that the form of the translated Chinese-only idioms is recognized. Further, translated L1-only idioms have been found to show a similar level of facilitation as congruent idioms relative to their corresponding matched controls. Using eye-tracking, Carrol et al. (2016) found that advanced Swedish-English bilinguals, but not English monolinguals, showed a processing advantage for translated Swedish-only idioms (*play monkey*) over controls (*taste monkey*) in terms of measures for the whole idiom (e.g., total reading time) and measures for its final words (e.g., likelihood of skipping). Importantly, the processing advantage for translated Swedish-only idioms over controls is comparable to that for congruent idioms (*lose your head*) over controls (*hurt your head*). This suggests that congruent idioms did not show more facilitation than translated L1-only idioms, despite (additional) L2 exposure. Carrol et al. (2016) thus concluded that it was L1, rather than L2, knowledge that determined the ease of processing of idioms in bilingual speakers. Together, these studies suggest that translated L1-only idioms activate L1 equivalents, leading to facilitation in L2 processing. They support the L1 MWE activation account, according to which known L1 MWEs are automatically activated in L2 processing. However, the present study did not find any facilitation for translated Chinese-only binomials by the Chinese-English bilinguals. This result does not show support for the L1 MWE activation account, which assumes the automatic activation of L1 MWEs in L2 processing.

7.4.2 L2 MWE Experience Account

Conversely, the finding of the present study is in line with studies on collocations which reported no facilitation for translated L1-only collocations over non-collocational combinations (Wolter & Yamashita, 2015, 2018). Using a double lexical decision task (2015) and a phrase acceptability judgment task (2018), Wolter and Yamashita found that Japanese-English bilinguals processed translated Japanese-only collocations as fast as non-collocational matched controls, irrespective of the type of collocations (adjective-noun or verb-noun) or the level of L2 proficiency of bilinguals (intermediate or advanced). Of note is that the non-collocational baseline items used in the two studies (e.g., *careful branch*, *dark guess*, *false months*) were less semantically plausible, which might have led to an increased processing cost (Sonbul & Siyanova-Chanturia, 2021). Nevertheless, despite the fact that the use of less semantically plausible controls may raise the processing cost, translated L1-only collocations are not processed faster than them. Similarly, the present study finds no evidence for the processing advantage for translated L1-only MWEs. Although the Chinese-English bilinguals showed robust priming for congruent binomials (*sun and moon*), they showed no priming effect for translated Chinese-only binomials (*wisdom and strength*). Both in the present study and in Wolter and Yamashita (2015, 2018), translated L1-only MWEs were processed similarly to unknown word combinations. These results suggest that although the L1 influences the processing of MWEs in the L2, this influence may not extend to translated L1-only MWEs.

Therefore, contrary to the prediction of the L1 MWE activation account, the results for the Chinese-English bilinguals suggest that known L1 binomials are not automatically activated in the processing of translated Chinese-only binomials. Instead, the absence of priming for translated Chinese-only binomials in the Chinese-English bilinguals is predicted by the L2 MWE experience account, which holds that prior exposure to a word sequence in an L2 is essential for it to be processed faster than its control, whether or not it has an equivalent form in the L1. Since there is no clear evidence of the simultaneous activation of L1 equivalents in L2 MWE processing, the finding of this study suggests that the greater processing advantage for congruent MWEs stems from their likely earlier acquisition and, thus, greater experience with their L2 form, compared with incongruent L2-only MWEs. In the L2 MWE experience account, congruent MWEs are assumed to be acquired earlier than incongruent MWEs, because acquisition is more straightforward when there is correspondence between the L1 and L2 (due to positive cross-language transfer). That is, acquiring congruent MWEs may be faster than acquiring incongruent, L2-only MWEs, because the former may take less time and require less exposure to the L2 than the latter. According to Ullman (2014), when a speaker or reader encounters a MWE in an L2 that also exists in their L1, an initial strong declarative memory trace is likely to be established, facilitating gradual acquisition of procedural knowledge from input. After sufficient experience with the language, procedural knowledge should tend to take precedence over declarative knowledge, resulting in a robust facilitation in the processing (Ullman, 2014, p. 143). A number of studies have shown that L1 plays an important role in the acquisition of L2 MWEs (Nesselhauf, 2005; Römer et al., 2014; Sonbul et al., 2020; Yamashita & Jiang, 2010). For example, Yamashita and Jiang (2010) found that congruent collocations are processed more accurately and/or more quickly than incongruent L2-only collocations by lower- and higher-proficiency Japanese-English bilinguals, whereas they were processed with no difference by monolingual English controls. To further attest to the tenet that congruent MWEs are acquired earlier than incongruent L2-only MWEs, future research may look at the processing of MWEs in children or L2 learners of different levels of proficiency (e.g., beginner, intermediate and advanced), which can provide more empirical evidence for the L2 MWE experience account.

If congruency does facilitate the acquisition of MWEs and, hence, congruent MWEs are acquired before incongruent L2-only MWEs in bilinguals, then congruent MWEs should be processed faster than incongruent MWEs due to the age/order-of-acquisition effect (e.g., Wolter & Gyllstad, 2013; Wolter & Yamashita, 2015, 2018). Akin to what has been reported for single words, age-of-acquisition also affects the processing of units longer than a word (e.g., phrases: Arnon et al., 2017; compound words: Juhasz, 2018). Arnon et al. (2017) found that early-acquired phrases/lexical bundles (*a good girl*) were responded to faster than late-acquired ones (*a good dad*) by L1 adults in a phrasal decision task. Similar results were reported in an eye movement study with L1 speakers. Juhasz (2018) found that the age-of-acquisition of compound words (e.g., *airport*, *bodyguard*), rather than that of the individual morphemes, affected the processing of compound words, as shown by the gaze durations and total fixation durations.

However, although the present study provides new evidence for the potential origin of the congruency effect in bilingual processing of MWEs (i.e., the likely earlier acquisition and, therefore, greater experience with their L2 form of congruent vs. incongruent MWEs), it is still insufficient to definitively support one of the two accounts, that is, the L1 MWE activation account or L2 MWE experience account. First, studies on idioms show a robust facilitation for translated L1-only idioms over novel L2 phrase controls, suggesting that there might be automatic activation of known L1 MWEs, as predicted by the L1 MWE activation account (Carrol & Conklin, 2014, 2017; Carrol et al., 2016). In contrast, the present study and studies on collocations have not found facilitation in the processing of translated L1-only binomials or collocations (Wolter & Yamashita, 2015, 2018).

It is hard to reconcile the seemingly contrasting results. One possibility is that these different findings are related to the type of MWEs being processed, i.e., literal versus figurative expressions. It is possible that the facilitation for translated L1-only idioms is driven by the conceptual L1-L2 overlap, in which case, the figurative meaning of idioms is activated even if they have never been encountered in the L2 before (Carrol et al., 2016, but see Carrol and Conklin, 2017). However, this hypothesis needs further verification, because evidence is mixed with regard to whether or not there is meaning activation of translated L1-only idioms, with meaning activation reported in Carrol et al. (2016) but not in Carrol and Conklin (2017). In addition, there is evidence for form activation in the processing of translated L1-only idioms, suggesting that the facilitation for translated L1-only idioms may also be driven by form overlap (Carrol & Conklin, 2017; Carrol et al., 2016).

Another possible reason for the discrepancy between the present results and those of Carrol and colleagues is the use of different methods. Most idiom studies employed eye-tracking and looked at idioms embedded in context, whereas collocation and binomial studies used behavioural measures such as phrase acceptability judgement tasks and primed lexical decision tasks and investigated stimuli out of context. This makes it hard to compare their findings. To address this issue, future studies may employ similar methodology such as eye-tracking and explore the processing of literal MWEs embedded in context to better compare with previous research on idioms.

Second, while the absence of priming for translated Chinese-only binomials for the Chinese-English bilinguals suggests that known L1 MWEs are not automatically activated in L2 processing, it may be too early to abandon the L1 MWE activation explanation entirely. For example, the original L1 MWEs may have been activated when the bilinguals read their L2 translation equivalents, but this activation may have been counteracted by the need to inhibit the non-target language (here, the participants' L1, Chinese), since the task was completed entirely in the L2 (Green, 1998). Additionally, the Chinese-English bilinguals' L1 may be inhibited, at the whole language level, in the context of their L2 immersion (Linck et al., 2009). Therefore, the L1 inhibition necessitated by the experimental task and the country of residence contexts may have cancelled out the possible activation of the L1 MWEs, resulting in no priming for translated Chinese-only binomials.

Further, the two accounts of the congruency effect are both theoretically and empirically motivated. The L1 MWE activation account is aligned with the non-selective lexical access account of bilingual processing (Dijkstra & Van Heuven, 2002; Van Heuven et al., 1998), supported by empirical evidence for masked translation priming, which has been observed even when the two languages do not share the same writing system (e.g., Hebrew-English: Gollan et al., 1997; Chinese-English: Wang & Forster, 2010) and in purely monolingual contexts (Thierry & Wu, 2007; Zhang et al., 2011). The L2 MWE experience account, on the other hand, is aligned with usage- and exemplar-based acquisition and processing accounts (Bybee, 2006; Langacker, 2000), supported by empirical evidence for phrase frequency effects in language processing (e.g., Arnon & Snider, 2010; Siyanova-Chanturia, Conklin, & Van Heuven, 2011).

Last but not least, one may wonder whether or not the two accounts have to be mutually exclusive. Is it possible that the congruency effect in MWE processing is due to the automatic activation of corresponding L1 MWEs *as well as* the earlier acquisition of congruent MWEs than incongruent L2-only MWEs? For example, in the case of translated L1-only MWEs, there may be some level of L1 activation, but the degree of L1 activation may not be sufficient for any significant processing advantage to be observed. In the case of congruent MWEs, however, both the activation of known L1 MWEs and likely earlier acquisition of congruent MWEs may contribute to the processing advantage for them over incongruent L2-only MWEs.

In summary, the present study replicates the findings of previous research that congruent MWEs show a processing advantage in the L2 over matched controls and extends the finding to a new type of MWEs, binomials. In addition, the present study did not find any processing advantage for translated L1-only binomials over matched novel controls, suggesting that there may be no automatic activation of the L1 MWEs. Thus, the congruency effect may be attributed to the cross-language MWE transfer at the learning/acquisition point, resulting in an order-of-acquisition effect, in that, congruent binomials may be acquired earlier than incongruent L2-only binomials. However, the discrepancy between studies on idioms and those on collocations and binomials are hard to reconcile, making it impossible to definitively support one or the other of the two accounts. Therefore, this important line of enquiry should continue, as more evidence is needed to tip the balance towards one or the other account of the congruency effect in L2 MWE processing.

7.5 Cross-language Influence in the L2→L1 Direction

The English-Chinese bilinguals showed some numerical facilitation for congruent binomials (*sun and moon*) compared to control phrases (*star and moon*), although they showed no facilitation in the processing of English-only binomials (*bread and butter*) compared to control phrases (*toast and butter*). They processed congruent binomials faster than their controls (mean difference = 20 ms, model estimate = 12 ms), although this facilitation did not reach statistical significance after applying the Bonferroni adjustment. Similar to the Chinese-English bilinguals in Experiment 1a, the English-Chinese bilinguals also showed a congruency advantage (marginally significant) in the processing of congruent English binomials over English-only binomials. This suggests that the L2 may also influence the processing of binomial expressions in the L1 and that cross-language influences in MWE processing may occur in both directions. What mechanism, however, underlies the L2 influence on the processing of L1 binomials?

One possibility is that L2 binomial equivalents are activated during the L1 task, leading to the faster processing of MWEs that exist in both languages. This result is not unlike the findings of an automatic activation of single words in the weaker L2 when stimuli are shown in mixed languages (Dijkstra et al., 2000) or in the L1 only (e.g., Van Hell & Dijkstra, 2002). In bilingual word processing, a number of studies have found that the weaker

L2 also affects native language processing, even when the bilinguals are performing in their native and dominant language and in a purely native language context (e.g., Van Hell & Dijkstra, 2002). For instance, Van Hell and Dijkstra (2002) found that L1 words that are cognates with their L2 translations (e.g., Dutch-English: *bakker-baker*) lead to faster lexical decisions than L1 control non-cognate words. In addition, studies using cross-language priming paradigm have also found priming from L2 to L1, with cognate translations (e.g., Gollan et al., 1997; Sánchez-Casas & García-Albea, 2005) and non-cognate translations (e.g., Schoonbaert et al., 2009). However, this interpretation should be viewed with caution, because whether or not corresponding L1 MWEs are automatically activated in L2 processing (i.e., the L1 MWE activation account) is still an open question. If known L1 MWEs are not simultaneously activated in L2 processing, it is unlikely that known L2 MWEs are simultaneously activated in L1 processing. Thus, future research should investigate the potential activation of known L2 MWEs in L1 processing in bilinguals.

