THE PROCESSING OF MULTI-WORD SEQUENCES IN FIRST AND SECOND LANGUAGE

by
Wanyin Li

A thesis submitted to
the University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

School of Education

College of Social Science

University of Birmingham

May 2023

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

Sentence reading not only involves recognising single words but also requires recurrent word sequences to be comprehended together. This thesis investigates the processing of different types of recurrent word sequences during sentence reading in both L1 and L2 speakers. Following a general introduction (Chapter 1) and a theoretical overview (Chapter 2), Chapter 3 followed a phraseological approach to define word combinations and compared the processing of free word combinations (of which all components are used in their literal sense and can be freely substituted, e.g., *affect one's mind*), restricted collocations (which have one lexical part that is usually used in a de-lexical or figurative sense and which cannot be freely replaced by a word with a similar semantic meaning, e.g., *enter one's mind*), and idioms (which have a conventionalised figurative interpretation, e.g., *blow one's mind*). Chapter 3 used a self-paced reading paradigm paired with a stop-making-sense task. L1 speakers read idioms more quickly than free word combinations, while L2 proficiency and the type of word combinations interacted to affect the processing of these word combinations. Chapter 4 used the same experimental paradigm and addressed idiomatic ambiguity by comparing ambiguous idioms (i.e., those that can also interpreted literally, e.g., *grey area*) to corpus-defined collocations (e.g., *vast area*) and novel literal expressions (i.e., newly coined word combinations, e.g., *rare area*), with word-level and phrase-level frequencies strictly matched. L1 speakers read idioms used in their figurative interpretation more quickly than collocations. However, this idiom advantage was not found in L2 speakers. Both L1 and L2 speakers demonstrated meaning dominance and semantic decomposability interacted to resolve the idiomatic ambiguity. Chapter 5 used eye-tracking-while-reading to investigate the change in word order in variation collocations by presenting collocations in passive constructions. The change in word order did not eliminate the collocation processing advantage in both L1 and L2 speakers.

A man is known by his friends.
A word is known by its combinations.

Acknowledgements

First, I would like to express my sincerest gratitude to my supervisors Dr. Bene Bassetti and Dr. Steven Frisson, for their invaluable mentorship and professional supervision throughout my doctoral studies. This thesis would not been in its shape without their guidance, insightful advice, and timely comments along the way. Without Bene, I would never have embarked on this exciting journey. Special thanks for Steven's support and encouragement, as well as the occasional pushes and "complaints." Not only has he always been there to offer advice and guidance, but he has helped me to believe that there is a light at the end of the tunnel.

Thanks to two dedicated examiners who have improved this thesis so much. I truly enjoyed the viva, and the discussion with Kathy was valuable.

Thanks to my "academic sibling" Dr. Mahmoud Medhat Elsherif. I have received lots of support, advice, and help from him in doing research and being a good researcher. His wit and charm have made my PhD study journey less stressful. I would like to take this opportunity to thank my PhD mate, Yifan, who made me believe I am not alone on this journey. Thanks also go to the past and present residents of 52 Pritchatts 405. Thank all of you, Zlati, Yifan, Paulina, Mahmoud, Iris, Freya and Daniel.

My family deserves a special mention. I would like to thank my family for their constant support, encouragement, understanding, and everything else. Thank my family for believing in me and being there for me every second.

谢谢我的妈妈，爸爸，还有鸿振，永远爱 他们。

Declaration

I declare that the work presented here is my own and was conducted during my time as a PhD student at the University of Birmingham. Two chapters of this thesis was submitted or is being submitted to peer reviewed journals.

The study in Chapter 5 has been submitted to *Bilingualism: Language and Cognition.*

The study in Chapter 4 is being submitted.

Table of Contents

CHAPTER 1 GENERAL INTRODUCTION 1

CHAPTER 2 GENERAL THEORETICAL REVIEW 10

2.1 Definitions of Multiword Sequences 11

2.2 Pervasiveness of MWSs 13

2.3 Diversity of MWSs 14

2.3.1 Idioms 15

2.3.2 Collocations 18

2.3.3 Binomial Expressions 19

2.3.4 Lexical Bundles 22

2.4 Characteristics of MWSs 24

2.4.1 Corpus Frequency and Familiarity 24

2.4.2 Predictability 25

2.4.3 Semantic Compositionality 26

2.4.4 Fixedness and Variation 27

2.5 Questions of Holistic Storage 29

2.6 Processing Morphologically Complex Words 33

2.7 Lexical Quality Hypothesis 37

2.8 Usage-based Models and Theories of Language 38

2.8.1 The Link between Usage Events and the Linguistic System 39

2.8.1.1 The Importance of Frequency 39

2.8.1.2 Construction 40

2.8.2 Empirical Evidence 40

2.9 Statistical Learning Theory 41

2.9.1 Definition 41

2.9.2 Empirical Evidence for Statistical Learning 42

2.9.3 Learning Non-adjacent Patterns 44

2.10 Second Language Acquisition and MWS 45

2.10.1 Difficulties in the Learning and Use of MWSs in L2 46

2.10.2 Number and Exposure of MWSs to L2 Learners 46

2.10.3 The Influence of L1 47

2.10.4 Low Salience 49

2.11 Summary 50

CHAPTER 3 COMPARING MULTI-WORD SEQUENCE PROCESSING 52

Abstract 53

3.1 Introduction 54

3.1.1 Effects of Frequency on Collocations Processing 54

3.1.2 A Frequency-based Approach to Defining Collocations 57
3.1.3 A Phraseology-based Approach to Defining Collocations 58
3.1.4 Collocation Processing under the Phraseological View 62
3.1.5 Idiom-Superiority Effect 64
3.1.6 Collocation and Idiom Processing in the Second Language 66
3.1.7 The Present Study 68

3.2 Method 69
3.2.1 Design 69
3.2.2 Participants 70
3.2.3 Materials 73
3.2.3.1 Selection of Items 73
3.2.3.2 Word-level and Phrase-level Length and Frequency 74
3.2.3.3 Context Sentences 77
3.2.4 Procedure 79
3.2.5 Measurements 82
3.2.6 Statistical Analysis 82

3.3 Results 84
3.3.1 Pre-processing Data 84
3.3.2 L1 Speakers 86
3.3.2.1 Critical Region 86
3.3.2.2 Spill-over Region 86
3.3.3 L2 Speakers 88
3.3.3.1 Critical Region 88
3.3.3.2 Spill-over Region 91
3.3.4 Comparison of L1 and L2 Speakers 91
3.3.4.1 Critical Region 91
3.3.4.2 Spill-over Region 91

3.4 Discussion 96
3.4.1 Summary of Findings 96
3.4.2 Restricted Collocations vs. Free Word Combinations 96
3.4.3 Idiom Superiority Effect 97
3.4.4 L2 Proficiency Effect 98
3.4.5 Limitations 99

CHAPTER 4 IDIOMATIC AMBIGUITY RESOLUTION 103

Abstract 104

4.1 Introduction 105
4.1.1 Multi-word Sequences 105
4.1.2 Collocations 106
4.1.3 Models of Idiom Processing 108
4.1.3.1 Non-compositional Models 108
4.1.3.2 Hybrid Models 109
4.1.4 Ambiguity Resolution 111
4.1.4.1 Lexical Ambiguity Resolution 111
4.1.4.2 Idiomatic Ambiguity Resolution 112
4.1.4.3 Additional Factors in Resolving Idiomatic Ambiguity 113

4.1.4.4 Decomposability Effect 114
4.1.5 Collocation and Idiom Processing in the Second Language 115
4.1.6 The Present Study 118

4.2 Method 119
4.2.1 Design 119
4.2.2 Participants 120
4.2.3 Materials 121
4.2.3.1 Word-level and Phrase-level Length and Frequency 124
4.2.3.2 Idioms: Idiomatic Meaning Dominance and Decomposability 125
4.2.3.3 Context 129
4.2.4 Procedure 130
4.2.5 Measurements 132
4.2.6 Statistical Analysis 132

4.3 Results 135
4.3.1 Pre-processing Data 135
4.3.2 L1 speakers 136
4.3.2.1 L1 Speakers' RTs on the Critical Region across Four Conditions 136
4.3.2.2 L1 Speakers' RTs on the Spill-over Region across Four Conditions 136
4.3.2.3 L1 Speakers: Decomposability and Meaning Dominance in Resolving Idiom Ambiguity 140
4.3.3 L2 Speakers 145
4.3.3.1 L2 Speakers' RTs on the Critical Region across Four Conditions 145
4.3.3.2 L2 Speakers' RTs on the Spill-over Region across Four Conditions 145
4.3.3.3 L2 Speaker: Decomposability and Meaning dominance on Resolving Idiom Ambiguity 147
4.3.4 Comparison L1 and L2 Data 151
4.3.4.1 Across Four Conditions 151
4.3.4.2 Decomposability and Meaning Dominance on Resolving Idiom Ambiguity 151

4.4 Discussion 153
4.4.1 Collocation Advantage in L1 and L2 speakers 154
4.4.2 Accessing the Idiomatic Meaning 154
4.4.2.1 Distributional Properties 154
4.4.2.2 Beyond the Distributional Properties 155
4.4.2.3 Similarities and Differences between Lexical and Idiomatic Ambiguity Resolution 157
4.4.2.4 The Role of Decomposability 159
4.4.2.5 Meaning Dominance and Decomposability 160

CHAPTER 5 PROCESSING NON-CANONICAL, PASSIVIZED COLLOCATIONS 162

Abstract 163

5.1 Introduction 164
5.1.1 Collocation Processing 165
5.1.2 Variation of Collocations and L1 Processing 169
5.1.3 Processing Collocations in L2 172
5.1.4 The Present Study 173

5.2 Method 174
5.2.1 Participants 174
5.2.2 Materials 174
5.2.2.1 Materials Preparation 175

5.2.2.2 Predictability 176
5.2.2.3 Frequency and Length 176
5.2.2.4 Apparatus and Procedure 177
5.2.2.5 Pre-processing of Data 178
5.2.3 Statistical Analysis 179

5.3 Results 181
5.3.1 L1 Speakers Results 183
5.3.1.1 Phrase Analysis 183
5.3.1.2 Final Word Analysis 183
5.3.2 L2 Speakers Results 183
5.3.2.1 Phrase Analysis 183
5.3.2.2 Final Word Analysis 184
5.3.3 Comparison of L1 and L2 Speakers 190
5.3.3.1 Phrase Analysis 190
5.3.3.2 Final Word Analysis 190

5.4 Discussion 192
5.4.1 Collocation Facilitation Effect 192
5.4.2 Passivized Collocations Facilitation Effect 194
5.4.3 The Question of Holistic Storage 195
5.4.4 Similarities between L1 and L2 Speakers 196
5.4.5 The Proficiency Effect on Collocation Processing in L2 Speakers 198

CHAPTER 6 GENERAL DISCUSSION 200

6.1 Summary of Findings 201
6.1.1 Effects of Phrase-level Frequency 201
6.1.2 Beyond Distributional Information 203
6.1.3 Resolving Idiomatic Ambiguity 205
6.1.4 Semantic Decomposability 206
6.1.5 The Processing Advantage in Reversed Collocations 207
6.1.6 Processing L2 MWSs 209

6.2 Limitations 210

6.3 Future Research and Implications 216

6.4 Conclusion 218

REFERENCES 219

APPENDICES 235

LIST OF FIGURES

Figure 3.1 Cowie and Howarth's Continuum Model of Classifying Four Types of Word Combinations 59

Figure 3.2 Presentation of the Procedure from One Trial..... 80

Figure 3.3 The Interaction between Types-of-MWSs and Proficiency for L2 Speakers..... 89

Figure 3.4 The Interaction between Types-of-MWSs and Language-background for the Critical Region 92

Figure 3.5 The Interaction between Types-of-MWSs and Language-background on the Spill-over Region 94

Figure 4.1 Meaning Dominance Norming Example..... 125

Figure 4.2 Local Semantic Decomposability Norming Example..... 127

Figure 4.3 The Interaction between Context and (Adjectives) Decomposability on the Critical Region for L1 Speakers..... 140

Figure 4.4 The Interaction between Context and (Nouns) Decomposability on the Critical Region for L1 Speakers..... 141

Figure 4.5 The Interaction between Meaning Dominance and (Adjectives) Decomposability on the Critical Region for L1 Speakers..... 142

Figure 4.6 The Interaction between Context and (Noun) Decomposability on the Critical Region for L2 Speakers..... 147

Figure 4.7 The Interaction between Meaning Dominance and (Adjectives) Decomposability on the Spill-over Region for L2 Speakers..... 148

Figure 4.8 The Interaction between Types-of-MWSs and Language-background on the Spill-over Region 150

Figure 5.1 Example of a Sentence with a Collocation and its Control in Canonical and Reversed Order 173

Figure 5.2 L2 Speakers: Interaction between LexTALE Score and Phrase Type in Total Fixation Count for Canonical Constructions..... 187

Figure 5.3 Total Phrase Reading Time for Collocations and Novel Word Combinations by Group..... 189

LIST OF TABLES

Table 3.1 Summary of Participants' Language Usage..... 71

Table 3.2 Summary of Items Characteristics..... 75

Table 3.3 An Example of One Set of Items and Their Context Sentences..... 77

Table 3.4 Descriptive Statistics: Mean RT (ms) and Standard Deviations for the Critical and Spill-over Regions across Three Conditions for Both L1 and L2 Speakers..... 84

Table 3.5 Summary of the Best-fitting Models of MWSs on the Critical Region for L1 Speakers..... 86

Table 3.6 Summary of the Best-fitting Models of MWSs for the Critical Region for L2 Speakers..... 88

Table 3.7 Summary of the Best-fitting Models of MWSs for the Critical Region for L1 and L2 Speakers 91

Table 3.8 Summary of the Best-fitting Models of MWSs on the Spill-over Region for L1 and L2 Speakers 93

Table 4.1 Summary of Language Proficiency Data..... 120

Table 4.2 An Example of One Set of Items and Context Sentences..... 121

Table 4.3 Summary of Items Characteristics..... 124

Table 4.4 Summary of Characteristics of Idioms..... 127

Table 4.5 Descriptive Statistics: Mean RT (ms) and Standard Deviations for the Critical Region and Spill-over Region across Four Conditions for Both L1 and L2 Speakers..... 136

Table 4.6 Summary of the Best-fitting Models of MWSs on Critical and Spill-over Regions for L1 Speakers 137

Table 4.7 Summary of the Best Fitting Models of Idioms on the Critical and Spill-over Regions for L1 Speakers..... 139

Table 4.8 Summary of the Best-fitting Models of MWSs for the Critical and Spill-over Regions for L2 Speakers..... 144

Table 4.9 Summary of the Best Fitting Models of Idioms on the Critical and Spill-over Regions for L2 Speakers..... 146

Table 5.1 Summary of Measures in L1 and L2 Speakers..... 180

Table 5.2 L1 Speakers: Summary Statistics for the Phrase Region..... 183

Table 5.3 L1 Speakers: Summary Statistics for the Final Word Region..... 184

Table 5.4 L2 Speakers: Summary Statistics for the Phrase Region..... 185

Table 5.5 L2 Speakers: Summary Statistics for the Final Word Region..... 186

CHAPTER 1 GENERAL INTRODUCTION

When a reader opens an English book, each page is filled with individual words. This fact leads previous researchers to assume that individual words are fundamental in building our language (with syntax providing the rules in which these words can be arranged to form language-appropriate sentences). It is, therefore, not surprising that for a long time, the focus in language processing has been the recognition of individual words in English. There is also a large amount of evidence that language users are sensitive to distributional information at the word level. For example, the print frequency of a word is a significant factor in the word recognition (Forster & Chambers, 1973; Grainger, 1990; Rayner & Raney, 1996). The frequency effect is not just apparent between very frequent and very infrequent words (e.g., *year* vs. *heresy*) but also between frequent and less frequent words (e.g., *rain* vs. *puddle*). These findings in word recognition have also shaped and influenced language acquisition theories and models, such as the usage-based approaches to language acquisition (Bybee, 1998; Ellis, 2008; Tomasello, 2009), wherein experience (as indexed by frequency) has an important role in such approaches to language acquisition. Regardless of the size and complexity of linguistic units, all linguistic units are supposed to be similarly affected by experience. The equal role of frequency in different linguistics units supports the single-system model of the language (Goldberg, 2006).

Contrary to the usage-based approaches listed above, a words-and-rules language theory (Pinker, 1991; Pinker & Prince, 1988; Pinker & Ullman, 2002) draws a clear distinction between mental lexicon (i.e., stored information of a word containing its meaning, pronunciation, syntactic features and so on) and mental grammar (i.e., a productive system that is used to assemble simple words into complex words, phrases, and sentences). This approach also claims that individual words are stored in the mental lexicon, while linguistic units larger than simple words (e.g., complex words, phrases, and sentences) are computed based on grammar and rules. Importantly, by following this approach, the processing of single words is expected to be affected by frequency. However, the computation of bigger linguistic constructions would not be affected.

When reading sentences in a book, readers assemble the meaning of each word according to syntactic rules, consistent with the words-and-rules language theory (e.g., Pinker & Ullman, 2002). However, sometimes we need to bundle a few words together to successfully resolve the meaning, e.g., *at the end of the day, my point of view, in order to, blow the whistle, pull string, by and large*. In the case of these examples, it is hardly possible to derive the correct meaning from their lexical parts. With this renewed understanding, the traditional emphasis on single words has changed, and there is increasing theoretical emphasis on, and empirical evidence for, multi-word linguistic sequences that are now considered equally fundamental in building our written and spoken linguistic knowledge as well (Arnon et al., 2017; Arnon & Snider, 2010). This thesis adopts and follows Conklin and Carrol's (2021) multi-word sequences (MWSs) definition, which states that MWSs are conventionalised word strings and sequences that native speakers tend to use together (Conklin & Carrol, 2021). MWSs are seen as familiar phrases that exhibit a certain degree of fixedness and are recognised as conventional by native speakers. MWSs include various types of word combinations, such as idioms, collocations, and binominals (Wray, 2008). Idioms are characterised as a string of words whose figurative meaning cannot be retrieved directly from the literal computation of their components (e.g., *blow your own trumpet*). Thus, idioms are broadly seen to be non-decomposable and fixed to some degree. While familiarity and frequency of co-occurrence are characteristic of MWSs, an exact definition of MWS remains elusive. First, what counts as "frequently co-occurring" is an open question (Carrol & Conklin, 2020). Second, idioms challenge the view of frequency, as some idioms are not very frequent in natural language compared to other types of MWSs. Collocations are broadly defined as pairs of words that occur together more frequently than we expect by chance (Biber et al., 1999) (e.g., *develop software* and *provide information*). Importantly, two main methods can be used to identify collocations. From a phraseological view (Howarth, 1998), one lexical part of the collocation is not used in its literal sense (e.g., *enter* in *enter one's mind*), making the whole collocation partially figurative. Under the

phraseological view, collocations are typically referred to as restricted collocations. From a frequency-based view, collocations are identified according to certain corpus-derived metrics, such as mutual information (MI). MI measures the exclusivity of the co-occurrence of two words. Any word combination with an MI score larger than 3 typically is taken as a collocation (Hunston, 2022). However, the two methods above can cause conflicts in identifying collocations. For example, the frequency-based view identifies *provide information* as a collocation, while the phraseological view identifies the same phrase as a free word combination as the two lexical components are used in their literal sense. In free word combinations, each component can be substituted with another word without affecting the meaning of the other. For example, in *blow a trumpet*, *blow* and *trumpet* are used literally and can be substituted by other words, e.g., *bubble*, *shofar*, as long as the requirements of semantics and grammar are satisfied. Furthermore, a given phrase that is identified as a restricted collocation (e.g., *build a career*) might not meet the requirements of corpus-derived metrics (e.g., the MI score might not cross the threshold). A full discussion of these two approaches to identifying collocations will be presented in Chapter 3. Similar to a free word combination, a binominal expression (e.g., *fish and chips*) has literal components. However, different from a free word combination, a binominal expression has a stable syntactic structure and is typically formed as a sequence of x and y that is connected by *and/ or* (e.g., *short and sweet*, *now or never*). Binominal expressions have two shared features (Carrol & Conklin, 2020; Masini, 2006): (1) the lexical parts of a binominal expression are from the same word class; (2) a specific word order of a binominal expression is highly preferred.

This thesis investigates the processing of English MWSs in first language (L1) speakers and second language (L2) speakers. Existing research has demonstrated that language users are sensitive to the frequency of multi-word sequences (Carrol & Conklin, 2020; Li et al., 2021; Öksüz et al., 2021; Sonbul, 2015; Vilkaitė, 2016; Vilkaitė & Schmitt, 2019; Wolter & Gyllstad, 2013). Such sensitivity to multi-word level frequency also exists in L2 speakers (Li et al., 2022; Öksüz et al., 2021; Sonbul, 2015;

Vilkaitė, 2016; Vilkaitė & Schmitt, 2019; Wolter & Gyllstad, 2013). However, it remains unclear whether factors other than distributional information, such as semantic decomposability, conventionality, and syntactic modification, also affect the processing of multi-word sequences. There is a dearth of research examining how these additional factors are implicated in the processing advantage shown for MWSs. Assessing these additional factors provides a valuable window into language since MWSs tap into both processing (i.e., it highlights that when encountering MWSs, the individual words are retrieved from memory and combined into a unified unit) and storage (i.e., MWSs are stored and accessed not just as individual words but also at a conceptual level).

The current thesis addresses the knowledge gap highlighted above by conducting, to my knowledge, the first systematic comparisons of different types of MWSs (i.e., corpus-based collocations, restricted collocations, idioms, and free word combinations; see Chapters 2 and 3 for detailed descriptions of these different types of MWSs). This comparison allows us to assess the effect of semantic decomposability and conventionality on MWSs processing.

This thesis also addresses the knowledge gap described above by investigating whether a change in the syntactic structure of MWSs, concretely reversing word order, affects the processing advantage. Although MWSs were traditionally seen to be conventionalised and "rigid" (Chafe, 1968; Weinreich, 1969), MWSs seem to possess a certain degree of flexibility, allowing a greater level of modification than previously thought (Kyriacou et al., 2020; Moon, 1998; Schmitt, 2006). By manipulating their syntactic structure, this thesis tests whether the processing advantage is retained for reversed MWSs even though this form hardly appears in corpora. This is an important question as it can tell us whether the processing advantage is merely driven by distributional information and whether MWSs are processed and stored holistically.

This thesis has six chapters: an introduction chapter (the present one), a general theoretical review chapter (Chapter 2), three empirical chapters and a general discussion chapter (Chapter 6). Each experimental chapter (Chapter 3, Chapter 4, and Chapter 5) has been written as an individual paper, and each chapter begins with a short abstract and a thorough review of the relevant experimental studies and theoretical literature. Each experimental chapter has been written in 1st person plural as they have been submitted or written in preparation for submission.

Chapter 2 summarises the important characteristics of MWSs and provides a broader theoretical review of the research of MWSs. This overview helps to provide general knowledge from which to develop empirical studies. First, the various definitions of MWSs are outlined. Then, the pervasiveness of MWSs and the importance of the use of MWSs in natural language are discussed. Different types of MWSs are then discussed, and the four widely studied MWSs, idioms, collocations, lexical bundles, and binomial expressions are reviewed. Following that, several features of MWSs, including frequency and familiarity, predictability, semantic compositionality, and fixedness and variation, are summarised. Next, theories and models of language acquisition (i.e., usage-based theories to language acquisition, statistical learning theories) and hypotheses (holistic storage of MWSs, models of the processing of complex words) related to the processing and learning of MWSs are reviewed. Finally, the chapter explains the difficulties in processing and using L2 MWSs.

Chapter 3 is the first empirical chapter investigating how different types of MWSs were processed in both L1 and L2 speakers. The comparison of types of MWSs is important, as the comparison can tap into whether additional factors (e.g., the degree of meaning conventionality and semantic decomposability) influence the processing of MWSs. To address this question, this study followed the phraseological approach to define MWSs (Cowie, 1998; Howarth, 1998) and three types of MWSs were examined: idioms, restricted collocations, and free word combinations. The three types of MWSs varied in semantic decomposability and conventionality, while distributional properties (i.e., the

frequency of the lexical parts of the MWSs and the frequency of the MWSs as a whole) were controlled for. The self-paced reading paradigm coupled with a concurrent stop-making-sense task (Boland et al., 1990; Gibson et al., 1996; Singh et al., 2016) was employed to measure processing difficulty. The following three questions were addressed in Chapter 3:

1. whether the idiom superiority effect existed when the phrase-level frequency was controlled for;
2. whether the processing cost for restricted collocations over free word combinations can be observed in a sentence reading task in order to extend a previous study (Gyllstad & Wolter, 2016);
3. whether L2 speakers processed the three types of MWSs in a similar manner to L1 speakers and whether L2 proficiency modulated MWSs processing.

Chapter 4 extended the previous investigation of different types of MWSs by considering additional idiomatic properties and addressing several limitations of the experiment reported in Chapter 3. Firstly, this study used ambiguous idioms that can be interpreted in either a literal or figurative sense (e.g., *grey area*; *sore spot*) with supporting prior context that led to either a figurative or a literal interpretation. Secondly, this study measured two extra properties of idioms: the degree of semantic decomposability (i.e., the degree to which the lexical parts of an idiom contribute to its overall figurative sense) and meaning dominance (i.e., the extent to which idioms are processed and interpreted figuratively as opposed to literally). Thirdly, contextual predictability and MWSs' familiarities and frequencies were more stringently controlled than in Chapter 3 and previous published studies. Therefore, it allowed us to address five questions:

1. whether the processing advantage in idioms is retained in the absence of higher familiarity and frequency;
2. whether the resolution of idiomatic ambiguity and single words were similar;

3. whether meaning dominance affected idiom processing;
4. whether semantic decomposability played a role in comprehending idioms;
5. whether L1 and L2 speakers processed MWSs and resolved idiomatic ambiguity in a similar fashion.

Chapter 5 shifted the attention from comparing different types of MWSs to examining statistically defined collocations exclusively. Unlike the previous experiments, this experiment employed the eye-tracking-while-reading paradigm. The chapter specifically assessed whether a change in canonical word order affected the processing of collocations in both L1 and L2 speakers. Previous literature shows a collocation advantage in L1 and L2 speakers (i.e., collocations are processed more rapidly relative to novel word combinations). The higher phrase-level frequency (i.e., the frequency of the entire expression) of collocations drives this advantage compared to novel word combinations that rarely appear in a corpus. This advantage is also mostly found with collocations in their canonical form (e.g., *provide information*). However, most previous studies of collocations have failed to control for contextual predictability. In addition, the inherent flexibility of collocations has hardly been examined (e.g., *providing information*, *information provided*). By reversing the word order of collocations, the following questions can be addressed:

1. whether the collocation advantage was still apparent for reversed collocations;
2. whether L1 and L2 speakers processed reversed collocation qualitatively differently;
3. whether the collocation advantage was independent of context and whether this advantage emerged during an early or later stage of processing.

In closing, Chapter 6 summarises the findings of the thesis. The concluding chapter brings together the results from the empirical chapters and aims to draw overarching conclusions relating to

the processing of MWSs in both native and non-native speakers. Finally, some limitations of this thesis are highlighted and future research into MWSs is proposed.

CHAPTER 2 GENERAL THEORETICAL REVIEW

This chapter provides a broad overview of the research on MWSs. The first section discusses the different definitions that have been used when referring to MWSs. The next section explains the pervasiveness and diversity of MWSs. To better understand the big picture of MWSs, a few sub-types of MWSs are introduced and reviewed. Then, as the thesis focuses on MWS processing, I will discuss models of processing complex words, followed by a section on relevant language acquisition theories as they pertain to MWSs. The last section looks at the MWSs from the view of second language processing.

2.1 Definitions of Multiword Sequences

In recent years, a growing number of studies have been conducted on multi-word sequences (MWSs) in different fields of linguistics. In general, MWSs are referred to as a string of words or word combinations that are recurrent and familiar to language users. However, researchers from different disciplines tend to use different terms to refer to MWSs. These terms include formulaic language (Conklin & Schmitt, 2008, 2012; Wray, 2002, 2008), prefabricated language (Erman & Warren, 2000), fixed expressions (Moon, 1998), chunked sequences (Ellis, 1996) and multiword unit (Arnon & Christiansen, 2017; Christiansen & Arnon, 2017; Geeraert et al., 2017; Jolsvai et al., 2020). More example terms can be found in Wray (2002).

Although these terms refer to a similar concept of multiple words, each term might place emphasis on different aspects of word combinations, such as its co-occurrence frequency, conventionalisation, or storage in the mental lexicon. One of the most cited and influential terms is "formulaic language." Wray (2002) claims formulaic language is a sequence of continuous or discontinuous words stored, retrieved, and processed as a whole chunk. Additionally, formulaic language is not subject to linguistic analysis and computation. Thus, according to Wray (2002), the computation of the meanings of the individual words does not play much of a role in defining formulaic language. It is emphasised that word combinations are holistic in terms of storage, representation, and processing. However, Wray (2012) later proposes that there may be a different quality to holistic

storage and access of MWSs. First, the processing of MWSs can entail accessing individual base components. Second, a direct mapping between the form and meaning can be involved, which results in MWSs' processing advantage. On the other hand, terms such as multi-word units or sequences do not attach importance to holistic retrieval and storage.

The same terms sometimes are interpreted differently or even incorrectly in the literature. One of the main reasons for a wrong interpretation of terms is that there are two major distinct approaches to the study of word combinations: the *phraseological* approach (Cowie, 1998; Howarth, 1998; Nesselhauf, 2003) and the *distributional or frequency-based* approach (Sinclair, 1991). Researchers following the phraseological approach have adopted a top-down method, using certain linguistic criteria (e.g., semantic compositionality or sense restriction; Chapter 3 will provide an in-depth description of these linguistic criteria) to distinguish different types of MWSs. Researchers following the distributional approach have used a bottom-up, corpus-driven method of identifying co-occurring words and placed less emphasis on distinguishing between different categories of word combinations. The details and operationalisations of the two approaches in defining, identifying, and categorising MWSs will be discussed and exemplified further in Chapter 3.

Despite many existing terms, the present thesis chooses "multi-word sequences" (MWSs) to describe this linguistic phenomenon. The MWSs are understood as two or more words recurring in near proximity to each other. There are a few reasons for choosing MWSs. First, MWSs is an inclusive term that can capture the prominent feature of having more than a single word. Second, MWSs is a neutral term and does not show a preference for either of the two approaches to defining word combinations. Finally, this term does not presuppose that MWSs are analysed, computed from their lexical parts, and processed and retrieved holistically.

2.2 Pervasiveness of MWSs

Before discussing the details of how MWSs are processed, it is important to know why the study of MWSs is necessary for language processing and learning. One reason is that MWSs are widespread in both spoken (Altenberg, 1998; Foster, 2001; Pawley & Syder, 1983) and written language (Biber & Conrad, 1999; Erman & Warren, 2000). More than half of written and spoken language consists of recurrent strings of words (Erman & Warren, 2000). According to Howarth's analysis of the academic corpus (Howarth, 1998), MWSs account for 31-40 percent of language. Biber et al. (1999) estimated that 30 percent of language in their conversational corpus comprises MWSs. Additionally, in the early speech production of children, 25 percent of what is produced consists of recurrent word combinations (Van Lancker Sidtis, 2012). The results of these studies show a convergent finding that using MWSs of L1 speakers is common.

MWSs are a phenomenon in the first language (L1) and the second language (L2). Studies also found L2 speakers develop the usage of MWSs. To illustrate, in L2 spoken language, 20 percent consists of recurrent MWSs (Foster, 2001). Additionally, in L2 written language, Howarth (1998) found that 25 percent of the text is conventional. Although these results demonstrate that L2 speakers use MWSs, much research on L2 MWSs has suggested that L2 speakers find the acquisition and the correct use of MWSs challenging (Fioravanti et al., 2021; Schmitt, 2012).

Corpus analysis has revealed a clear picture that MWSs are ubiquitous in language, which leads us to ask why MWSs are widespread. One reason is that MWSs might tend to rely on long-term memory, thus reducing the processing cost of working memory (Conklin & Schmitt, 2008). The evidence for reducing the cost of working memory comes from the eased processing loads for frequent MWSs (Jiang & Nekrasova, 2007; Tremblay & Baayen, 2010). MWSs might rely less on the online linguistic computation route via word selection and grammatical rules and more on the direct retrieval route from

long-term memory, which arguably is less cumbersome: reduced online computation is thought to require less cognitive demand (Ji et al., 2011).

Another reason for the ubiquity of MWSs is that they can provide useful functions in the field of communication (Conklin & Schmitt, 2012; Schmitt, 2012). For example, MWSs can be used to maintain conversations or small talk (e.g., *Nice to meet you*; *How are you; Let's keep in touch*), to express a message or an idea (e.g., *Don't cry over spilled milk*; *Better safe than sorry*; *it never rains it pours*), and to signpost in speech and writing (e.g., *I am going to cover*...; *on the other hand*; *in contrast*).

2.3 Diversity of MWSs

Regardless of whether the phraseological or the distributional approach is followed, MWS is an umbrella term and includes many types of word combinations (Carrol & Conklin, 2020), for instance, idioms, collocations, binomial expressions, clichés, and lexical/N-gram bundles. These types of MWSs have not generated equal interest in the literature. Idioms, as part of general figurative language, have received the most attention in the psycholinguistics literature (Cacciari, 2014; Siyanova-Chanturia, 2013; Titone et al., 2015). Previous studies have been interested in comparing the processing of idioms to matched literal expressions (Cacciari & Tabossi, 1988; Conklin & Schmitt, 2008; Gibbs, 1980; M. R. Libben & Titone, 2008; Tabossi et al., 2009), or resolving the ambiguity between their figurative interpretation and literal interpretation (Carrol & Conklin, 2017; Milburn & Warren, 2019; Siyanova-Chanturia, Conklin, & Schmitt, 2011). In contrast, collocations are mainly studied in corpus linguistics in terms of identifying collocations (Gries, 2013; Pecina, 2010; Wiechmann, 2008) and the use patterns of collocations (Altenberg, 1998; Biber et al., 1999; Erman & Warren, 2000; Foster, 2001; Glucksberg, 1989; Pawley & Syder, 1983).

Several typical types of MWSs and their characteristics will be discussed in the sections below. Idioms and collocations are the focus of this thesis, and hence, these two types of MWSs are reviewed in greater detail (though a more in-depth discussion of collocations will be postponed until Chapter 3). Binomial expressions and lexical bundles are also introduced to provide a broader overview of MWSs. While the general theories are covered in this introduction chapter, empirical studies of the processing of idioms and collocations will be discussed in subsequent experimental chapters.

2.3.1 Idioms

In this section, general and narrow definitions of idioms are introduced. Why idioms have been seen as fixed expressions is discussed. Finally, as idioms belong to the larger category of figurative language, different processing models and hypotheses of figurative language are reviewed.

Idioms are typical examples of the MWSs (Cacciari, 2014; Siyanova-Chanturia, 2015), and there has been a long history in the study of idioms. Idiomaticity broadly refers to the tendency of a string of words to take on meanings beyond the meaning of their lexical parts. Idiomaticity goes hand in hand with compositionality (i.e., the meaning of a word string as derived from the meanings of the words in the word string) (Wulff, 2008). Idioms are also narrowly defined as non-decomposable and lexically fixed word strings (Swinney & Cutler, 1979). The use of idioms delivers a conventionalised and non-literal meaning (e.g., *a piece of cake*, *kick the bucket*). One of the most striking features of idioms is that the relationship between the lexical parts of an idiom and the conventional meaning can be arbitrary in many cases.

Idioms have posed a major challenge to word-centered linguistic theories and models since idioms are traditionally seen as frozen and fixed word strings. Idioms are fixed in two levels. First, components of the idiom (i.e., its lexical parts) cannot be freely changed or replaced. Although idioms might be less fixed than people expect (Cacciari & Glucksberg, 1991; Konopka & Bock, 2009), the

substitution of lexical parts of an idiom can disrupt the figurative interpretation of an idiom. For example, if *break* from *break a leg* is substituted by a verb with a similar meaning (e.g., *smash*, *crack)*, the figurative interpretation of break a leg (i.e., good luck) is not maintained. Second, the conventionalised meaning of an idiom would not be changed by different contexts (Cacciari, 2014). In other words, the figurative sense of an idiom remains "frozen." In contrast, other types of figurative language, such as a metaphor (i.e., a word or phrase denoting one kind of object or idea by mentioning another; a metaphor usually provides a similarity and likeness, e.g., *people are ants, John is a wild cat*), can express different meanings depending on the context. For example, "From the top of the hill, the people are ants." describes the people as small as ants. However, "This is a busy road. The people are ants." describes the people as many as ants.

Due to the figurative interpretation of idioms, idioms also fall under the big category of figurative language (e.g., metaphor, metonym, proverbs, and irony). Early studies of figurative language hold that the analysis of the literal interpretation of a figurative term is completed before deriving the figurative interpretation (Grice, 1989). In a neutral or non-informative context, it is more difficult to comprehend the figurative interpretation than the literal one. Hence, it seems to be reasonable to assume that the literal interpretation is activated first and is compulsory (Inhoff et al., 1984; Ortony et al., 1978). However, more experimental data suggests that a figurative interpretation can be processed as fast as a literal interpretation and does not require more cognitive effort (Blasko & Connine, 1993). For example, in a metaphor study using a cross-modal priming task, Blasko and Connine (1993) show that both figurative-related (e.g., *secure*) and literal-related targets (e.g., *hard*) were primed after listening to familiar metaphoric items (e.g., *the family is a rock*) compared to the controls.

Although understanding figurative language does not require more cognitive demand than literal language (Blasko & Connine, 1993), the question of whether the literal analysis of lexical parts is active when comprehending figurative language remains to be answered. The graded salience

hypothesis (Giora, 2002) holds that context influences the processing of figurative senses when the activation of salient meaning is complete. This hypothesis does not specify exactly what a salient meaning is, but the salient meaning of a figurative term typically reflects the most common and conventionalised meaning that is strongly encoded and consolidated in the lexicon; thus, the salient meaning of a figurative term can be either a literal meaning or a figurative meaning. Additionally, it is hypothesised that the salient meaning of figurative language is always activated first, regardless of disambiguating context.

Some evidence supporting the graded salience hypothesis comes from studies investigating familiar (e.g., *very funny*) and less familiar (e.g., *thanks for your help*) ironies relative to literal interpretations of the same expressions (Giora & Fein, 1999). Participants were asked to judge an irony-related target (e.g., *angry*) or literal target (e.g., *useful*) after reading an expression either in its figurative use or literal use. When participants read fewer familiar ironies, the literal words were responded to more quickly than the irony-related words. However, there was no difference in response times when reading familiar ironies. This means the literal sense was the salient meaning and was activated first when processing fewer familiar ironies.