An alternative, but not *mutually exclusive*, hypothesis is that the English-Chinese bilinguals may have extensive practice in regulating the co-activation of L1 when learning congruent L2 binomials that have L1 translations and, thus, they may experience less processing costs when switching back to the L1 when it comes to congruent binomials relative to English-only binomials. Past studies have found that learning cognate L2 words that share sounds and semantic concepts with the L1 is faster than learning L2 non-cognate words (e.g., Kaushanskaya et al., 2013). According to Bogulski et al. (2019), faster learning for cognate words may be due to the fact that bilinguals have extensive practice in regulating the co-activation of L1 translations. Greater skill in regulating the L1 may lead to a bilingual advantage in word learning such that bilinguals outperform monolinguals in vocabulary learning, especially learning vocabulary through L1 translations (for a review, see Hirosh & Degani, 2018). For example, Bogulski et al. (2019) found that English-Spanish bilinguals were better able to learn novel Dutch vocabulary than English monolinguals, when both groups of participants learn the new words through L1/English translations. They argued that this finding is due to the possibility that bilinguals have extensive experience in regulating their L1 (English) and, thus, they have greater L1 regulatory skill that creates a foundation for learning foreign words through the L1. At the phrase level, greater skill in regulating the L1 is also found to lead to more successful second language learning. For example, Pulido (2021b) investigated how practice difficulty during second language learning predicts learning outcomes of verb-noun collocations. In the practice session, one group of English-Spanish bilinguals was presented with L1-related distractors that would be acceptable in the L1 but not in the L2 (e.g., *jugar* 'play' – *gastar* 'spend' – *partidos* 'matches', with *gaster* as the distractor), whereas the other group was presented with unrelated distractors (e.g., *jugar* 'play' –*poner* 'put' – *partidos* 'matches', with *poner* as the distractor). Then they were asked to select the correct verb which could collocate with the noun (e.g., *jugar* 'play'). The group of bilinguals who were presented with L1 distractors, the more difficult condition, showed higher rates of learning success than the group that was presented with unrelated distractors, as measured in subsequent tests. The author interpreted the results as evidence that more successful L2 learning is achieved through inhibition of interference from the native language. If learning new vocabulary through L1 translation improves L1 regulatory skill, it is plausible to hypothesize that learning congruent L2 MWEs that have L1 translations may also improve skill in regulating L1, because bilinguals may have had extensive practice in regulating their L1 when learning congruent L2 MWEs. Thus, when bilinguals switch back to their L1, congruent MWEs may experience less processing costs and exhibit faster access to their L1 than incongruent L1-only MWEs. That may explain why the English-Chinese bilinguals showed greater processing advantage in the processing of congruent L1 binomials over English-only binomials.

Another alternative is that when bilinguals learn congruent binomials in the L2, they may refer to the knowledge of the equivalent binomials in the L1. That is, they may learn L2 binomials via L1 translation equivalents (i.e., the cross-language transfer). This represents an additional access to the L1 binomials, making these L1 binomials more frequently co-activated in the course of L2 learning, compared with L1-only binomials. Consequently, it strengthens the knowledge of the L1 binomials that are congruent with L2 binomials, rendering these L1 binomials more resilient to the L1 inhibition in the immersion context.

Nevertheless, the congruency effect for the English-Chinese bilinguals was weaker compared with the Chinese-English bilinguals who showed a significant priming effect in the processing of congruent binomials and a clear congruency advantage over English-only binomials. This finding is in line with previous research that found that L1 typically has a higher impact on L2 processing than the reverse (Jiang, 1999; Schoonbaert et al., 2009). Many cross-language priming studies have found strong priming from L1 to L2 and weaker or no priming from L2 to L1 (Finkbeiner et al., 2004; Gollan et al., 1997; Jiang, 1999; Schoonbaert et al., 2009). For instance, Schoonbaert et al. (2009) showed that translation priming with noncognate translation pairs in unbalanced Dutch(L1)-English(L2) bilinguals was significantly stronger from L1 to L2 (*meisje-GIRL*) than from L2 to L1 (*girl-MEISJE*), although priming from L2 to L1 was also significant. This cross-language influence asymmetry has also been reported in eye-tracking studies (Marian & Spivey, 2003; Weber & Cutler, 2004). Taken together, these studies suggest that cross-language influences in single word processing are possible in both directions, from L1 to L2 and from L2 to L1, but cross-language influences may be stronger and more reliable from L1 to L2 than vice versa.

Finally, the English-Chinese bilinguals showed no facilitation for translated Chinese-only binomials over controls. The same pattern of results was observed in the English monolinguals and Chinese-English bilinguals. This indicates that prior exposure to the L1 sequences is necessary for them to show any facilitation in processing relative to control novel phrases in bilinguals. Since the Chinese-English bilinguals showed no significant priming for translated Chinese-only binomials, it is not surprising that there was no activation of translated Chinese-only MWEs for the English-Chinese bilinguals either, because cross-language influence in the L1-L2 direction is normally stronger than that in the opposite direction. That is, cross-language influence in the L2→L1 direction were less likely to take place when no such influences were observed in the L1→L2 direction.

7.6 Conclusion

In this chapter, I have discussed the effects that phrase frequency, L2 immersion, and cross-language congruency can have on the processing of MWEs in bilinguals. Phrase frequency effects have been consistently reported in earlier research. This study reaffirms previous findings by showing that monolingual and bilingual speakers are sensitive to the distributional properties of large language units: they showed significant priming for high frequency binomials (in the case of congruent binomials), but no priming for low frequency phrases (in the case of translated Chinese-only binomials). However, L2 immersion may negatively affect L1 processing such that the English-Chinese bilinguals showed no significant priming for congruent or English-only binomials. Their L1 may have been inhibited in L2 immersion context, which likely attenuated the activity of the L1 and resulted in processing costs when switching back to the L1 again. The Chinese-English bilinguals, however, were more balanced between the two languages and, thus, they may not have needed to inhibit their L1 as strongly. Their L1/Chinese significantly affected the processing of binomials in the L2. The results from the Chinese-English bilinguals in Experiment 1a show that congruent L2 binomials showed a significant priming effect while English-only binomials did not. In contrast, the monolingual English controls in Experiment 1b showed comparable priming effects for them. Thus, this study replicated previous studies and reaffirmed that congruency facilitates L2 MWE processing. In addition, translated Chinese-only binomials showed no priming effect in the Chinese-English bilinguals, suggesting that there may be no automatic activation of the L1 as predicted by the L1 activation account. Given no direct evidence supporting the L1 MWE activation account, the underlying mechanism behind the congruent effect in L2 MWE processing may be the L2 MWE experience account, according to which the greater processing advantage for congruent MWEs stems from their likely earlier acquisition (and therefore, greater experience with their L2 form), compared with incongruent L2-only MWEs.

By comparison, the English-Chinese bilinguals in Experiment 2 showed no priming for English-only binomials, but a weak priming for congruent binomials. No facilitation for English-only binomials suggests that L1 may be inhibited in L2 learning and immersion contexts and, thus, switching back to L1 may come at a cost. Interestingly, congruent binomials seem to experience less processing costs due to L1 inhibition. Similar to the Chinese-English bilinguals, the English-Chinese bilinguals also exhibited greater processing advantage for congruent over English-only binomials, suggesting that L2 may influence the processing of binomials in the L1. (However, as mentioned in section 6.4, this interpretation needs to be viewed with caution, because the greater processing advantage for congruent over English-only binomials in English-Chinese bilinguals might result, at least in part, from differences in the responses to the control items). Based on these results, I argue that crosslinguistic influence may occur from the non-dominant L2 to the dominant L1, even in an entirely within-L1 task. Thus, crosslinguistic influences in the processing of binomials may be bi-directional, although the influence in the L1→L2 direction is stronger than in the L2→L1 direction. This conclusion is in line with studies on bilingual word processing suggesting that crosslinguistic influences are bi-directional.

Chapter 8 Conclusion

8.1 General Conclusions

With the increase of research on MWEs, MWE processing has attracted much interest from psychologists, neurolinguists, and cognitive scientists. The research on MWE processing has centred on how L1, as well as L2, speakers process frequent versus novel linguistic information, or figurative versus literal language (Siyanova-Chanturia & Van Lancker Sidtis, 2019). The evidence on MWE processing converges on the conclusion that MWEs are processed faster than matched novel strings of language by L1 speakers (Arnon & Snider, 2010; Tremblay et al., 2011; Vilkaite, 2016), and L2 speakers (Hernández et al., 2016; Jiang & Nekrasova, 2007; Siyanova-Chanturia, Conklin, & Van Heuven, 2011). The processing advantages for MWEs that appear to exist among L1 and (proficient) L2 speakers are found to be driven by their phrase frequency, familiarity and, by extension, predictability (Siyanova-Chanturia, 2015).

While much of the research has centred on the role of factors such as frequency, familiarity, and predictability in the processing of MWEs by L1 and L2 speakers, how the knowledge of one language of a bilingual speaker affects the processing of MWEs in the other language has received limited attention, with only a handful of studies investigating this issue (Carrol et al., 2016; Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2018; Yamashita & Jiang, 2010). Nevertheless, these studies provide converging evidence suggesting that congruent MWEs are processed faster and more accurately than incongruent L2-only MWEs. This finding is known as the congruency effect and interpreted as a significant marker of the presence of CLI in L2 MWE processing.

What remains unclear, however, is what mechanisms underpin the congruency effect in the processing of L2 MWEs and whether or not the knowledge of an L2 can affect the processing of L1 MWEs, i.e., whether or not CLI can also occur in the direction of L2→L1. The aim of this research has been to shed light on the two issues. To accomplish this goal, two experiments were conducted to examine cross-language influences in MWE processing in both directions, L1→L2 and L2→L1. The first experiment was designed to investigate the role of L1 in L2 MWE processing, and the second experiment to examine the influence of L2 on L1 MWE processing. The goal of the present chapter is to summarize and further discuss the findings that emerged in the two experiments in this research. It also discusses the implications of this research for understanding the underlying mechanisms of CLI in MWE processing.

In Experiment 1, this research investigated the processing of binomials (*knife and fork*) with a group of Chinese-English bilinguals (Experiment 1a) and English monolingual controls (Experiment 1b). Using a primed lexical decision task, I tested how Chinese-English bilinguals and English monolingual controls processed three types of binomials relative to their corresponding controls, including congruent (*sun and moon* vs. *star and moon*), English-only (*bread and butter* vs. *toast and butter*), and translated Chinese-only binomials (*wisdom and strength* vs. *exercise and strength*). The congruency effect, as found in earlier studies (Carrol et al., 2016; Wolter & Gyllstad, 2011, 2013), was replicated in the present study. The Chinese-English bilinguals in Experiment 1a showed a priming effect for congruent binomials but not for English-only binomials, whereas the English monolingual controls in Experiment lb showed comparable priming effects for congruent and English-only binomials. Critically, the Chinese-English bilinguals processed translated Chinese-only binomials in a similar way to control novel phrases, as predicted by the L2 MWE experience account while contrary to the predictions of the L1 MWE activation account. It suggests that prior exposure to a word sequence in an L2 is essential for it to be processed faster than its control, in the case of binomials. The congruency effect in the processing of L2 binomials may thus be attributed to the likely earlier acquisition of congruent L2 binomials (and therefore, greater experience with their L2 form), compared with incongruent L2-only binomials.

In Experiment 2, I tested the processing of binomials in another group of bilinguals, English-Chinese bilinguals, with the same task and materials as Experiment 1. Contrary to the English monolingual controls in Experiment lb, the English-Chinese bilinguals in Experiment 2 showed no priming effect for English-only binomials and a weak priming effect for congruent binomials, presented in their L1 (English). Given the compelling evidence that L1 speakers show a robust processing advantage for binomials relative to control novel phrases (Siyanova-Chanturia et al., 2017; Siyanova-Chanturia, Conklin, & Van Heuven, 2011), I conclude that this result may be attributed to the immersion context where the English-Chinese bilinguals were. The L1 of the English-Chinese bilinguals may have been strongly inhibited in the L2 immersion context. Bilinguals may subconsciously suppress their L1, in order to better control L1 lexical competition and facilitate L2 learning (Linck et al., 2009), especially when their L1 is stronger than their L2 (Green, 1998). When they had to switch back to their L1 to complete the experiment, their L1 processing could have been impaired, resulting in weakened L1 performance. However, the English-Chinese bilinguals processed congruent binomials faster than their controls (mean difference = 20 ms). They showed a clear trend toward priming for congruent binomials, even when the overall L1 performance was weakened in an L2 immersion context. It suggests that L2 knowledge may affect the processing of L1 MWEs, indicating the likely presence of CLI in the direction of L2→L1.

Although CLI in MWE processing can occur in both directions, the underlying mechanisms involved may be different. The CLI in MWE processing in the L1→L2 direction may be due to the order-of-acquisition effect (as a result of the L1→L2 transfer). That is, congruent MWEs are likely to be acquired before incongruent L2-only MWEs and, thus, they are processed faster than later-acquired L2-only MWEs. However, the CLI in MWE processing in the L2→L1 direction may be due to better L1 regulation for congruent than English-only binomials as a result of learning congruent L2 binomials through L1 translations. Greater skill in regulating the L1 may lead to less switching or processing costs when switching back to the L1. Alternatively, referring to equivalent forms in the L1 when learning congruent L2 binomials may result in additional access to congruent binomials in the L1 when compared with L1-only (here, English-only) binomials, leaving congruent L1 binomials more frequently co-activated and, thus, less vulnerable to L1 inhibition. While these explanations may account for CLIs in MWE processing in each direction separately, one may wonder whether or not a unitary explanation can explain CLIs in MWE processing in both directions. If MWEs in both languages are simultaneously activated, this may explain why congruent MWEs show a processing advantage over incongruent MWEs. This cross-language activation account may explain CLIs in MWE processing in both directions. This is not unlike the cognate facilitation effect, wherein cognates are processed faster and with less effort than noncognate control words (de Groot et al., 2000, Experiment 2; Dijkstra, Van Heuven, et al., 1998; Libben & Titone, 2009). The cognate facilitation effect is generally interpreted as evidence that the two languages of a bilingual are activated simultaneously or non-selectively, and this cross-language activation influences processing (for a review, see Van Hell & Tanner, 2012). Similar to cognates, congruent MWEs are similar in form and meaning, i.e., they share form (translation equivalents), structure (same word order), and referential meaning (same construct). It follows that when processing congruent MWEs, the two languages of a bilingual may be activated simultaneously, leading to facilitation for congruent MWEs as for cognates.