Parallels can be drawn between the graded salience hypothesis and frequency effects. The literal-figurative sense ambiguities might be similar to the meaning ambiguities of ambiguous words (i.e., a word that has two semantically unrelated meanings, e.g., *letter*, *bank*). The relative frequencies of the alternative senses of an ambiguous word determine the access to the meaning of the ambiguous word (Duffy et al., 1988; Sereno et al., 1992). Likewise, the salient meaning in a figurative expression refers to the most common/frequent sense that is activated first. Contrary to the graded salience hypothesis that denies the role of context, accessing the meaning of ambiguous words depends on the disambiguating context constraints that interplay with meaning frequency (Duffy et al., 1988; Vu et al., 1998).

Another general model of figurative language is described by the underspecification model (Frisson & Pickering, 2001; Pickering & Frisson, 2001). This model assumes that initially, the language user activates an interpretation that fits both figurative and literal senses, instead of one that corresponds to a specific sense. The model hypothesises that the language user can use context to pin down the contextually appropriate sense after activating the underspecified meaning. Therefore, there is no immediate commitment to a single fully specified sense. Furthermore, if the prior context biases towards one sense, the language user might resolve the meaning rapidly. If the prior context is neutral, there might be a delay in resolving the specific sense. The evidence for the underspecification model came from an eye-tracking study examining the processing of ambiguous verbs (Frisson & Pickering, 2001) and studies of metonymic expressions such as *talking to the school* and *reading Dickens* (Frisson, 2015; Frisson & Pickering, 1999, 2007).

2.3.2 Collocations

Unlike idioms[1], defining a collocation and drawing the differences between collocations and other types of word combinations is still an open issue. There are two approaches to define and identify collocations: corpus-based statistical approaches (Sinclair, 1991) and phraseology-based typological approaches (Cowie, 1998; Howarth, 1998; Nesselhauf, 2003). A phraseology-based typological approach attaches the importance of selection restrictions on the lexical parts of collocations (e.g., in *make a friend*, *make* appears in a de-lexical configuration), but does not place emphasis on the frequency of occurrence. In contrast, a corpus-based approach defines collocations as statistically significant co-occurrences, regardless of the semantic bonds between the lexical parts of the collocations (e.g., *provide*

[1] There are still some disagreements on what idioms are, but it is less disputed than is the case for collocations (Granger & Meunier, 2008).

information, lower risk). The distinctions between these two approaches will be fully discussed and explored in Chapter 3.

The collocations that this thesis refers to are lexical collocations. Lexical collocations include two lexical words (e.g., *heavy rain*, *enter my mind*), which contribute to the collocational meaning (Granger & Meunier, 2008). In contrast, grammatical collocations are restricted word combinations that contain a lexical word and a grammatical word. One of the typical examples of grammatical collocation is adjective/verb/noun + preposition (e.g., *independent of, depend on, the answer to*). Grammatical collocations are beyond the scope of this thesis.

2.3.3 Binomial Expressions

This section covers the definition and features of another type of MWSs – binomial expression, before discussing their processing and learning.

Binomial expressions have a preferred syntactic structure and are typically formed as sequences of x and y (e.g., *bread and butter, short and sweet,* and *slice and dice*). Although the majority of x and y sequences in binominal expressions are connected by *and*, there are other types of binominal coordinators (e.g., *or*, *to*; *heart to heart; now or never*). According to these examples, it is apparent that binominal expressions have two features (Carrol & Conklin, 2020; Malkiel, 1959; Masini, 2006). Firstly, the lexical parts of a binominal expression are from the same word class (e.g., both *bread/butter* are nouns; *short/sweet* are adjectives; and *slice/dice* are verbs). Secondly, a specific word order of a binominal expression is highly preferred. Language users usually say *short and sweet*, *bread and butter*, instead of *sweet and short*, *butter and bread.* The reversed form of a binominal expression (e.g., *sweet and short*, *butter and bread*) is much less frequent than the normal form (*short and sweet*, *bread and butter*). In an eye-tracking study (Siyanova-Chanturia, Conklin, & van Heuven, 2011), the average frequency of the binominal expressions used was 247.3 in the British National Corpus (BNC), while the

average frequency of the reversed order was 27.4. On the one hand, the word order change in some of the binomial expressions does not affect understanding (e.g., *bride and groom*, *king and queen*, *pepper and salt,* and *food and drink*), but the change might slightly alter the emphasis of the expression. In other cases (e.g., *bread and butter*, *black and white*), the reversed word order might disrupt the conventionalised meaning (e.g., *butter and bread* might lose the conventionalised meaning of *bread and butter* i.e., *the way people earn money needed to live*).

Empirical evidence suggests that binomial expressions are processed more quickly than their reversed forms. In an eye-tracking study, Siyanova-Chanturia, Conklin, and van Heuven (2011) investigated binomial expressions and their reversed forms. They observed that both L1 and L2 speakers were sensitive to the phrase-level frequency (i.e., the frequency of the entire expression), with more frequent expressions being processed more quickly than less frequent ones. Additionally, L1 speakers and highly proficient L2 speakers processed binomial expressions faster than the reversed forms, while less proficient L2 speakers read binomial expressions and their reversed forms at a similar speed. The different findings with respect to phrase-level frequency and word order are because not all binomial expressions have higher frequencies than all reversed forms. For example, the binomial expression *east and west* appears 373 times in BNC, and the reversed form *west and east* occurs 61 times, which is more frequent than the binomial expression *sweet and sour,* which only appears 35 times. Overall, the results of the eye-tracking study suggest that the higher frequency of binomial expressions leads to the form becoming entrenched in the mental lexicon. In the case of less proficient L2 speakers, binomial expressions are not entrenched in memory due to the lack of encounters. However, their experiment included binomial expressions such as *black and white and heart and soul*, which have stronger figurative interpretations than other items such as *east and west and sweet and sour*. It has been known that figurative language and literal language are processed differently, and so the lack of control of the

figurativeness of the binomial expressions in their experiment might have introduced a confound in their design.

Another eye-tracking study demonstrated that sensitivity to novel binomial expressions arises in natural reading with L1 speakers after a small number of exposures (Conklin & Carrol, 2021). Participants were asked to read existing and novel binomial expressions in short stories (novel means the frequency of the expression was below ten times in BNC). Novel binomial expressions were invented by conjoining two words from the semantic network with the coordinator *and* (e.g., *stone and wood, grass and leaves* and *paint and glue*). The first word of each novel binomial expression did not come from the first word of existing binomial expressions to avoid predicting and activating the existing expressions. In the short stories, existing expressions and the reversed forms of existing expressions appeared one time only. Novel expressions appeared from one time up to five times, but the reversed form for the novel expressions appeared one time. They observed faster reading of forward novel binomial expressions compared to avoid predicting and activating the existing expressions. In the short stories, existing expressions and the reversed forms of existing expressions appeared one time only. Novel expressions appeared from one time up to five times, but the reversed form for the novel expressions appeared one time. They observed that increasing the number of encounters with a novel binomial expression sped up the reading of the novel binomial expression. Contrary to the existing binomial expressions where the reversed forms were processed more slowly, the reversed forms for the novel expressions were read more quickly than the initial encounter of the non-reversed novel expressions. In other words, repeatedly seeing *grass and leaves* sped leaves and grass, making the reversed form faster than the initial encounter, probably due to the semantic level retrieval of the binomial. Moreover, after 4-5 times exposure to a binomial expression, the novel forward expression was read faster than the reversed one. The more times the participants encountered the novel binomial expression, the slower they read the reversed expression.

2.3.4 Lexical Bundles

This section starts with a definition of lexical bundles. Following the features of lexical bundles, several important studies that investigate the processing of lexical bundles are reviewed.

Lexical bundles have been defined as frequent and recurring sequences of three or more words, for example, *as a result of*, *in front of*, and *in the middle of*. The identification of lexical bundles in a corpus is based on two aspects: the frequency of occurrence and the type of text. For example, a four-word lexical bundle has to appear at least 20 times per million words in five different types of texts (Biber et al., 2003; Biber & Conrad, 1999).

Although frequency-based collocations and lexical bundles are extracted from the corpus and rely on distributional information, two different distributional analyses are used. For frequency-based collocations, various co-occurrence analyses are used to test whether two words co-occur statistically more often than two words appearing together randomly (Biber et al., 1999). For example, mutual information (MI) analysis measures the extent to which words co-occur relative to the number of times these words appear separately (see Chapter 5 for more discussion on MI). When the MI score is larger than 3, the word combination is considered a collocation (Hunston, 2022). For lexical bundles, the identification is based on the frequency of occurrence alone, without the specific requirement to detect recurrent continuous word sequences.

One feature of lexical bundles is that lexical bundles are not necessarily semantically and structurally complete, such as *you want me to*, *as a result of*, compared to idioms and collocations. However, the structure of lexical bundles still places certain constraints on the incoming words, for example, there has to be a verb after *you want me to* and there has to be a noun-based structure after *as a result of*. Another feature is that lexical bundles are compositional (i.e., lexical bundles do not have

idiomatic meanings). In other words, unlike idioms, the meaning of a lexical bundle can be derived by adding up the meaning of each individual lexical part.

Previous studies have shown that lexical bundles are processed faster than novel expressions. A self-paced reading study (Tremblay et al., 2011) compared four- and five-word lexical bundles (e.g., *I sat in the middle of the bullet rain*) to non-lexical bundles (e.g., *I sat in the front of the bullet rain*). The same experimental sentences appeared in three different ways: word-by-word, chunk-by-chunk, and as part of the whole sentence. In the word-by-word presentation, participants read a lexical bundle one word at the time. In chunk-by-chunk presentation, a whole lexical bundle appeared together. In a sentence presentation, the whole sentence appeared at the center of the screen. The three self-paced reading experiments with different presentation methods showed that lexical bundles were read faster than non-lexical bundles, likely due to benefitting from stronger memory traces. However, the single word frequency and the length of the whole expression were not matched between non-lexical and lexical bundles.

A study using the plausibility task addressed the limitations above and observed a sensitivity to the frequency of four-word sequences (Arnon & Snider, 2010). They tested four-word sequences with different chunk frequencies and matched for the frequency and length of the lexical parts by comparing word sequences that differed only on the last word (e.g., *don't have to worry* vs. *don't have to wait*; *I don't know why* vs. *I don't know who*). Plausibility was also controlled between four-word sequences with high and ones with low frequencies. The results showed that the frequency of compositional four-word sequences affects processing. Similar to the processing of single words, higher frequency sped up the processing of four-word sequences or lexical bundles. Additionally, the speeding-up effect was captured better when frequency was treated as a continuous variable rather than being operationalised as a categorical variable with two levels (i.e., high frequency vs. low frequency).

2.4 Characteristics of MWSs

In the previous section, the definitions, features, and important studies of several types of MWSs have been reviewed. The different types of MWSs vary in different aspects, such as the degree of compositionality, conventionality, and idiomaticity. Despite these differences, the types of MWSs share similarities, such as a high degree of frequency, familiarity, and predictability. In this section, several key characteristics of MWSs are reviewed.

2.4.1 Corpus Frequency and Familiarity

As reviewed in the previous section 2.2, MWSs are widespread in language (e.g., Altenberg, 1998; Biber et al., 1999; Erman & Warren, 2000; Foster, 2001; Glucksberg, 1989; Pawley & Syder, 1983). In addition, MWSs are recurrent in language. Take one example: *in the middle of*. This lexical bundle appears 2807 times in BNC. Although MWSs appear more frequently than novel word combinations, the highly frequent MWSs tend to be much lower in frequency than highly frequent single words. Take another example: *provide information*. This collocation appears 339 times in BNC, which is considered to be a highly frequent word combination. In contrast, *information* is considered a frequent noun, which appears 37862 times in BNC. The difference in the high-frequency threshold between single words and MWSs is simply because the likelihood of co-occurring words is much lower.

There is a substantial amount of evidence that the frequency of co-occurrence can modulate the access and processing of MWSs. Akin to single words, high co-occurrence frequency speeds up processing of MWSs, including binomial expressions (Siyanova-Chanturia, Conklin, & van Heuven, 2011), collocations (Li et al., 2021, 2022; Öksüz et al., 2021; Sonbul, 2015; Vilkaitė, 2016; Vilkaitė & Schmitt, 2019; Vilkaitė-Lozdienė, 2022), idioms (Conklin & Schmitt, 2008; Gibbs, 1980; Gibbs et al., 1989; M. R. Libben & Titone, 2008; Tabossi et al., 2009), and lexical bundles (e.g., Arnon & Snider, 2010; Tremblay et

al., 2011). Additionally, the co-occurrence frequency effect is robust and emerges consistently, regardless of treating the frequency as a continuous or categorical predictor.

Language users have gained familiarity with MWSs due to the high frequency of MWSs. However, the corpus frequency is not identical to the degree of familiarity. Prior research with single words has shown that the frequency of single words is only an approximation of the familiarity ratings of single words (Gernsbacher, 1984; Kuperman & Van Dyke, 2013), though the average familiarity rating scores have a strong positive correlation (e.g., $r = 0.78$) with frequencies (Balota et al., 2001; Zeno et al., 1995). Frequency approximation tends to be less accurate for single words in the lower frequency range. There is greater variability in the lower range of frequency within and between corpora and subjects. Additionally, the variability is negatively correlated with the lower frequency range (Kuperman & Van Dyke, 2013). It is therefore reasonable to assume that MWSs with lower frequencies would also involve more variability within individual language users. This probably explains why language users can still be familiar with an individual idiom, despite it not being frequent in a corpus (Moon, 1998).

2.4.2 Predictability

A word can be predicted from its prior context, which is defined as the word's contextual predictability (Ehrlich & Rayner, 1981). Contextual predictability is often measured by a word's cloze probability (Taylor, 1953). Participants are asked to fill in the next word given the prior context. The proportion of participants providing that word is taken as contextual predictability for the word. Hence, the contextual predictability of a word is often taken as the contextual constraint on the word. When a word is highly predictable from its prior context, the likelihood of the word following the preceding context is high. For example, *umbrella* is highly predictable from the context *It starts to rain heavily, she opened her* __; it is less likely to predict *umbrella* from the context *She is walking down the street without an* __. Similar to lexical frequency, the predictability of a word also has an important influence

on how a word is recognised. When a word is highly predictable from the prior context, it is more likely to be spent less time and to be skipped; namely, the word is easier to be processed. These effects have been confirmed in a large amount of research (Balota et al., 1985; Brysbaert & Vitu, 1998; Ehrlich & Rayner, 1981; Kutas et al., 1984; Rayner et al., 2011; Rayner & Well, 1996).

For MWSs such as *a piece of cake*, *grey area, in the middle of*, the words up to the final word can provide certain constraints on the final word. The predictability of MWSs is the probability of completing an incomplete MWSs (for idioms, completing an idiom idiomatically; For example, participants are asked to complete the sentence fragment *This week I lost my job, my dog died, and my friend is in hospital. It never rains but it* __. The probability of completing the incomplete sentence with *pours* is the idiom's predictability.). In general, the longer the MWS, the fewer the available alternatives to continue in a syntactically well-formed and semantically meaningful way. Hence, there is a positive correlation between the length of MWSs and the predictability of MWSs (Fanari et al., 2010; Tabossi et al., 2011). Longer MWSs tend to have higher predictability.

2.4.3 Semantic Compositionality

Dating back to Frege's principle, compositionality is one of the central principles in language. Semantic compositionality is defined as the meaning of a sequence of words derived from the meanings of words in the sequence (Katz & Giesbrecht, 2006; Lin, 1999; Pagin & Westerståhl, 2010; Pelletier, 1994). For instance, the meaning of the sequence of words *a stressed research student* is obtained by assembling the meanings of *a*, *stressed*, *research*, and *student.*

MWSs display a continuum of compositionality (McCarthy et al., 2003). Some MWSs are compositional, such as lexical bundles (*in the middle of*, *at the same time*) and certain collocations (*provide information, promote development, fixed price, strong belief*). In contrast, some MWSs are non-compositional, such as idioms (*pull strings*, *kick the bucket*) and binominal expressions (*by and large*,

bread and butter – in its figurative *money* sense). There are MWSs that are in between compositional and non-compositional, such as certain other collocations (*make a mistake*, *take a photo*).

Idioms have been seen as a typical example of non-compositionality (Chomsky, 1980; Fraser, 1970) because the overall idiomatic interpretation of an idiom often cannot be derived from its lexical parts. However, even with idioms, the lexical parts can still contribute to the idiomatic meaning, which is referred to as the decomposability of idioms (Gibbs et al., 1989; Nunberg, 1978; Nunberg et al., 1994). For example, the meaning of lexical parts (*pop, question*) is related to the idiomatic meaning i.e., "propose marriage." Therefore, *pop the question* is seen as decomposable. In contrast, the lexical parts of *to kick the bucket* are not relevant to the idiomatic meaning "to die"; thus, *to kick the bucket* is considered to be non-decomposable. Compositionality was introduced to idiom processing and is reflected in the Idiom Decomposition Hypothesis (e.g., Gibbs et al., 1989). In Chapter 4, the Idiom Decomposition Hypothesis will be reviewed, along with other idiom processing models.

2.4.4 Fixedness and Variation

Contrary to free word combinations, the structure and meaning of MWSs can be restricted or even fixed to some degree. MWSs are thought to occur on a continuum of fixedness and restrictedness, in which different types of MWSs vary. There are also variations within the same MWS, with corpus studies showing that 40% of MWSs have variations in their expression (Moon, 1998). As discussed in the section 2.3, idioms are characterised by their fixedness in form and meaning and allow very limited degrees of modifications and substitution of their lexical parts (e.g., Sinclair, 1991). On the contrary, many MWSs can tolerate a certain degree of modification. To illustrate, a word or a few words can be inserted into a collocation or a lexical bundle (e.g., *provide information – provide <u>a lot of</u> information*; *in the hands of – in the <u>firm</u> hands of*). According to Moon (1998) and Schmitt (2006), there are five

common ways of modifying MWSs: lexical variation, grammatical variation, transposition, omission and amplification. These variations in MWSs are demonstrated in the following paragraphs.

Lexical variation is a type of modification by replacing components of MWSs. Alternation barely affects the conventionalised meaning of MWSs. For example, *tough/hard nut*, *bad/rotten apples*, *bend/stretch the rules*, *stand/fight shoulder to shoulder*.

Grammatical variation happens when there is a change in the grammatical functions of the lexical parts of MWSs (more likely applied to nouns and verbs). Thus, this type of modification does not change the meanings of MWSs and is only reflected in the forms of MWSs. This inflection modification can involve a change of tense of a component verb (e.g., *provide information -> provided information*; *write a letter* -> has *written/wrote a letter*), or pluralization (e.g., *plant a seed -> plant seeds*; *an ideal solution -> ideal solutions*).

When the order of the lexical parts of MWSs is reversed, such a modification is referred to as transposition. For example, the reversed word order can be achieved by passivisation (e.g., *provided information -> information was provided*). The reversing word order in collocation processing is investigated in Chapter 5 of this thesis. Changing the word order can sometimes alter the meaning of the MWSs. For example, the binomial expression *bread and butter* might lose the conventionalised meaning i.e., *the thing that gives you your main income* after changing the expression to *butter and bread*.

Omission modification involves simplifying and shortening traditional sayings, proverbs, and idioms. Importantly, the omission does not affect the completeness and preciseness of the meaning of the MWS. Put simply, the meaning of a reduced unit is equivalent to the meaning of the original one. For example, *birds of a feather flock together -> birds of a feather.*

Amplification and omission are two sides of the same coin. Amplification modification extends the form and interpretation of an MWS in order to add emphasis or precision. Some MWSs allow a word or a few words to be added, though where the amplification occurs could vary. As already listed at the beginning of this section, words can be inserted between components of an MWS, such as *provide information -> provide lots of information*. The amplification can happen at the end of an MWS as well, such as *a tough nut -> a tough nut to crack*.

2.5 Questions of Holistic Storage

One of the questions in MWSs processing is whether MWSs are stored, represented, and retrieved as a holistic unit. The assumption of holistic storage and retrieval most likely comes from the faster processing of MWSs. There is mounting evidence that language users respond to, judge, and read MWSs more quickly than matched novel word combinations (Arnon & Snider, 2010; Hernández et al., 2016; M.R. Libben & Titone, 2008; Öksüz et al., 2021; Siyanova-Chanturia, Conklin, & van Heuven, 2011; Sonbul, 2015; Tabossi et al., 2009; Vilkaitė, 2016; Wolter & Yamashita, 2018). Therefore, the holistic storage has been relied upon to explain the processing advantage of MWSs in several studies (e.g., Altenberg, 1998; Jiang & Nekrasova, 2007; Underwood et al., 2004).

If the hypothesis that MWSs are stored and retrieved as one unit is correct, then the representation, storage, and processing of one MWSs should be independent of their individual lexical parts. For example, the presentation and storage of MWSs such as *in the middle of* (lexical bundles), *black and white* (binomial expressions), *provide information* (frequency-method defined collocation), and *kick the bucket* (idioms) are independent of their lexical parts *middle, black/white, provide/information, kick/bucket*. In other words, these lexical parts should not be activated when processing *provide information* and *kick the bucket*. The next two paragraphs consider the paradigms and findings from a few studies that claim MWSs are stored and represented holistically (Jiang & Nekrasova, 2007; Kim & Kim, 2012; Underwood et al., 2004).

Underwood et al. (2004) conducted an eye-tracking experiment and compared reading a word in an idiomatic context to the same word in a literal context (e.g., *met the deadline by the skin of his teeth vs. the dentist looked at his teeth, t*he target word is underlined). The eye-tracking study observed that L1 and highly proficient L2 speakers fixated less often on the final word in the idiomatic context and read the final word in the idiomatic context more quickly relative to the same final word in the literal context. Therefore, Underwood et al. (2004) concluded that their findings provided evidence for formulaic sequences being stored and processed holistically. The study has limitations, such as unmatched predictability and varied syntactic structure, which might have affected their results, which will not be further discussed.

Furthermore, a grammaticality judgement study compared recurrent lexical bundles (e.g., *as soon as*, *out of work*, *to tell the truth*) with non-lexical bundles (e.g., *as mean as*, *out of state*, and to *tell the price*) and ungrammatical sequences (e.g., *when of at*, *bus on go*, *of like much a*) (Jiang & Nekrasova, 2007). The finding was that participants judged frequent lexical bundles more quickly and with fewer errors than non-lexical bundles. Therefore, Jiang and Nekrasova concluded that lexical bundles are recognised holistically. In a self-paced reading task, Kim and Kim (2012) investigated the role of phrasal frequency in phrasal verbs (e.g., *find out, sort out*). The shorter reading times of more frequent phrasal verbs led the authors to claim that language users processed those phrasal verbs without linguistic computation.

When looking closely at the design of the experiments discussed above, they share a common feature: comparing one type of MWS to novel matched word combinations or less frequent MWSs using a reading or judgment paradigm. The findings of faster processing of MWSs can be taken as evidence that language users store frequency information beyond single words. Importantly, the role of the lexical parts of an MWS compared to the whole MWS is the determining factor in how likely the MWS is to be processed holistically. Therefore, to assess whether the lexical parts of MWSs are still activated

and contribute to the processing of the MWS, it is necessary to investigate the frequency of their lexical parts. However, none of the studies reviewed above assessed the frequency of the lexical parts. Hence, the lexical parts could still have affected processing time, which would make it difficult to accommodate a holistic storage account.

More recent work on collocations supports parallel and interactive influences of the word-level and phrase-level frequencies and has shown that the frequency of the lexical parts also speeds up the processing of MWSs (Öksüz et al., 2021; Wolter & Yamashita, 2018). This finding provides evidence against the holistic view. In one collocation study, Öksüz et al. (2021), using an acceptability judgement task, examined how word-level frequency and phrase-level frequency affect the processing of adjective-noun collocations (e.g., *strong coffee*, *senior officer*, *long time*). Although a higher phrase-level frequency led to a reduced effect of word-level frequency, participants were still sensitive to the word-level frequency, with higher word-level frequency boosting the processing of collocations. If the hypothesis of holistic storage and representation were true, one would not expect to see word-level frequency contributing to the processing of collocations.

More evidence against the holistic view comes from studies of idioms. Idioms, which are considered the least decomposable MWSs, demonstrate that the individual lexical parts are activated as well (Cacciari & Tabossi, 1988; Cutting & Bock, 1997). For example, in a cross-modal priming paradigm, participants made a lexical decision to target words with three conditions: figuratively related to the idiom, literally related to the idiom, and control words (Cacciari & Tabossi, 1988). After listening to sentences containing idioms (e.g., *the actor asked the producer to put in a good word for him with the director*), participants were immediately presented with target words. For predictable idioms (i.e., expressions were more likely to be completed idiomatically, e.g., *put in a good word*), participants responded to target words that were figuratively related to idioms (e.g., *favour*) more quickly than control words (e.g., *glove*) and target words that literally related to the final word of the idiom (e.g.,

language). Targets literally related to the final word of idioms did not differ from controls. This suggests that predictable idioms were immediately processed idiomatically. In contrast, when idioms were less predictable (e.g., *go to the devil*) from the preceding context, target words literally related to the idiom (Cacciari & Tabossi, 1988) were judged more quickly than control words and target words figuratively related to the idioms. There was no difference between control words and target words figuratively related to the idioms. This suggests that with unpredictable idioms and without biased contextual cues, idiomatic strings were immediately processed literally. Additionally, target words, both figuratively and literally related to the unpredictable idioms, were responded to more quickly than control words when target words were presented 300 ms after participants heard the complete idioms. This suggests that a 300 ms delay allowed people to activate the idiomatic interpretation, even when the preceding context is semantically neutral. Taken together, this cross-modal priming study supports the view that the literal meaning of the lexical parts of an idiom is activated and provides evidence against the holistic view. If idioms were processed holistically, one would not expect to see semantic priming of the literal target words.

To sum up, language users retrieve and process familiar MWSs more quickly than novel word combinations, suggesting that the frequency effect can be extended to larger linguistic constructions. However, this does not deny that language users access the lexical parts of the MWSs. Even if language users tend to treat MWSs as one unit, language users are still sensitive to the frequency of the lexical parts. Hence, compared to a whole MWS, the lexical parts of the MWS play a larger role in determining whether the MWS is to be processed holistically (Blumenthal-Dramé, 2018). Access to the whole meaning can be achieved by the bottom-up route (i.e., the analysis and integration of lexical parts) and the top-down route (i.e., the direct access to the meaning of an MWS) interactively. These two routes operate on a continuum scale. When language users encounter one MWS many times, the top-down route takes precedence over the bottom-up route. When language users are less familiar with one

MWSs, the bottom-up route is more likely to take priority. Thus, instead of the holistic access and storage alone, the two routes co-exist and compete, with one route winning out.

2.6 Processing Morphologically Complex Words

Research in processing morphologically complex words could shed light on how MWSs are processed and stored in the lexicon. The full-listing hypothesis assumes that language users list all derivations of a word in the lexicon (e.g., *write, writes, wrote, written, writing*) and directly access complex words (Butterworth, 1983). On the contrary, the obligatory decomposition hypothesis proposes that language users store the stem and general rules (cf. Taft, 2004). For regular complex words, language users apply these rules when processing or decoding inflected words (e.g., adding *ed* to form the past tense). For irregular words, language users store all irregular forms in their lexicon (e.g., *ate*, *came*, *people*).

According to the obligatory decomposition hypothesis, when recognising complex words, affixes are stripped off. A series of studies provided evidence for the obligatory decomposition. In a lexical decision task, participants spent more time recognising words that have a fake prefix (e.g., *result, devout*) than prefixed words (e.g*., replica*, *delete*) (Taft, 1979, 1981). The cost in recognising those words came from the fact that language processors stripped off the pseudo prefix (e.g., *re* from *resul*t and *de* from *devout*) first and failed to match the resulting parts (e.g., *sult*, *vout*) and existing stems in the mental lexicon. In addition, words that are similar in their surface frequency (e.g., *seeming* vs. *mending*) but differ in their base frequency (i.e., *seem* is more frequent than *mend*) showed base frequency effects, with participants responding more quickly and more accurately to words with high base frequency (e.g., *seem*) than low base frequency (e.g., *mend*) (Taft, 2004). The base frequency effect indicates that complex words are decomposed during recognition.

Another hypothesis related to complex word recognition is the dual-pathway hypothesis (Sandra, 1990), which lies between the full-listing and the obligatory decomposition hypothesis. While regular words can be decomposed according to the general rule, highly frequent regular words are also listed in the lexicon, allowing for direct lexical access. The evidence for the dual-pathway hypothesis came from the absence of a base frequency effect in highly frequent words (Baayen et al., 1997; Bertram et al., 2000).

Ample evidence has shown that lexical processing of complex words is sensitive to the characteristics of both the complex word as a whole (e.g., *government*) and their morphological parts (e.g., *govern-ment*) (Holmes & O'regan, 1992; Niswander et al., 2000; Schreuder et al., 2003). For example, in an eye-tracking study by Niswander et al. (2000) with derived words (i.e., a word formed by a stem/root combined with a prefix or suffix, e.g., *familiarity*, *unfamiliar*, *appointment*, *traditional*), root morpheme frequency affected first fixation processing times while whole word frequency affected second fixation processing times. Hence, the frequency of the root affected the lexical access earlier than the frequency of the whole word. For inflected words (i.e., a word that is modified to show grammatical categories, such as number, tense, and voice, e.g., *flowers*, *watering*, *pushed*), word frequency affected processing on the first fixation, while root frequency effects were only found for inflected nouns.

Compound words (e.g*., teacup*, *buttercup*) are a special category of complex words and demonstrate the effects of both lexical access and morphological computation. On the one hand, it has been shown that the frequency of the lexical parts (e.g., *tea-cup*; *butter-cup*) affects the encoding of compound words. This is taken as evidence for the decomposition of the morphological parts (Bertram & Hyönä, 2003; Hyönä & Pollatsek, 1998; Juhasz et al., 2005; Pollatsek et al., 2000). On the other hand, the frequency of the whole compound word has also been shown to impact encoding (e.g., Bertram & Hyönä, 2003; Pollatsek et al., 2000).

If compounds are decomposed into their morphemes during lexical access, it is expected to observe the role of semantic transparency in the processing and representation of compounds. Semantic transparency of compound words is defined as the degree to which the meanings of the morphemes relate to the meaning of the compound (Pollatsek & Hyönä, 2005). The meanings of both morphemes are related to the meaning of the semantic transparent compound, e.g*., teacup and gingerbread*. In contrast, the meanings of both morphemes are not related to the meaning of semantically opaque compounds, e.g., *cocktail*, *deadline.* Compounds can be transparent, partially transparent/opaque, or opaque. For partially transparent compounds, the meaning of the first part of the compound may be related to the meaning of the compound (e.g., *heirloom*, *chatterbox*) or the meaning of the second part of the compound may be related to the meaning of the compound (e.g., *butterfly*, *dragonfly).*

While one would expect that semantically transparent compounds would be the easiest to process, the findings related to semantic transparency are mixed. Priming studies in which one component of the compound was either presented as a prime or a target have shown priming effects regardless of transparency (Jarema et al., 1999; G. Libben et al., 2003). This indicates that both transparent and opaque compounds undergo decomposition. However, there is also evidence that opaque compounds do not undergo decomposition. For example, Pollatsek and Hyönä (2005), using eye-tracking, did not observe a semantic transparency effect with fully and partially opaque compounds in Finnish. In addition, another eye-tracking study did not find evidence of a transparency effect with fully and partially opaque compounds when the compounds were presented unspaced (Frisson et al., 2008). However, the transparency effect emerged when a space existed between the components of the compound (e.g., *tea cup*; *dragon fly*), indicating that transparency played a role in encoding when participants were forced to assemble the components. The inconsistent findings of a semantic transparency effect between eye-tracking (e.g., Frisson et al., 2008; Pollatsek & Hyönä, 2005) and

priming (Jarema et al., 1999; G. Libben et al., 2003) studies could be attributed to the different stages of decomposition. The priming studies demonstrated the morpho-orthographic level decomposition, while the eye-tracking reading experiments tapped into morpho-semantic level decomposition that did not emerge.

When it comes to the temporal order of activation of compound words, there are several possible architectures. The sublexical model hypothesises that it is obligatory to undergo decomposition for compound words and that lexical access proceeds via morpheme (Taft, 1991). Therefore, under this model, initially, the first component of the compound is activated, then the second component is activated, and finally, the whole compound is activated. Take *teacup* as an example. First, *tea* is activated, followed by *cup*, resulting in the recognition of *teacup*. The supralexical model, by contrast, holds that the morpheme access happens only after the recognition of the compound word (Giraudo & Grainger, 2001). According to the supralexical model, the initial step is the activation of the full form of the compound, followed by the simultaneous activation of both components. In this case, the recognition of *teacup* happens initially, then *tea* and *cup* are activated at the same time.

An eye-tracking study has shown that multiple routes are taken to encode compounds and language users use these routes interactively at different times (Kuperman et al., 2009), who found an early effect of whole compound frequency on the first fixation measure, and the frequency effect lingered on through the entire compound encoding. This indicates that the activation of the whole compound did not happen until after having accessed the components of the compounds. Additionally, there was an interaction between the compound frequency and the frequency of the first word of the compound. This shows that the decomposition and the full-form route are not independent, and the morphological processing and the activation of the whole compound are not isolated from each other.

A fundamental underlying assumption behind the decoding of morphologically complex words is that extra effort at processing comes with the decomposition-based route (i.e., accessing the morphological parts). In contrast, the direct access route (i.e., accessing the whole word) is assumed to be faster and more economical since the computational load is reduced. Therefore, the direct route is taken to be the preferable route (Bertram & Hyönä, 2003). However, studies have consistently shown that the computation route is often taken in the recognition and decoding of complex words (Andrews et al., 2004; Pollatsek & Hyönä, 2005), even though this route is associated with longer processing and increased errors. Multiple routes available to decode complex words have shed light on the processing of MWSs. Similarly, MWSs could be accessed via multiple routes simultaneously and different routes can compete, with the more efficient route winning out.

2.7 Lexical Quality Hypothesis

The Lexical Quality Hypothesis (LQH) (Perfetti & Hart, 2002) and the Reading Systems Framework (Perfetti & Stafura, 2014) are leading theories which aim to explain individual differences in reading. According to the LQH, words represented and stored in the mental lexicon vary in detail and quality. High-quality representations of words are precise in orthographic and phonological representation, which are connected to each other and are connected to the higher level of semantic knowledge. High-quality representations of words produce synchronous activation between the different levels of representation of the word. For example, phonology information is stored as part of the lexical representation and will be retrieved during the identification of the word. With the increased use of words, the phonological and orthographic information is getting connected, and the connection is thought to become stronger over time.

In other words, a high-quality lexical representation includes sub-lexical features, such as letters and sounds related to the formation of a word, and semantic information of a word, such as multiple meanings of a word and related meanings under a semantic network. Additionally, grammatical

information (e.g., grammatical classes of the word represented; morpho-syntactic inflections represented) is specified in the high-quality representation as well. Therefore, the high-quality representation of a word can receive activation from different sources of representation, enabling quick and easy access to the word. In contrast, any information missing in the high-quality representation of a word can lead to low-quality representations of the word, and thus leading to difficulties in decoding the word. Moreover, individual differences in reading arise from differences in the individual quality of lexical presentation. The LQH posits that the mental lexicon of more skilled readers contains more high-quality representations of words than the lexicon of less skilled readers.

Previous research shows that reading comprehension in L2 speakers is generally lower than in L1 speakers (Melby-Lervåg & Lervåg, 2014). The comprehension problems in L2 speakers have been found to be linked to low and poor L2 vocabulary (Burgoyne et al., 2009). It is sensible to assume that L2 speakers have lower quality of lexical representation of words, compared to L1 speakers, due to the lack of exposure to the L2 language. For example, the lexical representation in L2 speakers might miss the information about the inflection of *write* (i.e., *written*), which slows down word identification. However, LQH is under-investigated for L2 reading comprehension. It should be noted that the LQH is focused on the representation of single words and has not addressed MWSs. However, it is theoretically possible that a high-quality representation includes multiple words statistically associated with a word.

2.8 Usage-based Models and Theories of Language

In this section, usage-based models of language are covered. Usage-based models and theories of language were first introduced by Langacker (1987), and there are many influential proponents of usage-based models. Under this framework, the processing and acquisition of language are shaped by actual usage and are the result of the interaction between cognition and use (Bybee, 1998, 2006; Ellis, 2008; Langacker, 1987; Tomasello, 2009).

2.8.1 The Link between Usage Events and the Linguistic System

In general, usage-based models share several characteristics. One of their characteristics is that the linguistic system is tightly linked to the usage events at three different levels. First, these usage events (i.e., instances of a language user's producing and understanding language to convey meaning in a specific situation) are the foundations of language learning and forming the linguistic system and structure. Second, the abstract representations in language users' grammar are influenced by the usage events. Linguistic constructions (e.g., morphological and syntactic patterns) are gradually built up by abstracting representations from repeatedly specific usage events. Repetition alone does not necessarily lead to abstract representations, though the repeated encounter of a usage event leaves traces and gets entrenched in memory. To form general representations, our mind detects the shared features and non-recurrent patterns from those usage events. Our brain filters out the non-recurrent patterns and then forms analogies between these shared features. Eventually, a general and abstract representation is formed (Tomasello, 2003). Third, the structuring and operation of the linguistic system are continuously affected by usage events. From this point of view, language production is not only the product of a linguistic system but is also an input for other language users. Therefore, for this ongoing procedure, usage events are the products of the linguistic system, and they also continue to shape the linguistic system.

2.8.1.1 The Importance of Frequency

The second characteristic of usage-based models is the important role that frequency plays in language processing and learning. For example, the more often a language learner experiences a certain linguistic construction will affect how strongly that construction will become entrenched and, therefore, how accessible the construction will be. The frequency of a specific language pattern is the product of the linguistic system, which in turn shapes the system.

2.8.1.2 Construction

Usage-based theories go hand in hand with constructions (i.e., constructions are mapping between particular structures with particular meanings or functions (Diessel, 2017)). Learning a language involves learning these constructions that comprise multiple elements for the purpose of communicative intentions. A construction can have varying degrees of complexity depending on the number of linguistic elements (Tomasello, 2009). Morphology, word classes, lexical items, idioms, syntactic structures, and pairing of forms are all considered types of constructions (Langacker, 1987). For example, a simple construction can be the English regular plural (e.g., *noun+s*), and a more complicated construction can be the English passive structure or an idiom (e.g., *xx is verb+ed by, kick the bucket*).