However, to provide such a unitary account for CLIs in MWE processing in both directions, we need to argue that the L1 is activated in L2 MWE processing in the first place. While the results of no processing advantage for translated Chinese-only binomials in Chinese-English bilinguals suggest that known L1 MWEs are not automatically activated, I am in no position to abandon the L1 MWE activation explanation entirely. First, studies on idioms did find a robust processing advantage for translated L1-only idioms over control novel phrases in bilinguals, suggesting that known L1 idioms may be automatically activated in L2 MWE processing. Second, we cannot rule out the possibility that the original L1 binomials may have been activated when the Chinese-English bilinguals read their L2 translation equivalents, but this L1 activation may have been counteracted by L1 inhibition. When Chinese-English bilinguals performed the task in their L2/English, they may have inhibited their L1 subconsciously. This may have cancelled out the possible activation of the L1 MWEs, resulting in no priming for translated Chinese-only binomials in the Chinese-English bilinguals. Third, we also cannot rule out the possibility that the original L1 MWEs were activated, but this activation was not sufficient for any significant processing advantage to present. According to Wolter and Gyllstad (2011), the processing advantage for congruent over L2-only collocations may be due to the fact that congruent collocations are activated in both L1 and L2, i.e., doubled activation. The likely *doubled* activation leads to the congruency effect, rather than activation in the L1 solely. That is, the congruency effect may be a result of both L1 and L2 activation. Together, it is still possible that known L1 MWEs are simultaneously activated in L2 MWE processing.

In sum, the robust congruency effect in L2 binomial processing by the Chinese-English bilinguals, and a trend for the congruency effect in L1 binomial processing by the English-Chinese bilinguals, substantiate the main conclusion that cross-language influences in MWE processing can be bi-directional. Directionality may modulate CLIs in MWE processing in that CLI in the direction of L1→L2 is stronger than CLI in the reverse direction, L2→L1. The present study extended previous findings by showing that congruency facilitates the bilinguals' processing of binomials, a less commonly studied type of MWEs. Importantly, rather than merely documenting the presence of such CLIs, the present study further probed the underlying mechanisms associated with such CLIs. While some studies proposed that congruent MWEs show greater processing advantages than incongruent L2-only MWEs because known L1 MWEs are activated simultaneously, the present study argues that it may not necessarily be the case. The congruency effect in L1 MWE processing may be due to the possibility that congruent MWEs are acquired earlier than incongruent L2-only MWEs as a result of positive L1 transfer and, consequently, congruent MWEs are processed faster due to the order-of-acquisition effect. The present study also contributed to the current literature by testing the CLI in MWE processing in the L2→L1 direction, an issue that has not been sufficiently addressed in the literature so far. A weak congruency effect shown by the English-Chinese bilinguals suggests that the non-dominant L2 may also affect the processing of L1 binomials. The underlying mechanisms, however, may be different from CLIs in the direction of L1→L2. The mechanisms may be the automatic co-activation of known L2 MWEs, or the better L1 regulation or greater frequency of access to (congruent English binomials) and, consequently, less L1 inhibition for congruent English binomials than English-only binomials in the course of L2 acquisition. However, a unitary account for CLIs in MWE processing in both directions should be explored in future research.

8.2 Limitations

The present study has a number of limitations. First, the Chinese-English bilinguals and the English-Chinese bilinguals were not fully matched in their L2 proficiency level. L2 proficiency level was measured in terms of three variables: standardized test scores (e.g., IELTS, TOEFL, HSK) or, if not available, self-rated ability, daily usage of L2, and years of living in an L2-speaking country. The Chinese-English bilinguals were advanced bilinguals, roughly equivalent to the levels B2-C1 of the Common European Framework of Reference, whereas the English-Chinese bilinguals were intermediate+ bilinguals. That is, the L2 proficiency level for the Chinese-English bilinguals should have been higher than that for the English-Chinese bilinguals. This may explain why we saw a robust congruency effect in L2 MWE processing for the Chinese-English bilinguals, but only a trend for the congruency effect in L1 MWE processing for the English-Chinese bilinguals. However, it was not easy to recruit sufficient number of high-proficiency English-Chinese bilinguals. Almost half of the English-Chinese bilinguals self-rated themselves as intermediate bilinguals and half as advanced bilinguals. Nevertheless, I found no effects of proficiency for the target word when a separate model was fitted to assess the contribution of Chinese proficiency level for the English-Chinese bilinguals, with a high/low proficiency group manipulation.

Second, I did not consider participants' proficiency in data analysis in Experiment 1a with Chinese-English bilinguals. While the issue of how proficiency modulates CLIs in MWE processing was not the aim of this study, there could be differences in the participants' performance based on their proficiency. Because participants provided scores of different standardized tests (e.g., IELTS, TOEFL, etc) and not every participant could provide such a test score, there was no uniform objective measure to assess participants' proficiency. It was thus not possible to address the issue of how proficiency affected CLIs in the MWE processing. However, the present study adopted a within-participants experimental design and used mixed effects regressions, which provided control over the variability in processing due to individual differences, including L2 proficiency. To address this issue of L2 proficiency more explicitly, future research could use vocabulary test scores to measure participants' proficiency.

Finally, the association strength score between the constituents of some translated Chinese-only binomials and their controls was missing. Since the USF is based on English, it is not surprising that the association strength for these items was not attested in the USF norm database. Because no comparable Chinese database exists for the Chinese language, it was not possible to obtain association strength from the Chinese database. Therefore, I only obtained the association strength for items which existed in the USF database and assigned the value of zero to those items which did not exist in the USF database. Translated Chinese-only binomials showed no processing advantage over control novel phrases. Future research, however, would need to find a way to better control this factor.

8.3 Future Directions

The present study has raised a number of important questions. The present study found L1 influence in the processing of congruent binomials, but not in the processing of translated Chinese-only binomials for the Chinese-English bilinguals, supporting one side of the conflicting findings in previous studies with idioms and collocations. Studies with idioms found that translated L1-only idioms are processed faster than control novel phrases, as the L1 MWE activation account would predict. In contrast, studies with collocations and the present study found that translated L1-only collocations or binomials and control novel phrases are processed in a similar way, as the L2 MWE experience account would predict. The underlying mechanisms behind congruency effects in the processing of L2 MWEs thus remains to be clarified in future research. I argue that this important line of enquiry should continue, as more evidence is needed to help discriminate between the two accounts of the congruency effect in L2 MWE processing. For example, does the nature of CLI differ in accordance with the type(s) of MWEs being processed? With regard to idioms, their processing advantage can be described in terms of two processes: form activation and meaning activation (Carrol et al., 2016). In contrast, the processing advantage for collocations and binomials originates mainly from form activation as a result of phrase frequency effects. For translated L1-only MWEs, there may be no form activation due to their very low frequency of occurrence in the L2, but there may still be meaning activation due to the conceptual overlap across languages. That may explain why facilitation was observed for translated L1-only idioms, but not for translated L1-only collocations or binomials. In the case of idioms, a whole-phrase representation (i.e., the figurative meaning of an idiom) is activated, which may underpin facilitation observed for translated L1-only idioms. To test this hypothesis, future work could compare CLIs in the processing of different types of MWEs with fundamentally different properties and consider whether CLI may differ in accordance with the type(s) of MWEs.

In addition, the present study employed a primed visual lexical decision task and presented each prime word for 250 ms and set inter-stimulus interval (ISI) at 150 ms, with prime-target stimulus onset asynchrony (SOA) being as long as 800 ms (the prime contains two words, e.g., knife and). A related question is how variation in SOA would affect the result. The present study found no priming for translated Chinese-only binomials. Longer SOA, however, may lead to greater facilitation for translated Chinese-only binomials, because Chinese-English and English-Chinese bilinguals may have a chance to use strategic priming. On the other hand, the present study found a robust priming for congruent binomials in Chinese-English bilinguals. Would reducing SOA for congruent binomials lead to greater facilitation or the opposite? To investigate how changes in SOA may affect the result, future research may run a series of experiments with different SOAs (for example, one with the original SOA, one with reduced SOA, and one with longer SOA).

Further, future research could use eye-movement method and ERPs to investigate CLIs in the processing of collocations, binomials, and other types of literal MWEs. Existing studies with collocations and binomials have used behavioural methods (e.g., phrase-acceptability judgment tasks and primed lexical decision tasks) to investigate how the L1 influences the processing of L2 MWEs, whereas most studies with idioms used eye-movement method to investigate the same topic. Different methods make it hard to reconcile the contrasting findings of these studies. Adopting the same paradigm, however, may help us better understand the origins of the contrasting findings. In addition, eye-tracking permits reading that is as close to normal reading as possible in an experimental setting, without participants having to perform secondary task or make a strategic or metalinguistic response (Duyck et al., 2007; Roberts & Siyanova-Chanturia, 2013).

Furthermore, future research can also investigate the issue of whether or not congruent MWEs are indeed acquired earlier than incongruent L2-only MWEs. Studies on collocations and the present study suggest that there should be no automatic activation of known L1 MWEs in L2 processing because translated L1-only MWEs and control novel phrases are processed in a similar way. We then need to consider other potential origins of the congruency effect in the processing of L2 MWEs. One possibility is that congruent MWEs are acquired earlier than L2-only MWEs, as a result of cross-language transfer. The rate of acquisition or the solidifying of acquisition process may, for example, be faster for congruent MWEs than for L2-only ones. They are thus likely to be processed faster than L2-only MWEs due to the order-of-acquisition effect (i.e., L2 MWE experience account). However, there has been no empirical evidence yet that shows that congruent MWEs are indeed acquired before incongruent L2-only MWEs. Does congruency facilitate MWE learning and processing at the very early stage of L2 acquisition? Looking at the processing of MWEs in L2 learners of varying proficiency can attest to the tenet that congruent MWEs are acquired earlier than incongruent L2-only MWEs and, hence, provide more empirical evidence for the L2 MWE experience account.

Additionally, it is important to note that some researchers proposed that L1 may also play an inhibitive role where there is incongruency between the L1 and the L2 – in the case of incongruent MWEs (Pulido & Dussias, 2020). Rather than interpreting the relative processing advantage for congruent relative to incongruent MWEs as an effect of facilitation for congruent MWEs, it is also possible to interpret it as an effect of L1 interference for incongruent MWEs where competing L1 MWEs (e.g., the Spanish equivalent of "run a business" is "llevar un negocio", which literally translates as 'carry a business') are automatically activated and interfere with L2 processing. This L1 interference account is not mutually exclusive with the L1 facilitation account. That is, L1 facilitation (in the case of congruent MWEs) as well as L1 interference (in the case of incongruent MWEs) could take place at the same time (Pulido & Dussias, 2020). However, the L1 interference account, like L1 MWE activation account, also assumes the automatic activation of L1 MWEs during processing in an L2. In this account, both languages become activated in parallel, and so bilinguals have to select among competing alternatives available in both their languages (*kill time*, whose Japanese equivalent literally translates as 'crush/break time'). This will result in some processing cost. Given the scarcity of such research with respect to L1 interference, future studies should further this line of enquiry.

The language context (immersion) arose from the present study as an important modulator of CLIs. The English-Chinese bilinguals were found to process English-only binomials and control novel phrases in a similar way, whereas the English monolinguals processed English-only binomials significantly faster than control phrases. The lack of binomial priming effect in the English-Chinese bilinguals may have been due to the language context that they were in – being immersed in an L2-speaking country. Their L1 may have been inhibited to control L1 interference and facilitate L2 learning. When they needed to complete the experiment in their L1, it may have produced the switching costs and impaired their L1 performance. To test this hypothesis, future research would need to compare the processing of L1 MWEs by bilinguals in an L2 immersion context and bilinguals in their L1 context. This would allow us to examine how, if any, immersion impacts the L1 MWE processing in bilinguals. Likewise, immersion may also play a role in the processing of binomials for the Chinese-English bilinguals. It has been argued that L1 inhibition may have counteracted the activation of known L1 MWEs, resulting in no significant binomial priming effect observed for the translated Chinese-only binomials in the Chinese-English bilinguals. Future research could also investigate the role of immersion in L2 MWE processing by comparing the processing of L2 MWEs by bilinguals in an L2 immersion context and those in their L1 context.

Further research is also needed to investigate L2 influence on L1 MWE processing. The present study is the first study, to my knowledge, that have attempted to address this issue. More research is thus needed to see whether the results reported in the present study can be replicated. The present study only found a weak (trending significant) congruency effect in L1 binomial processing by the English-Chinese bilinguals. One potential explanation for the weak CLI on L1 MWE processing is the relative language fluency of the English-Chinese bilinguals in the two languages. They were intermediate bilinguals, with their L1 being their dominant language. According to Van Hell and Dijkstra (2002), relative language fluency affects a bilingual's sensitivity to the interference from the nontarget language when processing in a target language. In a group of Dutch(L1)-English(L2)-French(L3) trilinguals that were highly proficient in L2 and relatively low in proficiency in L3, Van Hell and Dijkstra (2002) observed that processing in L1 was influenced by L2 *but not* L3 nontarget language. However, for another group of trilinguals that were equally fluent in L2 and L3, they observed that processing in L1 Dutch was influenced by L2 *and* L3 nontarget language. This indicates that a certain level of nontarget language proficiency is required for an influence of nontarget language knowledge on target language processing to become noticeable. The lower level of L2 proficiency may be the reason why the English-Chinese bilinguals only showed a trend towards L2 influence on L1 MWE processing. In contrast, the Chinese-English bilinguals were advanced bilinguals and they have been found to show a significant congruency effect in L2 MWE processing. The stronger language (L1) always has more influence on the weaker language (L2) but not vice versa. That might explain why L1 has more impact on L2 processing than L2 on L1 processing in the present study and previous studies (Jiang, 1999; Schoonbaert et al., 2009). To verify the role of L2 proficiency on L2 influence in L1 MWE processing, future research could compare bilingual speakers with the same L1, but with different levels of L2 proficiency, to examine how L2 proficiency may modulate CLI in the direction of L2→L1.