The usage-based way of looking at language acquisition stands in contrast to the view that children learn individual words as the first stage and learn words in an isolated way, then assemble individual different words together. Instead, what children learn is not limited to single words (Arnon et al., 2017; Bannard & Matthews, 2008; Lieven et al., 1992) and learning a language is the learning of constructions and pairing of forms, meaning and functions that are extracted from usage events. Therefore, children begin their production and comprehension with different sizes of constructions.

2.8.2 Empirical Evidence

Though there is ongoing debate, mounting evidence exists to support usage-based theories. Studies have shown that language users are sensitive to the distributional information at different language levels, such as single words (Forster & Chambers, 1973; Grainger, 1990; Rayner & Raney, 1996), and phonemes (Bell et al., 2009; Gregory et al., 1999). This means that language users can extract linguistic patters based on the repeated exposure to language. Hence, the linguistic system is shaped by usage events and experience.

Furthermore, the sensitivity to the distributional information extends to larger linguistic constructions. The evidence comes from research in the processing of MWSs such as binomial expressions (Conklin & Carrol, 2021; Siyanova-Chanturia, Conklin, & van Heuven, 2011), collocations (Durrant & Doherty, 2010; Gyllstad & Wolter, 2016; Öksüz et al., 2021; Sonbul, 2015; Vilkaitė, 2016), and lexical bundles (Arnon & Snider, 2010; Christiansen & Arnon, 2017; Hernández et al., 2016; Jolsvai et al., 2020; Tremblay & Baayen, 2010). This supports the view of usage-based language models that hold that construction is the building blocks of language. All linguistic constructions no matter their size are thought to be similarly affected by usage events.

The extension of the distributional information to larger linguistic constructions stands in contrast to a words-and-rules language theory (Pinker, 1991, 2015; Prince & Pinker, 1988; Pinker & Ullman, 2002). According to the words-and-rules theory, a clear distinction is made between the mental lexicon and mental grammar. The mental lexicon consists of a list of memorised forms, while the mental grammar contains rules and constraints used to assemble the content from the mental lexicon. Under this view, larger linguistic constructions are computed based on rules. Therefore, it is expected that single words, but not larger constructions, are shaped by experience that is indexed by frequency.

2.9 Statistical Learning Theory

In this section, the statistical learning theory is discussed from a few perspectives. The origin and definitions of the statistical learning theory are introduced. Then, important studies supporting this theory are discussed. Following this, the idea of statistical learning is extended from adjacent language patterns to non-adjacent ones.

2.9.1 Definition

We need to understand the nature of the input before we can understand the nature of processing (Gibson et al., 1996), which reflects the core of the statistical learning theory. This theory is

not specific to language, on the contrary, it originates from machine learning and is prominent in cognitive and developmental science (Saffran & Kirkham, 2018).

In general, statistical learning can be seen as humans being able to extract statistical regularities from the patterned input. To be more specific, learning is based on finding and generalising patterns from the input (Onnis, 2012). In the field of first language acquisition, statistical learning has received lots of attention because of the rapidity at which children acquire language. In general, statistical learning in language acquisition can be understood as the ability of language users to make use of information around them to facilitate language learning (Rebuschat & Williams, 2012).

On the one hand, narrowly speaking, statistical learning in the area of language acquisition tends to be closely associated with tracking sequential information and statistics. Transitional probability is one of the typical statistics and is defined as a conditional probability of Y given X in the sequence of XY. Initially, transitional probability is calculated over segments of syllables, and words of speech.

2.9.2 Empirical Evidence for Statistical Learning

Many studies have shown that transitional probabilities can be tracked by infants and adult language users. For example, one of the first studies investigated whether transitional probabilities could be used by language users in identifying words in continuous speech (Saffran, Newport, et al., 1996). The transitional probability from one sound to the next sound is higher when the two sounds follow each other within a word than when one sound is positioned at the end of a word and the other sound is positioned at the start of another word (e.g., pretty baby, the transitional probability from [p] [r] [i] to [t] [i] is much higher than [t] [i] to [b] [a]). In the study, young adult participants were exposed to continuous speech, which was concatenated by using six trisyllabic artificial words (e.g., *dutaba*, *pidabu*) repeatedly. The continuous speech did not have word boundaries and only included transitional

probabilities as the segmentation cue. After listening to the continuous speech, participants were asked to indicate which of the two strings sounded more like a word. One string was from one of the six artificial words, and the other was from a non-word comprising the same set of syllables but was never presented adjacently. Therefore, the transitional probabilities of the non-words were zero. In contrast to non-words, the transitional probabilities within the artificial words were high. Participants judged strings from the six artificial words as more word-like than non-words, indicating that language users could use the transitional probability to segment sounds and discover word boundaries.

On the other hand, broadly speaking, the statistical information can be extended to the frequency of individual linguistic units, co-occurring frequency of larger linguistic units, and mutual information (i.e., mutual information calculates the number of times when words co-occur against the number of times when words appear separately, the equation can be found from this link[2]). Learners can potentially make use of this statistical information regarding language patterns and structure to help discover word categories, grammar, and word meaning.

Though initially statistical learning was mainly focused on "lower" levels of language structure (i.e., Saffran, Aslin, et al., 1996; Saffran, Newport, et al., 1996), statistical learning has been expanded to multiple and higher levels of linguistic structure, such as words and grammar. For example, Saffran and Wilson (2003) tested whether infants could make use of multi-level statistical cues. Infants were exposed to one of two artificial languages. Each language consisted of 16 sentences, and each sentence consisted of five disyllabic artificial words. Language A and language B were ordered according to a finite-state grammar, but the two languages were different in how the five disyllabic artificial words were ordered. One participant was only exposed to one of the languages at the familiarisation stage, thus the other language became ungrammatical for the participant at the test stage. Additionally, the

[2] http://www.collocations.de/AM/index.html
https://www.english-corpora.org/mutualInformation.asp

transitional probabilities were identical between syllables for grammatical and non-grammatical sentences, which excluded the transitional probabilities as a cue to distinguish (un)grammatical structures. The study observed that infants listened to grammatical sentences longer than ungrammatical ones at the test stage, which was taken as evidence that infants learned the grammatical patterns characterised in the sentences. The results suggest that infants are able to use syllable-related statistics initially to segment sounds from continuous speech and then to track the order of words in sentences.

2.9.3 Learning Non-adjacent Patterns

Regardless of age or bilingual status, there is a clear picture of the ability of language users to track different types and levels of distributional regularities in language. However, all the findings described above are based on adjacent language, i.e., linguistic units occur without intervening units. Natural language is more complex and includes non-adjacent patterns as well. Non-adjacent patterns are defined as units that co-occur over one or more intervening units and are known as non-adjacent dependencies. To illustrate, collocations (e.g., *provide information*) can occur non-adjacently (*provide some information*, *information was provided*) and morphosyntactic agreement between verb tense and auxiliary in SOV languages can be non-adjacent as well (e.g., *the cat always does that*).

Much research has provided evidence that infants have the ability to detect nonadjacent regularity in the phonological domain (Gonzalez-Gomez & Nazzi, 2012; Nazzi et al., 2009) and with morphosyntactic dependencies (Santelmann & Jusczyk, 1998). Santelmann and Jusczyk (1998) conducted a study to test whether infants were sensitive to the morpheme *is* and the inflection *-ing*. Infants were able to discriminate between grammatical and ungrammatical sentences (e.g., *grandma is singing* vs. *grandma can singing*) when the morphemes and verbs appeared close together, up to a maximum of three syllables (e.g., *grandma is singing* or *grandma is quite often singing*). However,

infants failed to cue the presence or absence of –*ing* inflections when the morpheme *is* and the inflection were separated by four syllables (e.g., *someone is/*can quite skilfully adding salt*). Similarly, in a series of studies, Gonzalez-Gomez and Nazzi (2012) investigated French infants tracking non-adjacent consonant patterns within words. In French, labial consonants (e.g., [p] and [b]) preceding coronal consonants (e.g., [d] and [t]) are more frequent than coronal consonants preceding labial consonants. Additionally, there is often a vowel between the labial consonants and the coronal consonants (e.g., [pat], [tap]). Their results revealed that infants showed a preference for the labial-vowel-coronal sounds (e.g., [pat]) than coronal-vowel-labial sounds (e.g., [tap]), as the labial-vowel-coronal sounds are more frequent. This finding indicates that infants were able to track the non-adjacent phonological regularities.

Studies have also shown that adults can have difficulties detecting non-adjacent regularities but are selective in the types of non-adjacent patterns that they detect (Newport & Aslin, 2004; Onnis et al., 2005). One study (Newport & Aslin, 2004) consisted of 3-syllable artificial words. The statistical relation between two non-adjacent syllables was manipulated and an unrelated syllable was inserted in-between. For example, the first and third syllables of the trisyllabic non-word [ba] [di] [te] had a high probability of appearing together (non-adjacently) in a string, but the second syllable did not. Adult learners were poor at acquiring statistical information among non-adjacent syllables. In contrast, a follow-up experiment from the same study showed that adult participants were able to learn the statistical regularities when the non-adjacent patterns were less complicated and came from the same category (i.e., the transitional probability information appeared among consonants (e.g., [d]_[k]_[b]) and the intervened units were vowels (e.g., [o][u][e]), or the transitional probability information appeared among vowels and the intervened units were consonants.

2.10 Second Language Acquisition and MWS

This section focuses on L2 MWSs. First, the section outlines the challenges of learning and using MWSs in L2. Then, a few possible reasons for the challenges are discussed.

2.10.1 Difficulties in the Learning and Use of MWSs in L2

Mastery of L2 MWSs has been a hallmark of high proficiency of L2 (Ellis et al., 2008). Prior research has suggested that more proficient L2 speakers have a better command of MWSs (Durrant & Schmitt, 2009; Garner & Crossley, 2018). Despite the importance of the acquisition and (correct) use of MWSs in L2, it is generally agreed that the knowledge of L2 MWSs does not grow linearly with L2 vocabulary and tends to develop more slowly than L2 lexical knowledge (Bahns & Eldaw, 1993). Additionally, L2 language learners use far fewer MWSs than L1 speakers in both spoken and written languages (Granger, 1998; Nesselhauf, 2003). In the next section, a few potential reasons for the slow development of MWSs in L2 will be discussed.

2.10.2 Number and Exposure of MWSs to L2 Learners

A large number of MWSs exist in natural languages. To illustrate, Corpus analysis based on BNC has found that there are more than 12000 different phrasal verbs (Gardner & Davies, 2007). The Oxford collocations dictionary for students of English (Ayto, 2009) has more than 800 pages. Another corpus-based study has compiled more than 9000 idioms and the number for the most frequent idioms is over 300 (Liu, 2003). It is obvious that the large number of MWSs challenges L2 learners of English to master and use them correctly.

Furthermore, learning an L2 is very different from learning the L1 i.e., L2 learners do not receive as much exposure to the target language as L1 speakers. Additionally, individual MWSs are not very frequent, though MWSs are widespread in the language (Grant & Nation, 2006; Moon, 1998). As discussed in 2.4.1, the frequency threshold associated with single words is much higher than MWSs.

Therefore, L2 learners do not encounter individual MWS as frequently and it is challenging to acquire L2 MWSs from exposure.

2.10.3 The Influence of L1

The findings of the effect of L1 on L2 MWSs processing are mixed, depending on the specific type of MWSs. With respect to L1's influence on collocation processing, the standard paradigm in this line of research is the semantic acceptability judgement task (Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2015, 2018). These studies consistently found that congruent collocations (i.e., collocations in L2 that have a semantically equivalent collocation in L1) were responded to more quickly than incongruent collocations (i.e., not having equivalent L1 constructions), indicating that L1 exerts an influence on L2 collocation processing. For example, Wolter and Gyllstad (2011) found that Swedish-English bilinguals responded to congruent collocations (e.g., *give an answer* has the word-to-word translation in Swedish, "ge ett svar") more quickly than English-only collocations (e.g., *pay a visit*), indicating that the activation of L1 might have a facilitation effect on processing congruent collocations or that an extra cost is generated when translating incongruent collocations into L1 and searching for the equivalents in L1.

Wolter and Yamashita (2015) tested English native speakers and Japanese-English bilinguals with three types of collocations: (1) English translations of Japanese collocations (e.g., far eye, buy anger) were translated from bilingual participants' L1 (i.e., Japanese) but were unattested in L2 (i.e., English); (2) collocations that were only acceptable in English but not in Japanese (e.g., *low speed*, *catch breath*); and (3) and non-collocational control expressions (e.g., *give legs*, *crash mistakes*) which were implausible in both languages. The study found that Japanese-English bilinguals did not respond to English translations of Japanese collocations more quickly than non-collocational control expressions, which seems to indicate that Japanese-English bilinguals did not activate their L1 Japanese collocational

patterns in the processing of collocations in their L2 English. This evidence suggests that it is less likely that the activation of L1 leads to faster processing in congruent collocations in L2 speakers. Wolter and Yamashita explained that congruent collocations tend to be acquired earlier than incongruent collocations for L2 speakers, which contributed to the faster processing of congruent collocations. In other words, the age of acquisition (i.e., the age at which language users first learn a word or expression) (Carroll & White, 1973; Elsherif et al., 2023; Gilhooly, 1984) contributes to the processing advantage of congruent collocations in L2 speakers. However, age of acquisition has been known to be confounded with frequency as more frequent language units are more likely to be encountered at earlier ages compared to less frequent ones (Carroll & White, 1973). In Wolter and Yamashita's study, the age of acquisition and frequency were not manipulated to be distinguished. Therefore, the results from Wolter and Yamashita's study (2015) might still indicate that faster processing of congruent collocations stems from congruent collocations being more frequent than incongruent ones in the mental lexicon of L2 speakers.

Similar to collocations, L2 speakers also demonstrate that L1 lexical knowledge affects the processing of L2 idioms (Carrol et al., 2016; Carrol & Conklin, 2014b, 2017). In a lexical decision task (Carrol & Conklin, 2014), primes were read in a self-paced manner and consisted of an idiomatic expression up to the final word (e.g., *spill the*). The target was either the final word of an idiom or a matched literal string (e.g., *beans* from *spill the beans*; *chips* from *spill the chips*; *feet* from *draw a snake and add feet*; *hair* from *draw a snake and add hair*). There were two types of idioms: English idioms (e.g., *spill the beans*) and translated Chinese idioms (e.g., *draw a snake and add feet*, the form configuration in Chinese: "画蛇添足," which means spoiling something by overdoing it). After reading the primes, a line of asterisks appeared for 250 ms before the target was presented. Both Chinese-English bilingual and English monolingual participants responded to the targets that formed idioms in their respective L1 more quickly than matched literal strings. In other words, Chinese-English bilinguals

responded to *feet* from *draw a snake and add feet* faster than *hair*, while English monolinguals judged *beans* from *spill the beans* more quickly than chips. Though Chinese idioms were presented in less familiar L2, Chinese-English bilinguals still demonstrated an idiom advantage over matched literal strings. Since the translation of the Chinese idioms was unattested, it is likely that Chinese-English bilinguals never encountered them in English. Thus, the idiom advantage shown by Chinese-English bilinguals was unlikely to have originated from translated Chinese idioms. Rather, it seems that encountering translated Chinese idioms in English (e.g., *draw a snake and add feet*) activated underlying idiomatic concepts in L1 (i.e., spoiling something by overdoing it). This has been taken as evidence to support that idioms are activated at a conceptual level (Carrol & Conklin, 2014).

2.10.4 Low Salience

Another factor responsible for the slow development of L2 MWSs is their low salience. When encountering MWSs, L2 learners might not notice them (Bishop, 2004). This suggests that L2 leaners fail to comprehend or learn MWS due to MWSs without be noticed in the first place. Laufer (1989) stated that some linguistic constructions are deceptively transparent for L2 learners. Those linguistic constructions seem to be transparent, but they are not. To illustrate, *shake a leg* might be interpreted by L2 leaners as *move the leg from side to side* instead of the correct use i.e., tell someone to hurry. One study (Martinez & Murphy, 2011) showed that comprehension of a short text with MWSs was lower than text without MWSs for L2 learners. However, L2 learners overestimated how much they understood MWSs that went unnoticed or were misunderstood. One explanation for the low salience of MWSs is that L2 learners prefer to operate on the open-choice principle by selecting and combining single words (Sinclair, 1991). One study demonstrated that L1 and L2 speakers were different in how they perceived the compositionality in MWSs (Fioravanti et al., 2021). In this study, participants were asked to judge the acceptability of three types of Italian verb-noun word combinations: idioms (*ha perso il filo,* 'he lost the thread of the conversation'), restricted collocations (*ha mantenuto la promessa,* 'he

kept a promise'), free word combinations (e.g., *spedisce una lettera*, 'he posts a letter') and their variant controls that were created by replacing the verb of the combinations with a synonym (e.g., *ha perso il filo -> ha smarrito il filo*; *ha mantenuto la promessa -> ha conservato la promessa*; *spedisce una lettera -> invia una lettera*). Idioms, restricted collocations, and free word combinations in the study differed in the degree of restrictedness in the substitution of the components. Free word combinations allowed the substitution of all components, and restricted collocations allowed limited substitution of components. Unlike restricted collocations and free word combinations, idioms were highly fixed and did not allow lexical substitution. L1 speakers judged the use of synonyms in restricted collocations less acceptable than in free word combinations, and synonyms in idioms were judged as the least acceptable. The findings of L1 speakers were in line with the phraseological approach to define MWSs (Cowie, 1998; Howarth, 1998), whereby different types of MWSs are categorised according to the semantic transparency and lexical restriction of components of MWSs, with free word combinations being the least restricted, idioms being the most restricted and collocations being semi-restricted. For L2 speakers, idioms with synonyms were also judged as least acceptable, but they judged altered collocations as more acceptable than altered free word combinations. Restricted collocations were less conventionalised and less salient compared to idioms. Therefore, L2 speakers found it challenging to detect the degree of semantic/lexical restrictedness of collocations.

2.11 Summary

To sum up, MWSs are not only widespread and recurrent in natural language but also encompass different types of MWSs. All types of MWSs share common features such as high frequency/familiarity and predictability, while MWSs are operated on a continuum scale of compositionality and fixedness, with certain types of MWSs more restricted (e.g., idioms) and other types of MWSs more composable (e.g., collocations). There is a well-documented processing advantage in MWSs whereby MWSs are processed more quickly than novel word combinations in both L1 speakers

and L2 speakers[3]. This processing advantage has been taken as evidence for the usage-based model of language acquisition and statistical learning theory in language acquisition. However, faster processing could not support the holistic storage and representation of MWSs. Unlike L1 speakers, L2 speakers have faced challenges acquiring and using MWSs. A few factors, such as the large number of MWSs, a lack of exposure to L2 MWSs, and a lack of awareness of L2 MWSs can explain the slow development of L2 MWSs.

[3] L2 speakers do not demonstrate the processing advantage with all types of MWSs. At least, findings in idioms processing in L2 speakers are mixed.

CHAPTER 3 COMPARING MULTI-WORD SEQUENCE PROCESSING

Abstract

The present study investigated the processing of three types of multi-word sequences (MWSs): free word combination, restricted collocations, and idioms, which were defined by following the phraseological approach. A self-paced reading paradigm paired with a stop-making-sense task measured reading times for the three types of MWSs in 84 L1 speakers and 73 L2 speakers. In addition, L2 proficiency was assessed to see whether it mediated the processing of the different types of MWSs. We observed that L1 speakers read idioms faster than free word combinations, but there was no processing difference between free word combinations and restricted collocations, while the word-level and phrase-level frequency were matched. L2 proficiency and types of MWSs interplayed to affect processing. Less proficient L2 speakers processed free word combinations faster than the other two types of MWSs, while the faster processing of free word combinations decreased with the increase of L2 proficiency.

Keywords: multi-word sequences, collocation processing, phraseological approach, idiom processing

3.1 Introduction

3.1.1 Effects of Frequency on Collocations Processing

Multi-Words Sequences (MWSs) have been shown to contribute to language processing in a similar way as single words (Arnon et al., 2017; Arnon & Snider, 2010; Bybee, 1998; Christiansen & Arnon, 2017; A. E. Goldberg, 2006). One type of MWSs that has received detailed scrutiny is collocation, broadly defined as co-occurring words (e.g., *build a house*, *provide information*, *fatal mistake, strong tea*) (Biber et al., 1999). A collocation (e.g., *lower risk*) consists of a node (i.e., words that are searched for in corpora, e.g., *risk*) and a collocate (i.e., words that co-occur with the node, e.g., *lower*). Definitions of collocations are discussed in detail in the subsequent paragraphs.

Studies of collocations (Durrant & Doherty, 2010; Siyanova & Schmitt, 2008; Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2015, 2018) suggest that collocations (e.g., *provide information, fatal mistake*) are processed more quickly than novel word combinations that infrequently co-occur (e.g., *compare information, extreme mistake*). In this chapter, the literature review focuses on studies of collocations using the acceptability judgement task and masked priming paradigm with a lexical decision task, while the eye-tracking collocation studies will be reviewed in Chapter 5.

In one of the earliest collocation studies, Siyanova and Schmitt (2008) asked participants to make semantic judgements about common (i.e., frequent) and uncommon (i.e., infrequent) collocations. The authors found that more frequent collocations were judged more quickly than novel word combinations. Furthermore, Durrant and Doherty (2010) questioned whether the faster processing advantage is limited to collocations which contain words that are associated (e.g., *parish* and *church,* as in the collocation *parish church*) according to word association norms (e.g., the Edinburgh Association Thesaurus (Kiss et al., 1973)) and, thus, that the collocation advantage is driven by associative priming. To address this issue, Durrant and Doherty manipulated the association between nodes and collocates

using an unmasked priming task. They examined collocations without associated nodes and collocates (e.g., *lower risk*), collocations with associated nodes and collocates (e.g., *parish church*) and non-collocational and implausible control expressions (e.g., *narrow risk; feature church*). The collocations had either high or low phrase-level frequency. The target was the node of the collocation (*risk, church*). The prime was the collocate or control word (*lower, narrow, parish, feature*). The prime was presented for 60ms before being replaced by the target, which stayed visible until the participant made a response. Participants responded faster to the targets following primes from both associated and non-associated collocations, but only when the collocational (phrase-level) frequency was high. With lower phrase-level frequency, no priming was observed for either type of collocation. These results indicate that the collocation advantage extends beyond associative priming and that phrase-level frequency plays a crucial role in driving this advantage. Less frequent collocations may not show a priming effect in Durrant and Doherty's study, possibly because they are less entrenched in memory.

Öksüz et al. (2021) used an acceptability judgement task to test the effect of word-level and phrase-level frequency on the processing of adjective-noun collocations (e.g., *senior office*, *long time*). They found that both noun frequency and phrase-level frequency were associated with faster responses in both L1 and L2 speakers. This suggests that both L1 and L2 speakers are sensitive not only to the word-level but also to the phrase-level frequency when processing adjective-noun collocations. This result is in line with other findings in psycholinguistics, where frequencies at both the word and the supra-word level contribute to processing, with evidence coming from the processing of idioms (Kessler et al., 2021; Milburn et al., 2021), lexical bundles (Arnon & Snider, 2010; Hernández et al., 2016; Tremblay et al., 2011) and compound words (Pollatsek et al., 2000). Furthermore, Öksüz et al. (2021) found that, for both L1 and L2 speakers, the larger the collocational (i.e., phrase-level) frequency effect, the smaller the word frequency effect. This indicates that there is a shift away from relying on single words to a holistic configuration when processing collocations with high phrase-level frequency. The

shift from single-word frequency to collocational frequency suggests that collocations become more entrenched and lexicalised in memory. Of note, the lexicalisation does not indicate that collocations are only processed as whole chunks without an underlying compositional analysis, as language users remain sensitive to word-level frequency, even in highly frequent collocations. The high frequency of nouns facilitated the processing of collocations. In contrast, the increased frequency of adjective resulted in a slowing down in processing collocations. When participants encounter a highly frequent adjective, such as *long* in *long time*, it is unlikely that participants can predict the following noun given the many possible continuations (e.g., *time, hair, way, walk, journey, history, list, story*). On the other hand, when the adjective is less frequent (e.g., *senior*), there tend to be fewer plausible continuations, leading to a stronger prediction.

These results indicate that both L1 and L2 language users show a processing advantage for frequently co-occurring words, indicating that these combinations are becoming entrenched in memory. Thus, the collocation advantage mainly stems from the higher phrase-level frequency and greater familiarity of collocations, which generalises the frequency effect beyond single words to multi-word sequences. These findings also support language models that highlight the role of experience (i.e., frequency is seen as an index of experience) in the processing and entrenchment of language, such as the usage-based model to language acquisition (Bybee, 1998; A. E. Goldberg, 2006; Tomasello, 2003). However, the finding (i.e., sensitivity to phrasal frequency) is more difficult to be accommodated in the words-and-rules model. According to the words-and-rules model, the frequency effect is attached to the mental lexicon that stores memorised single words, but compositional phrases, such as collocations, are assumed to be computed according to grammar and rules. Therefore, under this model, language users are not expected to be sensitive to distributional information of larger linguistic constructions (e.g., MWSs).

3.1.2 A Frequency-based Approach to Defining Collocations

It is worth noting that the studies displaying a collocation advantage (Carrol & Conklin, 2020; Durrant & Doherty, 2010; Li et al., 2021; Öksüz et al., 2021; Siyanova & Schmitt, 2008; Sonbul, 2015; Vilkaitė, 2016) employed a frequency-based approach (Sinclair, 1991) defining and identifying collocations. However, there is another common approach to defining collocations, the phraseology-based approach (Cowie, 1998; Howarth, 1998). We will turn to the phraseology-based approach in the next section.

By following the frequency-based approach, collocations are defined as words that co-occur more frequently than can be expected by chance (Biber et al., 1999). Identifying frequency-based collocations involves the extraction of collocations from a corpus based on specific association measures that provide a measure of the strength of co-occurrence between two words (Hunston, 2022). There are several different association measures (see Gablasova et al. 2017 for a discussion); for example, Log Dice, the Mutual Information score (MI score), T-score, and Delta P. AMs do not solely rely on phrase-level frequency, and different AMs attach importance to different properties of collocations (Gablasova et al., 2017). To illustrate, Delta P captures the directionality of a collocation, which is very similar to, and highly correlated with, transitional probability (i.e., the statistical likelihood that a word will precede or follow another word). For example, *Christmas decorations* is imbalanced in terms of the directionality: *Christmas* primes *decoration* stronger than vice versa. The MI score[4] is a widely used association measure which highlights the exclusivity of a collocation. Word combinations with an MI score larger than three are seen as collocations (Hunston, 2022). The higher MI score a collocation has, the more exclusively associated the words in the collocation are, and the less likely components of the collocations form other collocations. In addition, the more frequently and exclusively word

[4] MI scores operate on a scale without a minimum and maximum (Gablasova et al., 2017).

combinations are used together, the more likely they are lexicalised (Conklin & Schmitt, 2008, 2012; Siyanova-Chanturia, Conklin, & van Heuven, 2011). Of note, studies following the frequency approach tend to select collocations that generally have transparent and literal meanings (Carrol & Conklin, 2020) (e.g., *celebrate victory*, *fatal mistake*, *detailed map*).

According to the thresholds of the phrase-level/collocational frequency and MI score, phrases such as *cut corners* (MI score = 5.71, collocational frequency[5] = 54), *full circle* (MI score = 5.2, collocational frequency = 147), *run a mile* (MI score = 3.77, collocational frequency = 38), and *plant a seed* (MI score = 5.28, collocational frequency = 21) can be identified as collocations. However, the conventionalised meaning of the word combinations listed above cannot be purely inferred from the lexical parts (e.g., the assembled meaning of *cut* and *corners* is very different from the conventionalised meaning of cutting and corners i.e., *save time, money, or energy by doing things quickly and not as carefully as you should*). Moreover, these word combinations (e.g., *cut corners*, *full circle*) are often defined as idioms that are fixed and non-compositional. Taken together, the frequency-based approach to identifying collocations emphasises frequency, but disregards other features of word combinations, such as semantics, which causes the boundaries between collocations and other MWSs to be unclear.

3.1.3 A Phraseology-based Approach to Defining Collocations

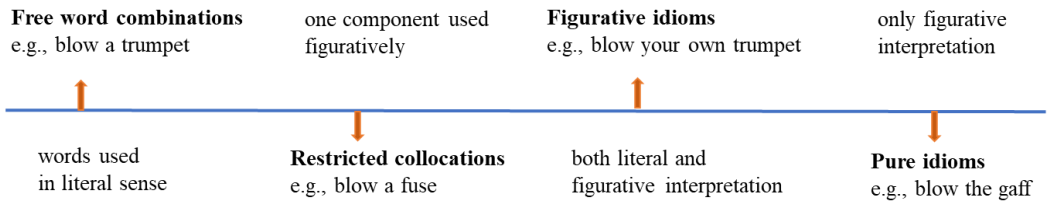
The alternative way of defining collocations is the phraseology-based approach (Cowie, 1998; Howarth, 1998). Though many different phraseological typologies exist, most of the classifications give prominence to one or more of the following features: semantic compositionality, syntactic restriction, internal structure, and discourse function. Hence, while a frequency-based approach has been criticised for certain characteristics of word combinations, a phraseological approach aims to take on these neglected characteristics.

[5] Frequency is expressed as raw frequency from British National Corpus (BNC).

According to one of the most influential phraseological word combinations models, Howarth's (1998) and Cowie's Continuum Model (1998), restricted collocations (in this model, collocations are classed as "restricted collocations") and three other categories of word combinations (free word combinations, figurative idioms, and pure idioms) are classified on a continuum scale (see Figure 3.1). The most literal and transparent word combinations, free word combinations (e.g., *blow a trumpet, under the table, easy task*), are on the left-hand side of the scale. All components (i.e., lexical parts) from free word combinations are used in their literal sense and can be freely substituted (e.g., *blow a trumpet*, *blow* and *trumpet* are used literally in this word combination and can be replaced by other words, e.g., *saxophone*). The most figurative, opaque, and non-decomposable word combinations are idioms. Idioms can be further classified as figurative idioms (e.g., *pay the price, grey area, in deep water*) and pure idioms (e.g., *pay the piper*). A figurative idiom preserves a possible literal meaning (e.g., the literal interpretation of *pay the price i*s plausible), while a pure idiom does not have a plausible literal interpretation. Importantly, as Figure 3.1 shows, restricted collocations (e.g., *under attack*, *blow a fuse*) are placed between free word combinations and figurative idioms on Howarth's continuum scale. One component of a restricted collocation is used in a limited number of contexts with specialised senses, while the other component of the restricted collocation is used in a literal sense. Compared to free word combinations, restricted collocations are more restricted in terms of word substitution. There can be disagreement between the phraseological-based and the frequency-based approaches on which one MWS should be classed as a collocation. A collocation identified by the frequency-based approach could be classified as a free word combination through the phraseological-based approach, depending on their semantic transparency and restriction (e.g., *provide information*, *lower risk* would be classed as collocations according to the frequency-based approach, but these two phrases would be grouped into free word combinations according to the phraseological-based approach).

Figure 3.1

Cowie and Howarth's Continuum Model of Classifying Four Types of Word Combinations



Another widely accepted phraseological definition of collocations draws on "restricted sense" (Nesselhauf, 2004), i.e., the sense of one component from one collocation is specific so that the component only allows the combination of a limited number of words to form collocations. In other words, when the component is used in the given sense, it cannot be combined with all other words that are syntactically and semantically plausible. In contrast, the senses of components from a free word combination are not restricted, so that components can be freely substituted as long as they respect syntactic and semantic restrictions. For example, the primary sense of *read* (i.e., *look at a written text and understand what it means*) can be combined with many other words that are syntactically and semantically plausible (e.g., *read books*/*newspapers*). Thus, *read books* is a free word combination. However, *take* from *take a picture* is used in the restricted sense because *take* can only be combined with a small set of other nouns. In terms of idioms, both components are used in their restricted sense (e.g., *sweeten the pill*). Differences exist in two phraseological typologies, but both typologies draw on the restriction of lexical parts of the word combinations. To sum up, the phraseology-based approach addresses the issue of semantic transparency/compositionality and attaches importance to the restriction of lexical parts of MWSs. However, the approach still suffers from subjective and arbitrary in the categorisation of word combinations as certain types of MWSs (Nesselhauf, 2004).

Both approaches to the definition and identification of collocations have advantages and disadvantages. The choice of approach should be methodical as the approach one takes to defining and identifying collocations can directly affect what items are selected for studies (Gablasova et al., 2017). In addition, it is important to be clear and consistent which approach is followed in the study. For example, in one study (Webb et al., 2013), collocations were defined from the statistical frequency-based approach and included items such as *cut corners* and *pull strings*. Under the phraseological view, these items would not be defined as collocations as they are semantically opaque and not compositional, introducing potential confounds in their study. Taken together, the selection of different collocational definitions, reflecting different linguistic approaches, could potentially lead to conflicting results. We decided to adopt the phraseology-based approach and follow Howarth's Continuum Model in the current study. There are two rationales behind this decision. First, research that follows a frequency-based approach to identify collocations has demonstrated a well-established collocation processing advantage over novel word combinations as a result of high phrase-level frequency (e.g., Carrol & Conklin, 2020; Li et al., 2021; Öksüz et al., 2021; Sonbul, 2015; Vilkaitė, 2016; Vilkaitė & Schmitt, 2019; Vilkaitė-Lozdienė, 2022; Wolter & Gyllstad, 2013; Wolter & Yamashita, 2018). However, by solely relying on frequency, such an approach disregards the internal semantics of collocations and, therefore, might underestimate the contribution of a distinction that other (phraseology-minded) researchers believe to be important. To our knowledge, only two existing studies followed the phraseology-based approach to assess the effect of semantics. Furthermore, these two studies do not show a consistent finding for L1 speakers (Gyllstad & Wolter, 2016; Shi et al., 2023). Second, while the phraseology-based approach postulates a continuum scale including four types of MWSs, the existing research only focused on collocations and free word combinations. The current chapter extends the scope of research to include idioms and aims to replicate Gyllstad and Wolter's study (2016) in a reading task. Though we followed

the phraseology-based approach, frequencies of all types of word combinations were checked with BNC to ensure selected word combinations are not rare.

3.1.4 Collocation Processing under the Phraseological View

While many studies that followed the frequency-based approach observe a consistent collocation advantage in both L1 and L2 speakers, two studies that followed the phraseology-based approach found conflicting results in L1 speakers (Gyllstad & Wolter, 2016; Shi et al., 2023). In a sensicality judgement task, Gyllstad and Wolter (2016) included one three-level categorical variable: restricted collocations (e.g., *seize the opportunity*), free word combinations (e.g., *pay tax*), and implausible controls (e.g., *stop eggs*). They found that both restricted collocations and free word combinations were judged more quickly than controls in both L1 and L2 speakers. Importantly, restricted collocations were processed more slowly than free word combinations in both L1 and L2 speakers, indicating that there was an extra cost in processing restricted collocations compared to free word combinations. The processing cost might stem from restricted collocations' components being used in a figurative or de-lexical sense, leading to competition between the figurative and literal sense (e.g., *seize* in *seize the opportunity*) or between the concrete and abstract sense (e.g., *take* from *take shape*). Though items in Gyllstad and Wolter's study (2016) were matched on phrase-level length and frequency across the three levels, items were not matched pairwise. Additionally, the methodology used does not measure natural reading and cannot provide an indication of how MWSs are processed in sentence context.

Shi et al. (2023) addressed these shortcomings by using a self-paced reading task. Their study compared the differences between experimental pairs (figurative collocations, e.g., *built a career*, vs. literal collocations, e.g., *built a house)* and control pairs (literal collocations, e.g., *chose a career*, vs. literal collocations e.g., *chose a house*). Their results showed a processing cost for the figurative-literal

collocation pairs (i.e., experimental pairs) when compared to the literal-literal collocation pairs, in L2 speakers. This suggests that L2 speakers are more likely to activate the literal interpretations of the figurative collocations initially. In the processing of a restricted collocations, the extra processing cost might then result from the rejection of the literal interpretation or the competition between the literal and figurative interpretations. Unlike L2 speakers, the descriptive results for L1 speakers showed faster reading times for figurative-literal collocation pairs than for literal control pairs, though the trend was statistically not significant. The contradictory findings for L1 speakers in the two studies (Gyllstad & Wolter, 2016; Shi et al., 2023) leave us uncertain about the processing of restricted collocations, as defined under the phraseological view, and it leaves open the possibility that other properties of MWSs beyond distributional information can contribute to the processing of MWSs.

One study suggests that meaningfulness exerts influence on the processing of MWSs in the absence of distributional information (Jolsvai et al., 2020). Using a judgement task, participants responded to three types of MWSs: idiomatic expressions (e.g., *play the field*, *up the creek*), compositional phrases (e.g., *nothing to wear*, *get a certificate*), and non-sensical phrasal fragments (e.g., *without the primary*, *because it lets*). These three types of MWSs were matched for word-level and phrasal-level frequency. What set phrasal fragments apart from idiomatic expressions and compositional phrases was that phrasal fragments were less meaningful than the other two types. They found that meaningfulness sped up the processing of WMSs. Idiomatic expressions and compositional phrases were responded to more quickly than phrasal fragments, but there was no processing difference between idiomatic and compositional expressions. However, we should be careful about generalising these findings due to a few limitations. Like Gyllstad and Wolter's study, Jolsvai et al. (2020) failed to match items pairwise, even though frequency was controlled on word level and phrase level. Additionally, their items varied widely in syntactic structure (e.g., *nothing to wear*, *get a certificate, without the primary*, *because it lets*), which could add potential confounds.

3.1.5 Idiom-Superiority Effect

The idiom category in Howarth's Continuum Model has received the most scrutiny. The idiomatic meaning of an idiom cannot be determined through an analysis of its individual word meanings. Thus, traditionally, idioms are characterised by non-compositionality and are seen as "fixed" or "dead" word strings to be represented, stored, and retrieved as a whole in the mental lexicon. In addition, early idiom models (e.g., the Direct Access Model, Gibbs, 1980; the Idiom List Hypothesis, Bobrow & Bell, 1973); the Lexical Representation Hypothesis, Swinney & Cutler, 1979) assume that a literal analysis of an idiom has a limited role in accessing the idiomatic meanings, or that there is no activation of the literal meaning when comprehending an idiom.

The historical assumption (i.e., the limited role of literal analysis in idioms processing) is most likely due to the well-documented idioms' superiority effect where idioms are consistently processed more quickly than novel literal word combinations (Cacciari & Tabossi, 1988; Conklin & Schmitt, 2008; Gibbs, 1980; Gibbs et al., 1989; M. R. Libben & Titone, 2008; McGlone et al., 1994; Ortony et al., 1978; Siyanova-Chanturia, Conklin, & Schmitt, 2011; Swinney & Cutler, 1979; Tabossi et al., 2009). A standard paradigm in this line of research used the semantic judgement task to compare idiomatic strings (e.g., *break the ice*) to matched literal strings (e.g., *break the glass*), and to implausible filler items (e.g., *wish out table*). This line of research observed a robust effect, with idiomatic strings responded to more quickly than matched literal controls. Additionally, the idioms' superiority effect has been replicated in reading paradigms as well. For example, an eye-tracking study replicated the idioms' superiority effect in L1 speakers (Siyanova-Chanturia, Conklin, & Schmitt, 2011). In this study, ambiguous idioms (i.e., ones that have both a plausible literal and an idiomatic interpretation, e.g., *at the end of the day*) and novel word combinations (e.g., *at the end of the war*) were embedded in a story context. For idioms, the prior contexts were biased towards either a literal interpretation or a figurative interpretation of the idiom. L1

speakers read both idiomatic strings in their literal and figurative use more quickly than free word combinations, while reading idioms used figuratively was not faster than ones used literally.

However, much research has shown that the higher familiarity of the idiomatic constructions compared to that of the novel literal word combinations can account for the idioms' superiority effect (e.g., Carrol & Conklin, 2017; Libben & Titone, 2008; Tabossi et al., 2009). Later findings demonstrate that syntactic analysis occurs during the processing of idioms (Cacciari & Corradini, 2015; Holsinger, 2013; Holsinger & Kaiser, 2013), which would argue against idioms being represented as single words/constructs. As a result, idioms are thought to have a hybrid representation, operating partly as lexical items and partly as compositional strings (Cacciari & Tabossi, 1988; Sprenger et al., 2006). Idioms are known and familiar to language users, and so the idiomatic meanings can be recognised and retrieved more quickly, and less compositional analysis is involved, compared to novel word combinations. In a semantic judgement task, Tabossi et al. (2009) asked participants to judge three different conventional word strings: decomposable idioms (i.e., the lexical parts of one idiom contribute to the idiomatic sense), non-decomposable idioms (i.e., the lexical parts of one idiom do not contribute to the idiomatic sense) and clichés (i.e., frequent multiword conventional expressions) with their correspondingly matched literal strings. Compared to idioms, clichés are compositional, which are expected to have less processing advantage over the matched literal strings. All types of conventional word strings were processed faster than their matched literal strings. In addition, the lack of a significant interaction between the type of expression (conventional or literal strings) and the type of conventional word strings (non-decomposable idioms, decomposable idioms, clichés) suggests that the processing advantages of the conventional word strings over the matched literal strings did not differ between the types of conventional word strings. These findings are taken as evidence for familiarity being the contributing factor for the idioms' superiority effect, and as evidence against the role of compositionality in the processing of idioms. Nevertheless, their post-hoc analysis showed that the

clichés in their study were much longer than the idioms, arguably resulting in compositional clichés to be more predictable than idioms. Therefore, the negligible difference in processing advantage between clichés and idioms compared to literal controls could stem from the higher predictability of the clichés. Additionally, using an arbitrary cut-off point to distinguish decomposable and non-decomposable idioms, as was done in this study, has been argued to be dichotomizing a continuous variable, which is statistically problematic (Cohen et al., 2013). In summary, the previous findings leave the role of de-compositionality in the processing of MWSs inconclusive and suggests that familiarity is the main driver for the faster processing of idioms.

Based on the discussion above, there is nothing inherent to idioms that contributes to the idiom superiority effect. The main contrasting difference between idioms and novel word combinations is that idioms are better known to language users, while novel word combinations are unknown. However, idioms do not simply fit in the usage-based approach, as individual idioms are not frequent in language (at least, idioms are infrequent in corpora) (Moon, 1998). This discrepancy leads us to question whether idioms still demonstrate the superiority effect in comparison to other word combinations when controlling for distributional information.

3.1.6 Collocation and Idiom Processing in the Second Language

As discussed above, whether L2 speakers process collocations in a similar way to L1 speakers depends on how collocation is defined (frequency-based versus phraseological-based). Following the frequency-based approach to defining collocations, much research on L2 collocation processing has demonstrated a facilitation effect for collocations over novel word combinations (Li et al., 2022; Öksüz et al., 2021; Siyanova & Schmitt, 2008; Sonbul, 2015; Vilkaitė & Schmitt, 2019). As is the case for L1 speakers, L2 speakers are also sensitive to phrase-level frequency. However, two studies (Gyllstad &

Wolter, 2016; Shi et al., 2023) show that there is an extra processing cost when a phraseological approach is followed to identify collocations.

The finding of an extra processing cost in restricted collocations in L2 speakers resembles what is sometimes found for L2 idioms processing. Unlike what was found for L1 speakers, there are mixed findings of L2 idiom processing. Some studies demonstrated that L2 speakers read or processed idioms more quickly than matched literal strings (Conklin & Schmitt, 2008), whereas other studies failed to observe the idioms' superiority effect (Carrol et al., 2016; Carrol & Conklin, 2017; Siyanova-Chanturia, Conklin, & Schmitt, 2011). However, it has consistently been found that L2 speakers take longer time to comprehend the idiomatic interpretation of an idiom than its literal interpretation (e.g., Carrol & Conklin, 2017; Siyanova-Chanturia et al., 2011).

Dual-route models (Van Lancker Sidtis, 2012; Wray, 2012; Wray & Perkins, 2000) can explain the lack of the idioms' superiority effect, the slow-down effect on restricted collocations and the facilitation effect on reading the literal interpretation of an idiom in L2 speakers. According to dual-route models, the direct access route is only available when encountering known and familiar MWSs. This route allows language users to recognise the canonical form and directly access its conventional (figurative) meaning. In contrast, when MWSs are unknown/less familiar to the language user, a decomposition route is taken whereby the whole meaning is computed online via assembling the meaning of lexical parts. In practice, when the MWS is known, both the direct access route and the decomposition route will compete or interact, with the direct route generally winning out. When the MWS is unknown, only the (slower) computational route is available. Due to the infrequent exposure to MWSs, L2 speakers are more likely to take the computational route. Thus, L2 speakers initially compute the literal meaning of an idiom and a restricted collocation. In line with this, according to the L2 Literal Salience model (Cieślicka, 2006), L2 speakers mainly rely on a literal analysis to comprehend figurative strings. Taken together, the priority of a literal analysis requires L2 speakers to take an additional step of rejecting or suppressing the literal

interpretation, which slows down processing the figurative interpretations of idioms and restricted collocations compared to free word combinations.

In addition, L2 collocation processing is also affected by other factors, such as the characteristics of the L1 language, word-level, and phrase-level frequency (Vilkaite & Schmitt, 2019; Öksüz et al., 2021), L2 proficiency, and their interactions (Wolter & Yamashita, 2018). In the case of frequency, similar to L1 speakers, L2 speakers demonstrate a sensitivity to both word-level and phrase-level frequencies. Additionally, compared to low phrase-level collocations, there is a reduced word-level effect on processing collocations with high phrase-level frequencies (Öksüz et al., 2021). L2 proficiency not only speeds up the processing of collocations, L2 proficiency, but also interacts with word-level and phrase-level frequency. With an increase in L2 proficiency, there is a shift away from relying on word-level frequency to relying on phrase-level frequency when processing collocations. However, compared to L1 speakers, L2 speakers still tend to rely on word-level frequency more than collocational frequency (Wolter & Yamashita, 2018). Despite this, Öksüz et al. (2021) did not observe an interaction between group and word-level frequency, thus L1 and L2 speakers were comparably sensitive to word-level frequency. The inconsistent findings between these two studies (Öksüz et al., 2021; Wolter & Yamashita, 2018) could be related to the type of L2 speakers used in these studies, with Öksüz et al. (2021) relying on highly proficient L2 participants while Wolter and Yamashita (2018) recruiting lower proficient L2 speakers.

3.1.7 The Present Study

Although specific types of MWSs have been well studied, there is a lack of comparison between the different types of MWSs. Following Howarth's phraseological continuum model (Howarth, 1998), the present study attempted to fill an empirical gap by using a self-paced reading task to investigate the processing of different types of MWSs. By employing MWSs in context sentences, this study compared

the processing of free word combinations, restricted collocations, and idioms in both L1 and L2 speakers. The first question of this study is whether there is a processing cost for restricted collocations over free word combinations in a reading task. The hypothesis corresponding to the first question is: if the findings of the present study mirror the previous findings from the sensicality judgement task (Gyllstad & Wolter, 2016), we expect to find an extra processing cost for restricted collocations in reading in both L1 and L2 speakers. The second question is whether idioms show a processing advantage over other types of MWSs when the phrase-level frequency between the three types of MWSs is controlled for. The hypothesis corresponding to the second question is: if the phrase-level frequency is the main driving factor of the idioms' superiority effect, no difference in processing time between idioms and other types of MWSs is expected in L1 speakers. If additional properties of MWSs, such as conventionality, are involved in the idioms' superiority effect, we expect to detect a difference in processing time between idioms and the other types of MWSs, with idioms reading more quickly in L1 speakers. The third question is to what extent L1 and L2 speakers' processing of MWSs overlap. The corresponding hypothesis is that compared to L1 speakers, L2 speakers are predicted to demonstrate extra processing costs in reading idioms compared to other types of MWSs due to L2 readers having to suppress the literal interpretation of the idioms. The fourth question is about the role of L2 proficiency in MWSs processing. The corresponding hypothesis is that when L2 speakers have lower L2 proficiency, they are predicted to read free word combinations faster than idioms. Increased L2 proficiency is expected to decrease the preference for free word combinations over idioms.

3.2 Method

3.2.1 Design

This study used the self-paced reading paradigm coupled with a concurrent stop-making-sense paradigm (Boland et al., 1990; Gibson et al., 1996). Participants were asked to read sentences at their

own pace and to decide whether what they read still made sense in English. This paradigm is potentially more sensitive in measuring processing than a standard moving-window self-paced reading task and can keep participants more engaged with reading sentences (Singh et al., 2016). In addition, it allows participants, particularly L2 participants, to terminate sentences when they have difficulty comprehending word combinations. Furthermore, the sentences were presented in a chunk-by-chunk manner, with the critical items appearing in the third region and the two words included in the spill-over region. The reasons for the chunk-by-chunk presentation are: (1) prior research has shown that the chunk-by-chunk presentation can detect the processing advantage of MWSs over controls as well as the word-by-word presentation (Tremblay et al., 2011) and (2) word-by-word presentation is more likely to induce rhythmic reading than when the length of the chunks is more variable (Boyce et al., 2020). Moreover, presenting words such as "the" or "a" by themselves looks more unnatural than when they are part of a larger chunk.

In this study, Language-background was a between-group factor with two levels: English native speakers (L1 speakers) and L1-Chinese L2-English speakers (L2 speakers). Types-of-MWSs was one three-level within-participant and between-item variable (idioms vs. restricted collocations vs. free word combinations). The dependent variable was reading times on the critical region and spill-over region.

3.2.2 Participants

Eighty-four native speakers of English (age: *M* =25.9, *SD* = 4.86) and 73 native speakers of Chinese with L2 English[6] (age: *M* = 28.2, *SD* = 6.37) took part in the study, and none of these participants participated in the pilot study described below. L2 participants were also asked to take the Vocabulary Levels Test of 3K and 5K (VLT-version 1, Schmitt et al., 2001) to estimate participants' language

[6]We recruited L2 participants online; thus, some participants had Cantonese as their first language due to the difficulty of reaching Mandarin-speaking participants in mainland China via Prolific.

proficiency. Following previous L2 collocation processing studies (Sonbul, 2015), this study also skipped 2K and 10K levels. The participant recruitment website, Prolific, is primarily accessible to people from the countries that join the Organisation for Economic Co-operation and Development. The nature of the recruitment website explains why L2 participants had lived in English-speaking countries for a relatively long time. Information on participants' language usage is summarised in Table 3.1. All participants had normal or corrected vision and no self-reported history of reading difficulties.

The experiment was conducted in accordance with the British Psychological Society ethical guidelines and was approved by the University of Birmingham's ethical committee (ERN_19-0936).

Table 3.1

Summary of Participants' Language Usage

	L1 speakers (*M*, *SD*)	**L2 speakers (*M*, *SD*)**
Age of English acquisition	1.29 (0.82)	6.75 (3.21)
Time lived in an English-speaking country (in years)	17.7 (512.4)	7.52 (9.11)
VLT score (out of 60)	59.6 (0.73)	56.7 (10.1)
Self-reporting of English proficiency (a 6-point scale)		
Proficiency in writing	4.89 (0.35)	4.06 (0.98)
Proficiency in speaking	4.92 (0.32)	4.06 (1.05)
Proficiency in listening	4.94 (0.28)	4.37 (0.91)
Proficiency in reading	4.94 (0.28)	4.42 (0.79)
Self-reporting of time using English (per day)		
Writing (in hours)	4.04 (4.06)	2.80 (2.75)
Speaking (in hours)	9.06 (5.54)	4.59 (4.20)
Watching/listening (in hours)	6.10 (4.29)	4.20 (3.12)
Reading (in hours)	3.93 (4.03)	3.38 (2.52)

Note. Standard deviations are in parentheses

3.2.3 Materials

There were 45 sets of items, with three conditions for each set (a triplet). The Types-of-MWSs was a three-level within-participant and between-item variable. Each item triplet included an idiom (e.g., *blows my mind*), a restricted collocation (e.g., *enters my mind*), and a free word combination as the baseline (e.g., *affects my mind*). To replicate and extend Gyllstad and Wolter's study (2016) using natural sentences, the selection of items was operationalised according to Howarth's continuum Model (1998). Therefore, free word combinations had their components interpreted literally, restricted collocations had one component used in a literal sense and the other component used in a de-lexical or figurative sense, and for idioms, both components were interpreted figuratively. All items are listed in Appendix 3.2.

3.2.3.1 Selection of Items

Initially, restricted collocations were selected from previous studies (Gyllstad & Wolter, 2016; Macis & Schmitt, 2017). Only those collocations were selected that satisfied the criteria established by the phraseological model in which one of the components was used in restricted contexts or figurative sense. This procedure yielded 65 candidate item sets.

To check whether candidate items were plausible and known to native English speakers, a semantic judgement task was conducted prior to the reading experiment. The task was programmed with E-prime 3.0. Semantically implausible fillers items were included, resulting in a total of 390 trials (half of which implausible filler items). We embedded target items (e.g., *enters my mind*, *earn his respect,* and *lose time*) and filler items (e.g., *fries one's mind*, *earn his mirror* and *cook time*) in neutral sentences providing minimum contextual information (e.g., *It enters/fries my mind*; *She is starting to earn his respect/mirror; We might lose/cook time.*). All candidate items and filler items are listed in Appendix 3.1.

Forty-seven native English participants (age: *M* = 19.81, *SD* = 1.3; gender: female, *N*= 34) were recruited, and they were asked to judge whether these word combinations made sense in English, with additional three participants removed (one was due to a technical issue and two was due to low accuracy rate < 80%). To begin, there were seven practice trials. At the start of each trial, a fixation string ("+++++") was presented in the middle of the screen for 500 ms. Then, it was replaced by a sentence fragment (e.g., *We might*). After participants pressed the space bar, the critical item (e.g., *lose time*) appeared. Participants were asked to press "K" to indicate that what they had read made sense in English and to press "D" to indicate what they had read did not make sense in English.

When more than half of the participants judged one critical item implausible in English (e.g., *mark time*), the full item triplet was removed from the final material (e.g., *lose/kill/mark time*). After removing less plausible items, 45 sets of items were kept for the study. The accuracy rates for idioms, restricted collocations and free word combinations were 90%, 92% and 79%, respectively.

3.2.3.2 Word-level and Phrase-level Length and Frequency

To compare restricted collocations to free word combinations, it is crucial to match item-related variables between these two types of constructions. Both lexical and phrasal frequency was checked using BNC. Phrasal frequency was checked using the "collocates" function with a maximum window size of ± 4 (e.g., *lose _ _ _ time* has a window size[7] of 4 when *time* occurs as the fourth word after *lose*; *lose time* has a window size of 1 when *time* occurs as the first word immediately after *lose*). Of note, this study checked the raw frequency of word combinations. One-way repeated ANOVA tests revealed that the three types of word combinations were not significantly different with respect to word-level

[7] The window size is the number of intervening words between the components of a collocation, plus the last component of the collocation.

frequency[8], phrase-level frequency[9], MI-score, phrase-level length, and word-level length (all *F*s < 1). Due to the item *last week* that had high raw phrase-level frequency (appearing 5125 times in BNC), the averages for the raw frequency and standard deviation in free word combination were high. We therefore used log frequencies to better account for outliers. (In eye-tracking, most researchers nowadays prefer to use log frequencies rather than raw frequencies, though both can be found, e.g., Schotter et al., 2014). Importantly, even when this outlier item was included, the statistical comparison with non-logged and logged frequencies generated a comparable result, i.e., the phrase-level frequency was matched and were the same between conditions (F (2, 132) = 0.821, p = .442; F (2, 132) = 0.06, p = .942). This suggests that the phrasal frequency was matched across conditions. The main characteristics of the experimental items are summarised in Table 3.2.

[8] Word-level contains the first word, second word and third word.

[9] For word-level and phrase-level frequency, these statistical analyses were performed with non-logged and log-transformed frequency data respectively. Comparable results were found as for the non-logged and log-transformed data: word-level and phrase-level frequencies were not statistically different in idioms, restricted collocations and free word combinations

Table 3.2

Summary of Items Characteristics

Items characteristics	**Free combinations (*M*, *SD*)**	**Collocations (*M*, *SD*)**	**Idioms (*M*, *SD*)**
Word1 length	4.31 (1.08)	4.22 (1.04)	4.29 (1.01)
Word2 length	3.89 (1.35)	4.18 (1.56)	3.98 (1.41)
Word3 length	3.2 (1.41)	3.42 (1.69)	3.18 (1.37)
Item length	9.55 (1.24)	9.45 (1.24)	8.85 (1.33)
Word1 frequency	104451 (395950)	104684 (395894)	104465 (395947)
Word1 frequency (log)	9.53 (1.5)	9.54 (1.52)	9.52 (1.52)
Word2 frequency	852095 (1771263)	848787 (1772802)	849151 (1772633)
Word2 frequency (log)	11 (2.57)	10.9 (2.67)	10.8 (2.84)
Word3 frequency	10249 (10042)	9171(9523)	12089 (12858)
Word3 frequency (log)	8.83 (0.91)	8.65 (1.00)	8.88 (1.06)
MI scores	3.33 (1.7)	3.56 (1.70)	3.57 (1.88)
Phrase-level frequency	158 (759)	51.7 (100)	52.0 (79.1)
Phrase-level frequency (log)	3.36 (1.37)	3.37(1.09)	3.29 (1.17)

Note. Frequency extracted from BNC with English Corpora and all frequency measures are expressed as raw frequency. Length is in number of characters.

Standard deviations were big due to including highly frequent words, e.g., *in, give, my, black.*

3.2.3.3 Context Sentences

Context sentence fragments were written for the 45 item triplets. Three different versions of context sentence fragments were created for each item triplet (see example in Table 3.3). Similar to previous studies (Carrol et al., 2016; Milburn & Warren, 2019; Siyanova-Chanturia, Conklin, & Schmitt, 2011; Tabossi et al., 2009), the pre-target contexts were kept as neutral and uninformative as possible to avoid the contextual predictability.

Each sentence consisted of six regions with the third region being the critical region, the fourth region being the spill-over region. The spill-over region was included because sometimes the differences in reading time can be delayed or spilled over to the next word (Jegerski, 2013). This way, the critical regions of interest never appeared at the beginning or end of the sentences, avoiding any wrap-up effects (Rayner et al., 2000).

Each item triplet appeared in three different sentence frames. Therefore, there were a total of 405 sentences (45 sets x 3 sentence frames x 3 item conditions) in this study (see Appendix 3.2). This design can increase statistical power. Table 3.3 demonstrates that each item triplet (*affect/enter/blows my mind*) is embedded within each member of the sentence frame. Following a Latin Square design, sentences were counterbalanced across three lists such that each list contained all three conditions of a triplet but presented in a different sentence frame (seen as in Table 3.3). For example, the list1 has "Walking down the street an idea affects my mind and I record in my diary"; "This quote, 'Racism is a scale', enters my mind and also helps me to understand this protest"; "This belief about social distancing blows my mind and I am not sure what to do now. The list2 has "Walking down the street an idea blows my mind and I record in my diary"; "This quote, 'Racism is a scale', affects my mind and also helps me to understand this protest"; "This belief about social distancing enters my mind and I am not sure what to do now.

Table 3.3

An Example of One Triplet of Items and Their Context Sentences

Regions of Context Sentences					
First	Second	Critical	Spill-over	Wrap-up	
Walking down the street,	an idea	affects/*enters*/**bows** my mind	and I	record it in my diary.	
This quote,	"Racism is a scale"	affects/*enters*/**bows** my mind	and also	helps me to understand this protest.	Does the quote influence me?
This belief about	social distance	affects/*enters*/**bows** my mind	and I	I am not sure what to do now,	

Note. The first words of the MWSs in the critical region are free combinations (underscored), restricted collocations (italicized), and idioms (bold).

For each item triplet, one comprehension question was created corresponding to one of three context sentences. Thus, the selected sentence frame was always followed by the same comprehension question in three lists. Each list had 45 comprehension questions, with 22 requiring a "No" response and 23 a "Yes" response.

Finally, 45 filler sentences were created, which constituted one-fourth of the trials in each list. This study used a self-paced reading paradigm coupled with a concurrent stop-making-sense task. To create the contrast between plausible and implausible in semantic meaning, filler sentences were designed to be semantically implausible, while context sentences were written to be semantically plausible and natural. One-third of the errors were placed at the beginning of filler sentences (e.g., *A piece of cake/will be interviewed soon, /so you can/write down/the number/and give us the information.*), another third in the middle of the sentences (e.g., *Primary school teachers suggest that/children go/to rock/concerts at/least five/times a week.*), and the rest of the errors were inserted towards the end of the sentences (e.g., *Before I realised/what was happening, /the man/had grabbed/my bag and/was talking with it.*). Following these steps, each list consisted of 135 critical sentences and 45 filler sentences.

3.2.4 Procedure

Participants were recruited online via Prolific (www.prolific.co), which is a web-based platform for recruiting participants for behavioural studies. Participants first completed ethics, including an informed consent form, and then completed a language background and usage questionnaire on Qualtrics. L2 participants were required to complete the VLT test after the questionnaire. At the end of questionnaire, participants were redirected to the online reading experiment.

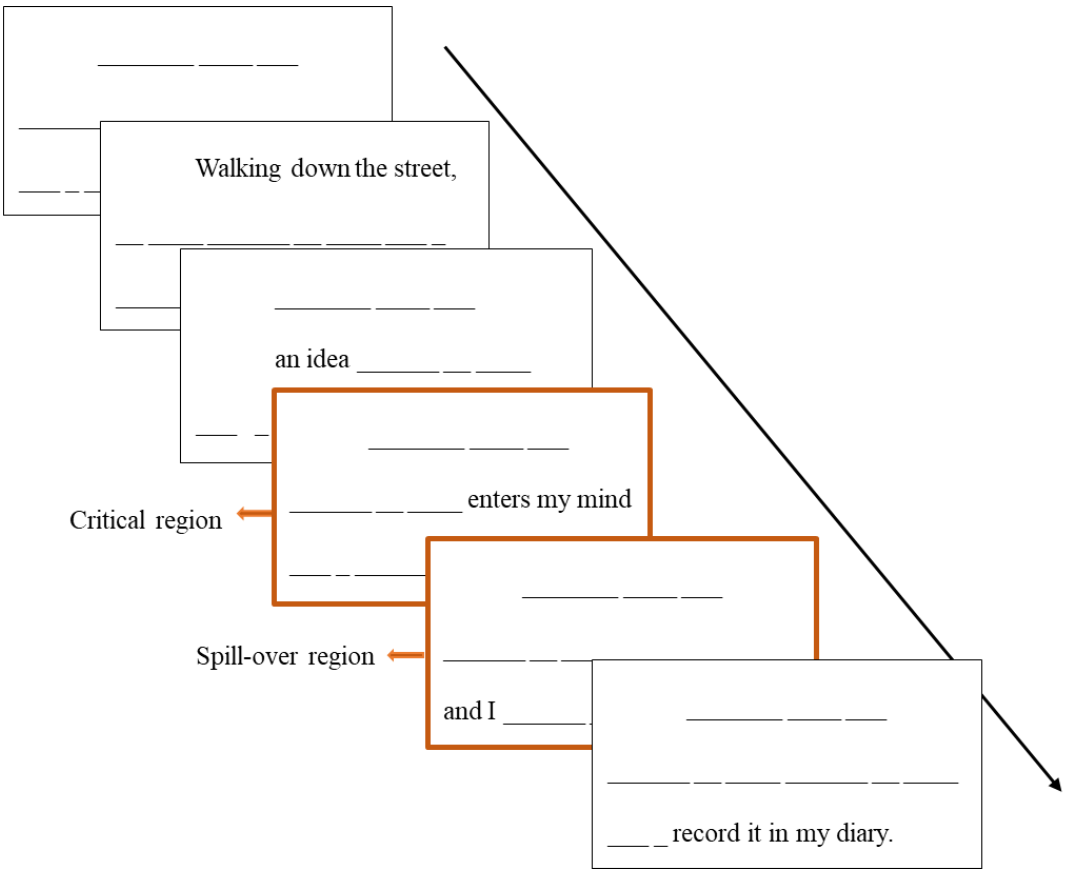
Participants first read instructions, then familiarised themselves with the procedure by means of 5 practice trials. During the practice session, participants were given their sentence reading times and their accuracy to the comprehension questions. Once participants were ready for the

main experiment, they were presented with 180 sentences in a randomized order. After finishing the reading experiment, participants were redirected to a Prolific page with a completion code.

The reading experiment used a self-paced reading paradigm coupled with a concurrent stop-making-sense task, programmed with PsychoPy software and hosted by Pavlovia. In this paradigm, self-paced reading was phrase by phrase and cumulative. Participants hit the space bar to reveal a trial. Each trial began with a dashed line on the screen, marking the length of the sentence. Each sentence was divided into six reading regions, and participants were instructed to press the "J" to reveal each region. As each region appeared, the preceding region was replaced by dashes (See Figure 3.2). Participants were asked to press the "J" if what they had read up to that still made sense. They were instructed to press a different key, "F", once what they read did not make sense to them. When "F" was pressed, the current trial was terminated, and the comprehension question was not displayed.

Figure 3.2

Presentation of the Procedure from One Trial



One-third of the context sentences were followed by yes/no comprehension questions within the experimental session. Participants were instructed to press "Y" or "N" to answer the questions. Participants were also informed that not all critical sentences were followed by comprehension questions. There were four opportunities for participants to take a break in the reading experiment. L1 participants required approximately 20 mins to complete the reading task, while L2 participants required around 30 mins.

3.2.5 Measurements

The region where MWSs were positioned was two to four words long. Given that self-paced reading effects are sometimes temporally delayed, a spill-over region was included, consisting of the two words following the critical region of interest. Hence, we analysed two regions of interest: 1) the critical region where MWSs appeared, and 2) the spill-over region.

3.2.6 Statistical Analysis

All analyses were performed using R version 4.2.1 (R Core Team, 2022) with two packages *lm4* (Bates et al., 2015) and *lmerTest* (Kuznetsova et al., 2020). As the dependent variable was continuous, RTs were log-transformed to fit linear mixed-effect models' assumptions (Brysbaert & Stevens, 2018; Lo & Andrews, 2015) by reducing skewness. First, models were fitted with L1 group and L2 group data separately, in order to see how L2 proficiency affected processing in the L2 speakers. Then, models were constructed with both L1 and L2 data combined to test whether there was a qualitive difference in processing of the different types of MWSs between L1 and L2 speakers.

For each dependent variable and each region of interest, the core model included Types-of-MWSs (treatment coded: free word combination as 0). A full random model was used whenever possible (Barr et al., 2013). If models failed to converge, "bobyqa" or "optimx" optimiser (Powell, 2009) was used to increase iterations, before simplifying random effects by removing interactions and then individual terms until the model converged.

The full L1 model included the core model and covariates including item length, individual word frequency, and phrase-level frequency. For L2 speakers, vocabulary score was entered into the models after taking care of the core model and the same covariates as in L1 models. For the combined L1 and L2 models, Language-background (treatment coded: L1 as 0) and the interactions with Language-background were entered into the models, after considering the core model and other covariates. All continuous predictors were centered by using the scale () function before adding into them to the models.

Covariates were subsequently reduced with the drop1 () function, which performs a model comparison with a chi-square test to select models. However, the experimental variables (Types-of-MWSs, vocabulary scores, Language-background) were always kept in the corresponding models to answer our research questions, even if these variables did not contribute to models significantly. Interactions with experimental variables were removed if they were not significant. Following these steps, the best-fitting models were obtained and variables with a *t*-value of greater than 2 were considered significant at the alpha = .05 level. The anova () function was further used to test the significance of interactions and main effects by comparing the reduced models with full models. When interactions were significant, they were further explored with plots using the plot_model () function from the *sjPlot package* (Lüdecke et al., 2023). Once the best-fitting models were obtained, the vif () function was used to check collinearity.

3.3 Results

3.3.1 Pre-processing Data

Participants were removed based on two criteria. First, participants with comprehension questions accuracy lower than 80% were removed from the final analysis. Low question accuracy implicitly reveals that participants did not pay enough attention during the experiment. At this stage, one L1 and seven L2 participants were removed. Second, we considered the percentage of pressing the "F" key. Participants who judged all semantically incorrect filler sentences plausible were excluded. This step removed five L1 participants. After removing these participants, the accuracy of comprehension questions showed that both groups of participants did not have difficulties understanding the sentences. The average accuracy rate for the comprehension questions of the remaining participants was 96% in the L1 group and 94% in the L2 group.

All filler sentences were removed, and the analyses were performed on the reading times for the remaining 135 critical items. Critical trials were removed if the associated comprehension question was answered incorrectly. Following Singh et al. (2016), critical sentences were removed if participants terminated the sentence before the end of the sentence (i.e., if they judged the sentence to have become non-sensical). For the remaining trials, RTs on the critical region larger than 2500ms and smaller than 200ms were removed. RTs on the spill-over region larger than 1500ms and smaller than 200ms were removed (the spill-over region was shorter than the critical region; hence the maximum cut-off was reduced). The outliers were detected by plotting boxplots (Howitt & Cramer, 2008; Levshina, 2015; Wilcox, 2012). We further excluded responses for which the RT was greater than 2.5 SD (Miller, 1991; Ratcliff, 1993) above and below the mean per participant in both critical and spill-over regions. After data cleaning, 19631 trials were retained for the critical region and 19685 trials for the spill-over region. The details of removing data for each group and conditions for each step are provided in Appendix 3.5.

Average reading times on the critical and spill-over regions are summarised in Table 3.4. Tables 3.5 and 3.6 report the best-fitting models for L1 and L2 speakers in the critical region, respectively. Table 3.7 and Table 3.8 summarise the combined models of L1 and L2 on the critical and spill-over regions, respectively.

Table 3.4

Descriptive Statistics: Mean RT (ms) and Standard Deviations for the Critical and Spill-over Regions across Three Conditions for Both L1 and L2 Speakers

group	**condition**	**critical regions**		**spill-over regions**	
		M	*SD*	*M*	*SD*
L1	free word combinations	781	311	593	181
L1	restricted collocations	764	305	580	179
L1	idioms	749	295	576	175
L2	free word combinations	892	398	637	209
L2	restricted collocations	899	407	628	208
L2	idioms	881	407	639	218

3.3.2 L1 Speakers

3.3.2.1 Critical Region

There was a main effect of Types-of-MWSs in the critical region ($\chi 2$ (3) = 6.20, p = .045), which indicates that L1 speakers processed idioms, restricted collocations, and free word combinations quantitatively differently. As shown in Table 3.5, idioms were read significantly faster than free word combinations (β = - 0.04, t = - 2.43, p = .015). Contradictory to the hypothesis, restricted collocations were read marginally faster than free word combinations (β = -0.03, t = -1.8, p = .072). However, there was no processing difference between idioms and restricted collocations (β = 0.009, t = 0.67, p = .506). Unsurprisingly, phrase length influenced reading times (β = 0.04, t = 7.19, p < .001), with longer items being read more slowly. Raw item frequency was entered into all models as a covariate. However, the covariate did not significantly contribute to modelling. The MWSs had similar effects in models with and without raw frequency.

3.3.2.2 Spill-over Region

There was no main effect of Types-of-MWSs on the spill-over region. Further analyses (see the best model in Appendix 3.3) showed that idioms were processed marginally faster than free combinations (β = -0.025, t = -1.75, p = .083)

Table 3.5

Summary of the Best-fitting Models of MWSs on the Critical Region for L1 Speakers

	L1 Critical Region			
Predictors	***β***	***SE***	***t***	***p***
(Intercept)	6.6	0.03	237.43	<.001
phrase length	0.04	0.01	7.19	<.001*
idioms	-0.04	0.01	-2.43	.015*
restricted collocations	-0.03	0.01	-1.8	.072
Random Effects				
σ^2		0.07		
$\tau_{00\ items}$		0		
$\tau_{00\ Subject}$		0.06		
ICC		0.46		
$N_{\ Subject}$		84		
$N_{\ items}$		133		
Observations		10531		
Marginal R^2 / Conditional R^2		0.016 / 0.470		

Note. The reference level is free word combinations, reading times are log-transformed, phrase length was centered.

3.3.3 L2 Speakers

3.3.3.1 Critical Region

There was no main effect of Type-of-MWSs; however, L2 speakers demonstrated an interaction between proficiency and Types-of-MWSs ($\chi 2(2) = 7.79$, $p = .020$), which is in line with the last hypothesis. The interaction is visualised in Figure 3.3. When L2 speakers were less proficient in L2, free word combinations were read faster than the other two types of MWSs. As the L2 proficiency increased, the faster reading of free word combinations compared to other types of MWSs gradually disappeared. In addition, the three types of MWSs were affected by L2 proficiency differently. Increased L2 proficiency sped up reading times for MWSs, but the effect was more distinctive for idioms than for the other two types of MWSs. Additionally, the significant effect of phrasal length showed that the longer items were read more slowly ($\beta = 0.05$, $t = 7.57$, $p < .001$).

Table 3.6

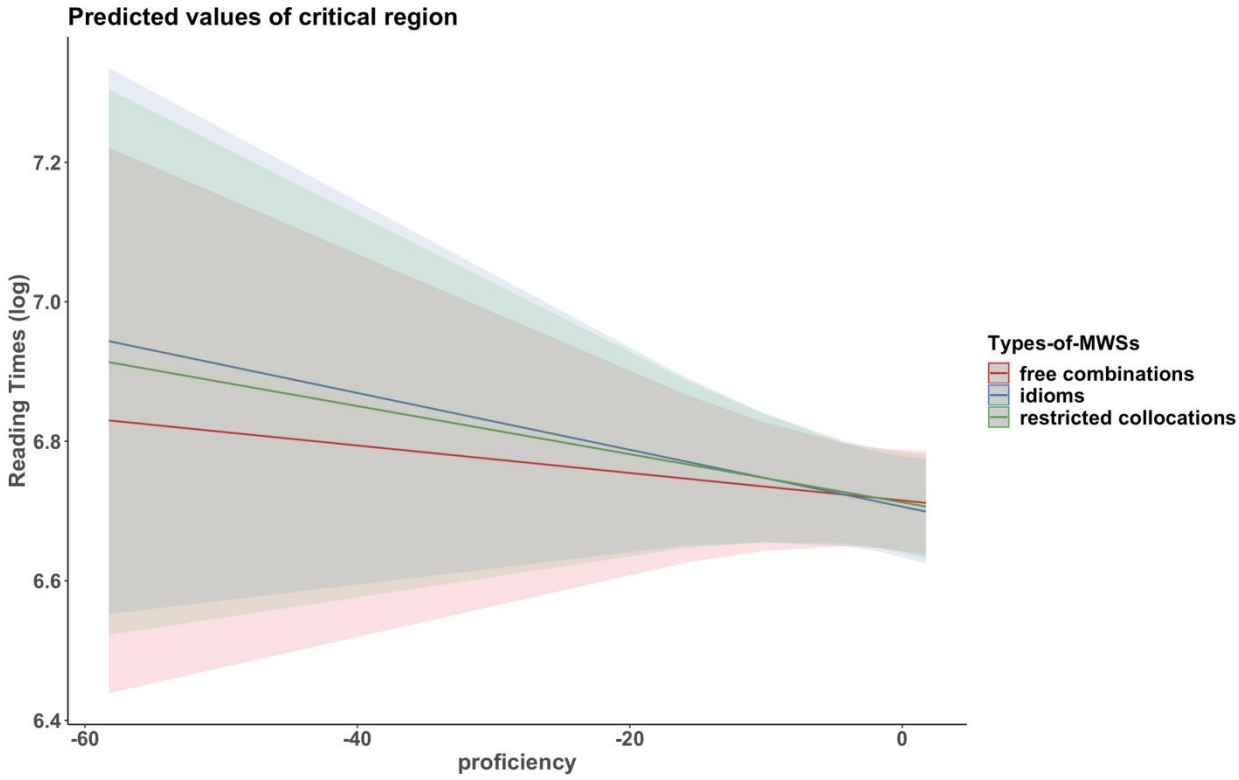
Summary of the Best-fitting Models of MWSs for the Critical Region for L2 Speakers

	L2 Critical Region			
Predictors	***β***	***SE***	***t***	***p***
(Intercept)	6.71	0.04	181.42	<0.001
idioms	-0.01	0.02	0.42	-.512
restricted collocations	-0.003	0.02	-0.155	.876
proficiency	-0.002	0.004	-0.57	.569
phrase length	0.053	0.007	7.57	<.001*
idioms * proficiency	-0.002	0.0008	-2.71	.007*
restricted collocations* proficiency	-0.002	0.0008	-1.93	.053
Random Effects				
σ^2		0.09		
$\tau_{00\ items}$		0.01		
$\tau_{00\ Subject}$		0.09		
ICC		0.51		
$N_{\ Subject}$		73		
$N_{\ items}$		133		
Observations		8863		
Marginal R^2 / Conditional R^2		0.021 / 0.517		

Note. The reference level is idioms, reading times are log-transformed, phrase length and proficiency were centered.