Factors in relation to individual differences in the domain of general cognitive abilities (e.g., chunking abilities) have been found to modulate CLIs in bilingual processing. A recent eye-tracking study by (Pulido, 2021a) has found that individuals' chunking ability – to what extent speakers are sensitive to the structural probabilities in the input, rather than L2 reading speed, predicts ease of processing. Importantly, this study found that the congruency effect examined in the processing of L2 verb-noun phrases was not homogeneously present in all bilinguals. Bilinguals with higher L1 and L2 chunking sensitivity and more knowledge of L2-specific multiword units showed the congruency effect, while bilinguals with lower chunking-ability and L2 multiword-based proficiency showed no congruency effect. This study suggests that future research may need to take individual differences such as chunking ability into account in the investigation of CLIs in MWE processing.

Last but not least, while the presence of CLI has been documented in the on-line *comprehension* of MWEs, future research could investigate CLIs in MWE *production*. Some production studies suggest that higher frequency phrases are articulated more quickly than lower frequency ones by L1 speakers (Arnon & Cohen Priva, 2013; Bannard & Matthews, 2008; Janssen & Barber, 2012). However, proficient L2 speakers showed no articulatory advantage for frequent phrases over less frequent ones (Siyanova-Chanturia & Janssen, 2018), suggesting that productive learning and use may be more difficult than receptive learning and use. As argued by Siyanova-Chanturia and Janssen (2018), more exposure to and experience with the language might be required to use MWEs productively than receptively. However, little evidence exists with respect to the role of cross-language overlap or congruency on the production of MWEs. If congruent MWEs are acquired earlier than L2-only MWEs as a result of positive L1 transfer, congruent MWEs might exhibit some processing advantage in their production (e.g., articulatory durations) over L2-only MWEs. This issue remains to be explored in future research.

In sum, future research should continue to explore the many and varied factors that can modulate the degree and nature of CLI in MWE processing, such as speakers' characteristics, including proficiency, language dominance, and the language context (immersion) in which speakers interact. In addition, individual differences in the domain of general cognitive abilities (e.g., chunking ability and working memory) can also modulate the degree of CLI evident in MWE processing (Pulido, 2021a). Further, the modality of language use (production vs. comprehension) should be investigated to see how it may modulate CLI in MWE processing.

References

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, *20*(3), 242-275. https://doi.org/10.1016/j.jneuroling.2006.10.003

Arnon, I., & Cohen Priva, U. (2013). More than words: The effect of multi-word frequency and constituency on phonetic duration. *Language and Speech*, *56*(3), 349-371. https://doi.org/10.1177/0023830913484891

Arnon, I., & Cohen Priva, U. (2014). Time and again: The changing effect of word and multiword frequency on phonetic duration for highly frequent sequences. *The Mental Lexicon*, *9*(3), 377-400. https://doi.org/10.1075/ml.9.3.01arn

Arnon, I., McCauley, S. M., & Christiansen, M. H. (2017). Digging up the building blocks of language: Age-of-acquisition effects for multiword phrases. *Journal of Memory and Language, 92*, 265-280. https://doi.org/10.1016/j.jml.2016.07.004

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, *62*(1), 67-82. https://doi.org/10.1016/j.jml.2009.09.005

Babaei, S., Taleb Najafabadi, F., & Fotovatnia, Z. (2015). Processing of lexical bundles by Persian speaking learners of English. *The Journal of Teaching Language Skills*, *6*(4), 1-18. https://doi.org/10.22099/jtls.2015.3229

Bannard, C., & Matthews, D. (2008). Stored word sequences in language learning: The effect of familiarity on children's repetition of four-word combinations. *Psychological Science*, *19*(3), 241-248. https://doi.org/10.1111/j.1467-9280.2008.02075.x

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255-278. https://doi.org/10.1016/j.jml.2012.11.001

Basnight-Brown, D. M., & Altarriba, J. (2007). Differences in semantic and translation priming across languages: The role of language direction and language dominance. *Memory & Cognition*, *35*(5), 953-965. https://doi.org/10.22099/jtls.2015.3229

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*.

Baus, C., Costa, A., & Carreiras, M. (2013). On the effects of second language immersion on first language production. *Acta Psychologica*, *142*(3), 402-409. https://doi.org/10.1016/j.actpsy.2013.01.010

Beck, S. D., & Weber, A. (2016). Bilingual and monolingual idiom processing is cut from the same cloth: The role of the L1 in literal and figurative meaning activation. *Frontiers in Psychology*, *7*, https://doi.org/10.3389/fpsyg.2016.01350

Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory, M., & Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *The Journal of the Acoustical Society of America*, *113*(2), 1001-1024. https://doi.org/10.1121/1.1534836

Benor, S. B., & Levy, R. (2006). The chicken or the egg? A probabilistic analysis of English binomials. *Language*, *82*(2), 233-278. https://www.jstor.org/stable/4490157

Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, *26*(3), 263-286. https://doi.org/10.1016/j.esp.2006.08.003

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at…: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, *25*(3), 371-405. https://doi.org/10.1093/applin/25.3.371

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Longman.

Bobrow, S. A., & Bell, S. M. (1973). On catching on to idiomatic expressions. *Memory & Cognition*, *1*(3), 343-346.

Boers, F. (2020). Factors affecting the learning of multiword items. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 143-157). Routledge. https://doi.org/10.4324/9780429291586

Bogulski, C. A., Bice, K., & Kroll, J. F. (2019). Bilingualism as a desirable difficulty: Advantages in word learning depend on regulation of the dominant language. *Bilingualism: Language and Cognition*, *22*(5), 1052-1067. https://doi.org/10.1017/S1366728918000858

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, *1*(1), 1-20. https://doi.org/10.5334/joc.10

Bybee, J. (2001). Frequency effects on French liaison. In J. L. Bybee & P. J. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 337-360). John Benjamins.

Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, *14*(3), 261-290. https://doi.org/10.1017/S0954394502143018

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, *82*(4), 711-733. https://www.jstor.org/stable/4490266

Bybee, J. (2010). *Language, usage and cognition*. Cambridge University Press.

Bybee, J., & Scheibman, J. (1999). The effect of usage on degrees of constituency: The reduction of don't in English. *Linguistics*, *37*(4), 575-596. https://doi.org/10.1515/ling.37.4.575

Cacciari, C. (2014). Processing multiword idiomatic strings: Many words in one? *The Mental Lexicon*, *9*(2), 267-293. https://doi.org/10.1075/ml.9.2.05cac

Cacciari, C., & Glucksberg, S. (1991). Understanding idiomatic expressions: The contribution of word meanings. In G. B. Simpson (Ed.), *Understanding word and sentence* (pp. 217-240). Elsevier Science Publisher. https://doi.org/10.1016/S0166-4115(08)61535-6

Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language*, *27*(6), 668-683. https://doi.org/10.1016/0749-596X(88)90014-9

Caillies, S., & Butcher, K. (2007). Processing of idiomatic expressions: Evidence for a new hybrid view. *Metaphor and Symbol*, *22*(1), 79-108. https://doi/10.1080/10926480709336754

Carrol, G., & Conklin, K. (2014). Getting your wires crossed: Evidence for fast processing of L1 idioms in an L2. *Bilingualism: Language and Cognition*, *17*(4), 784-797. https://doi.org/10.1017/S1366728913000795

Carrol, G., & Conklin, K. (2017). Cross language lexical priming extends to formulaic units: Evidence from eye-tracking suggests that this idea 'has legs'. *Bilingualism: Language and Cognition*, *20*(2), 299-317. https://doi.org/10.1017/S1366728915000103

Carrol, G., & Conklin, K. (2020). Is all formulaic language created equal? Unpacking the processing advantage for different types of formulaic sequences. *Language and Speech*, *63*(1), 95-122. https://doi.org/10.1177/0023830918823230

Carrol, G., Conklin, K., & Gyllstad, H. (2016). Found in translation: The influence of the L1 on the reading of idioms in a L2. *Studies in Second Language Acquisition*, *38*(3), 403-443. https://doi.org/10.1017/s0272263115000492

Casado, A., Szewczyk, J. M., Wolna, A., & Wodniecka, Z. (2021). The relative balance between languages predicts the magnitude of whole-language inhibition. *PsyArXiv*, 15. https://doi.org/10.31234/osf.io/2jsdg

Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain research*, *1147*, 192-208. https://doi.org/10.1016/j.brainres.2007.01.137

Cieślicka, A. (2006). Literal salience in on-line processing of idiomatic expressions by second language learners. *Second Language Research*, *22*(2), 115-144. https://doi.org/10.1191/0267658306sr263oa

Cieślicka, A. (2013). Do nonnative language speakers chew the fat and spill the beans with different brain hemispheres? Investigating idiom decomposability with the divided visual field paradigm. *Journal of Psycholinguistic Research*, *42*(6), 475-503. https://doi.org/10.1007/s10936-012-9232-4

Cieślicka, A., & Heredia, R. R. (2011). Hemispheric asymmetries in processing L1 and L2 idioms: Effects of salience and context. *Brain and Language*, *116*(3), 136-150. https://doi.org/10.1016/j.bandl.2010.09.007

Conklin, K. (2020). Processing single-word and multiword items. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 174-188). Routledge.

Conklin, K., & Carrol, G. (2018). First language influence on the processing of formulaic language in a second language. In A. Siyanova-Chanturia & A. Pellicer-Sánchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 62-77). Routledge.

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, *29*(1), 72-89. https://doi.org/10.1093/applin/amm022

Cook, V., Iarossi, E., Stellakis, N., & Tokumaru, Y. (2003). Effects of the L2 on the syntactic processing of the L1. In V. Cook (Ed.), *Effects of the second language on the first* (pp. 193-213). Multilingual Matters. https://doi.org/10.21832/9781853596346-toc

Cowie, A. P. (1998). Phraseological dictionaries: Some East-West comparisons. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 209-228). Oxford University Press.

Cutting, J. C., & Bock, K. (1997). That's the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blendsß. *Memory & Cognition*, *25*(1), 57-71. https://doi.org/10.3758/BF03197285

Davies, M. (2008). The corpus of contemporary American English (COCA): 520 million words, 1990-present.

de Cock, S. (1998). A recurrent word combination approach to the study of formulae in the speech of native and non-native speakers of English. *International Journal of Corpus Linguistics*, *3*(1), 59-80. https://doi.org/10.1075/ijcl.3.1.04dec

de Groot, A. M., Delmaar, P., & Lupker, S. J. (2000). The processing of interlexical homographs in translation recognition and lexical decision: Support for non-selective access to bilingual memory. *The Quarterly Journal of Experimental Psychology Section A*, *53*(2), 397-428. https://doi.org/10.1080/713755891

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. https://doi.org/10.1016/0749-596X(91)90012-9

de Groot, A. M., Thomassen, A. J., & Hudson, P. T. (1986). Primed-lexical decision: The effect of varying the stimulus-onset asynchrony of prime and target. *Acta Psychologica*, *61*(1), 17-36. https://doi.org/10.1016/0001-6918(86)90019-3

Declerck, M. (2020). What about proactive language control? *Psychonomic Bulletin & Review*, *27*(1), 24-35. https://doi.org/10.3758/s13423-019-01654-1

Declerck, M., Philipp, A. M., & Koch, I. (2013). Bilingual control: Sequential memory in language switching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(6), 1793-1806. https://doi.org/10.1037/a0033094

Degani, T., Prior, A., & Tokowicz, N. (2011). Bidirectional transfer: The effect of sharing a translation. *Journal of Cognitive Psychology*, *23*(1), 18-28. https://doi.org/10.1080/20445911.2011.445986

Dijkstra, A. (Ton), 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. https://doi.org/10.1017/S1366728918000287

Dijkstra, A. (Ton). (2003). Lexical processing in bilinguals and multilinguals: The word selection problem. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *The multilingual lexicon* (pp. 11-26). Kluwer.

Dijkstra, A. (Ton)., Timmermans, M., & Schriefers, H. (2000). On being blinded by your other language: Effects of task demands on interlingual homograph recognition. *Journal of Memory and Language*, *42*(4), 445-464. https://doi.org/10.1006/jmla.1999.2697

Dijkstra, A. (Ton)., & Van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger, & A. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 189-225). Lawrence Erlbaum Associates, Inc.

Dijkstra, A. (Ton)., Van Heuven, W. J. B., & Grainger, J. (1998). Simulating cross-language competition with the bilingual interactive activation model. *Psychologica Belgica*, *38*(3-4), 177-196. https://doi.org/10.5334/pb.933

Dijkstra, A. (Ton)., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175-197. https://doi.org/10.1017/S1366728902003012

Dijkstra, A. (Ton)., 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. https://doi.org/10.1017/S1366728998000121

Du, L., Elgort, I., & Siyanova-Chanturia, A. (2021). Cross-language influences in the processing of multiword expressions: From a first language to second and back. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2021.666520

Du, L., Siyanova-Chanturia, A., & Elgort, I. (Under review). In I. Elgort, A. Siyanova-Chanturia, & M. Brysbaert (Eds.), *Cross-language influences in second language acquisition and processing: Interdisciplinary insights and perspectives*. John Benjamins.

Durrant, P., & Doherty, A. (2010). Are high-frequency collocations psychologically real? Investigating the thesis of collocational priming. *Corpus Linguistics and Linguistic Theory*, *6*(2), 125-155. https://doi.org/10.1515/cllt.2010.006

Durrant, P., & Mathews-Aydınlı, J. (2011). A function-first approach to identifying formulaic language in academic writing. *English for Specific Purposes*, *30*(1), 58-72. https://doi.org/10.1016/j.esp.2010.05.002

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

Ecke, P. (2001). Lexical retrieval in a third language: Evidence from errors and tip-of-the-tongue states. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Cross-linguistic influence in third language acquisition* (pp. 90-114). Multilingual Matters. https://doi.org/https://doi.org/10.21832/9781853595509-007

Edmonds, A. (2014). Conventional expressions: Investigating pragmatics and processing. *Studies in Second Language Acquisition*, *36*(1), 69-99. https://doi.org/10.1017/S0272263113000557

Ellis, N. C. (2002). Frequency effects in language processing. *Studies in Second Language Acquisition*, *24*(2), 143-188. https://doi.org/10.1017/s0272263102002024

Ellis, N. C. (2012). Formulaic language and second language acquisition: Zipf and the phrasal Teddy Bear. *Annual Review of Applied Linguistics*, *32*, 17-44. https://doi.org/10.1017/s0267190512000025

Ellis, N. C., Frey, E., & Jalkanen, I. (2009). The psycholinguistic reality of collocation and semantic prosody (1): Lexical access. In U. R¨omer & R. Schulze (Eds.), *Exploring the lexis-grammar interface* (pp. 89-114). John Benjamins.

Ellis, N. C., Simpson‐Vlach, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly*, *42*(3), 375-396. https://doi.org/10.1002/j.1545-7249.2008.tb00137.x

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text & Talk*, *20*(1), 29-62. https://doi.org/10.1515/text.1.2000.20.1.29

Fanari, R., Cacciari, C., & Tabossi, P. (2010). The role of length and context in spoken idiom recognition. *European Journal of Cognitive Psychology*, *22*(3): 321-334. https://doi.org/10.1080/09541440902843866

Fellbaum, C. (2007). *Idioms and collocations: Corpus-based linguistic and lexicographic studies*. Continuum.

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

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*(1), 116-124. https://doi.org/10.3758/BF03195503

Frisson, S., Rayner, K., & Pickering, M. J. (2005). Effects of contextual predictability and transitional probability on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(5), 862-877. https://doi.org/10.1037/0278-7393.31.5.862

Gabrieli, J. D. (1998). Cognitive neuroscience of human memory. *Annual Review of Psychology*, *49*(1), 87-115. https://doi.org/10.1146/annurev.psych.49.1.87

Ghyselinck, M., Lewis, M. B., & Brysbaert, M. (2004). Age of acquisition and the cumulative-frequency hypothesis: A review of the literature and a new multi-task investigation. *Acta Psychologica*, *115*(1), 43-67. https://doi.org/10.1016/j.actpsy.2003.11.002

Gibbs, R. (1986). Skating on thin ice: Literal meaning and understanding idioms in conversation. *Discourse Processes*, *9*(1), 17-30. https://doi.org/10.1080/01638538609544629

Gibbs, R. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory & Cognition*, *8*(2), 149-156.

Gibbs, R., Nayak, N. P., Bolton, J. L., & Keppel, M. E. (1989). Speakers' assumptions about the lexical flexibility of idioms. *Memory & Cognition*, *17*(1), 58-68.

Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, *8*(3), 183-206. https://doi.org/10.1515/cogl.1997.8.3.183

Giora, R. (2003). *On our mind: Salience, context, and figurative language*. Oxford University Press.

Goldberg, A. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press.

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew–English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(5), 1122-1139. https://doi.org/10.1037/0278-7393.23.5.1122

Gollan, T. H., & Silverberg, N. B. (2001). Tip-of-the-tongue states in Hebrew–English bilinguals. *Bilingualism: Language and Cognition*, *4*(1), 63-83. https://doi.org/10.1017/S136672890100013X

Grant, L., & Bauer, L. (2004). Criteria for re-defining idioms: Are we barking up the wrong tree? *Applied Linguistics*, *25*(1), 38-61. https://doi.org/10.1093/applin/25.1.38

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*(2), 67-81. https://doi.org/10.1017/S1366728998000133

Gregory, M. L., Raymond, W. D., Bell, A., Fosler-Lussier, E., & Jurafsky, D. (1999). The effects of collocational strength and contextual predictability in lexical production. *CLS* (35).

Gyllstad, H., & Wolter, B. (2016). Collocational processing in light of the phraseological continuum model: Does semantic transparency matter? *Language Learning*, *66*(2), 296-323. https://doi.org/10.1111/lang.12143

Hallin, A. E., & Van Lancker Sidtis, D. (2017). A closer look at formulaic language: Prosodic characteristics of Swedish proverbs. *Applied Linguistics*, *38*(1), 68-89. https://doi.org/10.1093/applin/amu078

Hernández, M., Costa, A., & Arnon, I. (2016). More than words: Multiword frequency effects in non-native speakers. *Language, Cognition and Neuroscience*, *31*(6), 785-800. https://doi.org/10.1080/23273798.2016.1152389

Hirosh, Z., & Degani, T. (2018). Direct and indirect effects of multilingualism on novel language learning: An integrative review. *Psychonomic Bulletin & Review*, *25*(3), 892-916. https://doi.org/10.3758/s13423-017-1315-7

Hoey, M. (2003). Why grammar is beyond belief. In J.-P. van Noppen, C. Den Tandt, & I. Tudor (Eds.), *Beyond: New perspectives in language, literature and ELT (Belgian Journal of English Language and Literature, new series*, 1, pp. 183-196). Academia Press.

Hoey, M. (2004). The textual priming of lexis. In S. Bernadini, G. Aston, & D. Stewart (Eds.), *Corpora and language learners* (pp. 21-41). John Benjamins.

Hoey, M. (2005). *Lexical priming: A new theory of words and language*. Routledge.

Hoey, M. (2013). Lexical priming. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1-6). Blackwell Publishing. http://doi.org/10.1002/9781405198431.wbeal0694

Hoshino, N., & Kroll, J. F. (2008). Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition*, *106*(1), 501-511. https://doi.org/10.1016/j.cognition.2007.02.001

Howarth, P. (1998a). Phraseology and second language proficiency. *Applied Linguistics*, *19*(1), 24-44. https://doi.org/10.1093/applin/19.1.24

Howarth, P. (1998b). The phraseology of learners' academic writing. In A. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 161-186). Oxford University Press.

Howell, D. C. (2012). *Statistical methods for psychology*. Cengage Learning.

Huang, M. Z. (2009). Solving the riddle of metaphor: A salience-based model for metaphorical interpretation in a discourse context. In S. Pourcel, & V. Evans (Eds.), *New directions in cognitive linguistics* (pp. 107-126). John Benjamins.

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.

Jackson, G. M., Swainson, R., Cunnington, R., & Jackson, S. R. (2001). ERP correlates of executive control during repeated language switching. *Bilingualism: Language and Cognition*, *4*(2), 169-178. https://doi.org/10.1017/S1366728901000268

Jackson, G. M., Swainson, R., Mullin, A., Cunnington, R., & Jackson, S. R. (2004). ERP correlates of a receptive language-switching task. *The Quarterly Journal of Experimental Psychology*, *57A*(2), 223-240. https://doi.org/10.1080/02724980343000198

Jackson, M. D., & McClelland, J. L. (1975). Sensory and cognitive determinants of reading speed. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 565-574. https://doi.org/10.1016/S0022-5371(75)80044-2

Jacobs, C. L., Dell, G. S., Benjamin, A. S., & Bannard, C. (2016). Part and whole linguistic experience affect recognition memory for multiword sequences. *Journal of Memory and Language*, *87*, 38-58. https://doi.org/10.1016/j.jml.2015.11.001

Janssen, N., & Barber, H. A. (2012). Phrase frequency effects in language production. *PloS ONE*, *7*(3), Article e33202. https://doi.org/10.1371/journal.pone.0033202

Jarvis, S. (1998). *Conceptual transfer in the interlingual lexicon*. Indiana University Linguistics Club Publications.

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. Routledge.

Jiang, N. (1999). Testing processing explanations for the asymmetry in masked cross-language priming. *Bilingualism: Language and Cognition*, *2*(1), 59-75. https://doi.org/10.1017/S1366728999000152

Jiang, N., & Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *The Modern Language Journal*, *91*(3), 433-445. https://doi.org/10.1111/j.1540-4781.2007.00589.x

Juhasz, B. J. (2018). Experience with compound words influences their processing: An eye movement investigation with English compound words. *The Quarterly Journal of Experimental Psychology*, *71*(1), 103-112. https://doi.org/10.1080/17470218.2016.1253756

Kapatsinski, V., & Radicke, J. (2009). Frequency and the emergence of prefabs: Evidence from monitoring. In R. Corrigan, E. Moravcsik, H. Ouali, & K. Wheatley (Eds.), *Formulaic Language* (pp. 499-522). John Benjamins. https://doi.org/10.1075/tsl.83.14kap

Kaushanskaya, M., Yoo, J., & Van Hecke, S. (2013). Word learning in adults with second-language experience: Effects of phonological and referent familiarity. *Journal of Speech, Language, and Hearing Research*, *56*(2), 667-678. https://doi.org/10.1044/1092-4388(2012/11-0084)

Keatley, C. W., Spinks, J. A., & de Gelder, B. (1994). Asymmetrical cross-language priming effects. *Memory & Cognition*, *22*(1), 70-84. https://doi.org/10.3758/BF03202763

Kemmer, S., & Barlow, M. (2000). Introduction: A usage-based conception of language. In M. Barlow & S. Kemmer (Eds.), *Usage based models of language* (pp. vii-xxviii). CSLI pubilications, Center for the Study of Language and Information.

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

Kim, S. H., & Kim, J. H. (2012). Frequency effects in L2 multiword unit processing: Evidence from self-paced reading. *TESOL Quarterly*, *46*(4), 831-841. http://www.jstor.org/stable/43267892

Konopka, A. E., & Bock, K. (2009). Lexical or syntactic control of sentence formulation? Structural generalizations from idiom production. *Cognitive Psychology*, *58*(1), 68-101. https://doi.org/10.1016/j.cogpsych.2008.05.002

Kroll, J. F., & de Groot, A. M. B. (1997). Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (Vol. 81, pp. 169-199). Erlbaum.

Kroll, J. F., & Dussias, P. E. (2012). The comprehension of words and sentences in two languages. In T. K. Bhatia & W. C. Ritchie (Eds.), *The handbook of bilingualism and multilingualism:*

Second Edition (pp. 216-243). John Wiley and Sons. https://doi.org/10.1002/9781118332382.ch9

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

Kuiper, K., & Haggo, D. (1984). Livestock auctions, oral poetry, and ordinary language. *Language in Society*, *13*(2), 205-234. https://doi.org/10.1017/S0047404500010381

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13).

Langacker, R. (2000). A dynamic usage-based model. In M. Barlow & S. Kemmer (Eds.), *Usage based models of language* (pp. 1-64). CSLI pubilications, Center for the Study of Language and Information.

Larson, M. J., Clayson, P. E., & Clawson, A. (2014). Making sense of all the conflict: A theoretical review and critique of conflict-related ERPs. *International Journal of Psychophysiology, 93*(3), 283-297. https://doi.org/10.1016/j.ijpsycho.2014.06.007

Lenth, R. (2019). emmeans: Estimated Marginal Means, aka Least-Squares Means. *R package version 1.4.1*.

Levy, B. J., McVeigh, N. D., Marful, A., & Anderson, M. C. (2007). Inhibiting your native language: The role of retrieval-induced forgetting during second-language acquisition. *Psychological Science, 18*(1), 29-34. https://doi.org/10.1111/j.1467-9280.2007.01844.x

Li, K., & Ho, Y. (2016). A corpus-based comparative study on George Orwell's 1984 Chinese translation strategies. *Studies in Literature and Language*, *13*(2), 26-33. https://doi.org/10.3968/8671

Libben, M. R., & Titone, D. A. (2008). The multidetermined nature of idiom processing. *Memory & Cognition*, *36*(6), 1103-1121. https://doi.org/10.3758/MC.36.6.1103

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

Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech*, *6*(3), 172-187. https://doi.org/10.1177/002383096300600306

Lin, P. M. S. (2010). The phonology of formulaic sequences: A review. In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 174-193). Continuum.

Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing access to the native language while immersed in a second language: Evidence for the role of inhibition in second-language learning. *Psychological Science*, *20*(12), 1507-1515. https://doi.org/10.1111/j.1467-9280.2009.02480.x

Liontas, J. (2002). Context and idiom understanding in second languages. In S. H. Foster-Cohen, T. Ruthenberg, & M. L. Poschen (Eds.), *EUROSLA yearbook: Annual Conference of the European Second Language Association* (Vol. 2, pp. 155-185). John Benjamins.

Macizo, P., Bajo, T., & Paolieri, D. (2012). Language switching and language competition. *Second Language Research*, *28*(2), 131-149. https://doi.org/10.1177/0267658311434893

Manning, C., & Schutze, H. (1999). *Foundations of statistical natural language processing*. MIT Press.

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition*, *6*(2), 97-115. https://doi.org/10.1017/s1366728903001068

Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics*, *33*(3), 299-320. https://doi.org/10.1093/applin/ams010

Matlock, T., & Heredia, R. R. (2002). Understanding phrasal verbs in monolinguals and bilinguals. In R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 251-274). Elsevier Press.

Matsuno, K. (2017). Processing collocations: Do native speakers and second language learners simultaneously access prefabricated patterns and each single word? *Journal of the European Second Language Association*, *1*(1), 61-72. https://doi.org/10.22599/jesla.17

McDonald, S. A., & Shillcock, R. C. (2003a). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science*, *14*(6), 648-652. https://doi.org/10.1046/j.0956-7976.2003.psci_1480.x

McDonald, S. A., & Shillcock, R. C. (2003b). Low-level predictive inference in reading: The influence of transitional probabilities on eye movements. *Vision Research*, *43*(16), 1735-1751. https://doi.org/10.1016/s0042-6989(03)00237-2

McEnery, A., & Xiao, Z. (2004). The Lancaster Corpus of Mandarin Chinese: A corpus for monolingual and contrastive language study. *Religion*, *17*, 1-4.

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. Psychology Press.

Meunier, F., & Granger, S. (2008). *Phraseology in foreign language learning and teaching*. John Benjamins Publishing.