Figure 3.3

The Interaction between Types-of-MWSs and Proficiency for L2 Speakers



Note. Proficiency was centered and reading times were log-transformed.

3.3.3.2 Spill-over Region

L2 speakers did not show any effects on the spill-over region (see the best model in Appendix 3.4).

3.3.4 Comparison of L1 and L2 Speakers

3.3.4.1 Critical Region

As expected, the effect of Language-background was significant, with L2 speakers reading MWSs more slowly than L1 speakers (β = 0.12, t = 2.83, p = .005) (see Table 3.7). Furthermore, there was an interaction between Types-of-MWSs and Language-background ($\chi 2(2) = 8.88$, p = .012). As Figure 3.4 shows, the interaction was mainly driven by the difference between idioms and free word combinations in L1 speakers. L1 speakers processed idioms more easily than free word combinations, which was not found in L2 speakers. This finding is partially in line with the third hypothesis. L1 and L2 speakers indeed processed MWSs in different manners, with the faster processing of idioms only demonstrated with L1 speakers. However, different from the hypothesis, L2 speakers did not spend more time on idioms compared to other types of MWSs.

3.3.4.2 Spill-over Region

Likewise, Language-background was significant (β = 0.07, t = 2.33, p = .02), with L2 speakers reading MWSs more slowly than L1 speakers. The spill-over region also demonstrated a significant interaction between Language-background and the Types-of-MWSs ($\chi 2(2) = 6.74$, p = .034). Figure 3.5 shows that the interaction was driven by the different processing patterns within L1 and L2 speakers. It was more difficult for L1 speakers to process free word combinations than idioms and restricted collocations, while L2 speakers processed restricted collocations more easily than idioms and free word combinations. However, as separate analyses with L1 and L2 data showed, there was no significant processing difference within the three types of word combinations on the spill-over region.

Table 3.7

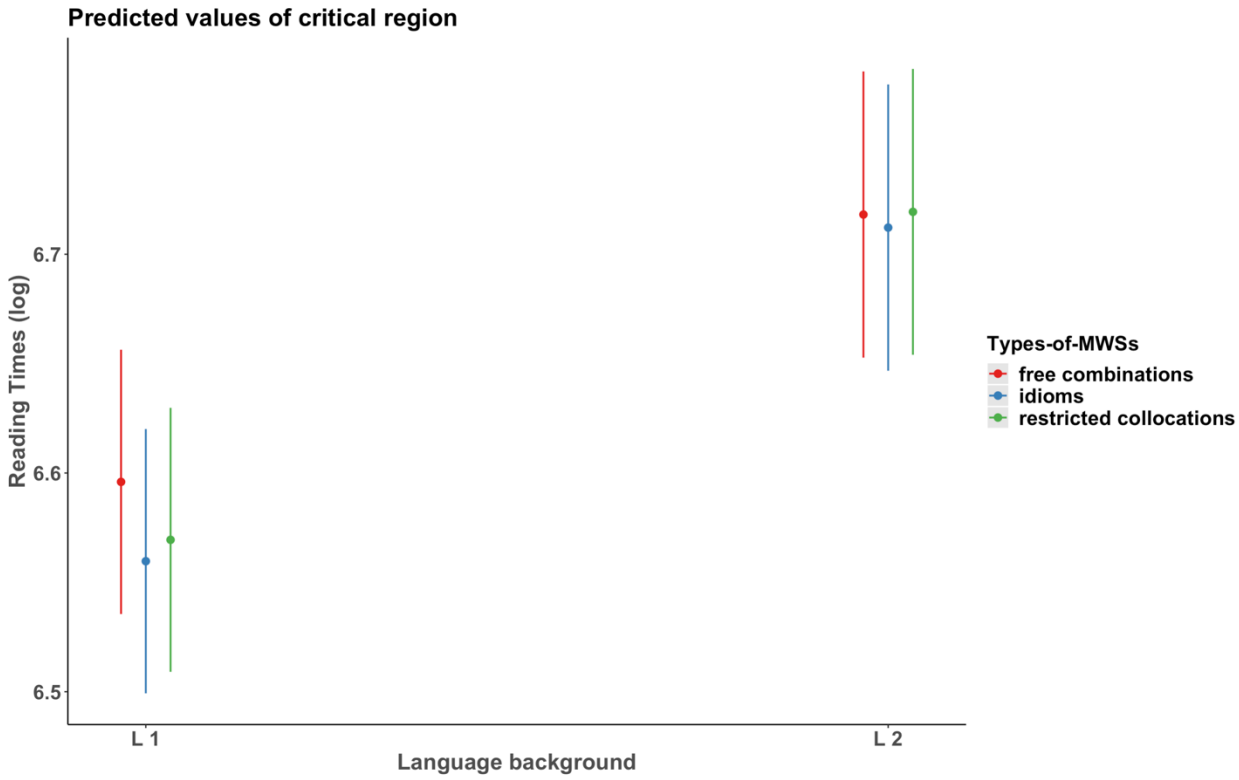
Summary of the Best-fitting Models of MWSs for the Critical Region for L1 and L2 Speakers

	L1 and L2 Critical Region			
Predictors	***β***	***SE***	***t***	***p***
(Intercept)	6.6	0.03	213.94	<0.001
language-background [L2]	0.12	0.04	2.81	.005*
idioms	-0.04	0.01	-2.42	.015*
restricted collocations	-0.03	0.02	-1.81	.07
phrase length	0.05	0.01	7.56	<0.001
language-background [L2] * idioms	0.03	0.01	2.87	.004*
language-background [L2] *restricted collocations	0.03	0.01	2.61	.009*
Random Effects				
σ^2	0.08			
$\tau_{00\ Subject}$	0.07			
$\tau_{00\ items}$	0			
$\tau_{11\ items.language_background\ L2}$	0			
$\rho_{01\ items}$	0.33			
ICC	0.49			
$N_{Subject}$	157			
N_{items}	133			
Observations	19394			
Marginal R^2 / Conditional R^2	0.05 / 0.51			

Note. Free word combinations and L1 group were the reference levels, reading times were log-transformed, the phrase length was centered.

Figure 3.4

The Interaction between Types-of-MWSs and Language-background on the Critical Region



Note. L1 refers to L1 speakers; L2 refers to L2 speakers; reading times were log-transformed.

Table 3.8

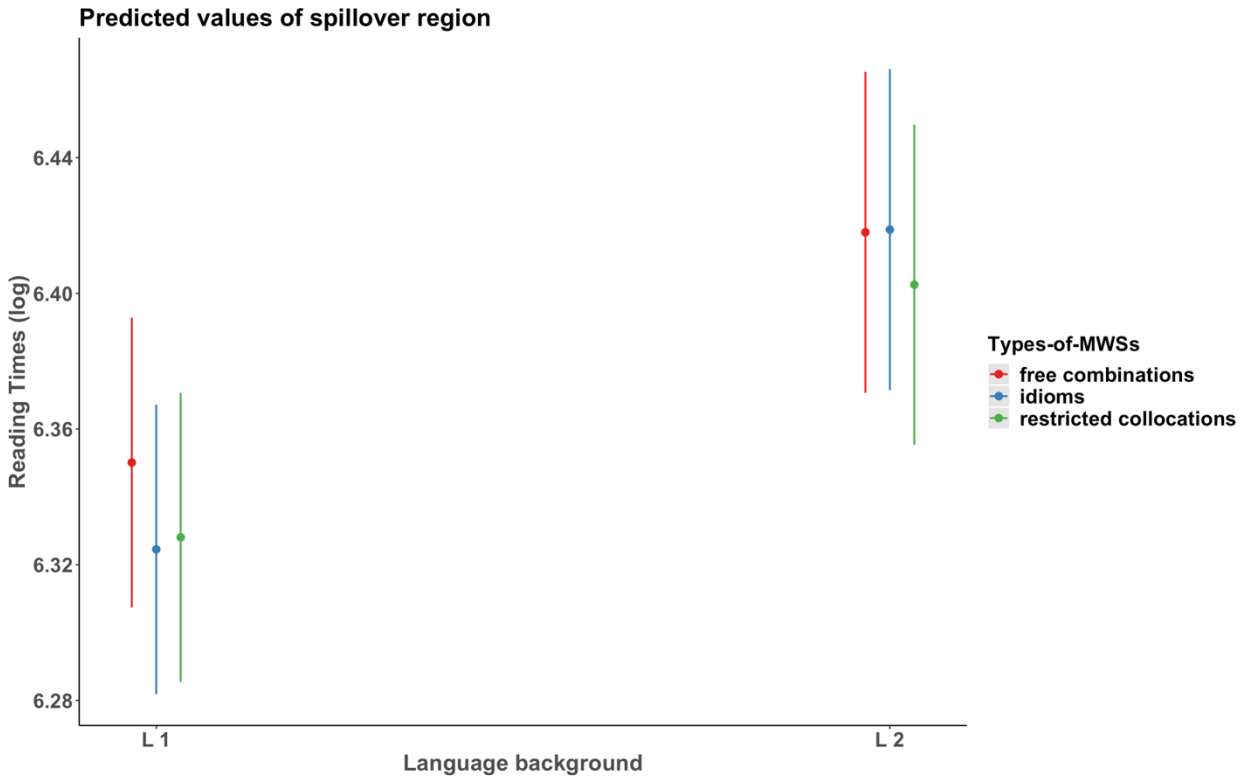
Summary of the Best-fitting Models of MWSs on the Spill-over Region for L1 and L2 Speakers

	L1 and L2 Spill-over Region			
Predictors	***β***	***SE***	***Statistic***	***p***
(Intercept)	6.3501	0.02	291.38	<0.001
language-background [L2]	0.068	0.03	2.33	0.02*
idioms	-0.02	0.02	-1.76	.079
restricted collocations	-0.02	0.01	-1.52	.129
phrase length	0.006	0.006	0.95	.341
language-background[L2] *idioms	0.0264	0.01	2.52	.012*
language-background [L2] *restricted collocations	0.0066	0.01	0.63	.526
Random Effects				
σ^2	0.05			
$\tau_{00\ Subject}$	0.03			
$\tau_{00\ items}$	0			
$\tau_{11\ items.language_backgroundL2}$	0			
$\rho_{01\ items}$	0.31			
ICC	0.42			
$N_{Subject}$	157			
N_{items}	133			
Observations	19160			
Marginal R^2 / Conditional R^2	0.019 / 0.430			

Note. Free word combinations and L1 group were the reference levels, reading times are log-transformed, phrase length was centered.

Figure 3.5

The Interaction between Types-of-MWSs and Language-background on the Spill-over Region



Note. L1 and L2 refers to L1 speakers and L2 speaker respectively; reading times were log-transformed.

3.4 Discussion

3.4.1 Summary of Findings

The current study used a self-paced reading task to investigate the processing of different types of MWSs to assess Howarth's phraseological continuum model (Howarth, 1998). Extending the study of Gyllstad and Wolter (2016), we manipulated the phrase-level frequency across free word combinations, restricted collocations, and idioms. In contrast to the results of Gyllstad and Wolter's study, the present study found no extra cost for L1 speakers in processing restricted collocations. Our results suggest a trend of restricted collocations to be processed faster than free word combinations (see Shi et al., 2023). We also found that L1 speakers read idioms faster than free word combinations in both the critical and the spill-over regions, which is in line with the hypothesis that the idiom superiority effect is the result of conventionality rather than the result of phrase-level frequency. For L2 speakers, the results were in line with the hypothesis that the processing of MWSs is dependent on L2 proficiency. Less proficient L2 speakers processed free word combinations the fastest and idioms the slowest. Increased L2 proficiency sped up reading MWSs, and the speed-up effect was more prominent in reading idioms. For the most proficient L2 speakers, all types of MWSs were processed similarly.

3.4.2 Restricted Collocations vs. Free Word Combinations

Our first research question was whether there was an extra processing cost for restricted collocations over free word combinations in natural sentence reading. Our results show that L1 speakers did not slow the reading of restricted collocations compared to free word combinations. This finding aligns with Shi et al. (2023), who also found a numeric difference in the processing of figurative and literal collocations in L1 speakers, with an advantage for figurative collocations. This finding indicates that the competition between the literal and figurative interpretations of the figurative component, or the competition between the concrete and the abstract senses of the de-lexical component, does not exert influence on the processing of restricted collocations. The current

finding contrasts with the findings of Gyllstad and Wolter (2016), who found that both L1 and L2 speakers processed free word combinations faster than restricted collocations. The current study and Shi et al.'s study used a self-paced reading task, whereas Gyllstad and Wolter employed a sensicality judgement task. The judgement task may conceivably give rise to ambiguities that lead to the processing cost for restricted collocations. The ambiguities can vanish when the context is present. One might argue that the self-paced reading task provided contextual information, which might not only diminish the ambiguities but also speed up the processing of restricted collocations. However, the current study and Shi et al.'s study ensured that the context was uninformative. Hence, the observed trend of faster processing of restricted collocations is less likely due to the context facilitation. The different findings between the two self-paced reading studies and the judgement study are more likely due to experimental items. The two self-paced reading experiments matched item-related variables on a pair-wise level (e.g., *build a house*/*career*; *blows*/*enters*/*affects one's mind*), which was not the case for Gyllstad and Wolter (e.g., *seize the opportunity*, *pay tax*).

3.4.3 Idiom Superiority Effect

The second research question was whether the idioms' superiority effect was still observed after phrase-level frequency between idioms and the other types of MWSs were controlled for. In the current study, L1 speakers demonstrated faster reading times for idioms compared to free word combinations, but there was no difference in processing time between idioms and restricted collocations. The faster processing of idioms compared to free word combinations is difficult to fit within frequency-based language models, as word-level and phrase-level frequencies were matched between idioms and free word combinations. Though we did not measure the familiarity of the idioms, it has been suggested that frequency is highly correlated to familiarity (Balota et al., 2001; Gernsbacher, 1984; Kuperman & Van Dyke, 2013). Hence, this finding suggests that additional properties of idioms contribute to the idioms' superiority effect and not just frequency of occurrence.

In the current study, an idiom differs from a free word combination in having a conventionalised and idiomatic meaning that cannot be derived from its components of the idiom. Due to the high degree of conventionality, when encountering idioms, language users might rely less on the computation of lexical parts (which is obligatory for free word combinations) but activate the idiomatic meaning more directly. Moreover, much research in psycholinguistics suggests that the computational/decomposition route and direct access route simultaneously and interactively contribute to recognising and decoding different levels of linguistic constructs e.g., inflexed words (Bertram et al., 2000; Sandra, 1990), compound words (Kuperman et al., 2009), idioms and (Cutting & Bock, 1997; Sprenger et al., 2006). The two routes are more likely to operate interactively on a continuous scale. In the processing of idioms, the direct access route that is associated with less cognitive cost takes precedence, due to the high conventionality. However, in the processing of free word combinations, the computational route that is seen to be less efficient takes priority. Therefore, the tendency of direct retrieval leads to the faster processing of idioms compared to free word combinations.

3.4.4 L2 Proficiency Effect

We detected an interaction between L2 vocabulary level (used as a proxy for proficiency) and the processing of the different types of MWSs. L2 speaker with low vocabulary level demonstrated a processing difference: free word combinations were read the fastest and idioms the slowest. With increased vocabulary level, the difference in processing times decreased. This result is in accordance with the interaction between proficiency and figurativeness suggested by Shi et al.'s study (2023). In their study, the slowing-down effect in processing figurative collocations diminished with increased L2 proficiency (measured by C-test[10]). Similar trends were observed in other studies on MWSs (Sonbul, 2015; Wolter & Yamashita; 2018), in which more proficient L2 speakers are more

[10] A paragraph-length text in which the second half of every second word is deleted. Participants reconstruct responses by completing the deleted parts of the words.

sensitive to phrase-level frequency or relied more on phrase-level frequency than word-level frequency. Taken together, there seems to be a shift away from the word-level analysis to the direct retrieval of larger phrases with increased L2 proficiency.

The shift described above is in line with the Dual-route Model (van Lancker Sidtis, 2012; Wray, 2002; Wray & Perkins, 2000) and the Literal Salience Model (Cieślicka, 2006). When L2 speakers are less proficient in their L2 and not familiar with L2 MWSs, the computational route is taken, in which L2 speakers mainly rely on a literal analysis. The initial integration of literal interpretations in figurative contexts is costly, as demonstrated by this study, with L2 speakers with low proficiency reading idioms the slowest. The integration cost decreases with gains in L2 proficiency. Though we did not observe the idioms' superiority effect in L2 speakers, more proficient L2 speakers can recover faster from the failure to integrate the literal interpretations into figurative contexts. In addition, it is more likely that the direct access route is available for more proficient L2 speakers, returning figurative meanings more quickly. These explanations are consistent with the observational evidence of the speeding up of processing idioms relative to free word combinations when proficiency increases.

3.4.5 Limitations

The present study has several limitations. First, similar to previous studies of idioms (Carrol et al., 2016; Milburn & Warren, 2019; Siyanova-Chanturia, Conklin, & Schmitt, 2011; Tabossi et al., 2009), the contextual predictability of the idioms was not measured. Though pre-target contexts were kept as neutral as possible in the present study, it is possible that there was variance in the predictability of the MWS. It should be noted that one could distinguish between the predictability of the MWS as a whole and the predictability of the individual words in an MWS. This will be addressed in Chapter 4 and Chapter 5.

Previous literature shows that contextual predictability is one of the fundamental linguistic variables that affect reading (Clifton et al., 2007; Rayner, 1998). Thus, prior contexts in this study

might have affected reading times for idioms, collocations, and free word combinations differently. In the next two experiments (Chapter 4 and Chapter 5), the effect of context was addressed, and cloze probability tests were carried out (Taylor, 1953). Before finalising experimental materials, a cloze probability test should be conducted by asking participants to fill in a word for an incomplete sentence up to the start of the MWSs (e.g., *This belief about social distancing* _____) and for the end of the MWSs (e.g., *This belief about social distancing blows/enters/affects* ____). To ensure that differences in reading time are not contaminated by differences in predictability, contextual predictability should be matched across different conditions.

Second, despite matching on both word- and phrase-level frequency between idioms, restricted collocations, and free word combinations were matched for, the subjective familiarity of MWSs was not measured. Although familiarity and frequency are often highly positively correlated, it is necessary to distinguish between these two factors in the studies of idioms. Brenner's study (2003) shows that idioms are widespread, at least in English, but individual idioms do not frequently occur in a corpus (Moon, 1998). In addition, research on single words has shown that corpus word frequencies are only an approximation of subjective familiarity (Gernsbacher, 1984). However, this approximation might be less accurate for low-frequency words since the familiarity of low-frequency words might be affected by individual differences in language experience. For example, it is likely that psycholinguists are more familiar with the word *stimuli* than corpus linguists because psycholinguists more often encounter and use this word than corpus linguists do. Similar to infrequent words, the corpus frequency of idioms might not be an accurate approximation of subjective familiarity ratings for idioms. The argument above suggests that language users can be more familiar with idioms even if idioms and collocations (and free word combinations) have a comparable frequency in corpora. To address this limitation, the familiarity of idioms should be normed with native speakers in future studies.

Third, for non-native speakers, the current study tried to ensure that L2 participants had no difficulties in comprehending idioms in two ways: (1) we selected familiar idioms and (2) the self-paced reading paradigm coupled with a concurrent stop-making-sense paradigm allowed L2 participants to terminate the sentence if they had difficulties in comprehending idioms. Sentences deemed acceptable provided implicit evidence that the idioms were known to the L2 speakers. In future studies, researchers can check whether L2 participants understand the selected idioms, after the reading experiment.

Fourth, the filler sentences were not pre-tested on how implausible they were. The present study used a self-paced reading paradigm coupled with a concurrent stop-making-sense task. Participants were instructed to press a key if what they had read up to that point still made sense and to press a different key to terminate a trial if what they had read up to that point did not make sense. We created filler sentences that we considered semantically implausible (to some degree), and we expected participants to press the different key. However, the results showed that many participants read implausible filler sentences to the end, which suggests that some filler sentences were still acceptable to participants or that participants used this "read-till-the-end" strategy, in which case they probably did the same with target sentences. For the experiment reported in Chapter 4, we ran a pre-test to assess the implausibility of the filler sentences.

Finally, the current study tested L2 speakers' proficiency with the 3K and 5K Vocabulary Levels Test (Schmitt et al., 2001) and observed an interaction between L2 proficiency and Type-of-MWSs. However, it is worth noting that the average vocabulary scores of L2 speakers (56) were similar to the scores of L1 speakers (59), both close to the full score (60). To avoid a potential ceiling effect, the Vocabular Level Test should include 10K, which would provide a more nuanced measurement of L2 proficiency.

To conclude, the present study compared reading patterns of different types of MWSs in L1 and L2 speakers. Contrary to previous work, we did not find a processing cost for restricted

collocations compared to free word combinations in a stop-making-sense sentence reading task. We did observe the expected idiom superiority effect, even after controlling for word-level and phrase-level frequency. This effect, which was only found for L1 speakers, suggests that other properties of MWSs beyond distributional information exert an influence on processing MWSs. Furthermore, we detected an interaction between L2 proficiency and different types of MWSs. Less skilled L2 speakers demonstrated a processing cost for idioms compared to free word combinations. The processing cost diminished when L2 proficiency increased.

CHAPTER 4 IDIOMATIC AMBIGUITY RESOLUTION

Abstract

Prior studies have shown that the processing advantages in MWSs over novel word combinations can be attributed to higher phrase-level frequency and familiarity. However, these studies typically focus on one type of MWS, and few studies have compared different types of MWS in one experiment. The current self-paced reading experiment assessed the processing of collocations (e.g., *vast area*), ambiguous idioms (e.g., *grey area*) used in a literal or figurative sense, and novel word combinations (e.g., *rare area*) in 87 L1 speakers and 114 L2 speakers. We observed faster reading times for idioms compared to collocations in L1 speakers, indicating that the idioms' processing advantage exists beyond mere distributional information and the resolution of idiomatic and lexical ambiguity overlaps. Moreover, both L1 and L2 speakers demonstrated an effect of semantic decomposability, which argues against the direct access of the idiomatic interpretations of idioms but supports the Idiom Decomposition Hypothesis. The decomposability effect was also mediated by the meaning dominance in both groups of speakers: more balanced literal-figurative interpretations for idioms were facilitated by the increasing semantic decomposability. In contrast, the faciliatory effect of semantic decomposability was diminished for idioms with highly frequent figurative interpretations.

Keywords: multi-word sequences, idiom processing, idiomatic ambiguity, semantic decomposability

4.1 Introduction

4.1.1 Multi-word Sequences

Reading requires processing and integrating the meaning of individual words in a sentence as the sentence unfolds. To some degree, this reflects semantic compositionality, where the meaning of a sequence of words is determined by the meanings of its sub-parts (Fodor & Lepore, 2002; Katz & Giesbrecht, 2006; Lin, 1999; Pelletier, 1994). There are, however, some exceptions. For instance, language users tend to bundle recurrent and familiar sequences of co-occurring words as a whole. These strings of recurrent word combinations are defined as Multi-word Sequences (MWSs) (Sag et al., 2002) and encompass a wide range of linguistic forms, for example, idioms, collocations, binominals, and lexical bundles (see below for discussion). There are two prevalent characteristics of MWSs (Carrol & Conklin, 2020). One characteristic is that MWSs are recurrent in language. The other characteristic is that language users read MWSs more quickly than novel word combinations.

Despite the two shared features mentioned above, MWSs vary in many aspects, such as the degree of conventionalization (i.e., the extent to which the form configuration or meaning of MWSs are static), semantic decomposability/transparency (i.e., the extent to which individual components of MWSs independently contribute to the overall meaning, e.g., *pop the question* is considered a decomposable idiom as *pop* and *question* semantically contribute to the idiomatic meaning; but *kick the bucket* is non-decomposable; some collocations are completely decomposable and transparent, such as *provide information*, *fresh apple*, but other collocations are less transparent and decomposable, such as *make friends*, *cast doubt*), and meaning dominance (i.e., the relative frequency of one meaning over others (Duffy, Morris, & Rayner, 1988). For idioms, meaning dominance is the extent that idioms are processed and interpreted figuratively as opposed to literally).

Due to the diversity and differences among types of MWSs, prior research has mainly focused on one type of MWSs and the effect of the distributional properties of MWSs on their

processing. To our knowledge, only one previous study investigated how and to what extent frequency and predictability can explain the processing advantage of MWSs over novel literal language (Carrol & Conklin, 2020). Their study tested the effect of phrase frequency and predictability on explaining the faster processing of idioms (e.g., *spill the beans*), collocations (e.g., abject poverty) and binominals (e.g., *king and queen*) over their corresponding controls (e.g., *drop the beans, prince and queen, total poverty*). Therefore, the design only allowed for a comparison of each type of MWSs to its respective control. Still, it did not allow for the comparison between types (e.g., idioms vs. collocations). The current study aimed to directly compare the processing of different types of MWSs with both L1 and L2 speakers when the distributional information was controlled.

4.1.2 Collocations

Collocations can vary in different aspects depending on different operational definitions. As discussed in Chapters 2 and 3, there are two distinct approaches to defining collocations: a corpus-based frequency method (Sinclair, 1991) and a phraseology-based method (Cowie, 1998; Howarth, 1998). Following the frequency method, frequently two co-occurring words than random chance, which is defined and identified as collocations (e.g., *provide information*, *fatal mistake*) that generally have transparent and literal meanings. In contrast, idioms have a figurative and opaque interpretation, and are therefore different from collocations.

Studies adopting a frequency method have shown a consistent collocation advantage in L1 speakers, with faster processing of collocations compared to novel word combinations (Durrant & Doherty, 2010; Li et al., 2021; Öksüz et al., 2021; Sonbul, 2015; Vilkaitė, 2016). However, when using Howarth's Continuum model (Howarth, 1998), the findings are less clear. Gyllstad and Wolter (2016) used a semantic judgement task with one three-level categorical variable: free word combinations (e.g., *write a letter*), restricted collocations (e.g., *break a promise*) and implausible word combinations (e.g., *cross a victim*). Though the phrase-level frequency and length were matched for

across the two critical conditions (i.e., free word combination and restricted collocation), this was not done pairwise. Both L1 and L2 speakers judged restricted collocations more slowly than free word combinations. The processing cost for semantically restricted collocations might have come from the component being used in a de-lexical or figurative sense, resulting in competition between the literal and figurative interpretation of the figurative component (e.g., *break* from *break a promise*) or between the more abstract and the more concrete meaning of the de-lexical component (e.g., *make* from *make progress*). In contrast, in a self-paced reading study, Shi et al. (2023), following Howarth's Continuum Model used collocations with one component used in a figurative sense (e.g., *build a career* where *build* is used figuratively). Shi et al. observed no significant difference between reading figurative collocations (e.g., *build a career*) and literal collocations (e.g., *build a house*) in L1 speakers. Together, these contrasting findings raise the question of whether additional properties, such as semantic properties indeed contribute to the processing of MWSs in the absence of differences in phrasal-level frequency.

Furthermore, one study suggests that meaningfulness contributed to the processing of MWSs (Jolsvai et al., 2020) when the distributional information was matched. In the phrasal-decision task, they compared idiomatic expressions (e.g., *play the field*, *on my mind*), compositional phrases (e.g., *nothing to wear*, *is really nice*), and phrasal fragments (e.g., *without the primary*, *because it lets*) while controlling phrase-level and word-level frequencies. The meaningfulness (was assessed by asking participants to rate each word sequences on a scale 1-7 according to how meaningful they were as a unit) demonstrates a facilitation effect, in which the more meaningful an MWS was rated to be, the faster participants were responded to. However, items from the study vary in syntactic structures (e.g., *play the field* is a verb + noun phrase, but *without the primary* is a prepositional phrase). Similar to Gyllstad and Wolter, though Jolsvai et al. matched items on phrase-level and word-level frequencies across three critical conditions, they did not match both frequencies pairwise. Any difference in performance can reflect either the experimental manipulation of

meaningfulness or the variants of the items. Thus, we need more evidence from comparing different types of MWSs when other properties of items are controlled for.

4.1.3 Models of Idiom Processing

One of the most studied MWS are idioms, which are characterised as a string of words whose figurative meaning cannot be derived from computing the individual components literally (Swinney & Cutler, 1979). Following this view, idioms tend to be fixed, semantically opaque, non-decomposable and have a "dead" or "frozen" figurative meaning (Moon, 1998). In comparison to idioms, frequency-based collocations tend to be transparent, and the collocational meanings can be deduced from their components (e.g., *provide information*).

4.1.3.1 Non-compositional Models

One of the central issues in idiom processing is whether language users activate both the literal and the figurative meaning of an idiom. Early non-compositional models suggest that idioms are represented as frozen strings and are retrieved directly from the mental lexicon (Bobrow & Bell, 1973; Gibbs, 1980; Swinney & Culter, 1979). Idioms, therefore, are represented as single and unanalysable units. The Direct Access Model (Gibbs, 1980) posits that figurative meaning can be directly accessed by by-passing the computation of the literal meaning (which is only computed if the figurative meaning turns out to be incongruent with the context). The Idiom List Hypothesis (Bobrow & Bell, 1973) assumes that idioms are processed and interpreted as "big words," thus no literal activation of lexical components is activated. The Lexical Representation Hypothesis (Swinney & Cutler, 1979) proposes the parallel processing of the literal and figurative interpretations of an idiom. Moreover, direct retrieval of figurative meanings is assumed to be faster than the computation of literal meanings. Though these early models differ in the order and time course of the retrieval of literal and figurative meanings, they agree that the computation of the literal meaning contributes a limited role in accessing the figurative meaning of an idiom and that the idiomatic process is independent of the literal one.

4.1.3.2 Hybrid Models

In contrast to early models, the Configuration Hypothesis (Cacciari & Tabossi, 1988) and the Hybrid Representation Hypothesis (Cutting & Bock, 1997; Sprenger et al., 2006) argue that idiom comprehension involves a compositional analysis of the literal meaning. According to the Configuration Hypothesis, the literal process progresses until the idiomatic key (i.e., the place where the expression can be recognised as an idiom) is reached. Once the idiom is recognised, the idiom might be accessed directly from the mental lexicon. Take this idiom (i.e., *see which way the* | *wind is blowing*. The vertical bar is where the idiomatic key is placed.) as an example. When "see" is encountered, the obligatory literal analysis and computation starts until reading up to the idiomatic key "the." After reaching the idiomatic key, the combination of the sequence "see which way the" activates the known configuration of the idiom (i.e., *see which way the wind is blowing*). The sequences after the idiomatic key (i.e., *wind is blowing*) can be automatically triggered/retrieved, which returns the idiomatic meaning more quickly than the computation of the rest of the lexical parts. The Hybrid Representation Hypothesis (Cutting & Bock, 1997) suggests that idioms resemble words and are directly associated with their unitary conceptual representations. At the same time, idioms resemble complex syntactic structures, subjected to compositional analysis. Sprenger et al (2006) extended the hybrid model by introducing the term "superlemma", a representation of the idiomatic phrase on the lexical-syntactic level. The "superlemma" contains the syntactic properties of the idiom and is connected to the components of the idiom. Moreover, the introduction of a "superlemma" allows idioms to compete with single words, such as "hit the road" can compete with "leave." When encountering an idiom, for example, "hit the road," the underlying "superlemma" would be activated. Thus, according to the hybrid models, idioms seem to be simultaneously compositional and non-compositional (Kuiper et al., 2007).

Much research has provided evidence supporting the view that the literal analysis of the components plays a role in comprehending idioms (Cacciari & Corradini, 2015; Cutting & Bock, 1997; Holsinger & Kaiser, 2013; Smolka et al., 2007; Titone et al., 2019). Using the self-paced reading

paradigm, Holsinger and Kaiser (2013) investigated the cost of meaning recovery in ambiguous phrasal verbs (e.g., *dig into*). Preceding contexts biased towards the literal, or the figurative interpretations were matched or mismatched to literal and figurative disambiguation contexts after the ambiguous phrases (e.g., ***The hungry waitress/The daring archaeologist*** *who had been working all day* <u>*dug into*</u> *the* ***sandwich/tomb*** *just after noon on Sunday*; the match or mismatch regions are in bold, the idioms underlined). Regardless of whether prior context biases towards a literal or figurative interpretation of the ambiguous phrase, the phrase was read equally fast when interpreted literally. This indicates that contextual bias does not play a strong role in determining the literal meaning of an idiom. On the contrary, contextual bias affects the resolution of the idiomatic meaning, with participants reading an ambiguous phrase faster when the preceding context supports the idiomatic interpretation compared to when it supports the literal interpretation. Taken together, literal interpretation of an ambiguous idiom is obligatory and is computed prior to the idiomatic retrieval.

However, the study (Holsinger & Kaiser, 2013) did not consider the frequency of the use of literal and figurative interpretations. It is possible that phrasal verbs in the study were used more frequently in their literal interpretations (e.g., *live on*, *come into*) and the idiomatic interpretations are less conventionalised, thus leading to the primary activation of literal interpretation. According to the graded salience hypothesis (Giora, 2002), the salient meaning (i.e., the most interpreted meaning) that is strongly associated with the expression is activated first regardless of the disambiguating contexts. If the literal use of an idiom is more frequent than the figurative use, the salient meaning could be the literal interpretation and be accessed first. Furthermore, research on ambiguous words (Duffy et al., 1988) also demonstrates that the most frequent meaning of an ambiguous word is more easily activated (the lexical ambiguity will be discussed in the next section). Therefore, it remains unclear whether a change in the idioms' meaning dominance would affect the ease of accessing the literal or figurative meaning.

4.1.4 Ambiguity Resolution

4.1.4.1 Lexical Ambiguity Resolution

Ambiguity is common in language and exists at both the single word level (lexical ambiguity, e.g., *bank*) and multi-word level (idioms, e.g., *grey area*). Since both lexical and idiomatic ambiguity have an expression linked to multiple concepts, it seems reasonable to draw a comparison between lexical and idiomatic ambiguity resolution. Earlier research on lexical ambiguity examined two key factors: the location of the context (before or after the ambiguous word) and the type of ambiguous words (balanced or unbalanced). Since the current study provided a disambiguating context prior to an idiom, the rest of the paragraph will only review the supporting context prior to unbalanced ambiguous words. First, the context can precede the target ambiguous word (prior biasing context) or follow the word (neutral prior context). Second, based on the relative frequency of the alternative meanings, ambiguous words can be grouped into unbalanced and balanced (Rayner & Duffy, 1986). An unbalanced, ambiguous word has one dominant meaning, which is more frequent than the other (subordinate) meaning. However, when the two meanings of the ambiguous word are relatively equal in frequency, we call them balanced ambiguities.

When an unbalanced ambiguous word follows a biasing context, reading times are longer when the prior context biases towards the subordinate meaning compared to the dominant meaning of that word or a frequency-matched control word (Colbert-Getz & Cook, 2013; Rayner et al., 1994; Sereno et al., 1992). To explain this processing difficulty, Duffy et al. (1988) argue that when preceding context biases towards the subordinate meaning, access to the subordinate meaning is facilitated, i.e., the subordinate meaning receives a contextual boost (top-down). The dominant meaning is still activated due to its higher frequency (bottom-up). The competition between these two meanings or the inhibition of the dominant meaning leads to difficulty processing an unbalanced, ambiguous word preceded by prior context supporting the subordinate meaning. Research also shows that the processing difficulty can be diminished when unbalanced ambiguous words have more frequent subordinate meanings (Wiley & Rayner, 2000). In addition, a

recent eye-tracking study shows that a strongly biasing context towards the subordinate meaning can generate strong facilitatory access to the subordinate meaning with only minimal interference from the dominant meaning (e.g., Colbert-Getz & Cook, 2013).

4.1.4.2 Idiomatic Ambiguity Resolution

Similar to unbalanced ambiguous words, most ambiguous idioms also have one dominant interpretation. Concretely, the figurative interpretation of the string tends to be more frequent than the literal interpretation (Cronk et al., 1993). The retrieval of the literal and figurative interpretation of an idiom can be guided by different contexts. Prior biasing contexts can drive an idiom's interpretation towards the literal or figurative interpretation (Holsinger, 2013; Titone & Connine, 1999). For instance, Holsinger (2013) found that after listening to an idiom in a figurative context, participants were more likely to look at a word figuratively related to the idiom. After listening to an idiom in a literal context, participants tended to look at a word literally related to the idiom.

However, there are inconsistent findings on whether an idiom following a prior figuratively biased context is processed faster than the idiom following a prior literally biased context. Two earlier studies observed that the figurative interpretation was faster than the literal interpretation (Gibbs, 1980; Ortony et al., 1978). Like unbalanced ambiguous words, the figurative interpretation is more easily activated by the supporting prior figurative context due to the higher frequency of the figurative interpretation. Prior literal context facilitates the retrieval of the literal interpretation, but the strong figurative interpretation is still activated. This results in a competition between two interpretations and an inhibition of the figurative interpretation. Hence, the literal interpretation of the idiom takes longer to process than the figurative one, even in a literally biased context. In contrast, two eye-tracking studies showed no reading differences between literal and figurative contexts (Carrol & Conklin, 2017; Siyanova-Chanturia, Conklin, & Schmitt, 2011). The different findings could be attributed to the strength of the context and meaning dominance. Research in lexical ambiguity shows that the processing cost can be reduced if the subordinate meaning from an unbalanced ambiguous word becomes more frequent (e.g., Wiley & Rayner, 2000).

Importantly, lexical ambiguity research compares reading times in three conditions: an ambiguous word following the prior context supporting either a dominant or a subordinate meaning and a frequency-matched control word. However, to our knowledge, no study has contrasted ambiguous idioms with frequency-matched controls (phrase-level frequency matched for).

4.1.4.3 Additional Factors in Resolving Idiomatic Ambiguity

Apart from the role of context in idiomatic ambiguity, many other factors can mediate the activation of the literal and figurative interpretations (Cacciari & Corradini, 2015; Milburn & Warren, 2019; Schweigert, 1986; Titone & Libben, 2014) The next section will discuss how familiarity, predictability, meaning dominance, semantic decomposability, and their interactions with context exert influence on the activation of the literal and figurative interpretations of an idiom.