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. https://doi.org/10.1006/JMLA.1998.2602

Millar, N. (2011). The processing of malformed formulaic language. *Applied Linguistics*, *32*(2), 129-148. https://doi.org/10.1093/applin/amq035

Molinaro, N., Canal, P., Vespignani, F., Pesciarelli, F., & Cacciari, C. (2013). Are complex function words processed as semantically empty strings? A reading time and ERP study of collocational complex prepositions. *Language and Cognitive Processes*, *28*(6), 762-788. https://doi.org/10.1080/01690965.2012.665465

Molinaro, N., & Carreiras, M. (2010). Electrophysiological evidence of interaction between contextual expectation and semantic integration during the processing of collocations. *Biological Psychology*, *83*(3), 176-190. https://doi.org/10.1016/j.biopsycho.2009.12.006

Moon, R. (1998). *Fixed expressions and idioms in English: A corpus-based approach*. Oxford University Press.

Morales, L., Paolieri, D., Cubelli, R., & Bajo, M. T. (2014). Transfer of Spanish grammatical gender to English: Evidence from immersed and non-immersed bilinguals. *Bilingualism: Language and Cognition*, *17*(4), 700-708. https://doi.org/10.1017/S1366728914000017

Moreno, E. M., Federmeier, K. D., & Kutas, M. (2002). Switching languages, switching palabras (words): An electrophysiological study of code switching. *Brain and Language*, *80*(2), 188-207. https://doi.org/10.1006/brln.2001.2588

Mosca, M., & de Bot, K. (2017). Bilingual language switching: Production vs. Recognition. *Frontiers in Psychology*, *8*, Article 934. https://doi.org/10.3389/fpsyg.2017.00934

Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford University Press.

Nesselhauf, N. (2005). *Collocations in a learner corpus* (Vol. 14). John Benjamins Publishing Company.

Northbrook, J., & Conklin, K. (2019). Is what you put in what you get out?—Textbook-derived lexical bundle processing in beginner English learners. *Applied Linguistics*, *40*(5), 816-833. https://doi.org/10.1093/applin/amy027

Nunberg, G., Sag, I. A., & Wasow, T. (1994). Idioms. *Language*, *70*(3), 491-538. https://doi.org/10.1353/lan.1994.0007

Paulmann, S., Ghareeb-Ali, Z., & Felser, C. (2015). Neurophysiological markers of phrasal verb processing: Evidence from L1 and L2 speakers. In R. Heredia & A. Cieślicka (Eds.), *Bilingual figurative language processing* (pp. 245-267). Cambridge University Press

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (Vol. 191, pp. 191-225). Longman.

Perea, M., & Rosa, E. (2002). Does the proportion of associatively related pairs modulate the associative priming effect at very brief stimulus-onset asynchronies? *Acta Psychologica*, *110*(1), 103-124. https://doi.org/10.1016/S0001-6918(01)00074-9

Pulido, M. F. (2021a). Individual chunking ability predicts efficient or shallow L2 processing: Eye-tracking evidence from multiword units in relative clauses. *Frontiers in Psychology*, *11*, Article 4004. https://doi.org/10.3389/fpsyg.2020.607621

Pulido, M. F. (2021b). Native language inhibition predicts more successful second language learning: Evidence of two ERP pathways during learning. *Neuropsychologia*, *152*, 107732. https://doi.org/10.1016/j.neuropsychologia.2020.107732

Pulido, M. F., & Dussias, P. E. (2020). Desirable difficulties while learning collocations in a second language: Conditions that induce L1 interference improve learning. *Bilingualism: Language and Cognition*, *23*(3), 652-667. https://doi.org/10.1017/S1366728919000622

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC nonword database. *The Quarterly Journal of Experimental Psychology Section A*, *55*(4), 1339-1362. https://doi.org/10.1080/02724980244000099

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372-422. https://doi.org/10.1037/0033-2909.124.3.372

Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*(8), 1457-1506. https://doi.org/10.1080/17470210902816461

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(5), 1188-1200. https://doi.org/10.1037/0096-1523.22.5.1188

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, *3*(4), 504-509. https://doi.org/10.3758/BF03214555

Reuterskiöld, C., & Van Lancker Sidtis, D. (2013). Retention of idioms following one-time exposure. *Child Language Teaching and Therapy*, *29*(2), 216-228. https://doi.org/10.1177/0265659012456859

Roberts, L., & Siyanova-Chanturia, A. (2013). Using eye-tracking to investigate topics in L2 acquisition and L2 processing. *Studies in Second Language Acquisition*, *35*(2), 213-235. https://doi.org/10.1017/S0272263112000861

Römer, U., O'Donnell, M. B., & Ellis, N. C. (2014). Second language learner knowledge of verb–argument constructions: Effects of language transfer and typology. *The Modern Language Journal*, *98*(4), 952-975. https://doi.org/10.1111/modl.12149

Sánchez-Casas, R., & García-Albea, J. E. (2005). The representation of cognate and noncognate words in bilingual memory: Can cognate status be characterized as a special kind of morphological relation? In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 226–250). Oxford University Press.

Schmitt, N. (2004). *Formulaic sequences: Acquisition, processing, and use*. John Benjamins Publishing.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan. https://doi.org/10.1057/ 9780230293977

Schoonbaert, S., Duyck, W., Brysbaert, M., & Hartsuiker, R. J. (2009). Semantic and translation priming from a first language to a second and back: Making sense of the findings. *Memory & Cognition*, *37*(5), 569-586. https://doi.org/10.3758/MC.37.5.569

Schwieter, J. W., & Sunderman, G. (2008). Language switching in bilingual speech production: In search of the language-specific selection mechanism. *The Mental Lexicon*, *3*(2), 214-238. https://doi.org/10.1075/ml.3.2.06sch

Shaughnessy, J. J., Zechmeister, E. B., & Zechmeister, J. S. (2000). *Research methods in psychology*. McGraw-Hill.

Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, *31*(4), 487-512. https://doi.org/10.1093/applin/amp058

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.

Siyanova-Chanturia, A. (2010). *On-line processing of multi-word sequences in a first and second Language: Evidence from eye-tracking and ERP*. PhD thesis, University of Nottingham.

Siyanova-Chanturia, A. (2013). Eye-tracking and ERPs in multi-word expression research: A state-of-the-art review of the method and findings. *The Mental Lexicon*, *8*(2), 245-268. https://doi.org/10.1075/ml.8.2.06siy

Siyanova-Chanturia, A. (2015). On the 'holistic' nature of formulaic language. *Corpus Linguistics and Linguistic Theory*, *11*(2), 285-301. https://doi.org/10.1515/cllt-2014-0016

Siyanova-Chanturia, A., Conklin, K., Caffarra, S., Kaan, E., & Van Heuven, W. J. (2017). Representation and processing of multi-word expressions in the brain. *Brain and Language*, *175*, 111-122. https://doi.org/10.1016/j.bandl.2017.10.004

Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research*, *27*(2), 251-272. https://doi.org/10.1177/0267658310382068

Siyanova-Chanturia, A., Conklin, K., & Van Heuven, W. J. (2011). Seeing a phrase "time and again" matters: The role of phrasal frequency in the processing of multiword sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 776-784. https://doi.org/10.1037/a0022531

Siyanova-Chanturia, A., & Janssen, N. (2018). Production of familiar phrases: Frequency effects in native speakers and second language learners. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(12), 2009-2018. https://doi.org/10.1037/xlm0000562

Siyanova-Chanturia, A., & Martinez, R. (2015). The idiom principle revisited. *Applied Linguistics*, *36*(5), 549-569. https://doi.org/10.1093/applin/amt054

Siyanova-Chanturia, A., & Omidian, T. (2020). Key issues in researching multi-word items. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 529-545). Routledge.

Siyanova-Chanturia, A., & Pellicer-Sánchez, A. (2019). Formulaic language: Setting the scene. In A. Siyanova-Chanturia & A. Pellicer-Sánchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 1-15). Routledge.

Siyanova-Chanturia, A., & Van Lancker Sidtis, D. (2019). What online processing tells us about formulaic language. In A. Siyanova-Chanturia & A. Pellicer-Sánchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 38-61). Routledge.

Siyanova-Chanturia, A., Canal, P., & Heredia, R. R. (2019). Event-related potentials in monolingual and bilingual non-literal language processing. In J. W. Schwieter & M. Paradis (Eds.), *The handbook of the neuroscience of multilingualism* (pp. 508-529). John Wiley & Sons Ltd. https://doi.org/10.1002/9781119387725.ch25

Siyanova-Chanturia, A., & Lin, P. M. S. (2018). Production of ambiguous idioms in English: A reading aloud study. *International Journal of Applied Linguistics*, *28*(1), 58-70. https://doi.org/10.1111/ijal.12183

Siyanova-Chanturia, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *Canadian Modern Language Review*, *64*(3), 429-458. https://doi.org/10.3138/cmlr.64.3.429

Snider, N., & Arnon, I. (2012). A unified lexicon and grammar? Compositional and non-compositional phrases in the lexicon. In S. T. Gries & D. Divjak (Eds.), *Frequency effects in language representation* (pp. 127-164). Mouton de Gruyter.

Sonbul, S. (2015). Fatal mistake, awful mistake, or extreme mistake? Frequency effects on off-line/on-line collocational processing. *Bilingualism: Language and Cognition*, *18*(3), 419-437. https://doi.org/10.1017/s1366728914000674

Sonbul, S., El-Dakhs, D. A. S., & Al-Otaibi, H. (2020). Productive versus receptive L2 knowledge of polysemous phrasal verbs: A comparison of determining factors. *System*, *95*, 102361. https://doi.org/10.1016/j.system.2020.102361

Sonbul, S., & Schmitt, N. (2013). Explicit and implicit lexical knowledge: Acquisition of collocations under different input conditions. *Language Learning*, *63*(1), 121-159. https://doi.org/10.1111/j.1467-9922.2012.00730.x

Sonbul, S., & Siyanova-Chanturia, A. (2021). Research on the on-line processing of collocation: Replication of Wolter and Gyllstad (2011) and Millar (2011). *Language Teaching*, *54*(2), 236-244. https://doi.org/10.1017/S0261444819000132

Sosa, A. V., & MacFarlane, J. (2002). Evidence for frequency-based constituents in the mental lexicon: Collocations involving the word of. *Brain and Language*, *83*(2), 227-236. https://doi.org/10.1016/S0093-934X(02)00032-9

Sprenger, S. A., Levelt, W. J., & Kempen, G. (2006). Lexical access during the production of idiomatic phrases. *Journal of Memory and Language*, *54*(2), 161-184. https://doi.org/10.1016/j.jml.2005.11.001

Su, I. R. (2001). Transfer of sentence processing strategies: A comparison of L2 learners of Chinese and English. *Applied Psycholinguistics*, *22*(1), 83-112. https://doi.org/10.1017/S0142716401001059

Swinney, D. A., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behaviour*, *18*(5), 523-534. https://doi.org/10.1016/S0022-5371(79)90284-6

Tabossi, P., Fanari, R., & Wolf, K. (2005). Spoken idiom recognition: Meaning retrieval and word expectancy. *Journal of Psycholinguistic Research*, *34*(5), 465-495. https://doi.org/10.1007/s10936-005-6204-y

Tabossi, P., Fanari, R., & Wolf, K. (2009). Why are idioms recognized fast? *Memory & Cognition*, *37*(4), 529-540. https://doi.org/10.3758/MC.37.4.529

Taft, M. (2002). Orthographic processing of polysyllabic words by native and nonnative English speakers. *Brain and Language*, *81*(1-3), 532-544. https://doi.org/10.1006/brln.2001.2545

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences*, *104*(30), 12530-5. https://doi.org/10.1073/pnas.0609927104

Titone, D., Columbus, G., Whitford, V., Mercier, J., & Libben, M. (2015). Contrasting bilingual and monolingual idiom processing. In R. R. Heredia & A. B. Cieślicka (Eds.), *Bilingual figurative language processing* (pp. 171–207). Cambridge University Press. https://doi.org/10.1017/CBO9781139342100.011

Titone, D., & Connine, C. M. (1999). On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics, 31*(12), 1655-1674. https://doi.org/10.1016/S0378-2166(99)00008-9

Titone, D. A., & Connine, C. M. (1994). Comprehension of idiomatic expressions: Effects of predictability and literality. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(5), 1126-1138. https://doi.org/10.1037/0278-7393.20.5.1126

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press.

Toomer, M., & Elgort, I. (2019). The development of implicit and explicit knowledge of collocations: A conceptual replication and extension of Sonbul and Schmitt (2013). *Language Learning, 69*(2), 405-439. https://doi.org/10.1111/lang.12369

Tremblay, A., & Baayen, R. H. (2010). Holistic processing of regular four-word sequences: A behavioral and ERP study of the effects of structure, frequency, and probability on immediate free recall. In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 151-173). The Continuum International Publishing Group.

Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing advantages of lexical bundles: Evidence from self‐paced reading and sentence recall tasks. *Language Learning, 61*(2), 569-613. https://doi.org/10.1111/j.1467-9922.2010.00622.x

Tremblay, A., & Tucker, B. V. (2011). The effects of n-gram probabilistic measures on the recognition and production of four-word sequences. *The Mental Lexicon, 6*(2), 302-324. https://doi.org/10.1075/ml.6.2.04tre

Ullman, M. T. (2014). The declarative/procedural model: A neurobiologically motivated theory of first and second language. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 135-160). Routledge.

Underwood, G., Schmitt, N., & Galpin, A. (2004). The eyes have it. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing, and use* (pp. 153-172). John Benjamins.