Familiar idioms are consistently easier to be comprehended than less familiar idioms (M. R. Libben & Titone, 2008; Schweigert, 1986). Though familiarity speeds up idioms processing, the familiarity effect does not interact with context. Regardless of idioms being interpreted literally or figuratively, more familiar idioms require less effort to be processed (Schweigert, 1986). In contrast, Cacciari and Corradini (2015) found an interaction between context and the idiom's predictability (i.e., the likelihood of completing the idiomatic fragment with the expected idiomatic word, e.g. the probability of saying *road* following *hit the*). For highly predictable idioms (those with a cloze probability of over 70%), a prior literal context slowed down processing compared to a prior figurative context. For less predictable idioms (with a cloze probability of less than 30%), a prior literal context resulted in faster processing than a prior figurative context. Their findings indicate that language users are more committed to the figurative meaning of predictable idioms. However, given that familiarity is highly positively correlated with predictability, it is unclear why context interacts with idiom predictability but not with idiom familiarity.

Furthermore, one eye-tracking study revealed that the dominance of one interpretation of an idiom (i.e., what percentage of the time participants expect an idiom to have either a literal or figurative interpretation) interacted with semantic relatedness of the alternative interpretations

(Milburn & Warren, 2019). First, idioms were read faster when their figurative and literal interpretations were more related. Second, idioms with a more dominant figurative interpretation were read faster than idioms with balanced literal and figurative interpretations. Importantly, high semantic relatedness of the two interpretations slowed down processing when idioms were more often used figuratively, while high semantic relatedness sped up processing when idioms when idioms appeared in literal contexts more frequently. This study sheds light on the role of semantic relation in comprehending idioms. As explained next, the most widely used method to assess semantic relation in idioms is semantic decomposition.

4.1.4.4 Decomposability Effect

It has been suggested that semantic decomposability might affect accessing idioms' meanings. Decomposability was introduced by Gibbs and Nayak (1989) and Gibbs et al. (1989) proposed the Idiom Decomposition Hypothesis i.e., the extent to which components of an idiom directly relate to the idiomatic interpretation. Under this definition, *pop the question* is a decomposable idiom, as the two components *pop* (is related to suddenly asking) and *question* (is related to a kind of question) are closely related to the idiomatic interpretation (i.e., *ask someone to marry you*). On the other hand, *kick the bucket* is a non-decomposable idiom, as *kick* and *bucket* are not related to the idiomatic interpretation (i.e., *die*). The Idiom Decomposition Hypothesis proposed the processing of decomposable and non-decomposable idioms in a different manner. Understanding decomposable idioms benefits from the compositional analysis of the components since the components' assembled meaning can be used quickly to activate and recognise the figurative meaning. In contrast, understanding a non-decomposable idiom is more difficult than a decomposable idiom since the analysis of the literal components is incongruent with the overall figurative meaning. In other words, the incongruence between the overall figurative meaning and the meaning that is computed based on the literal components leads to disruption and additional cost.

However, evidence of the role of decomposability in idiom processing is contradictory. Some studies reported a decomposability facilitation effect with decomposable idioms processed faster than non-decomposable ones (Caillies & Butcher, 2007; Gibbs & Nayak, 1989; Hamblin & Gibbs, 1999; M. R. Libben & Titone, 2008; Milburn et al., 2021; Titone et al., 2019). On the other hand, other studies failed to detect the effect (Carrol & Conklin, 2014b; Tabossi et al., 2008), while one study (Titone & Libben, 2014) observed a decomposability inhibition effect, suggesting that increased decomposability can even slow down idiom comprehension.

The conflicting results of semantic decomposability might come from the different operationalisations. Some studies categorise idioms into non-decomposable and decomposable binary groups based on a certain (arbitrary) cut-off point (Caillies & Butcher, 2007), while other studies measured semantic relatedness on a 1-7 scale (e.g., Milburn & Warren, 2019) or on a 0-100 scale (Titone & Libben, 2014) and treated it as a continuous variable. Apart from the difference discussed above, idiom decomposability has been mainly measured globally by assessing the extent of assembled literal meaning mapping to overall figurative meaning (e.g., assembled meanings of *pop* and *question* mapping to *propose marriage*), whereas the Idiom Decomposition Hypothesis highlights the contribution of the individual components of an idiom to its figurative meaning (e.g., the extent of *pop* and *question* related to the idiomatic interpretation). It has been criticised whether the use of a global measure is sensitive enough to capture the semantic decomposability and whether a more local measure is more appropriate (Libben & Titone, 2008).

4.1.5 Collocation and Idiom Processing in the Second Language

Much of the research on L2 processing of collocations follows the corpus-based approach to defining collocations and demonstrates a well-documented facilitation effect for collocations over novel word combinations. This has been found in judgement tasks (e.g., Siyanova & Schmitt, 2008), priming studies (e.g., Durrant & Doherty, 2010) and eye-tracking studies (e.g., Li et al., 2022; Sonbul, 2015; Vilkaitė & Schmitt, 2019). Like L1 speakers, L2 speakers are also sensitive to phrase-level

frequency. With enough exposure to collocations, L2 speakers also start reading collocations faster than novel word combinations. In addition, the collocation facilitation effect in L2 readers is also modulated by other factors, such as the L1 and L2 proficiency. Those factors have been discussed in Chapter 2 and Chapter 3.

Although much of the research has revealed a clear advantage for L1 speakers when processing idioms compared to control novel word combinations (van Lancker Sidtis, 2012; Wray, 2012), the findings are mixed for L2 speakers. Since idioms are familiar and known to L1 speakers, the overall higher familiarity contributes to L1 speakers' advantage in idiom processing (Siyanova-Chanturia & Martinez, 2015; Tabossi et al., 2009). In contrast, L2 speakers do not demonstrate a reliable advantage, possibly due to the absence of strong connections between the idiom's form and the idiomatic interpretation. For example, an eye-tracking study compared ambiguous idioms (e.g., *at the end of day*) to matched novel word combinations (e.g., *at the end of war*) in a short discourse (Siyanova-Chanturia, Conklin, & Schmitt, 2011). Prior context sentences were biased towards either a literal interpretation or a figurative interpretation of an idiom. L1 speakers read idioms in both literal and figurative interpretations more quickly than matched novel word combinations. However, L2 speakers did not demonstrate the faster processing of the idiomatic interpretations over matched control word combinations. Moreover, L2 speakers sped up reading the literal interpretations of idioms in comparison to the figurative ones. Of note, several limitations in the eye-tracking study might challenge the results. First, as mentioned in the limitation of Chapter 3, though the familiarity of idioms was assessed by rating on a scale of 1 to 4 prior to the reading experiment and frequent idioms were selected, idioms still were more likely to be unknown or unfamiliar to L2 participants. Low-frequency linguistics units were influenced more significantly by individual differences in language experience. Second, the eye-tracking study did not measure contextual predictability. Relatively long idioms in the study were embedded in long discourses. It has long been known that the longer a sentential fragment, the fewer the available alternatives in the language to continue

and/or conclude it in a syntactically well-formed and semantically meaningful fashion (Miller & Selfridge, 1950).

Two models (the Dual-route model and the Literal Salience model) and the hypothesis in decoding complex words can explain the lack of idiom advantage in the processing of L2 idioms. On the one hand, the Dual-route model (van Lancker Sidtis, 2012; Wray, 2002; Wray & Perkins, 2000) proposes that the direct retrieval route is available if idioms have been encountered previously. This route allows language users to recognise the canonical form and directly access the conventional (idiomatic) meaning. Similar to the processing of complex words, multiple routes are available, which interactively affect the processing (Kuperman et al., 2009). The direct access route is preferred as the computational load is reduced and the activation of a word is faster. On the other hand, the computation route is necessary when idioms are less known to the language user. For known idioms, both routes will compete and interact, with the direct route generally winning out. However, for unknown idioms, the (slower) computational route is the only route available. Due to the infrequent exposure to idioms, L2 speakers are more likely to take the computation route, leading to the absence of a processing advantage for idioms over control novel word combinations. Additionally, Cieślicka (2006) proposed the Literal Salience model, in which L2 speakers mainly rely on a literal analysis to compute idiomatic meanings. For example, when L2 speakers encounter *As the law is not totally transparent, there is a grey area that we haven't noticed*, they compute the literal meaning of *grey area* primarily. Only if the integration of the literal meaning fails does the computation of figurative meaning happens. Furthermore, when considering MWSs processing more generally in L2, L1 and L2 speakers seem to process MWSs fundamentally differently (Cieślicka, 2006; Wray, 2002), with L1 speakers mapping meaning to largely linguistic chunks while L2 speakers heavily relying on assembling the meanings of MWSs' components (Wray, 2002, 2008).

4.1.6 The Present Study

The present study used a self-paced reading paradigm paired with the stop-making-sense task to compare different types of MWSs i.e., collocations and idioms in literal and figurative interpretations to novel word combinations[11]. The first question is whether idioms have a processing advantage over other types of MWSs when the distributional information (i.e., the phrase-level frequency, familiarity, and contextual predictability) is controlled for. The hypothesis corresponding to the first question is: if the distributional information is the main driver of the idiom processing advantage, then we do not expect any processing differences between idioms and other types of MWSs. If additional properties of MWSs, such as conventionality, contribute to idiom processing, we expect to detect a processing advantage for idioms over collocations, at least in L1 speakers. The second question is to what extent lexical and idiomatic ambiguity overlap. The hypothesis corresponding to the second question is: if the processing of lexical and idiomatic ambiguities is similar, then we predict that there will be a processing cost for idioms used literally compared to collocations and idioms used figuratively, at least in L1 speakers. However, if lexical and idiomatic ambiguities are processed differently, then we might fail to detect that L1 speakers slow down reading idioms used literally relative to collocations and idioms used figuratively. The third question is how meaning dominance and semantic decomposability interact to resolve idiomatic ambiguity. If the current study mirrors the effect of semantic relatedness from Milburn and Warren (2019), then we expect an interaction between meaning dominance and semantic decomposability in L1 speakers at least. The fourth question is to what extent MWSs' processing overlaps between L1 and L2 speakers. This question is exploratory, and we do not have a directional hypothesis.

[11] Novel word combinations are not one type pf MWSs, they act as one comparison level in the current study.

4.2 Method

4.2.1 Design

This study used the self-paced reading paradigm coupled with a concurrent stop-making-sense paradigm (Boland et al., 1990; Gibson et al., 1996). To illustrate, participants read sentences at their pace region by region and decided on whether they read still made sense in English. This paradigm is potentially more sensitive in measuring processing than a standard moving-window self-paced reading task and can keep participants more engaged with reading sentences (Singh et al., 2016). In addition, it allows participants, particularly L2 participants, to terminate sentences when they had difficulty in comprehending them.

When items like idioms are infrequent in the language, there can be significant individual variability between the language users. Idioms pre-tested to be familiar with one group of L2 speakers might be unknown to other individual L2 speakers, such as those who participated in the reading experiment. To ensure L2 participants knew the idioms, they were instructed to press the "No" key when facing comprehension difficulties. After completing the reading experiment, L2 participants were also assigned a multiple-choice test to check whether they understood the idioms involved in the reading task. In the multiple-choice test (See Appendix 4.4), sentence fragments, e.g., *Since Lucas is always so active and energetic, we need to manage the live wire…*were provided. Participants were asked to choose the most likely interpretation of the last two words *live wire* in the sentence out of three interpretations. The three interpretations contained one literal interpretation of *live wire,* one figurative interpretation of *live wire* and one unrelated interpretation. Trials for which L2 participants gave the wrong answer in the multiple-choice test and those for which L2 participants stopped reading were removed from the final analysis.

Language-background was a between-group factor with two levels: English native speakers and L1-Chinese L2-English speakers. Types-of-MWSs was one four-level within-participant and between-item variable (idioms used figuratively vs. idioms used literally vs. collocations vs. novel

word combinations). For idioms, context was a two-level factor, and meaning dominance and decomposability were continuous predictors. The dependent variable was reading times on the critical region and spill-over region.

4.2.2 Participants

Eighty-seven native British English speakers (age: *M* = 31.9, *SD* = 5.25) and 114 Chinese-English bilingual speakers (age: *M* = 29.90, *SD* = 7.62) participated in the present study. These participants did not participate in the norming study described below. There is a larger number of Chinese-English bilinguals than native English speakers to ensure the dataset was balanced since L2 speakers were more likely to judge sentences as implausible and were more likely to provide wrong answers to the multi-choice questions in the questionnaire. The Vocabulary Levels (VLT) of 3k, 5k and 10k (Schmitt et al., 2001) were administered to measure L2 proficiency. In Chapter 3, we tested the English proficiency level for L2 participants with 3k and 5k, which shows that their proficiency scores were close to the maximum scores. For this reason, the 10k level was added to the test of this study and the eye-tracking study (Chapter 5). The average VLT score was 53.82 (out of 90). On average, L2 speakers started learning English at age eight and spent 7.62 years in an English-speaking country. Participants' language background and usage data are summarised in Table 4.1.

The experiment was conducted in accordance with the British Psychological Society's ethical guidelines and was approved by the University of Birmingham's ethical committee (ERN_19-0936).

Table 4.1

Summary of Language Proficiency Data

	L1 speakers (*M*, *SD*)	L2 speakers (*M*, *SD*)
Age of English acquisition	1.51 (3.02)	8.05 (3.96)
Time lived in an English-speaking country (in years)	30.5 (5.3)	7.62 (7.84)
VLT score (out of 90)	84.1 (4.49)	53.82 (17.82)
Self-reported English proficiency (a 6-point scale)		
Proficiency in writing	4.97 (0.18)	3.62 (0.83)
Proficiency in speaking	4.93 (0.25)	3.59 (0.90)
Proficiency in listening	4.98 (0.15)	3.89 (0.83)
Proficiency in reading	4.95 (0.21)	4.00 (0.71)
Self-reported English language frequency of use (per day)		
Writing (in hours)	2.7 (2.94)	2.22 (2.06)
Speaking (in hours)	8.41 (5.55)	3.57 (3.58)
Watching/listening (in hours)	4.24 (3.04)	4.22 (4.69)
Reading (in hours)	2.60 (2.20)	3.38 (2.57)

4.2.3 Materials

There were 40 sets of items, with four conditions for each (see Table 4.2). Types-of-MWSs was one four-level within-participant and between-item variable. Each set of items included an idiom used figuratively, an idiom used literally, a collocation, and a novel word combination (as the control). All items were adjective-noun pairs (e.g., *grey area*, *vast area*, *rare area*).

Table 4.2

An Example of One Set of Items and Context Sentences

Condition	Sentences			
	disambiguating region	**control region**	**critical and spill-over region**	**wrap-up region**
novel word combinations	Although most of it is known to us,	there is a	rare area **that we**	have not noticed.
collocations	As Russia is such a big country,		vast area **that we**	
idioms (in literal context)	As the wall is not painted properly,		grey area **that we**	
idioms (in figurative context)	As the law is not totally transparent,		grey area **that we**	

Note. Key regions consisted of critical items (underscored) and spill-over regions (in bold).

To arrive at the final list of candidate items, three steps were followed.

1. The most frequent 2000 lemmas of adjectives and nouns (Leech et al., 2001) in the British National Corpus (BNC) were consulted to make sure the components in the MWSs were frequent. Following this, 73 idioms were selected from previous research (Macis & Schmitt, 2017) and the Oxford Dictionary of English idioms (Ayto, 2009) (see Appendix 4.1 for 73 sets of items). All idioms were ambiguous between an idiomatic and a literal interpretation.
2. Collocations were then extracted from the BNC in a window size of ±1 (the adjective and the noun in one combination appeared adjacently in the corpus with no words in between). The same nouns used in the idioms were retained, so only adjectives were searched for. There are different metrics for extracting collocations from corpora (Gries, 2013). Following previous collocation research (Durrant & Doherty, 2010; Li et al., 2021; Öksüz et al., 2021; Sonbul, 2015; Vilkaitem, 2016), this study used MI-score and phrase-level frequency. A cut-off point of three of the MI scores was used to identify collocations (Hunston, 2022).
3. Forty-five native English speakers normed the 73 sets of items on familiarity. Following a Latin square design, items were grouped into three lists and each list was normed by 15 participants. Following prior experiments' measurement of familiarity (e.g., Carrol & Conklin, 2017; Kyriacou et al., 2020; Shi et al., 2023), each participant was asked how often they used or encountered 73 items (including idioms, collocations, and controls) on a scale of 1 (Never) to 7 (Extremely frequently). To match the items' variables including familiarity, phrase-level frequency, and MI-scores of the idioms and collocations, the item selection was based on collocation familiarity rank. The 48 most familiar collocations were selected first, and corresponding idioms and controls were finalised. Then eight sets whose idioms had the lowest familiarity scores were excluded to ensure accessibility of idiomatic meanings. Finally, 40 sets of items were selected, and each set had one ambiguous idiom, one collocation, and one control.

4.2.3.1 Word-level and Phrase-level Length and Frequency

The adjectives in each pair were closely matched for letter length and lexical frequency. Unfortunately, other item-related variables could not be matched entirely. The variables left uncontrolled were therefore entered into the linear mixed models analyses as covariates to take care of potential confounding effects. One-way repeated ANOVA tests showed that the adjective frequency was controlled across idioms, collocations and novel word combinations ($F(2, 117) = 1.22$, $p = .30$), but the adjective length was significant ($F(2, 117) = 5.29$, $p = .006$), and the Bonferroni post-hoc comparisons of the three groups indicated that the adjectives for the idioms were somewhat shorter than those from novel word combinations ($\beta = 0.58$, $t(117) = 3.01$, $p = .009$) and those from collocations ($\beta = 0.5$, $t(117) = 3.01$, $p = .03$). There was also a statistical difference in phrase-level's length ($F(2, 117) = 3.55$, $p = .032$), showing that idioms were somewhat longer than novel word combinations, based on Bonferroni post-hoc comparisons ($\beta = 0.7$, $t(117) = 2.46$, $p = .046$) (See Table 4.3 for an overview of the variables).

As one of the aims was to test whether the idiom superiority effect still existed when phrase-level frequency and familiarity were controlled, phrase-level frequency ($t(39) < 1$) and familiarity ($t(39) < 1$) between idioms and collocations were matched for. For each set of candidate items, the number of occurrences in the BNC for collocations and most idioms was larger than 10 (idioms generally are less frequent than collocations, and three idioms appeared 3, 4 and 7 times in the BNC), while novel word combinations only appeared 0-1 time in the BNC.

Table 4.3

Summary of Items Characteristics

Items characteristics	Novel word combinations (*M*, *SD*)	Collocations (*M*, *SD*)	Overall idioms (*M*, *SD*)	Idioms-literal (*M*, *SD*)	Idioms-figurative (*M*, *SD*)
Adjective length	5.1 (1.06)	5 (1.04)	4.42 (0.90)	---	---
Phrase-level length	9.55 (1.24)	9.45 (1.24)	8.85 (1.33)	---	---
Adjective frequency	10825 (8789)	13260 (13392)	15186 (14607)	---	---
Phrase-level frequency	---	51.7 (51.4)	54.7 (55.4)	---	---
Familiarity	---	3.29 (0.8)	3.45 (0.66)	---	---
Predictability_word1	0.003 (0.01)	0.006 (0.02)	---	0.009 (0.03)	0.025 (0.06)
Predictability_word2	0.116 (0.17)	0.078 (0.12)	---	0.068 (0.11)	0.048 (0.12)

Note. Frequency extracted from BNC with English Corpora and all frequency measures are expressed as the raw numbers of occurrences. Length is in number of characters. Data for overall idioms was regardless of the two interpretations of the idiom. Idioms-literal refers to idioms in the literal context; idioms-figurative refers to idioms in the figurative context.

4.2.3.2 Idioms: Idiomatic Meaning Dominance and Decomposability

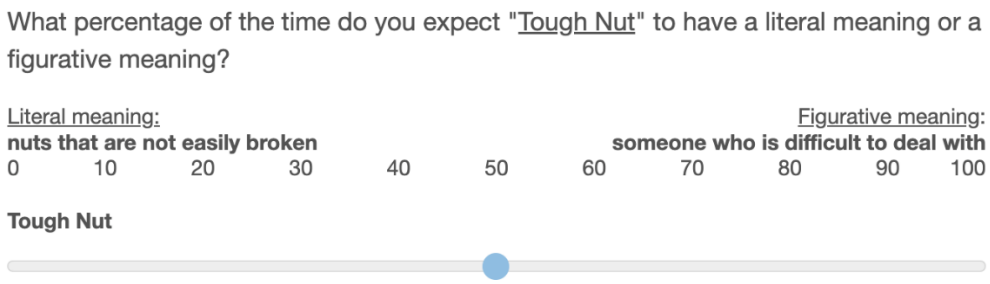
One of the research questions of this study is to assess how meaning dominance and semantic decomposability affect access to the idiomatic and literal meanings of an idiom. Familiarity, meaning dominance and local decomposability were tested with the same 45 native English participants (age: *M* = 30.78, *SD* = 11.33; gender: female, *N* = 31). Following prior research on idioms (Milburn & Warren, 2019), each idiom was presented with a 0% - 100% slider, with the literal meaning at the 0% end and the figurative meaning at the 100% end. For example, participants saw *tough nut* with its literal meaning (*nuts that are not easily broken*) on the left side of the slider and its figurative meaning (*someone who is difficult to deal with*) on the right side of the slider. The instructions for the meaning dominance norming included descriptions of what is meant with figurative and literal interpretations of a phrase, and examples of figurative and literal uses of phrases were provided. An example of how meaning dominance was measured can be found in

Figure 4.1. Participants moved the slider to indicate what percentage of the time they expected an idiomatic phrase to appear in a literal or figurative sense.

To ensure that the meaning dominance measurement of idioms is valid and reliable. First, internal reliability was checked with coefficient Cronbach's Coefficient Alpha (α = 0.912), indicating high internal consistency (Tavakol & Dennick, 2011). Second, meaning dominance was measured with another group of native English speakers (N = 20). The results from these two measures were highly correlated (r (16) = 0.93, p < .001), implying that people have a consistent intuition about the meaning dominance of idioms. In addition, we went through all instances of an idiom from the BNC if the number of instances was smaller than 50. If the number of instances was larger than 50, we randomly selected instances of idioms from the BNC. Then, we used the surrounding context to determine how often they were used figuratively or literally. Instances that could not be clearly classified were excluded. Meaning dominance was calculated by dividing the number of times the idiom was used figuratively by the total number of times the idiom appeared, or 50 in the BNC. The BNC frequency results correlated with the human ratings (r (38) = 0.4, p < .05).

Figure 4.1

Meaning Dominance Norming Example

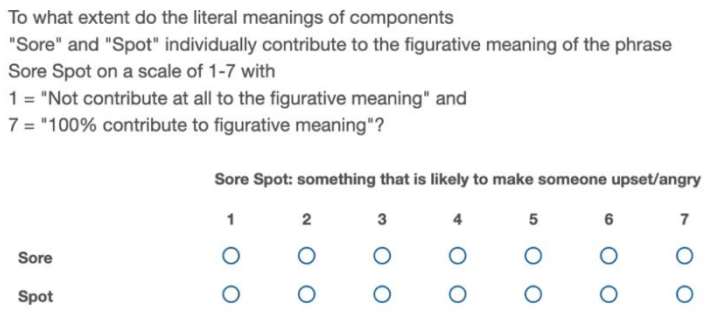


After the meaning dominance task, decomposability was measured following previous idiom studies (e.g., Gibbs et al., 1989). To gain a measure of “local” decomposability, participants were asked to what extent the literal meaning of the individual components of an idiom contributed to

the provided figurative meaning (e.g., participants rated to what extent the literal meanings of sore and spot individually contribute to the overall figurative meaning). They rated the contribution of both the adjective and the noun of the idiom to the whole figurative meaning on a scale of 1 (does not contribute at all) to 7 (100% contribution). To gain a "global" measure of decomposability, 15 new native English participants rated the contribution of the overall literal meaning of the expression to the figurative meaning of the idiom on a scale of 1 (does not contribute at all) to 7 (100% contribution) (e.g., participants rated to what extent the overall literal meaning of sore spot contributes to the figurative meaning of sore spot i.e., something that is likely to make someone upset/angry). An example of the local and global semantic decomposability task is presented in Figure 4.2. The values of the idioms' familiarity, decomposability, and idiomatic meaning dominance (see Table 4.4) were used as continuous predictors in the statistical analysis.

Figure 4.2

Local Semantic Decomposability Norming Example



Global Semantic Decomposability Norming Example

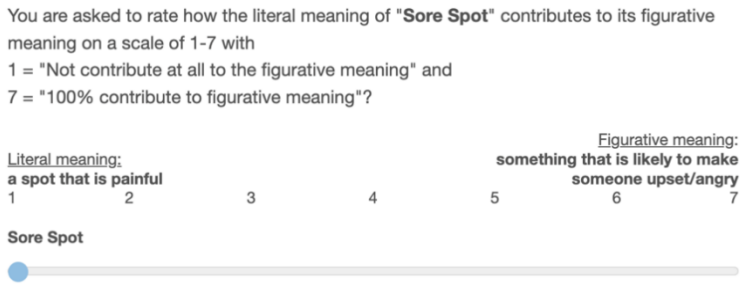


Table 4.4

Summary of Characteristics of Idioms

Norm (scale)	***Mean***	***SD***
Familiarity (1-7)	3.45	0.66
Decomposability: adjectives (1-7)	5.29	0.75
Decomposability: nouns (1-7)	4.15	0.89
Overall Decomposability (1-7)	4.48	0.63
Figurative Dominance (0-100%)	69%	13.4

4.2.3.3 Context

Forty sets of sentence fragments were written. Each sentence consisted of a disambiguating context, a control region, a critical region containing the MWS, a spill-over region (defined as the next two words following the critical region), and a wrap-up region (see Table 4.2). The control region and the spill-over region were identical for each condition of an item. For idioms, the prior contexts biased either the literal interpretation or the figurative interpretation.

Contextual predictabilities were measured on the first and the final words of a given MWS. The 40 sets of sentence fragments were counterbalanced over four lists by following a Latin square design, with an equal number of items per condition. An additional 80 native British English participants (age: *M* = 31.35, *SD* = 6.93; gender: female, *N*= 56) (20 participants per list) were tasked to write a continuation for the sentences truncated either immediately before the final word of the MWSs (*e.g., Although most of it is known to us, there is a rare* ____), testing the predictability of the noun, and another different 80 native British English participants (age: *M* = 31.22, *SD* = 6.71; gender: female, *N*= 57) continued the sentence fragments before the whole MWSs (*e.g., Although most of it is known to us, there is a* ____), testing the predictability of the adjective. Therefore, Predictability_word1 stands for the cloze probability of the whole MWS, followed by Li et al. (2021). Predictability_word2 stands for the cloze probability of the final word of MWS. Cloze probability was calculated as the percentage of participants who provided the correct (or intended) completion. Contextual predictability for the final words of MWSs was matched across conditions (F (3, 156) = 1.99, p = .118). However, there was a difference in contextual predictability for the whole MWSs (F (3, 156) = 3.14, p = .019), with idioms used figuratively somewhat more predictable than novel word combinations (β = -0.02, t (156) = -2.96, p = .021). See Table 4.3 for the reports of the predictability scores.

To make sure the sentences matched in terms of their naturalness, first, reading times up to the target regions were analysed and found no significant differences across conditions ($\chi 2$(1) = 3.99, p = .26). If sentence naturalness affected the reading times, we would expect to detect significant

differences across different types of MWSs. In addition, post-viva I asked 20 native English speakers (age: *M* = 33.1, *SD* = 6.12; gender: female, *N* = 11) who did not participate in the main study to rate the naturalness of the sentence frames (up to the spill-over region) on a five-point scale (1 = *very unnatural*, 5 = *very natural*). There was no significant difference across the different conditions (F (3, 156) = 1.67, p = .176).

Finally, fillers sentences were included to provide semantically non-natural contrast and mask the experimental manipulation, as this study used the self-paced reading paradigm paired with concurrent stop-making-sense task. To create the contrast between plausible and implausible in semantic meaning, filler sentences were semantically incorrect, varying in the location of implausibility. In half of the filler sentences, the implausibility was introduced close to the beginning the sentence (e.g., *For some reason we do not fully walk, she is the winner as she was doing so poorly in the competition*), while the implausibility happened at the end of filler sentences in the other half of the filler sentences (e.g., *There was a lot of pressure on the team tonight and that is why they were entirely relaxed*). An additional 20 English native speakers (age: *M* = 30.9, *SD* = 8.84) were asked to judge whether the filler sentences make sense in English. Another 20 grammatically correct and plausible sentences were included in the rating test to provide the natural contrast. A filler sentence was excluded if less than 80% of participants judged the sentence sensible.

After all pre-tests, 40 sets of items were retained, and they were counterbalanced into four lists by following a Latin Square design. Each list had 40 critical sentences containing MWSs and 40 filler sentences (See Appendix 4.2 for all experimental items and sentences).

4.2.4 Procedure

Participants were recruited via Prolific and Sona Systems (used for research participation at the School of Psychology, University of Birmingham). Participants first completed ethics on Qualtrics and were then directed to the online reading experiment. The reading experiment was programmed using PsychoPy software with version 2021.2.3 (Peirce, 2007) and hosted by Pavlovia. PsychoPy is

one of the best behavioural sciences research software and achieves high timing accuracy for online web studies (Bridges et al., 2020).

Participants read instructions on the computer screen and then completed eight practice trials. During the practice session, feedback for comprehension questions was provided. At the start of each trial, “A new sentence starts:” was presented and disappeared after “Space bar” pressed. For presentation purpose, each context sentence was divided into five reading regions, with the critical region always appearing third. A self-paced reading paradigm coupled with a concurrent stops-making-sense task was used as this paradigm is thought to be more sensitive to pick up plausibility effects than “regular” self-paced reading (Singh et al., 2015). Participants pressed the “Space bar” to start reading the first region. They then pressed the “J” key if what they had read up to that point still made sense. This made the current region disappear and revealed the next region. As long as the sentence continued to make sense to them, they continued to press the “J” key to reveal the next region. If the sentence did no longer make sense, participants pressed the “F” key. Pressing the “F” key terminated the whole trial and the next trial began. After each critical sentence containing an MWS, a comprehension question appeared. Only when participants thought the whole sentence was plausible, a following question was presented to them. In other words, pressing the “F” key terminated a trial and skipped the comprehension question.

All regions and questions were displayed at the centre of the screen in black in Courier font over a white background. The 80 trial sentences were randomised and presented in two blocks, with a self-paced break after 40 trials.

Once participants completed the online reading experiment, they were redirected to the Qualtrics page. Both L1 and L2 speakers had to indicate their language background and complete a language usage questionnaire. L2 speakers also took the vocabulary-level test as well as a multiple-choice test to check whether they understood the idioms involved in the reading task. On average, L1 speakers took 30 mins and L2 speakers took 40 mins to complete the whole study.

4.2.5 Measurements

For RT measurements, this study analysed two regions from the self-paced reading paradigms: (1) the critical region where the critical items (MWS and control) were positioned, and (2) the spill-over region, which was the next two words following the critical region. The spill-over region was included because the reading differences might be delayed and/or spilled over to the next words.

4.2.6 Statistical Analysis

The analysis has two steps: first including all types of MWSs and second, for idioms only. For each step, all analyses were carried out using R version 4.2.1 (R Core Team, 2022) with two packages *lm4* (Bates et al., 2015) and *lmerTest* (Kuznetsova et al., 2020). Following recommendations (Brysbaert & Stevens, 2018; Lo & Andrews, 2015) for analysing continuous data, a logged Gaussian distribution to the RTs was fitted.

For each step of the analysis, models were first fitted with L1 group and L2 group data separately to see how proficiency affected processing in the L2 group. Second, models were constructed with L1 and L2 data to test whether there was a qualitative difference in processing between L1 and L2 speakers.

For the MWSs analysis, the core model included Types-of-MWSs (treatment coded: novel word combinations as 0 for the critical region, and treatment code: idioms in figurative context as 0 for the spill-over region). For the idiom analysis, the core model included context (treatment coded: idioms in figurative context: 0 vs. idioms in literal context: 1). The choice of different baseline levels (novel word combinations treatment coded 0 in the critical region and figurative idioms treatment coded 0 in the spill-over region) was made due to the easier presentation of comparisons between different types of MWSs. When the baseline was novel word combinations, other types of MWSs were compared to novel word combinations in the regression table. When the baseline was idioms in a figurative context, other types of MWSs were compared to idioms used figuratively. I believe

that this strategy made the presentation of the key findings from the experiment easier to follow. Alternatively, pairwise comparisons with the Bonferroni correction could be used. Similar results (direction and significance) were indeed found using these pairwise comparisons with the Bonferroni correction. However, the more tests we conducted, the stricter the Bonferroni correction became, which led to a higher likelihood of committing a Type II error. A fully random model was used whenever possible (Barr et al., 2013). If convergence was not achieved, "bobyqa" or "optimx" optimiser (Powell, 2009) was used to increase iterations. If this did not work, the random effect was simplified by first removing interactions and then, if necessary, individual terms until the model converged.

Starting with the analysis of MWSs, for the L1 group, the full model of MWSs included the core model and the covariates trial number, individual word frequency, contextual predictabilities, and item length. For the L2 group, vocabulary score was entered as an additional fixed effect after constructing the core model with the other covariates. For the L1 and L2 groups, the full model included the core model, trial number, Language-background factor (treatment coded: L1: 0 vs. L2: 1), and all interactions with Language-background. When it comes to the analysis of idioms, the degree of decomposability and meaning dominance were entered as fixed factors after taking constructing the core model with covariates. Of note, all continuous variables were centered by the scale () function.

When there was no main effect of Types-of-MWSs, we still decided to compare different types of MWSs. Theoretically, if there are a priori reasons for expecting a difference in type, for example because it was found in previous literature (as was the case here), it is acceptable to look for different types of MWSs over novel literal controls, even though there was no interaction or main effect of MWS type. Second, it is also a philosophically controversial issue whether the significance of a specific binary contrast (Winter, 2019), such as collocation vs. its controls, should

depend on having performed a hypothesis test for unrelated comparisons, such as idioms in literal use versus idioms in figurative use.

All full models were subsequently reduced with the drop 1() function, which performs model comparisons with a chi-square test. Note however that the manipulated variables (Types-of-MWSs, context, decomposability, meaning dominance, vocabulary score and Language-background) were always kept in the corresponding reduced models to reflect the experimental design, even if their effects were not significant. The interactions with manipulated variables and covariates were removed if they were not significant. Based on these tests, reduced models were achieved to best fit the data. In the reduced models, variables with a t value of greater than 2 were considered significant at the alpha=.05 level (Baayen et al., 2008). The significance of interactions and main effects was assessed by model comparisons with the anova () function. For all, the vif () function from the *car* package (Fox et al., 2023) was used to check collinearity.

4.3 Results

4.3.1 Pre-processing Data

Participants were excluded based on the accuracy of both implausible filler sentences and comprehension questions. Participants who judged nearly all implausible filler sentences sensible and judged plausible critical sentences insensible were removed (*N* = 2 for each group of speakers). Then, participants whose accuracy for the comprehension questions was lower than 80% were excluded from the final analysis (*N* = 2 for L1 and *N* = 3 for L2). Following Singh et al. (2016), critical trials that participants thought implausible in English were removed. When the comprehension question was answered incorrectly, the concomitant trial was removed as well. Furthermore, RTs on critical regions larger than 2500ms (L1 speakers)/4000ms (L2 speakers) and smaller than 200ms were removed. RTs on spill-over regions larger than 2000ms and smaller than 200ms were removed (the spill-over region was shorter than the critical region; hence, the maximum cut-off was reduced). We further excluded responses for which the RT was greater than 2.5 SD (Miller, 1991; Ratcliff, 1993) above and below the mean per participant and per condition on critical and spill-over regions, respectively. After data cleaning, 78% (2706 observations) of the critical region data and 80% (2774 observations) of the spill-over region data in L1 speakers were entered in the analyses; 70.92% (3312 observations) of the critical region data and 71.05% (3318 observations) of the spill-over region data in L2 speakers were entered in the analyses. The details of removing data for each group and conditions for each step are provided in Appendix 4.7.

Average reading times on the critical and spill-over regions are summarised in Table 4.5. Tables 4.6 and 4.8 report the best fitting models for the different types of MWSs for L1 speakers and L2 speakers, respectively. Tables 4.7 and 4.9 summarise the best fitting models for the idiom-only analyses for L1 and L2 speakers, respectively.

4.3.2 L1 speakers

4.3.2.1 L1 Speakers' RTs on the Critical Region across Four Conditions

For L1 speakers, there was a main effect of type of MWSs at the critical region ($\chi2(3) = 21.2$, $p < .001$), with faster reading times for idioms used figuratively ($\beta = -0.10$, $t = -3.7$, $p < .001$), idioms used literally ($\beta = -0.06$, $t = -2.19$, $p = .029$), and collocations ($\beta = -0.09$, $t = -3.33$, $p < .001$), compared to novel word combinations. In addition, idioms used figuratively were read more quickly than when used literally ($\beta = -0.04$, $t = -2.91$, $p = .004$). To control unmatched item variables, the length of adjectives and contextual predictability for the whole MWS were initially entered into models. Neither contextual predictability ($\chi2(1) = 0$, $p = .96$) nor adjectives' length ($\chi2(1) = 0.45$, $p = .5$) was a significant predictor, suggesting that these variables did not confound the effects.

4.3.2.2 L1 Speakers' RTs on the Spill-over Region across Four Conditions

A main effect of type of MWSs was observed for the spill-over region ($\chi2(3) = 10.4$, $p = .015$). In Table 4.6, idioms in their figurative usage were read more quickly by L1 speakers than other types of MWSs: idioms used figuratively were faster than collocations ($\beta = 0.06$, $t = 2.83$, $p = .005$), than idioms used literally ($\beta = .02$, $t = 2.18$, $p = .029$), and than novel word combinations ($\beta = 0.04$, $t = 2$, $p = .045$). Like the critical region analysis, no effects of contextual predictability ($\chi2$ (1) = 0.46, $p = .5$) or word length were observed ($\chi2$ (1) = 0.82, $p = .36$).

For L1 speakers, when the phrase-level frequency was added into the model on the critical region, it did not significantly improve the model ($\chi2(1) = 0.30$, $p = .581$). Furthermore, phrase-level frequency and type-of-MWSs were highly overlapping because both variables represented similar distributional information. The overlap caused multi-collinearity in the model, which was tested by calculating Generalised Variation Inflation Factors, GVIFs, and GVIF^(1/2*df), where df is degrees of freedom. If GVIF^(1/2*df) is higher than 2, it would be considered an indication of multicollinearity. The GVIF^(1/2*df) value for phrase-level frequency was 2.67. Therefore, the phrase-level frequency was not added into the statistical modelling.