Upton, T. A., & Lee-Thompson, L.-C. (2001). The role of the first language in second language reading. *Studies in Second Language Acquisition*, *23*(4), 469-495. https://doi.org/10.1017/S0272263101004028

Valsecchi, M., Künstler, V., Saage, S., White, B. J., Mukherjee, J., & Gegenfurtner, K. R. (2013). Advantage in reading lexical bundles is reduced in non-native speakers. *Journal of Eye Movement Research*, *6*(5), 1-15. https://doi.org/10.16910/jemr.6.5.2

Van Hell, J. G., & Dijkstra, A. (Ton). (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review*, *9*(4), 780-789. https://doi.org/10.3758/BF03196335

Van Hell, J. G., & Tanner, D. (2012). Second language proficiency and cross-language lexical activation. *Language Learning*, *62*(s2), 148-171. https://doi.org/10.1111/j.1467-9922.2012.00710.x

Van Heuven, W. J. B., Dijkstra, A. (Ton)., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, *39*(3), 458-483. https://doi.org/10.1006/jmla.1998.2584

Van Lancker Sidtis, D., & Rallon, G. (2004). Tracking the incidence of formulaic expressions in everyday speech: Methods for classification and verification. *Language & Communication*, *24*(3), 207-240. https://doi.org/10.1016/j.langcom.2004.02.003

Van Lancker, D., & Canter, G. J. (1981). Idiomatic versus literal interpretations of ditropically ambiguous sentences. *Journal of Speech & Hearing Research*, *24*(1), 64-69. https://doi.org/10.1044/jshr.2401.64

Van Lancker, D., Canter, G. J., & Terbeek, D. (1981). Disambiguation of ditropic sentences: Acoustic and phonetic cues. *Journal of Speech & Hearing Research*, *24*(3), 330-335. https://doi.org/10.1044/jshr.2403.330

Van Lancker Sidtis, D. (2003). Auditory recognition of idioms by native and nonnative speakers of English: It takes one to know one. *Applied Psycholinguistics*, *24*(1), 45-57. https://doi.org/10.1017/s0142716403000031

Van Lancker Sidtis, D. (2004). When novel sentences spoken or heard for the first time in the history of the universe are not enough: Toward a dual-process model of language. *International Journal of Language & Communication Disorders*, *39*(1), 1-44. https://doi.org/10.1080/13682820310001601080

Van Lancker Sidtis, D. (2012). Two-track mind: Formulaic and novel language support a dual-process model. In M. Faust (Ed.), *The handbook of the neuropsychology of language* (pp. 342-367). Wiley-Blackwell. https://doi.org/10.1002/9781118432501.ch17

Vespignani, F., Canal, P., Molinaro, N., Fonda, S., & Cacciari, C. (2010). Predictive mechanisms in idiom comprehension. *Journal of Cognitive Neuroscience*, *22*(8), 1682-1700. https://doi.org/10.1162/jocn.2009.21293

Vilkaitė, L. (2016). Are nonadjacent collocations processed faster? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(10), 1632-1642. https://doi.org/10.1037/xlm0000259

Vilkaitė, L., & Schmitt, N. (2019). Reading collocations in an L2: Do collocation processing benefits extend to non-adjacent collocations? *Applied Linguistics*, *40*(2), 329-354. https://doi.org/10.1093/applin/amx030

Voga, M., & Grainger, J. (2007). Cognate status and cross-script translation priming. *Memory & Cognition*, *35*(5), 938-952. https://doi.org/10.3758/BF03193467

Wang, X., & Forster, K. I. (2010). Masked translation priming with semantic categorization: Testing the Sense Model. *Bilingualism: Language and Cognition*, *13*(3), 327-340. https://doi.org/10.1017/s1366728909990502

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*(1), 1-25. https://doi.org/10.1016/S0749-596X(03)00105-0

Winter. (2013). *Linear models and linear mixed effects models in R with linguistic applications*. Cornell University Library. https://arxiv.org/abs/1308.5499

Winter. (2019). *Statistics for linguists: An introduction using R*. Routledge.

Wodneicka, Z., Casado, A., Kalamala, P., Marecka, M., Timmer, K., & Wolna, A. (2020). The dynamics of language experience and how it affects language and cognition. In K. D. Federmeier & H.-W. Huang (Eds.), *Adult and second language learning*. Academic Press.

Wolter, B., & Gyllstad, H. (2011). Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Applied Linguistics*, *32*(4), 430-449. https://doi.org/10.1093/applin/amr011

Wolter, B., & Gyllstad, H. (2013). Frequency of input and L2 collocational processing. *Studies in Second Language Acquisition*, *35*(3), 451-482. https://doi.org/10.1017/s0272263113000107

Wolter, B., & Yamashita, J. (2015). Processing collocations in a second language: A case of first language activation? *Applied Psycholinguistics*, *36*(5), 1193-1221. https://doi.org/10.1017/s0142716414000113

Wolter, B., & Yamashita, J. (2018). Word frequency, collocational frequency, L1 congruency, and proficiency in L2 collocational processing. *Studies in Second Language Acquisition*, *40*(2), 395-416. https://doi.org/10.1017/s0272263117000237

Wood, D. (2020). Classifying and identifying formulaic language. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 30-45). Routledge.

Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics*, *21*(4), 463-489. https://doi.org/10.1093/applin/21.4.463

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.

Wray, A. (2008). *Formulaic language: Pushing the boundaries*. Oxford University Press.

Wray, A. (2012). What do we (think we) know about formulaic language? An evaluation of the current state of play. *Annual Review of Applied Linguistics*, *32*, 231-254. https://doi.org/10.1017/s026719051200013x

Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language & Communication*, *20*(1), 1-28. https://doi.org/10.1016/S0271-5309(99)00015-4

Wu, Y. J., Cristino, F., Leek, C., & Thierry, G. (2013). Non-selective lexical access in bilinguals is spontaneous and independent of input monitoring: Evidence from eye tracking. *Cognition*, *129*(2), 418-425. https://doi.org/10.1016/j.cognition.2013.08.005

Wu, Y. J., & Thierry, G. (2010). Chinese-English bilinguals reading English hear Chinese. *Journal of Neuroscience*, *30*(22), 7646-7651. https://doi.org/10.1523/JNEUROSCI.1602-10.2010

Yamashita, J. (2018). Possibility of semantic involvement in the L1-L2 congruency effect in the processing of L2 collocations. *Journal of Second Language Studies*, *1*(1), 60-78. https://doi.org/10.1075/jsls.17024.yam

Yamashita, J., & Jiang, N. (2010). L1 Influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly, 44*(4), 647-668. https://doi.org/10.5054/tq.2010.235998

Yang, S-Y, Ahn, J. S., & Van Lancker Sidtis, D. (2015). The perceptual and acoustic characteristics of Korean idiomatic and literal sentences. *Speech, Language and Hearing*, *18*(3), 166-178. https://doi.org/10.1179/2050572814Y.0000000061

Zeng, T., Branigan, H. P., & Pickering, M. J. (2020). Do bilinguals represent between-language relationships beyond the word level in their lexicon? *Journal of Neurolinguistics*, *55*, Article 100892. https://doi.org/10.1016/j.jneuroling.2020.100892

Zhan, W., Guo, R., & Chen, Y. (2003). *The CCL corpus of Chinese texts: 700 million Chinese characters, the 11th Century B. C. - present*. Available online at the website of Center for Chinese Linguistics of Peking University. http://ccl. pku. edu.cn:8080/ccl_corpus

Zhang, T., Van Heuven, W. J. B., & Conklin, K. (2011). Fast automatic translation and morphological decomposition in Chinese-English bilinguals. *Psychological Science*, *22*(10), 1237-1242. https://doi.org/10.1177/0956797611421492

Appendices

Appendix 1 Ethics approval

TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI

VICTORIA

UNIVERSITY OF WELLINGTON

Human ethics application approval 0000026162. Automated Email, Do Not Reply

researchmaster-help@vuw.ac.nz
Sat 16/06/2018 9:05 AM
To: Lingli Du (FHSS-Linguistics & App Language Studies) 1
Cc: Isobel Cairns; HEC

Dear Lingli,

Thank you for your application for ethical approval (On Language Non-selective Access in Visual Phrase Processing, reference 0000026162), which has now been considered by the Standing Committee of the Human Ethics Committee.

Your application is approved as of today. Your approval applies for three years from the date of this email.

If you would like to receive a formal letter please contact the HEC Administrator (ethicsadmin@vuw.ac.nz).

Best wishes with the research.

Judith Loveridge, Convenor
Human Ethics Committee

Appendix 2 Information sheet

Cross-language influences in the processing of multi-word expressions: from a first language to second and back

INFORMATION SHEET FOR PARTICIPANTS

You are invited to take part in this research. Please read this information before deciding whether or not to take part. If you decide to participate, thank you. If you decide not to participate, thank you for considering this request.

Who am I?

My name is Lingli Du and I am a Doctoral student in Linguistics and Applied Language Studies at Victoria University of Wellington. This research project is work towards my dissertation. My supervisors are Dr Anna Siyanova and Dr Irina Elgort in the School of Linguistics and Applied Language Studies.

What is the aim of the project?

This project looks at how bilinguals process multiword expressions. It aims to find that whether bilingual speakers who speak two languages are able to switch off one language when they read in the other language.

This research has been approved by the Victoria University of Wellington Human Ethics Committee (Approval number 0000025629).

How can you help?

You have been invited to participate because you meet the language requirements for my experiment. If you agree to take part you will be asked to come in for an experiment session at Victoria University/Beijing University [locations of the experiment have not been confirmed yet but will be given later]. In the experiment, you will read a two-word combination ending with “and” (such as “bride and”), and be asked to judge whether the following word is a real word or not by pressing the specified “yes” or “no” button. Your responses, and the time it takes you to respond, will be recorded. The experiment session will take about 30 minutes. You can withdraw from the study up to ___________[Date, which will be two weeks after the planned experiment session]. If you withdraw, the information you provided will be destroyed or returned to you. If you are a student, your decision whether or not to participate will not affect your grades in any way.

You will get a voucher of $10 in recognition your efforts in participating.

What will happen to the information you give?

This research is confidential. I will not name you in any reports, and I will not include any information that would identify you. Only my supervisors and I will have access to the data in a form where you can be identified. The data from the experiments will be anonymised. The anonymised data will be stored securely and kept indefinitely. It may be made available to any other researchers in the field at any time to inform future research about bilingual lexical processing.

What will the project produce?

The information from my research will be used in my PhD dissertation, academic publications and conferences presentations reporting the results of this research. I will take care not to identify you in any presentation or report.

If you accept this invitation, what are your rights as a research participant?

You do not have to accept this invitation if you don't want to. If you do decide to participate, you have the right to:

- choose not to answer any question;
    - ask to stop the experiment at any time;
    - ask to review the data you have provided;
- withdraw from the study before __________ [Date, which is two weeks after the experiment session];
- ask any questions about the study at any time;
- be able to read any reports of this research by emailing the researcher to request a copy.

If you have any questions or problems, who can you contact?

If you have any questions, either now or in the future, please feel free to contact me:

Student	Primary supervisor:	Secondary supervisor:
Lingli Du	Dr Anna Siyanova	Dr Irina Elgort
lingli.du@vuw.ac.nz	Senior Lecturer	Senior Lecturer
	School of Linguistics and Applied Language Studies, Victoria University of Wellington,	Centre for Academic Development, Victoria University of Wellington,
	Kelburn Parade	10 Waiteata Road, Kelburn,
	PO Box 600	PO Box 600
	Wellington 6140, NZ	Wellington 6140, NZ
	Phone: +64 (0)4 463 5922	Phone: +64 4 4635970
	anna.siyanova@vuw.ac.nz	irina.elgort@vuw.ac.nz

Appendix 3 Consent form

CONSENT TO COLLECT EXPERIMENT DATA

This consent form will be held until 01/11/2019.

Researcher: Lingli Du,
School of Linguistics and Applied Language Studies
Victoria University of Wellington.

- I have read the Information Sheet and the project has been explained to me. My questions have been answered to my satisfaction. I understand that I can ask further questions at any time.
- I agree to take part in this experiment.

I understand that:

- I may withdraw from this study at any point before ____________ *[DATE, which will be two weeks after the experiment session]*, and any information that I have provided will be returned to me or destroyed.
- The data from all participants will be kept securely in an anonymised form indefinitely. It may be made available to any other researchers in the field at any time to inform future research about bilingual lexical processing (as set out in the Information Sheet). Any information linking me to the data I provide in this experimental session will be destroyed at the conclusion of this research.
- Any information I provide which could identify me will be kept confidential to the researcher and the supervisors. I understand that the results will be used for a PhD dissertation, academic research and a summary of the results may be used in academic reports and/or presented at conferences.
- My name will not be used in reports, nor will any information that would identify me.

- I would like to receive a summary of the findings of this study and have added my email address below. Yes ☐ No ☐

Signature of participant: ________________________________

Name of participant: ________________________________

Date: ______________

Email (or other contact details): ________________________________

Appendix 4 Language background questionnaire for Chinese-English bilinguals and English monolinguals

Student:	Primary supervisor:	Secondary supervisor:
Lingli Du	Dr Anna Siyanova	Dr Irina Elgort
lingli.du@vuw.ac.nz	Senior Lecturer	Senior Lecturer
	School of Linguistics and Applied Language Studies,	Centre for Academic Development,
		10 Waiteata Road,
	Victoria University of Wellington	Victoria University of Wellington
	PO Box 600	PO Box 600
	Wellington 6140, NZ	Wellington 6140, NZ
	Phone: +64 4 463 5922	Phone: +64 4 4635970
	anna.siyanova@vuw.ac.nz	irina.elgort@vuw.ac.nz

Project: Cross-language influences in the processing of multi-word expressions: from a first language to second and back

Participant questionnaire

Gender: ___________ Age: ____________________

Place of birth: ___________________ (Country)

Occupation: ___________________

What is your first language?__________________

Mother's place of origin: Country: ______________Mother's first language: _______________

Father's place of origin: Country: ______________ Father's first language: _______________

What is the main language spoken in your home? ___________________

What other languages do you speak? Please state your proficiency for each language (native, near-native, fluent, intermediate, poor):

1. ____
2. ____
3. ____
4. ____

On average, what percentage of the time do you use each of the languages you listed above everyday with family, friends and classmates? (percentages should add to 100%):

English: _____________ Second language: _______ Third language: _______ Other:

Overall English test scores if available (e.g. IELTS, TOEFL): _____________________

Total length of time in New Zealand or any other English-speaking countries (for Chinese speakers of English as a second language), ___________________________

Eyesight (please tick one): Normal ______ Corrected to normal _______ Poor

Are you left-handed _______or right handed_______ ? (please tick one)

Appendix 5 Binomial and control items used in the experiment

English binomials

Congruent binomials	Controls	English-only binomials	Controls
Love and hate	Like and hate	Salt and pepper	Spice and pepper
Push and pull	Drag and pull	Hands and knees	Arms and knees
Knife and fork	Spoon and fork	Hunting and fishing	Sailing and fishing
Sun and moon	Star and moon	Drugs and alcohol	Grain and alcohol
Army and navy	Sailors and navy	Bread and butter	Toast and butter
Birth and death	Cancer and death	Bed and breakfast	Supper and breakfast
Facts and figures	Shapes and figures	Soap and water	Boat and water
Height and weight	Diet and weight	Pride and joy	Fun and joy
Wind and rain	Sunshine and rain	Milk and honey	Oats and honey
Vitamins and minerals	Vegetables and minerals	Apples and oranges	Carrots and oranges
Newspapers and magazines	Articles and magazines	Arts and crafts	Hobbies and crafts
Costs and benefits	Welfare and benefits	Sticks and stones	Bricks and stones
Theory and practice	Skill and practice	Car and truck	Van and truck
Song and dance	Show and dance	Eyes and nose	Face and nose
Tables and chairs	Stools and chairs	Cheese and crackers	Nuts and crackers
Master and slave	Servant and slave	Cold and flu	Fever and flu
Attitudes and behaviours	Manners and behaviours	Marriage and divorce	Wedding and divorce
Thick and thin	Skinny and thin	Cream and sugar	Honey and sugar
Science and technology	Computer and technology	Cops and robbers	Criminals and robbers
Deaf and dumb	Smart and dumb	Meat and dairy	Butter and dairy

Chinese-only binomials

Translated Chinese-only binomials	Controls
Poor and underdeveloped	Old and underdeveloped
Diligent and brave	Faithful and brave
Pots and bowls	Jars and bowls
Fish and shrimp	Crab and shrimp
Flowers and applause	Success and applause
Workers and peasants	Landlords and peasants
Chickens and ducks	Swans and ducks
Greetings and wishes	Prayers and wishes
Pigs and dogs	Bears and dogs
Wine and meat	Beer and meat
Drought and flood	Earthquake and flood
Dragon and phoenix	Bird and phoenix
Talent and beauty	Confidence and beauty
Cold and fever	Cough and fever
Knife and sword	Bow and sword
Wisdom and strength	Exercise and strength
Agricultural and rural	Industrial and rural
Experience and lessons	Knowledge and lessons
Peace and unity	Harmony and unity
Identity and status	Reputation and status

Appendix 6 Familiarity test for English binomials

How familiar are you with the following phrases?

Please give the following phrases a score based on your familiarity with them.

1 = Not familiar at all

2 = Slightly familiar

3 = Moderately familiar

4 = Very Familiar

5 = Extremely Familiar

For example: Boys and girls ___5___

love and hate __________

knife and fork __________

army and navy__________

facts and figures__________

wind and rain____________

newspapers and magazines__________

theory and practice___________

tables and chairs_____________

attitudes and behaviours__________

science and technology___________

salt and pepper____________

hunting and fishing__________

bread and butter_______________

soap and water_______________

milk and honey______________

arts and crafts_______________

push and pull_____________

sun and moon_____________

birth and death__________

height and weight________

vitamins and minerals_______

costs and benefits__________

song and dance____________

master and slave___________

thick and thin______________

deaf and dumb_____________

hands and knees____________

drugs and alcohol___________

bed and breakfast___________

pride and joy_______________

apples and oranges__________

sticks and stones____________

car and truck____________ eyes and nose_______________

cheese and crackers_____________ cold and flu________________

marriage and divorce_____________ cream and sugar___________

cops and robbers________________ meat and dairy_____________

Appendix 7 Familiarity test for Chinese-only binomials

Please give the following phrases a score based on your familiarity with them.

1 = Not familiar at all

2 = Slightly familiar

3 = Moderately familiar

4 = Very Familiar

5 = Extremely Familiar

贫穷、落后_____________	勤劳勇敢_____________
锅碗__________________	鱼虾__________________
鲜花和掌声______________	工人和农民______________
鸡鸭_________	问候和祝愿____________
猪狗____________	酒肉____________
旱涝____________	龙凤____________
才貌_____________	感冒、发烧__________
刀剑__________	智慧和力量___________
农业和农村的__________	经验和教训____________
和平和统一____________	身份和地位____________

Appendix 8 Filler and non-word items used in the experiment

Item type	Item	Prime 1	Target
filler	sky and ocean	sky	ocean
filler	business and soul	business	soul
filler	floor and legs	floor	legs
filler	chapter and mouse	chapter	mouse
filler	family and pain	family	pain
filler	crime and root	crime	root
filler	angel and verse	angel	verse
filler	wealthy and well	wealthy	well
filler	aches and hook	aches	hook
filler	far and late	far	late
filler	alive and high	alive	high
filler	fire and steel	fire	steel
filler	flesh and yummy	flesh	yummy
filler	early and mixed	early	mixed
filler	seal and pleasure	seal	pleasure
filler	flight and writing	flight	writing
filler	mood and carpet	mood	carpet
filler	food and stance	food	stance
filler	forgive and match	forgive	match
filler	fruit and gentlemen	fruit	gentlemen
filler	good and down	good	down
filler	chunk and distance	chunk	distance
filler	hope and punishment	hope	punishment
filler	hot and clean	hot	clean
filler	income and dream	income	dream
filler	intents and devils	intents	devils
filler	iron and friends	iron	friends
filler	ladies and drinks	ladies	drinks

filler	lock and ceiling	lock	ceiling
filler	pen and blood	pen	blood
filler	check and forget	check	forget
filler	neat and bad	neat	bad
filler	lamp and plant	lamp	plant
filler	taxi and curtain	taxi	curtain
nonword	number and yops	number	yops
nonword	street and plepped	street	plepped
nonword	ability and byled	ability	byled
nonword	seek and broursed	seek	broursed
nonword	fear and rhins	fear	rhins
nonword	purpose and zoys	purpose	zoys
nonword	beat and meeze	beat	meeze
nonword	object and dwould	object	dwould
nonword	federal and wronts	federal	wronts
nonword	academic and glorned	academic	glorned
nonword	subject and ginth	subject	ginth
nonword	property and floaned	property	floaned
nonword	surface and blize	surface	blize
nonword	collection and strurd	collection	strurd
nonword	cover and rhonth	cover	rhonth
nonword	value and flirred	value	flirred
nonword	press and greemb	press	greemb
nonword	private and flars	private	flars
nonword	stuff and jinth	stuff	jinth
nonword	agent and ined	agent	ined
nonword	modern and migns	modern	migns
nonword	sleep and skrymbs	sleep	skrymbs
nonword	risk and swield	risk	swield

nonword	medical and sheed	medical	sheed
nonword	fight and shrows	fight	shrows
nonword	adult and brores	adult	brores
nonword	exist and snorpse	exist	snorpse
nonword	yard and krirk	yard	krirk
nonword	report and gigns	report	gigns
nonword	building and naphed	building	naphed
nonword	church and phleffed	church	phleffed
nonword	top and lunks	top	lunks
nonword	shoulder and thoaned	shoulder	thoaned
nonword	pattern and sprighs	pattern	sprighs
nonword	positive and phrarned	positive	phrarned
nonword	price and smapps	price	smapps
nonword	recent and ghouled	recent	ghouled
nonword	future and gints	future	gints
nonword	bank and breld	bank	breld
nonword	mention and frurze	mention	frurze
nonword	finger and bryled	finger	bryled
nonword	painting and droaled	painting	droaled
nonword	refer and jepped	refer	jepped
nonword	describe and neffed	describe	neffed
nonword	wrong and shronck	wrong	shronck
nonword	rest and stighed	rest	stighed
nonword	detail and jide	detail	jide
nonword	camp and warced	camp	warced
nonword	budget and gwownse	budget	gwownse
nonword	heart and jonde	heart	jonde
nonword	product and knamped	product	knamped
nonword	involve and theps	involve	theps

nonword	performance and spict	performance	spict
nonword	challenge and flaced	challenge	flaced
nonword	battle and gnouled	battle	gnouled
nonword	agreement and splode	agreement	splode
nonword	return and claste	return	claste
nonword	situation and cleas	situation	cleas
nonword	defense and drand	defense	drand
nonword	author and spind	author	spind
nonword	method and chaumb	method	chaumb
nonword	reality and phlassed	reality	phlassed
nonword	civil and hamps	civil	hamps
nonword	explain and sorld	explain	sorld
nonword	court and gnants	court	gnants
nonword	sport and vempt	sport	vempt
nonword	focus and gherck	focus	gherck
nonword	heat and shans	heat	shans
nonword	violence and shrond	violence	shrond
nonword	mouth and durze	mouth	durze
nonword	response and gynx	response	gynx
nonword	shot and prunned	shot	prunned
nonword	consumer and brins	consumer	brins
nonword	threat and gwacs	threat	gwacs
nonword	victim and mieze	victim	mieze
nonword	kitchen and theph	kitchen	theph
nonword	brain and saumths	brain	saumths
nonword	spirit and kugns	spirit	kugns
nonword	judge and twund	judge	twund
nonword	travel and franns	travel	franns
nonword	track and triend	track	triend

nonword	client and zours	client	zours
nonword	annual and ghekked	annual	ghekked
nonword	professor and keffed	professor	keffed
nonword	vote and zows	vote	zows
nonword	born and phlince	born	phlince
nonword	prevent and rhurned	prevent	rhurned
nonword	plane and blapse	plane	blapse
nonword	variety and knirnde	variety	knirnde
nonword	neck and grourn	neck	grourn
nonword	employee and yompt	employee	yompt
nonword	fan and puiced	fan	puiced
nonword	senior and pald	senior	pald
nonword	forest and zorled	forest	zorled
nonword	species and pripped	species	pripped
nonword	nuclear and snoursed	nuclear	snoursed
nonword	literature and juild	literature	juild
nonword	replace and moost	replace	moost
nonword	video and smord	video	smord
nonword	admit and foosed	admit	foosed
nonword	credit and cield	credit	cield
nonword	freedom and scroars	freedom	scroars
nonword	thigh and throns	thigh	throns
nonword	aid and wofts	aid	wofts
nonword	possibility and pharned	possibility	pharned
nonword	global and pleaned	global	pleaned
nonword	citizen and driend	citizen	driend
nonword	corner and phraft	corner	phraft
nonword	effective and thraned	effective	thraned
nonword	puzzle and ghossed	puzzle	ghossed

nonword	crowd and drix	crowd	drix
nonword	customer and drouled	customer	drouled
nonword	reform and slaphed	reform	slaphed
nonword	key and noiced	key	noiced
nonword	critical and wext	critical	wext
nonword	strike and knigned	strike	knigned
nonword	neutral and wrunned	neutral	wrunned
nonword	gather and snild	gather	snild
nonword	complain and zacks	complain	zacks
nonword	access and skoursed	access	skoursed
nonword	score and zaft	score	zaft
nonword	recall and eed	recall	eed
nonword	labour and shraud	labour	shraud
nonword	dilemma and prersed	dilemma	prersed
nonword	classroom and phrence	classroom	phrence
nonword	stretch and darred	stretch	darred
nonword	option and grauced	option	grauced
nonword	debate and prold	debate	prold
nonword	stare and wriced	stare	wriced
nonword	concept and saphed	concept	saphed
nonword	comprise and neud	comprise	neud
nonword	complex and hoosed	complex	hoosed
nonword	fashion and thrinth	fashion	thrinth
nonword	restaurant and laffed	restaurant	laffed
nonword	front and scincs	front	scincs
nonword	publish and prins	publish	prins
nonword	touch and tews	touch	tews
nonword	ministry and sciled	ministry	sciled
nonword	bridge and creened	bridge	creened

nonword	consequence and drinned	consequence	drinned
nonword	release and ghewed	release	ghewed
nonword	mission and friled	mission	friled
nonword	troop and clurs	troop	clurs
nonword	announce and bloist	announce	bloist
nonword	liberal and twauve	liberal	twauve
nonword	section and dwood	section	dwood
nonword	chief and thwield	chief	thwield
nonword	faculty and forld	faculty	forld
nonword	review and fimps	review	fimps
nonword	tea and hect	tea	hect
nonword	gender and thruys	gender	thruys
nonword	bond and pleasure	bond	pleasure
nonword	slow and rhardes	slow	rhardes
nonword	wheel and broughg	wheel	broughg

Appendix 9 Language background questionnaire for English-Chinese bilinguals

TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI
VICTORIA
UNIVERSITY OF WELLINGTON

Participant questionnaire

Gender: ___________ Age: ____________________

Occupation: ___________________

Which university do you study in (if you are a student)? ___________________

What degree are you studying for? (Tick one) Undergraduate_______ Master _______

PhD _____

What's your major? _______________

Place of birth: ___________________ (Country)

What is your first language?

Mother's first language: _______________ Father's first language: _______________

What is the main language spoken in your home when you were a child?

What's the dominant language in your life? ____________________

What other languages do you speak and your proficiency (could be more than one)?

1. ____

2. ____

3. ____

When did you start to learn Chinese? _____________________________

What percentage of the time do you use Chinese every day with family, friends and classmates?

Overall Chinese test scores if available (e.g. HSK): ___________________________

Total length of time in China: ____year(s) _____ months

What type of exposure do you think you have when you learn Chinese (Please tick one)?

Native _______ Immersion exposure ________ Classroom exposure ___________

What's the level of your Chinese? Beginner ____ Intermediate _____ Advanced ______

My vision is: Normal _____ Corrected to normal (glasses/contacts) ______ Not normal ______(please tick one).

Are you left-handed _______or right handed_______ ? (please tick one)