Table 4.5

Descriptive Statistics: Mean RT (ms) and Standard Deviations for the Critical Region and Spill-over Region across Four Conditions for Both L1 and L2 Speakers

Group	Conditions	Critical region		Spill-over region	
		M	***SD***	***M***	***SD***
L1	novel word combinations	1133	452	747	263
	collocation	1019	399	759	273
	idiom in literal use	1074	447	733	255
	idiom in figurative use	1026	429	723	262
L2	novel word combinations	1332	574	770	277
	collocation	1226	522	771	301
	idiom in literal use	1260	539	738	84
	idiom in figurative use	1223	530	760	297

Table 4.6

Summary of the Best-fitting Models of MWSs on Critical and Spill-over Regions for L1 Speakers

Critical Region				
Predictors	***β***	***SE***	***t***	***p***
(Intercept)	6.98	0.03	225.01	<.001
trial number	-0.06	0.01	-7.62	<.001*
collocation	-0.09	0.03	-3.33	<.001*
idiom-figurative	-0.10	0.03	-3.7	<.001*
idiom-literal	-0.06	0.03	-2.19	.029*
word1 frequency	-0.02	0.01	-1.88	.06
word2 frequency	-0.04	0.01	-3.25	.001*
Random Effects				
σ^2	0.07			
$\tau_{00\ items.x}$	0.01			
$\tau_{00\ ID}$	0.05			
ICC	0.44			
$N_{\ ID}$	87			
$N_{\ items.x}$	120			
Observations	2691			
Marginal R^2 / Conditional R^2	0.055 / 0.473			

Spill-over Region				
Predictors	***β***	***SE***	***t***	***p***
(Intercept)	6.54	0.03	227.45	<.001
trial number	-0.07	0.01	-11.34	<.001*
collocation	0.06	0.02	2.83	.005*
novel word combinations	0.04	0.02	2.00	.045*
idiom-literal	0.02	0.01	2.18	.029*
Random Effects				
σ^2	0.04			
$\tau_{00\ items.x}$	0.01			
$\tau_{00\ ID}$	0.05			
ICC	0.60			
$N_{\ ID}$	87			
$N_{\ items.x}$	120			
Observations	2710			
Marginal R^2 / Conditional R^2	0.054 / 0.616			

Note. Reading times were log-transformed, idiom-literal refers to idioms in literal contexts and idiom-figurative refers to idioms in figurative contexts; novel word combinations were coded as 0 for the critical region and idiom-figurative was coded as 0 for the spill-over region.

4.3.2.3 L1 Speakers: Decomposability and Meaning Dominance in Resolving Idiom Ambiguity

We observed a significant interaction between context and local decomposability (i.e., whether the literal meaning of the adjective and the noun contribute to the overall figurative idiomatic meaning) on the critical region for both adjectives ($\chi 2$ (1) = 8.86, p = .003) and nouns ($\chi 2$ (1) = 11.5, p < .001), as shown in Figure 4.3 and Figure 4.4. Increasing decomposability of the adjective sped up reading times for idioms used figuratively but slowed down reading idioms used literally. Increasing decomposability of the noun sped up reading times for both idioms used literally and figuratively. In short, it was easier to activate the figurative meaning of an idiom when components of the idiom were more semantically decomposable. There was also an interaction between meaning dominance and decomposability of the adjective ($\chi 2$ (1) = 6.61, p = .01) on the critical region. As Figure 4.5 shows, increasing decomposability of the adjective sped up reading when the figurative meaning was less dominant. The effect of decomposability became weaker when idioms were used more frequently in figurative contexts. Moreover, the spill-over region also demonstrated a noun decomposability effect (β = -.04, t = -2.28, p = .028).

Table 4.7

Summary of the Best Fitting Models of Idioms on the Critical and Spill-over Regions for L1 Speakers

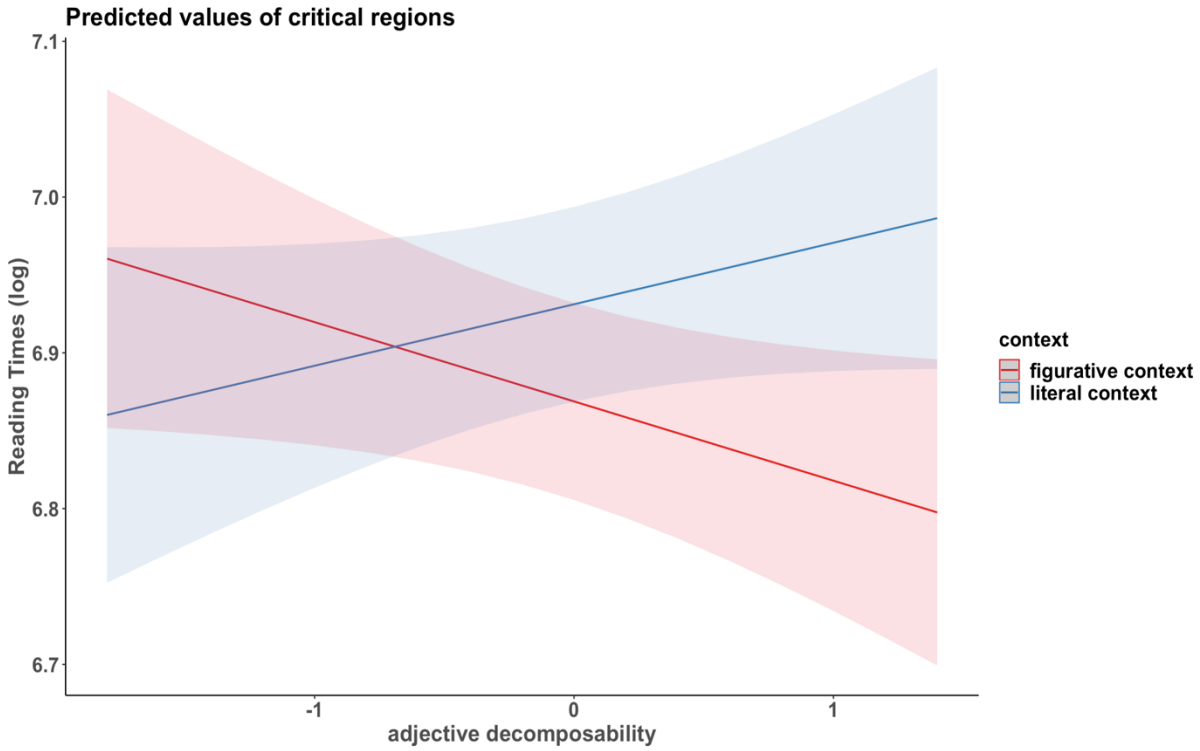
Critical Regions				
Predictors	***β***	***SE***	***t***	***p***
(Intercept)	6.87	0.03	213.9	<.001
trial number	-0.06	0.01	-5.19	<.001*
noun decomposability	-0.1	0.02	-4.79	<.001*
adjective decomposability	-0.05	0.03	-1.94	.052*
meaning dominance	0.0001	0.0006	0.36	.722
context [idiom-literal]	0.06	0.03	2.13	.034*
adjective decomposability* meaning dominance	0.001	0.0007	2.58	.01*
adjective decomposability* literal context	0.09	0.03	2.98	.003*
noun decomposability *literal context	0.06	0.02	3.41	.001*
Random Effects				
σ^2	0.08			
$\tau_{00\ ID}$	0.05			
$\tau_{00\ items.x}$	0.01			
ICC	0.43			
N_{ID}	87			
$N_{items.x}$	40			
Observations	1405			
Marginal R^2 / Conditional R^2	0.069 / 0.47			

Spill-over Regions				
Predictors	***β***	***SE***	***t***	***p***
(Intercept)	6.53	0.03	234.55	<.001
noun decomposability	-0.04	0.02	-2.28	.028*
idiom-literal	0.03	0.01	2.39	.017*
trial number	-0.08	0.01	-9.86	<.001*
Random Effects				
σ^2	0.04			
$\tau_{00\ ID}$	0.05			
$\tau_{00\ items.x}$	0.01			
ICC	0.6			
N_{ID}	87			
$N_{items.x}$	40			
Observations	1494			
Marginal R^2 / Conditional R^2	0.07 / 0.624			

Note. Reading times were log-transformed; decomposability and dominance were centered; Context has two levels: idiom-literal and idiom-figurative; idiom-literal refers to idioms in literal contexts; idioms in figurative contexts were coded as 0.

Figure 4.3

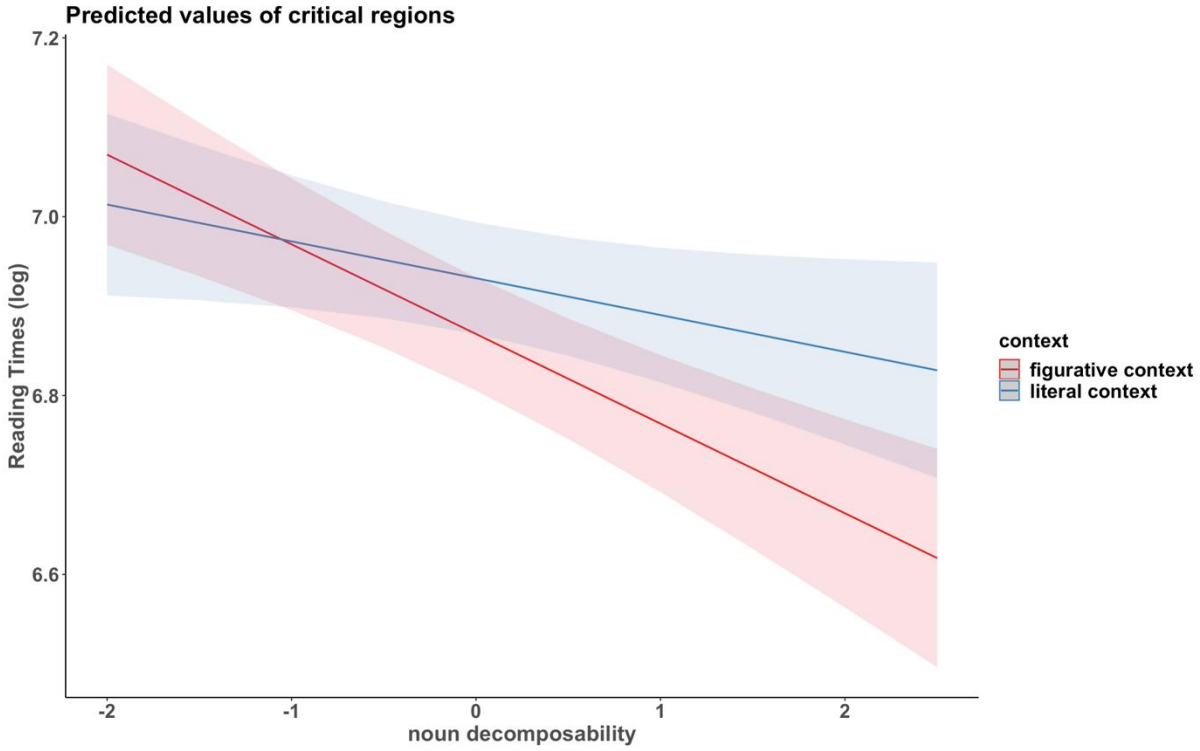
The Interaction between Context and (Adjectives) Decomposability on the Critical Region for L1 Speakers



Note. Decomposability was centered; reading times were log-transformed; context has two levels: figurative and literal.

Figure 4.4

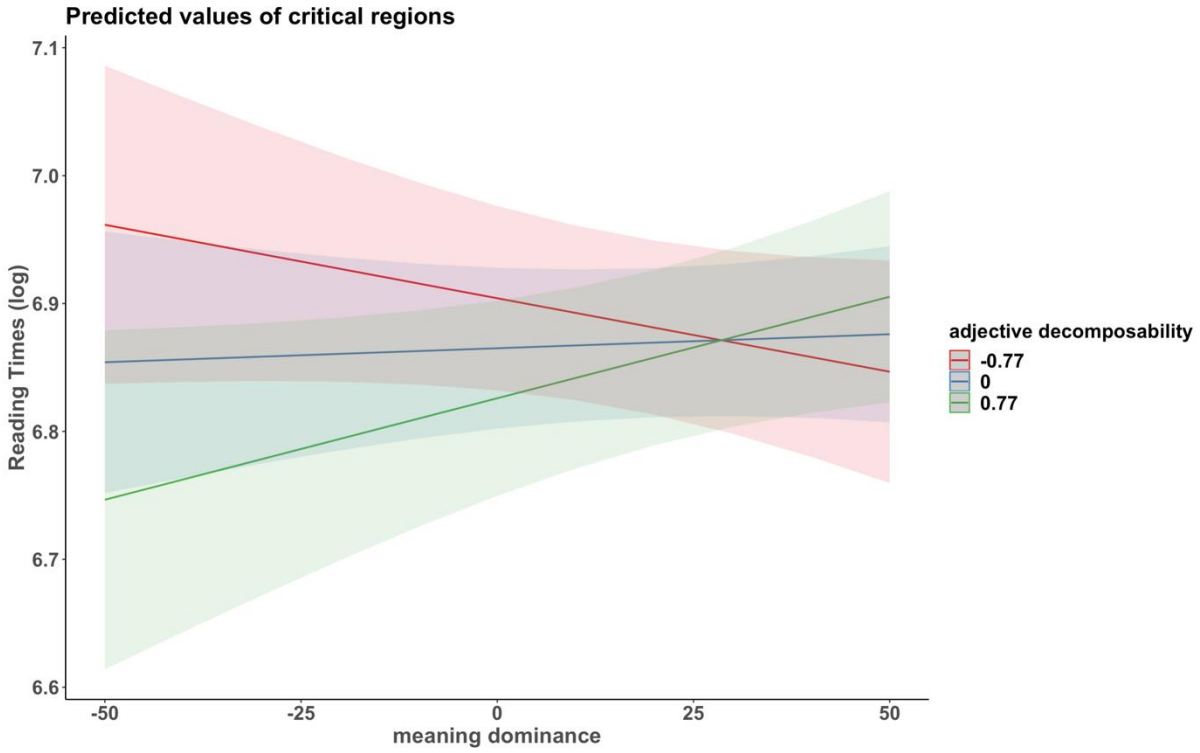
The Interaction between Context and (Nouns) Decomposability on the Critical Region for L1 Speakers



Note. Decomposability was centered; reading times were log-transformed; context has two levels: figurative and literal.

Figure 4.5

The Interaction between Meaning Dominance and (Adjectives) Decomposability on the Critical Region for L1 Speakers



Note. Decomposability and dominance were centered; reading times were log-transformed.

4.3.3 L2 Speakers

4.3.3.1 L2 Speakers' RTs on the Critical Region across Four Conditions

There was no significant main effect of type of MWSs for the critical region ($\chi 2(3) = 5.7$, $p = .13$). Further comparisons revealed that L2 speakers read collocations faster than novel word combinations ($\beta = -0.07$, $t = -2.34$, $p = .021$). Vocabulary size approached significance ($\beta = -0.004$, $t = -1.99$, $p = .049$), with higher vocabulary scores speeding up the reading of MWSs. Additionally, no interaction between vocabulary and the type of MWSs existed ($\chi 2(3) = 2.27$, $p = .52$). In addition, unsurprisingly, higher frequencies of the adjectives ($\beta = -0.02$, $t = -2.37$, $p = .02$) and nouns ($\beta = -0.04$, $t = -4.31$, $p < .001$) resulted in shorter reading times for MWSs. Like L1 speakers, there was no effect of word length ($\chi 2(1) = 0.68$, $p = .41$) or contextual predictability ($\chi 2(1) = 0.6$, $p = .44$).

4.3.3.2 L2 Speakers' RTs on the Spill-over Region across Four Conditions

We did not detect the main effect of type of MWSs for the spill-over region ($\chi 2(3) = 5.46$, $p = .14$). Further analyses showed that idioms used figuratively took longer to read those idioms used literally ($\beta = -0.03$, $t = -2.16$, $p = .031$). There was no effect of vocabulary size ($\beta = -0.001$, $t = -1.03$, $p = .303$) and no interaction between vocabulary size and type of MWSs ($\chi 2(3) = 7.42$, $p = .06$).

Similarly, for L2 speakers, the phrase-level frequency did not contribute to the model significantly ($\chi 2(1) = 0.28$, $p = .6$). Adding phrase-level frequency also led to multicollinearity whereby the GVIF^(1/2*df) value for phrase-level frequency was 2.67.

Table 4.8

Summary of the Best-fitting Models of MWSs for the Critical and Spill-over Regions for L2 Speakers

Critical Regions					Spill-over Regions				
Predictors	***β***	***SE***	***t***	***p***	***Predictors***	***β***	***SE***	***t***	***p***
(Intercept)	7.12	0.03	219.86	<.001	(Intercept)	6.6	0.03	242.21	<.001*
collocations	-0.07	0.03	-2.34	.021*	collocations	0.005	0.03	0.19	.851
idiom-figurative	-0.05	0.03	-1.68	.096	novel word combinations	0.01	0.03	0.29	.77
idiom-literal	-0.05	0.03	-1.61	.109	idiom-literal	-0.03	0.01	-2.16	.031*
vocabulary size	-0.004	0.002	-1.99	.049*	trial number	-0.06	0.01	-7.55	<.001*
word1 frequency	-0.02	0.01	-2.37	.02*	vocabulary size	-0.001	0.001	-1.03	.303
word2 frequency	-0.04	0.01	-4.31	<.001*					
trial number	-0.002	0.0004	-4. 68	<.001*					
Random Effects					**Random Effects**				
σ^2	0.09				σ^2	0.05			
$\tau_{00\ items.x}$	0.01				$\tau_{00\ items.x}$	0.01			
$\tau_{00\ ID}$	0.07				$\tau_{00\ ID}$	0.03			
ICC	0.5				ICC	0.49			
N_{ID}	117				N_{ID}	117			
$N_{items.x}$	120				$N_{items.x}$	120			
Observations	3502				Observations	3618			
Marginal R^2 / Conditional R^2	0.048 / 0.523				Marginal R^2 / Conditional R^2	0.035 / 0.509			

Note. Reading times were log-transformed, idiom-literal refers to idioms in literal contexts and idiom-figurative refers to idioms in figurative contexts; novel word combinations were coded as 0 for the critical region and idiom-figurative was coded as 0 for the spill-over region.

4.3.3.3 L2 Speaker: Decomposability and Meaning dominance on Resolving Idiom Ambiguity

For L2 speakers, both the critical region (β = -0.09, t = -4.3, p < .001) and spill-over region (β = -0.06, t = -2.38, p = .017) showed that local decomposability of the noun was significant. We also observed a significant interaction between context and local decomposability of the noun ($\chi 2$ (1) = 5.42, p = .02) on the critical region, as shown in Figure 4.6, with a higher degree of decomposability of the noun speeding up reading an idiom in the figurative interpretation more than in the literal one. L2 speakers also demonstrated an interaction between meaning dominance and local decomposability of the adjective (β = 0.001, t = 2.19, p = .029) on the spill-over region. As seen in Figure 4.7, increased decomposability slowed down reading when an idiom was less balanced with respect to its literal and figurative interpretations. In contrast, there was a facilitation effect of the increased decomposability when an idiom was more balanced with respect to its literal and figurative interpretations. As expected, higher vocabulary scores resulted in shorter reading times of idioms in the critical region (β = -0.004, t = -2.22, p = .027).

Table 4.9

Summary of the Best-fitting Models of Idioms on the Critical and Spill-over Regions for L2 Speakers

Critical Regions					Spill-over Regions				
Predictors	***β***	***SE***	***t***	***p***	***Predictors***	***β***	***SE***	***t***	***p***
(Intercept)	7.09	0.04	179.48	<.001	(Intercept)	6.61	0.03	226.38	<.001
adjective decomposability	0.003	0.02	0.13	.897	adjective decomposability	0.01	0.03	0,37	.708
noun decomposability	-0.09	0.02	-4.3	<.001*	noun decomposability	-0.06	0.02	-2.38	.017*
context [idiom-literal]	0.04	0.03	1.28	.199	context [idiom-literal]	-0.03	0.02	-1.42	.155
meaning dominance	0.0001	0.0001	1.36	.173	meaning dominance	-0.0001	0.0005	-0.31	.755
trial number	-0.001	0.0005	-2.14	.032	trial number	-0.05	0.01	-4.87	<.001*
vocabulary size	-0.004	0.002	-2.22	.027*	vocabulary size	-0.001	0.001	-0.92	.356
noun decomposability* context[idiom-literal]	0.04	0.018	2.33	.02*	adjective decomposability* meaning dominance	0.0001	0.0004	2.19	.029*
Random Effects					**Random Effects**				
σ^2	0.09				σ^2	0.05			
$\tau_{00\ ID}$	0.07				$\tau_{00\ ID}$	0.04			
$\tau_{00\ items.x}$	0.01				$\tau_{00\ items.x}$	0.01			
ICC	0.47				ICC	0.51			
N_{ID}	117				N_{ID}	117			
$N_{items.x}$	40				$N_{items.x}$	40			
Observations	1524				Observations	1600			
Marginal R^2 / Conditional R^2	0.049 / 0.498				Marginal R^2 / Conditional R^2	0.056 / 0.537			

Note. Reading times were log-transformed; decomposability and dominance were centered; Context has two levels: idiom-literal and idiom-figurative; idiom-literal refers to idioms in literal contexts; idioms in figurative contexts were coded as 0.

Figure 4.6

The Interaction between Context and (Noun) Decomposability on the Critical Region for L2 Speakers

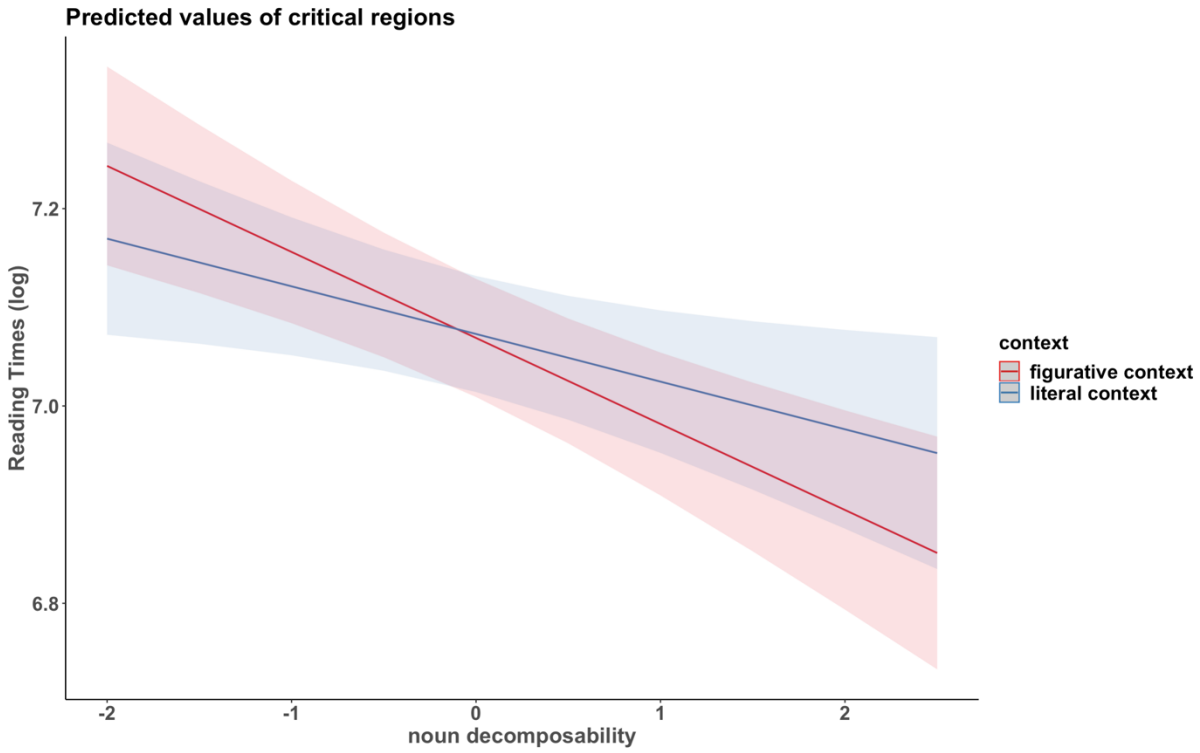
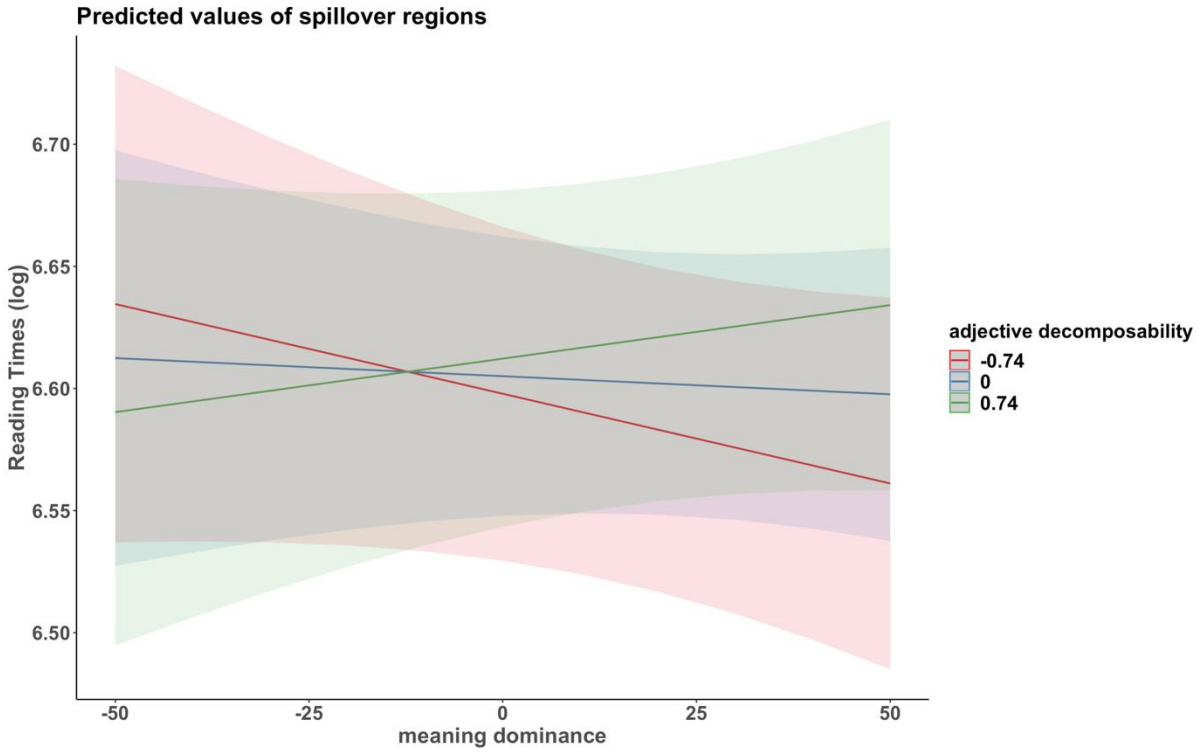


Figure 4.7

The Interaction between Meaning Dominance and (Adjectives) Decomposability on the Spill-over Region for L2 Speakers



Note. Decomposability and dominance were centered; reading times were logged.

4.3.4 Comparison L1 and L2 Data

4.3.4.1 Across Four Conditions

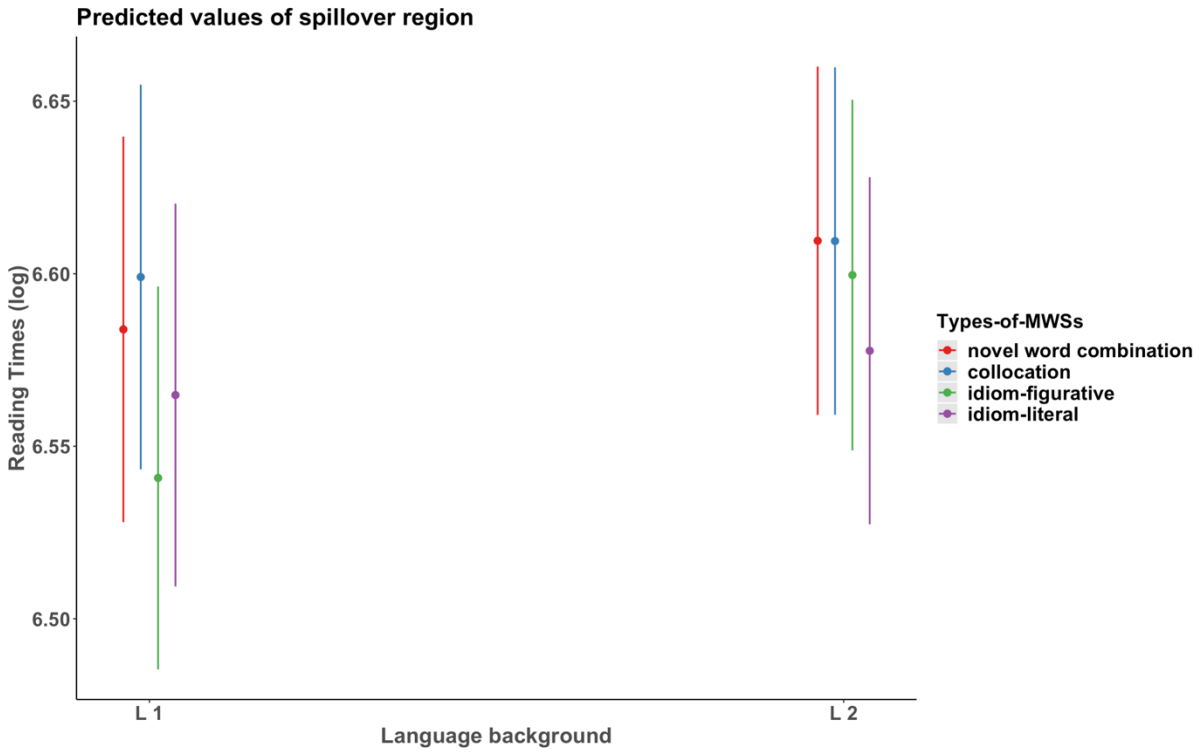
There was no interaction between Language-background and Types-of-MWSs ($\chi 2(3) = 5.21$, $p = .157$) in the critical region. The spill-over region demonstrated an interaction between Language-background and Types-of-MWSs ($\chi 2(3) = 10.20$, $p = .001$). Figure 4.8 demonstrated the interaction was mainly driven by different processing patterns of MWSs between L1 and L2 speakers. L1 speakers read idioms in figurative interpretations more quickly than literal ones, while an opposite pattern was found in L2 speakers. It also seemed that the interaction came from the differences between novel word combinations and collocations. However, a separate analysis of L1 and L2 data has shown no significant difference between collocations and novel word combinations on the spill-over region in L1 and L2 speakers. The full statistical reports can be found in Appendix 4.5.

4.3.4.2 Decomposability and Meaning Dominance on Resolving Idiom Ambiguity

Neither the critical region nor spill-over region demonstrated an interaction between Language-background and meaning dominance or decomposability ($p > 0.1$). The full statistical reports can be found in Appendix 4.6.

Figure 4.8

The Interaction between Types-of-MWSs and Language-background on the Spill-over Region



Note. L1 and L2 refer to L1 and L2 speakers respectively, reading times were log-transformed.

4.4 Discussion

This study investigated whether different types of MWSs were processed similarly when phrase-level frequency, familiarity, and contextual predictability were controlled for and to what extent lexical and idiomatic ambiguity resolution overlapped. Additionally, the ambiguity between the literal and figurative interpretations of idioms was assessed by considering the semantic decomposition and meaning dominance. Importantly, this study assessed whether both L1 and L2 speakers processed MWSs similarly. The present study showed that L1 speakers read idioms used in a figurative context faster than collocations, even when the distributional information (i.e., the phrase-level frequency, familiarity, and contextual predictability) was controlled. This is in line with the hypothesis that the idioms' processing advantage is the result of conventionality rather than distributional information. The current study also showed that L1 speakers processed idioms used literally slower than idioms used figuratively, though not slower than collocations. This finding indicates that lexical and idiomatic ambiguities overlap to some degree, whereby L1 speakers read idioms used literally longer than idioms used figuratively rather than collocations, partially in line with the second hypothesis. Furthermore, the present results also mirrored the effect of semantic relatedness found in Milburn and Warren (2019). Both L1 and L2 speakers demonstrated an effect of semantic decomposability whereby increased decomposability sped up reading an idiom in a figurative context and an interaction between semantic decomposability and meaning dominance. Increased decomposability sped up reading when the figurative interpretation of an idiom was less dominant, and increased decomposability slowed down reading when idioms were less balanced with respect to their literal and figurative interpretations. Finally, this study reproduced the finding that collocations were read more quickly than novel word combinations by both L1 and L2 speakers (e.g., Carrol & Conklin, 2020; Li et al., 2021; Öksüz et al., 2021; Sonbul, 2015; Vilkaitė, 2016; Vilkaitė & Schmitt, 2019; Vilkaitė-Lozdienė, 2022; Wolter & Gyllstad, 2013; Wolter & Yamashita, 2018), while idioms were only read more quickly by L1 speakers.

4.4.1 Collocation Advantage in L1 and L2 speakers

Compared to novel word combinations, the current study replicated the facilitation effect for collocations, and this was found on the critical region itself: *vast area* was read faster than *rare area* in both L1 and L2 speakers. This is in line with previous research (Carrol & Conklin, 2020; Li et al., 2021, 2022; Öksüz et al., 2021; Sonbul, 2015; Vilkaitė, 2016; Vilkaitė & Schmitt, 2019; Wolter & Gyllstad, 2013) and points towards the influence of phrase-level frequency and overall familiarity: as language users are exposed to certain language constructions more frequently, they can recognise their forms and meanings faster. This finding generalises the frequency effect beyond single words to multi-word units, which is taken as evidence for language models that emphasise the role of experience (i.e., experience is quantified by frequency). For example, the usage-based model of language acquisition (Bybee, 1998; Goldberg, 2006; Tomasello, 2003) suggests that every exposure to a language construction will reinforce it and leave a trace in memory. On the other hand, the current finding is against the words-and-rules approach (Pinker, 1999; Pinker & Ullman, 2002), which draws a clear distinction between the linguistic forms that are stored in the mental lexicon and those that are computed by grammar rules. According to the words-and-rules approach, frequency is expected to affect the stored linguistic forms, not the computed ones. However, the present finding that the processing of collocations is affected by frequency stands in contrast to the words-and-rules approach.

4.4.2 Accessing the Idiomatic Meaning

4.4.2.1 Distributional Properties

This study observed the processing advantage of idioms used figuratively compared to novel word combinations in L1 speakers in the critical region, while the same pattern was not found in L2 speakers. To illustrate, only L1 speakers read *grey area* in its idiomatic meaning more quickly than *rare area.* This finding is in line with the existing L1 literature (Carrol et al., 2016; Carrol & Conklin, 2017; Siyanova-Chanturia, Conklin, & Schmitt, 2011; Swinney & Cutler, 1979; Tabossi et al., 2009).

Like collocations, idioms also are more frequently than novel word combinations in language, possibly leading to faster processing.

On the other hand, L2 speakers did not show a reliable speed-up effect for idioms in figurative interpretations over novel word combinations. Although L2 speakers spent around 64 ms less time reading idioms in a figurative context than novel word combinations, the comparison was not significant in the model. The idiom effect vanished possibly due to the variations in participants' reading behaviours. Furthermore, the lack of the idioms' advantage might indicate that L2 speakers have not been exposed often enough to the idioms in their L2 to build the connection between the form and meaning of idioms. Our finding of idioms in L2 speakers contrasts with some previous studies showing L2 processing idioms in figurative interpretations faster than novel word combinations (Conklin & Schmitt, 2008; Underwood et al., 2004). There are two potential explanations for the different findings between the current study and Conklin and Schmitt's study. First, Conklin and Schmitt's study allowed participants to read the same idiom in different conditions, which introduced noise to the results. Second, each idiom and its control were embedded in substantially different passages, and contextual predictability was not controlled for. Thus, context might have imposed different constraints on comprehending the critical idioms and novel word combinations. In addition, the idioms used in Conklin and Schmitt's study (e.g., *one stone to kill two birds*; *his back against the wall)* were much longer than the ones used in the current study (always two words, e.g., *rotten apple*, *tough nut*). The predictability of longer idioms can be potentially higher than the predictability of short idioms since the longer a sentential fragment, the fewer the available alternatives in the language to continue syntactically and semantically (Miller & Selfridge, 1950).

4.4.2.2 Beyond the Distributional Properties

The spill-over region demonstrated a speed-up effect for idioms compared to collocations in L1 speakers, which suggests idioms' superiority effect is beyond the mere distributional information. Previous studies have demonstrated that idioms (e.g., *at the end of day*, break *the ice*) are processed

more quickly than novel literal word combinations (e.g., *at the end of* war, break *the glass*) and suggest that idioms have become familiar and predictable for language users as they encounter idioms multiple times (Cacciari & Tabossi, 1988; Gibbs et al., 1989; M. R. Libben & Titone, 2008; Siyanova-Chanturia, Conklin, & Schmitt, 2011; Swinney & Cutler, 1979; Tabossi et al., 2009). This frequent encounter leads to faster idiom processing when compared to novel word combinations. However, this explanation has difficulties in accommodating the current result as idioms were matched to collocations in terms of phrase-level frequency, familiarity, and contextual predictability. It is unlikely that this speed-up effect was driven by these three factors.

Importantly, our finding suggests that additional properties might contribute to the faster processing of idioms. In the current study, what sets idioms apart from collocations is the higher degree of conventionality and the lower degree of decomposability rather than differences in phrase-level frequency or subjective familiarity. The most prominent feature of idioms is that the idiomatic meaning cannot simply computed by assembling the meanings of idioms' lexical parts. This feature has made idioms less decomposable and more conventionalised. In contrast, collocations in the current study are more decomposable and less conventionalised, as the collocational meaning can be computed based on its lexical parts. Furthermore, according to the idiom Hybrid Representation Hypothesis (Kuiper et al., 2007; Sprenger et al., 2006), an idiom exists as individual words ("lemma") and as an overall figuratively conceptual entry ("super-lemma"), which means that an idiom can be activated via analysing the lexical parts and via direct access of the unitary meaning interactively and synchronously. In a similar vein, research in complex words has also suggested that multiple routes are available for decoding a complex word (i.e., a decomposition route and a direct access route) (Ji et al., 2011; Kuperman et al., 2009). It is believed that there is a processing cost associated with the decomposition-based route. Consequently, the direct access route in decoding complex words is preferred as this route is more efficient than the decomposition route. Due to the high degree of conventionality of an idiom, the direct activation of the idiomatic concept plays a more prominent role in comprehending the idiomatic meaning than the literal computation of lexical

parts, leading to a processing advantage. On the other hand, a collocation is more likely to be processed via the decomposition route, whereby the collocational meaning is computed online as collocations are less conventionalised than idioms. It is important to note that we do not want to argue that direct access to meaning is not available for collocations or that both routes could not be used simultaneously and interactively. We propose that idioms are *more* likely to be activated by the direct access of the idiomatic concept compared to collocations.

4.4.2.3 Similarities and Differences between Lexical and Idiomatic Ambiguity Resolution

To some degree, idioms and ambiguous words are similar in resolving their inherent ambiguity. There was a context effect in the critical region for L1 speakers with idioms in figurative contexts being read more quickly than idioms in literal contexts. Hence, L1 speakers read *grey area* preceded by a disambiguating figurative context faster than *grey area* preceded by a disambiguating literal context. This suggests that L1 speakers preferred to interpret the MWS figuratively despite the fact that the literal interpretation was perfectly acceptable in the literal context. The current findings support the idea that the figurative interpretation of an ambiguous idiom is easier to be accessed than the literal interpretation, if accompanied by the support of prior context. This finding is similar to what is found for ambiguous words (Duffy et al., 1988; Rayner et al., 1994; Sereno et al., 1992). The reading time of an unbalanced ambiguous word is longer when the prior context supports the subordinate meaning of the ambiguous word, compared with the context supporting the dominant meaning. When the preceding context supports the subordinate meaning, access to the meaning is facilitated. However, due to the high frequency of the dominant meaning, the strong dominant meaning is still activated in the inappropriate context. In the pre-test of the present study, the figurative meanings were the dominant meanings for most idioms. It is therefore possible that there might be an automatic activation of the figurative meaning when L1 speakers read an idiom in the literal context. This competition between the two interpretations slowed down reading idioms in literal contexts.

This study did not detect a significant processing cost for idioms used literally compared to collocations. However, in lexical ambiguity, an ambiguous word following the context supporting the subordinate meaning is also processed more slowly than a frequency-matched control word. More balanced ambiguous idioms involved in the current study might explain the different findings. Though the figurative interpretations of the idioms on average were more frequent than their literal interpretations, the literal interpretations of the idioms were still frequent (the average frequencies of literal interpretations were 31%), compared to the subordinate meanings of unbalanced ambiguous words (< 7%). Research has shown that that the processing difficulty can be diminished when unbalanced ambiguous words have more frequent subordinate meanings (8% - 30%) (Wiley & Rayner, 2000).

The processing cost existed in idioms used literally compared to idioms used figuratively, which contrasts with previous findings (Carrol & Conklin, 2017; Siyanova-Chanturia, Conklin, & Schmitt, 2011). These studies did not observe any processing difference between these two context conditions. One explanation may be the different strengths of context employed (providing very strong or weak biasing context). More recent research of ambiguous words (Colbert-Getz & Cook, 2013; Leinenger & Rayner, 2013) has shown an interplay of context and meaning frequency, with a strongly biasing context for the subordinate meaning leading to similar reading times compared to a biasing context for the dominant meaning. However, when the biasing context is less strong, the dominant meanings are read faster than the subordinate ones (Martin et al., 1999). It is possible that the supporting context in the present study was not as strong as the one used in two prior studies (Carrol & Conklin, 2017; Siyanova-Chanturia, Conklin, & Schmitt, 2011). These studies used a long discourse as a prior disambiguating context, while the present study only had a short clause. One reason for our shorter contexts was the need to control for contextual predictability (which was not controlled in the other studies), which is more difficult with long informative contexts. Therefore, it remains unclear whether differences in contextual strength can affect resolving the interpretation of ambiguous idioms.

In contrast to L1 speakers, L2 speakers showed shorter reading times on the spill-over region for idioms in a literal context than for idioms in a figurative context. This finding is in line with previous studies (Carrol & Conklin, 2017; Shi et al., 2023; Siyanova-Chanturia, Conklin, & Schmitt, 2011) and suggests that L2 speakers have not yet developed strong connections between forms and idiomatic meanings. Even with the support of a biasing figurative context, idiomatic meanings were not activated quickly by L2 speakers. Instead, the computation of the literal meaning led to incongruency between the literal interpretation and the preceding, figurative-related context, which slowed down reading.

4.4.2.4 The Role of Decomposability

There was a significant interaction of context and local decomposability (of adjectives and nouns) (i.e., to what extent each component of an idiom contributed to the figurative meaning) in L1 speakers, and there was an interplay of context and local decomposability (of nouns) in L2 speakers. Increased decomposability sped up reading idioms in a figurative context in both L1 speakers and L2 speakers. However, global decomposability (i.e., to what extent the assembled literal meaning of an idiom contributed to the figurative meaning) did not affect reading times for either group. The inconsistency between local and global decomposability in this study might suggest that local measurements are more likely to detect the fine-grained decomposability effect. This might provide a clue as to why some studies that used global measurement failed to demonstrate a decomposability effect (Carrol & Conklin, 2014; Tabossi et al., 2009).

Importantly, the current finding provides evidence for the facilitatory role of decomposability in the processing of idioms, supporting the Idiom Decomposition Hypothesis (Gibbs & Nayak,1989; Gibbs et al., 1989). Regardless of L1 and L2 speakers, the increased decomposability sped up reading idioms in the figurative contexts. When idioms became more decomposable, individual components of the idioms contributed to overall figurative processing more. Thus, the literal analysis of individual components was more congruent with the overall figurative meaning, leading to a facilitation effect on reading idioms in figurative contexts.

The effect of decomposition suggests that the semantic meanings of the components of an idiom facilitate access to the idiomatic interpretation. On the one hand, this adds evidence against these hypotheses in which the idiomatic meaning of an idiom is accessed directly, and the literal computation is not part of the processing of the idiomatic meaning (the Direct Access Model, Gibbs, 1980; the Idiom List Hypothesis, Bobrow & Bell, 1973; the Lexical Representation Hypothesis, Swinney & Culter, 1979). If an idiom is accessed as a whole unit, one does not expect to observe any influence of semantic decomposability and literal analysis. On the other hand, if two routes (i.e., the direct conceptual access and decomposition routes (Cacciari & Tabossi, 1988; Cutting & Bock, 1997; Sprenger et al., 2006) are taken simultaneously, one would expect to observe the role of semantic decomposition.

It is not surprising that L2 speakers benefited from the decomposability in understanding idiomatic meanings. According to the dual-route model (van Lancker Sidtis, 2012; Wray, 2002; Wray & Perkins, 2000), when idioms are unknown or less familiar for language users, it is more likely that L2 speakers prioritise the computational route where each component of the idiom is analysed and integrated. The Literal Salience model (Cieślicka, 2006) also holds that the literal analysis of an idiom's components is the obligatory route for L2 speakers. Off-line rating tasks (Abel, 2003; Hubers et al., 2020) show that L2 speakers tend to judge idioms as more decomposable and transparent than L1 speakers do. This indicates that L2 speakers tend to treat all types of MWSs in a compositional manner. Unlike L1 speakers, L2 speakers are more prone to rely on an obligatory literal analysis of idioms, likely due to the lack of exposure to idioms (Cieślicka, 2006). This, in turn, allows the decomposability to affect the processing of idioms more directly.

4.4.2.5 Meaning Dominance and Decomposability

When an idiom was more predominantly used in a figurative interpretation, the facilitation effect of semantic decomposability decreased and increased semantic decomposability even started to slow down reading idioms for both L1 (observed in the critical region) and L2 (observed in the spill-over region). Both groups of speakers demonstrated a similar pattern and used a unified

method of processing ambiguous idioms. These findings suggest a shift from the priority of taking the decomposition route towards the priority of taking the direct access route in the processing of an idiom's figurative sense. The shift is dependent on the dominance of the meaning of the idiom. On the one hand, the increased decomposability elicited a facilitation effect on reading idioms when these idioms appeared in their subordinate interpretation more frequently. This indicates that semantic decomposability contributed a more significant role in the processing of idioms, and the decomposition route was more prominent than the direct conceptual activation route when idioms were more likely to be used literally. On the other hand, as idioms were more likely to be used figuratively, both groups of speakers were less likely to perform a literal analysis and gave priority to the direct retrieval of the figurative meaning route. This finding is in line with the eye-tracking study (Milburn & Warren, 2019), showing that increased meaning relatedness between the figurative and literal interpretation (arguably, greater decomposability) slowed down reading more when idioms were dominant in their figurative sense. Although the current and Milburn and Warren's studies showed this interaction effect, Milburn and Warren did not ask participants to rate to what extent the literal meaning of the components of the idiom contributed to the overall figurative meaning but instead asked how related the figurative and literal senses of an idiom are. In addition, Milburn and Warren only reported global measurements.

To summarise, the present study compared reading patterns for different types of MWSs. This allows us to assess to what degree resolving lexical and idiomatic ambiguity overlapped and investigate how meaning dominance and decomposability affected the reading of idioms in different contexts for both L1 and L2 speakers. First, this study replicated the collocation advantage for both L1 and L2 speakers and the idiom superiority effect for L1 speakers. Second, we found that the resolution of lexical and idiomatic ambiguity overlapped, and the degree of conventionality played a role in the ease of reading MWSs in L1 speakers. In addition, our results show that both L1 and L2 speakers relied on decomposability to access the idiomatic interpretations of an idiom, especially when the literal and figurative interpretations of the idiom were more balanced in nature.

CHAPTER 5 PROCESSING NON-CANONICAL, PASSIVIZED COLLOCATIONS

Abstract

Collocations, defined as frequently co-occurring words, have shown a processing advantage over novel word combinations in both L1 and L2 speakers. This collocation advantage is mainly observed for default configurations (e.g., *provide information*), however, collocations can also occur in variation configurations (e.g., *provide some of the information*). These variation collocations still show a processing advantage in L1 speakers, but not in L2 speakers. The present eye-tracking reading experiment investigated word order variation by passivizing collocations (e.g., *quality was improved*), while controlling for contextual predictability. Our results revealed that changing the word order of collocations did not eliminate the collocation advantage in either L1 or L2 readers, and that this advantage was independent of contextual predictability.

Keywords**:** collocations, variation collocations, multi-word sequences, bilingualism, eye-tracking

5.1 Introduction

One type of MWS that has received considerable attention is collocation. A collocation (e.g., *lower risk*) consists of a node (i.e., words that are searched for in corpora, e.g., *risk*) and a collocate (i.e., words that co-occur with the node, e.g., *lower*). There are two different approaches to defining and identifying collocations. The phraseology-based method (Cowie, 1998; Howarth, 1998) emphasises the importance of semantic transparency and of restrictions on lexical substitution. Thus, collocations are not fully compositional or semantically transparent (e.g., *take pictures* and *draw conclusions*). The frequency-based method (Sinclair, 1991) broadly defines collocations as two words that co-occur more frequently than we expect by chance (Biber et al., 1999).

In keeping with previous eye-tracking studies of collocations (e.g., Vilkaitė, 2016), we followed the frequency-based method in this study. Following this method, collocations generally have transparent and literal meanings (Carrol & Conklin, 2020). Identifying collocations involves the extraction of collocations from a corpus based on specific association measures, which provide a measure of the strength of co-occurrence between two words (Hunston, 2022). Association measures do not solely rely on phrase-level frequency but can take other properties of collocations into account as well, such as exclusivity. The Mutual Information (MI) score[12] is a widely used association measure which highlights the exclusivity of a collocation. The higher MI-score a collocation has, the more exclusively associated the words in the collocation are, and the less likely components of the collocations form other collocations. Typically, word combinations with an MI-score larger than 3 are considered collocations (Hunston, 2022). For our study, we selected collocations with an MI-score of >3 and which were not listed in the Oxford Dictionary of English idioms (Ayto, 2009).

[12] MI scores operate on a scale without a minimum and maximum (Gablasova et al., 2017).

5.1.1 Collocation Processing

Studies of collocation processing (Durrant & Doherty, 2010; Siyanova & Schmitt, 2008; Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2015, 2018) show that collocations (e.g., *provide information*) are processed faster than novel word combinations (e.g., *compare information*). Siyanova and Schmitt (2008) asked participants to make sensicality judgements about frequent and infrequent collocations and found that frequent collocations were judged faster than novel word combinations.

However, tasks that involve judgements, such as the ones above, do not reflect natural reading and only measure the endpoint of processing. Eye-tracking provides a more ecological approach to researching collocation processing. It also provides measures tapping into earlier stages of processing (e.g., first fixation duration, gaze duration) and those that capture reanalysis and more in-depth processing (e.g., total reading time, fixation counts) (Clifton et al., 2007; Holmqvist et al., 2011; Rayner, 1998). Eye-tracking studies have tended to use different approaches to identify collocations. Early studies (Frisson et al., 2005; McDonald & Shillcock, 2003b, 2003a) used transitional probabilities (TP, the statistical likelihood that a word will precede or follow another word), while more recent studies (Li et al., 2021; Sonbul, 2015; Vilkaitė, 2016) used collocational frequency and MI-score that took into account a more flexible collocation window. Both methods are frequency-based and draw on quantitative evidence of word co-occurrence in corpora but result in different selections of collocations (Gablasova et al., 2017). To illustrate, word pairs that were selected based on an MI score of 3 (Vilkaitė, 2016) had a much higher phrase-level frequency and TP than the word pairs that were selected based on TP (Frisson et al., 2005) (360 vs. 60, and .0127 vs. .00677, respectively).

Early eye-tracking studies using TP found that collocations are read faster than control (low-TP) sequences, but also that this advantage may result from contextual predictability. McDonald & Shillcock (2003a, 2003b) found that verb-noun pairs with high TP (e.g., *accept defeat*) had shorter reading times than ones with low TP (e.g., *accept losses*). This manifested itself in early eye-

movement measures, suggesting that TP affects immediate processing. However, Frisson et al. (2005) extended this research by introducing a contextual manipulation that made the high TP and low TP verb-noun pairs either predictable or not from the preceding context and found strong effects of contextual predictability but no independent effect of TP. Frisson et al. suggested that contextual predictability rather than TP affects reading and that TP is one of the sources readers can draw on to guide prediction. Hence, it can be challenging to disentangle contextual predictability and TP because contextual predictability, as measured by cloze probabilities (see below), is likely to capture some degree of TP. This raises questions of whether collocation effects might be driven by contextual predictability.

Recently, eye-tracking research on collocations has mainly adopted phrase-level frequency and MI (score >3) as the selection method and consistently found that collocations are read faster than novel word combinations. Sonbul (2015) examined adjective-noun word combinations with phrase-level frequencies ranging from low (e.g., *extreme mistake*), middle (e.g., *awful mistake*), to high (e.g., *fatal mistake*). Collocations with higher frequency were processed more quickly than novel word combinations with lower frequency, indicating that language users are sensitive to phrase-level frequency. It is worth noting that the effect was only observed in the early (first pass/run reading time) but not the late eye-tracking measures (total reading time and fixation count). The early effect in Sonbul's study was potentially driven by uncontrolled adjective's frequency. In order to examine this possibility, we reanalysed Sonbul's (2015) data by splitting items into five quantiles based on the phrase-level frequency (the way of splitting items followed the original experimental design). The first quantile included novel word combinations with zero or one occurrence in the corpus, and the fifth quantile included the most frequent collocations. A regression model with logged adjectives' frequencies as the dependent variable and phrase-level frequency as a categorical predictor with five levels showed that the most common collocations contained more frequent adjectives than novel word combinations (β = 0.69, t = 3.52, p < .001) and the second quantile (β = 0.47, t = 2.40, p = .018). In addition, the contextual predictability was not

measured, likely leading to unmatched predictability across different conditions. Individual word frequency and predictability have been reported as robust predictors of reading times (Clifton et al., 2016; Rayner, 1998). Similarly, computational models of eye movements in reading (e.g., E-Z Reader, Reichle et al., 1999, 2003; SWIFT, Engbert et al., 2005) have treated predictability and word frequency as two fundamental linguistic variables that affect eye movements. Hence, Sonbul's conclusions can potentially be challenged by these uncontrolled variables.

Li et al. (2021) improved on Sonbul's (2015) study by controlling adjectives' length and frequency. Additionally, they used a two (collocation: strong collocations vs. weak collocations) by two (context: predictable vs. neutral context) design. This design enabled them to test whether the effect of collocation strength can be independent from contextual predictability. Two main effects, collocation strength and predictability were observed, which did not interact, indicating that the collocation strength effect found in both early and late eye-tracking measures did not depend on the context. Li et al. (2021) operationalised contextual predictability in terms of cloze probability (Taylor, 1953), though only for the first word in the MWS. The cloze probability test is commonly used to estimate predictability and involves asking participants to provide the next word or words to continue a given sentence fragment. Li et al. asked participants to fill in a word for an incomplete sentence up to the start of the collocation (e.g., *My friend and I made a* _____), which provides a predictability value for the first word of the collocation. However, no cloze test was carried out for the second word (as would be the case for e.g., *My friend and I made a fatal*_____), and if, as argued by Frisson et al. (2005), TP can affect cloze, the predictability of the second word is likely higher for the collocations than the control condition. Hence, it remains unclear to what extent collocation effects are driven by contextual predictability.

It is unclear whether Sonbul's (2015) findings with adjective-noun collocations generalise to verb-noun collocations. These two types of collocations may be processed differently, as Wolter and Yamashita (2015) found that native English speakers judged Japanese verb-noun collocations (collocations that exist only in Japanese and are translated into English) faster than non-collocations,

but this was not the case for adjective-noun collocations. In addition, Frisson et al.'s eye-tracking study did not observe the collocation effect from their verb-noun pairs. Vilkaitė (2016) examined whether adjacent verb-noun collocations (e.g., *provide information*) were processed differently from non-adjacent verb-noun collocations (e.g., *provide some of the information*). A non-adjacent collocation means there are words intervening between the node and the collocate (here, between the verb and the noun). Non-adjacent collocations are relatively common in natural language use and reflective of a certain flexible span for the co-occurrence of words (Krishnamurthy et al., 2004; Stubbs, 2002). Vilkaitė observed that collocations were consistently processed faster than novel word combinations in both early and late measures, irrespective of being adjacent or non-adjacent collocations. This supports the view that non-adjacent collocations also enjoy processing advantages. Vilkaitė (2016) further claims that adjacent collocations have greater processing advantages than non-adjacent collocations, as interactions between the collocational status and adjacency nearly reached significance in late measures and larger facilitation was observed for the final word from adjacent collocations than from non-adjacent ones.

Though Vilkaitė's study (2016) seems to detect a processing advantage for both adjacent and non-adjacent collocations, it suffered from the same limitation as previously discussed, i.e., inadequately controlled contextual predictability. In Vilkaitė's study, participants filled in a word for an incomplete sentence up to the final word of the collocation. Contextual predictability was numerically higher in both adjacent (4%) and non-adjacent collocations (7%) than in adjacent (0%) and non-adjacent controls (1%). As the study did not provide a statistical comparison, we selected the 40 (out of 48) adjacent pairs of items and related sentences from Vilkaitė's study. Each pair had one collocation and one novel word combination (e.g., *provide information* vs. *compare information*). Sixty native English speakers, split over four lists, were asked to provide written continuations for the sentences truncated immediately before the final word of the collocations or novel word combinations. This post-hoc test showed that contextual predictability was not well matched between collocations and novel word combinations, with collocations being significantly

more predictable than novel word combinations ($t(584) = 5.00$, $p < .001$). The significant difference in contextual predictability between the two conditions could explain why predictability was a significant predictor in their early measures' models. It also raises concerns about the findings of the collocation effect in both early and late measures. To summarise, while more recent eye-tracking studies (Li et al., 2021; Sonbul, 2015; Vilkaitė, 2016) showed a processing advantage for collocations over novel word combinations, the results were potentially confounded by a lack of contextual predictability control. Hence, it is unclear whether the collocation effect would still hold in both early and late eye-tracking measures when contextual predictability is strictly controlled.

5.1.2 Variation of Collocations and L1 Processing

Most reading studies discussed above presented collocations in their canonical forms. MWSs are conventionally treated as sequences of words with a canonical form and a commonly agreed interpretation. However, at other times, MWSs can tolerate some degree of internal structure variation while maintaining their meaning (e.g., *under the control of the manager* vs. *under his control*). There are different ways in which the form of collocations can be modified: 1) inserting words between the node and the collocate, making them non-adjacent (e.g., *provide information* vs. *provide some of the information*), 2) varying the collocation's morphological forms (*provide information* vs. *provided/providing information*), and 3) varying the collocation's word order (*provided information* vs. *information was provided*).

A few studies looked at the effect of variation on MWSs processing and there is inconsistent evidence that modified MWSs retain online processing advantages. In a study combining neural and behavioural measures, Molinaro et al. (2013) tested variations of lexical bundles by inserting an adjective (*capable*) before the first noun (*hands*) (e.g., *in the hands of* (lexical bundle) vs. *in the capable hands of* (variation)). The self-paced reading behavioural results showed that reading times were longer for the noun (*hands*) and the following word (*of*) in the variation condition compared to lexical bundles in their canonical forms. The longer RTs in the variation condition suggest that extra

processing is required when the canonical form is disrupted. In contrast, there was a reduced N400[13] amplitude for the noun (*hands*) in the variation condition compared to the lexical bundles, indicative of less semantic processing effort and easier contextual integration. According to the authors, the reduced N400 may show that inserting an adjective does not disrupt lexical bundle processing. However, Left Anterior Negativity[14] (LAN) was found for the second nouns (e.g., *restorers*) that followed variation lexical bundles (e.g., *in the (capable) hands of restorers*), suggesting additional difficulty in processing and integrating non-canonical MWSs. In short, the mixed findings of Molinaro et al.'s study do not provide reliable evidence for processing (dis)advantages of variation MWSs.

As discussed above, Vilkaitė (2016) found shorter reading times and processing advantages not only in adjacent collocations but also in non-adjacent collocations where words had been inserted between the node and the collocate. Recently, Vilkaitė-Lozdienė (2022) tested whether collocations retained their processing advantage when their morphological forms were changed. The researcher presented native Lithuanian readers with stories containing verb-object collocations and novel word combinations, each in three morphological forms: the canonical form, namely infinitive + accusative (e.g., *švęsti pergalę,* '*to celebrate the victory*'), past tense third person + accusative (e.g., *šventė pergalę,* '*celebrated the victory*'), and attributive passive participle (e.g., *švęsta pergalę,* '*the victory one celebrated*'). At the phrase level (i.e., reading times for the entire phrase), collocations were read faster than controls across morphological form variations. The collocation advantage was found only in late (i.e., total reading times) but not early, eye-tracking measures. Additionally, collocations in their canonical form were read the fastest, while past tense collocations were read the slowest, though still faster than their corresponding controls. At the final word level (i.e., reading times for the final word of collocations), both early and late measures revealed faster reading times and collocation variations were fixated less often than controls. Taken together, this evidence seems to indicate that morphological variation retains the collocational advantage. However, it is important

[13] N400 is closely related to the semantic processing of a word and the ease of a word's integration to the sentence meaning (Holcomb, 1988).

[14] LAN is often evoked by anomalies or increased working load (Kutas & Federmeier, 2000).

to be cautious about generalising these findings since Vilkaitė-Lozdienė (2022) also did not control for possible contextual predictability effects.

To summarise, previous studies on the variation of collocations reveal that collocations may maintain their processing advantage even with inserted words and in different morphological forms. This shows that collocations are flexible and can be modified without negative impacts on processing. This may mean that they are not stored as a whole chunk, which would contradict the holistic storage hypothesis of MWSs (Wray, 2012). This hypothesis was initially proposed to explain the faster recognition and processing of MWSs (Goldinger, 1996; Jiang & Nekrasova, 2007; Wray & Perkins, 2000) and holds that MWSs are processed as wholes without activating and combining individual words.

However, to our knowledge, there have been no studies that investigated effects on processing when the word order of the collocation is changed (e.g., *provide information* vs. *information was provided*). An eye-tracking study examined word order change for idioms, using passivisation to reverse them. They compared canonical idioms (e.g., *he kicked the bucket*) and idioms in their reversed forms (*the bucket was kicked*) to their counterpart controls (e.g., *he kicked the apple* vs. *the apple was kicked*) in English. Late eye-tracking measures for both the phrase-level and word-level regions showed that word order (canonical or reversed) and idiomaticity (idiom or control) were significant predictors, and there was no interaction. Idioms were read faster than controls both in canonical and reversed word order, and canonical idioms were read faster than reversed ones. The authors further claimed that reversed idioms retain the conventional figurative meaning that is associated with the idioms' canonical form and order. However, this conclusion is debatable. The idioms in Kyriacou et al. (2020) were predictable from context, as evidenced by the results of the cloze test. Furthermore, active idioms (0.38) were significantly more predictable than reversed idioms (0.28). Assuming that the predictability of the control items was low (no cloze results were reported for these items), the pattern of results could be explained by the differences in predictability.

5.1.3 Processing Collocations in L2

The collocation processing advantages discussed above were found in L1 speakers. Several studies suggest that proficient L2 learners of English also demonstrate the collocation processing advantage driven by the phrase-level frequency (Li et al., 2022; Öksüz et al., 2021; Siyanova & Schmitt, 2008; Sonbul, 2015; Vilkaitė & Schmitt, 2019; Wolter & Gyllstad, 2011, 2013). Siyanova and Schmitt (2008) asked L1 and L2 speakers to judge frequent and infrequent collocations and found that both L1 and L2 speakers judged more frequent collocations faster than less frequent ones. Sonbul (2015) recorded the eye movements of participants and found that more frequent L2 collocations showed an advantage in early eye-tracking measures, indicating that L2 speakers are also sensitive to phrase-level frequency.

While these studies found similarities between L1 and L2 speakers in processing canonical collocations, to our knowledge, only one study investigated the effects of variation on L2 collocation processing. Vilkaitė and Schmitt (Vilkaitė & Schmitt, 2019) presented L2 speakers with the same canonical and non-adjacent collocations used by Vilkaitė (2016) with L1 speakers. Though L2 speakers, similar to L1 speakers, read collocations more quickly than novel word combinations, L2 speakers did not show a processing advantage for non-adjacent collocations. However, the results are difficult to interpret for two reasons. First, as explained above, Vilkaitė (Vilkaitė, 2016; Vilkaitė & Schmitt, 2019) did not control contextual predictability. Second, L2 speakers in the Vilkaitė (2019) study had various first languages. This may be an issue, because Wolter and Gyllstad (2011, 2013) found that the L1 exerts an influence on L2 collocation processing. In these studies, L2 speakers responded more quickly to L2 collocations that were congruent (i.e., having a semantically equivalent collocation in the L1), compared with incongruent collocations. To address these two issues, the present study controlled contextual predictability and restricted the sample to a single L1 (Chinese).

5.1.4 The Present Study

The present study consists of an eye-tracking experiment exploring the effect of changing the word order of collocation on processing during sentence reading. The first research question is whether collocations still maintain their processing advantage when word order is changed (e.g., she brought peace -> peace was brought). The second research question is whether similar effects are observed for both L1 and L2 speakers. The corresponding hypotheses to the two questions above are: L1 speakers will still show the collocation processing advantage, though it is unclear whether L2 speakers will show a similar effect. The third research question concerns the relationship between the collocation effect and a contextual predictability effect, and whether these are separate effects. In addition, given that eye-tracking can provide evidence for different stages of processing, we will examine whether the collocation advantage shows up during an early stage of the reading process or only at a later processing stage. We are not certain whether the collocation effect is separate from the contextual predictability effect since earlier and recent studies have inconsistent findings.

5.2 Method

5.2.1 Participants

Thirty-nine Chinese speakers of L2 English (age: *M* = 23.8, *SD* = 3.06) and 40 native speakers of British English (age: *M* = 22, *SD* = 3.93) participated in the study (one additional L2 speaker was removed due to major track loss). Since this was, to our knowledge, the first eye-tracking study investigating word order effects on collocation processing, it was difficult to conduct an appropriate power analysis for mixed effects models. To achieve a reliable effect size, the present study had more observations (1920) than recent eye-tracking collocations studies (1536 in Li et al., 2021; 1120 in Vilkaitė, 2016). L2 participants' English proficiency was estimated using the Lexical Test for Advanced Learners of English (LexTALE) (Lemhöfer & Broersma, 2012) and Vocabulary Levels Test of 3k, 5k and 10k (VLT, Schmitt et al., 2001). The average LexTALE score was 66.21 (out of 100), and the average VLT score was 53.82 (out of 90). On average, L2 speakers started learning English at age 7.42 and spent 2.5 years in an English-speaking country. Participants' language background and usage data are summarised in Appendix 5.1. All participants had normal or corrected vision and no self-reported history of reading difficulties.

The participants gave informed consent. The experiment was conducted in accordance with the British Psychological Society ethical guidelines and was approved by the University of Birmingham's ethical committee (ERN_19-0936).

5.2.2 Materials

Materials consisted of 192 verb-noun phrases. There were 48 sets of four phrases, corresponding to four conditions. Items were embedded within sentences consisting of four sections (see Figure 5.1). The pre-target contexts and spill-over regions were the same across the four conditions. The end of sentence region was kept the same whenever possible, but some minor changes were sometimes required to ensure sentences sounded natural (note that the end-of-sentence region was not analysed further). The pre-target contexts were kept as neutral and

uninformative as possible, and collocations were placed in the middle of sentences to avoid line break issues and sentence wrap-up effects (Rayner et al., 2000). All experimental materials can be found in Appendix 5.5.

Figure 5.1

Example of a Sentence with a Collocation and its Control in Canonical and Reversed Order

Word order	Pre-target context	Target region	Spill-over region	End of sentence
Canonical		they brought peace/meat		
Reversed	Margaret is not sure that	peace/meat was brought	to the	village.

Note. Collocations are underlined. Collocations and novel word combinations are separated by a slash.

5.2.2.1 Materials Preparation

Target phrases were selected from two previous studies (Frisson et al., 2005; Vilkaitė, 2016) and contained words included in the list of the first 2000 most frequent lemmas of verbs and nouns in the British National Corpus (BNC). Only collocations and novel word combinations that could be used in their passive voice were selected. To control for L1-L2 congruency effects, no collocation had a semantic equivalent in Chinese, as assessed by five native Chinese speakers. Seventy-nine collocation – novel word combination pairs were selected for further testing.

To ensure that the areas of interest had similar lengths in the canonical and reversed conditions, a subject was added before canonical phrases (e.g*., he improved quality*; cf. Kyriacou et al., 2020). For example, *he improved quality* and *quality was improved* both consist of three words. Additionally, the tense of the canonical and reversed constructions was kept the same (e.g., *improved quality -> quality was improved; had prevented damage -> damage had been prevented*).

5.2.2.2 Predictability

To assess predictability in the context sentences, 79 sentence frames were created, into which a collocation and its corresponding novel word combination could be inserted interchangeably (see Figure 1). The sentence frames were split into four lists, with an item only appearing once in each list and each list containing a different condition of the same item. Sixty native British English participants (age: *M* = 29.15, *SD* = 7.4; gender: female, *N*= 36), 15 per list, provided written continuations for sentences truncated immediately before the final word of the collocation or novel word combinations (e.g., *Margaret is not sure that they brought* ____ or *Margaret is not sure that meat was* _____).

Thirty-one pairs had high context predictability scores and were eliminated. In the remaining 48 pairs, predictability was not statistically different between canonical collocations and canonical novel word combinations ($t(719) = 1.34$, $p = .18$), and between the reversed collocations and reversed novel word combinations ($t(719) = 1.15$, $p = .25$). The mean cloze values for the final words of the 48 collocations (and their controls) selected for the present study were below or near 1%. Hence, the preceding contexts provided little information regarding the sentence continuation and collocations and novel word combinations were equally non-predictable.

Another 15 native British speakers (age: *M* = 34.8, *SD* = 6; gender: female, *N*= 1) were asked to continue the selected 48 sentences truncated before the whole collocation/control (e.g., *Margaret is not sure that they* _____). Predictability for these sentence fragments was zero.

5.2.2.3 Frequency and Length

The 48 pairs of collocations and novel word combinations were closely matched for individual word length and word frequency (see Appendix 5.2). All frequency information, including individual word frequency and phrasal-level frequency with different morphological forms, was obtained from the BNC using Sketch Engine (Kilgarriff et al., 2004). The verbs in each pair did not differ in lemma frequency ($t(47) = -0.08$, $p = .93$), past participle frequency ($t(47) = 0.19$, $p = .84$), and root frequency ($t(47) = -0.4$, $p = .64$). The nouns also did not differ in lemma frequency ($t(47) = -0.23$,

p = .81) or form frequency (t(47) =1.50, p = .13). The verbs,' nouns,' and items' lengths were also matched within each pair (t(47) = 0.44, p = .66; t(47) = -0.57, p = .57; t(47) = -0.70, p =.49).

5.2.2.4 Apparatus and Procedure

Participants were tested individually in a quiet laboratory using an Eyelink 1000 SR research eye-tracker (SR Research, Ontario, Canada) with a tower mount recording the right eye gaze location with 1000 Hz during reading. All sentences were displayed either on a single line or on two lines, and the regions of interest never appeared at the end of a line. When there were two lines, double line spacing was used. Sentences were presented in 14 Courier New font as black text with white background. At the beginning of the experiment, a 5-point calibration procedure was carried out, and a re-calibration was performed whenever necessary. Between each trial, a drift correction check was performed, and the presentation of the sentence was automatically triggered by a fixation-cross placed just the left of the first word of a sentence.

The reasons for a 5-point calibration grid are: (1) the 9-point calibration grid provides four extra points positioned in the four corners of the screen where our experimental sentences would not be displayed. If the experiment ran with longer paragraphs, it would be more reasonable to use the 9-point calibration. (2) Given that it is more difficult to obtain accurate calibrations for those four corners, resulting in larger errors between the estimated gaze positions and actual target positions, this would degrade average calibration accuracy.

Participants were instructed to read silently for comprehension at a normal pace. At the start of each trial, a fixation cross appeared on the left of the screen. After participants fixated on the cross, the cross was replaced by a sentence. Once having finished reading a sentence, participants pressed the spacebar to make the sentence disappear. To check that participants were reading for meaning, half of the sentences were followed by a comprehension question. Half of the questions required a "Yes" answer and half a "No" answer. The reading experiment took approximately 20 minutes for L1 participants and 30 minutes for L2 participants. After participating

in the reading experiment, participants filled out a questionnaire about their linguistic background survey and the two vocabulary tests. This took around 15 minutes.

5.2.2.5 Pre-processing of Data

Prior to analyses, sentences with major track loss, e.g., as a result of blinking, were removed (1% of the trials in total). Following software standard fixation cleaning procedures, short fixations below 80ms were merged into nearby longer fixations. Fixations under 100ms and above 1000ms were trimmed from the data. Participants generally did not have difficulties answering comprehension questions, with an average accuracy of 91% (92% for L1 speakers and 90% for L2 speakers). Fixation times that were more than 2.5 *SDs* from the mean for each participant were excluded as outliers (Miller, 1991; Ratcliff, 1993. For L1 data, 43 trials (2.37%) were removed, consisting of 13 trials (0.71%) from canonical collocations, 12 trials (0.66%) from canonical controls, 9 trials (0.49%) from reversed collocations and 9 trials (0.49%) from reversed controls. For L2 data, 36 trials (2.08%) were removed, consisting of 10 trials (0.57%) from canonical collocations, 7 trials (0.40%) from canonical controls, 6 trials (0.34%) from reversed collocations and 13 trials (0.75%) from reversed controls.

Most previous eye-tracking studies focused on the processing of individual words, and it is challenging to apply the same areas of interest to MWSs. Carroll and Conklin (2014) suggested a hybrid method: analysis of the MWS region and the final word of the MWS region, for two reasons. First, the effect is more likely to be observed on the final word after encountering the whole MWS. Second, the effect might occur not only in the final word but also when re-reading the MWS. Following Carroll and Conklin's (2014) recommendation and previous collocation studies (Li et al., 2021; Sonbul, 2015; Vilkaitė-Lozdienė, 2022; Vilkaitė, 2016), we report two regions of interest: the whole phrase region and the final word of the phrase region.

Several reading measures are reported. For the whole collocation region to capture both early and late processing, early processing measures consisted of first run reading time (sum of fixations on a region before exiting the region for the first time); later processing measures included

regression path reading time (sum of all fixation durations from when the current interest area was entered until moving further in the text) and selective regression path reading time (sum of fixation durations of the current interest area before moving further in the text); and comprehensive measures included total reading time (sum of all fixation durations in a region of interest) and total fixation count (sum of the number of fixations in a region of interest). For the final word analysis, measures were: first fixation reading time (first fixation on a word before exiting the region for the first time), first run reading time, regression path reading time, selective regression path reading time, total reading time, and total fixation count.

5.2.3 Statistical Analysis

All analyses were carried out using R version 4.2.1 (R Core Team, 2022) with packages *lme4* (Bates et al., 2015) and *lmerTest* (Kuznetsova et al., 2020). Reading time measures were analysed with linear mixed-effect models, and fixation counts were analysed with generalized linear mixed effect model (*glmer*()). Language background was a between-group factor with two levels (English native speakers vs. L1-Chinese L2-English speakers). First, models were fitted with L1 and L2 speakers' data separately to see how proficiency affected processing in the L2 speakers. Then, models were constructed with both L1 and L2 speakers to test for differences in collocation processing between L1 and L2 speakers. For each dependent variable and each region of interest, a core model included phrase type (sum coded: novel word combination: -1 vs. collocation: 1), word order (sum coded: reversed: -1 vs. canonical: 1) and their interaction as fixed effects. A full random model was used whenever possible (Barr et al., 2013). If convergence did not occur, "bobyqa" or "optimx" optimiser (Powell, 2009) was used to increase iterations before simplifying random effects by removing interactions and then individual terms until the model converged. The trial number was centred and entered into models as a covariate.

For the L1 speakers, the full model included the core model and trial number. For the L2 speakers, vocabulary score and the interactions with vocabulary score were entered as additional

fixed effects after entering a core model and trial number. Both LexTALE scores and VLT scores measured L2 speakers' English proficiency, and these two scores were highly correlated ($r(42) = 0.72$, $p < .001$). The two scores were separately entered into two models corresponding to each reading measurement. The VLT score was added to L2 speakers' models, and similar patterns in terms of significance and direction were found when the LexTALE scores were added. The decision to retain LexTALE scores instead of VLT scores is based on the BIC and AIC values, which indicated that the models with LexTALE scores provided a better fit for the data. For the combined L1 and L2 speakers' data, the full model included a core model, trial number, language background factor (sum coded: L1 speakers: -1 vs. L2 speakers: 1), and all interactions with language background. All full models were subsequently reduced with the drop 1() function, which performs model comparisons with a chi-square test. Note, however, that the two manipulated factors (phrase type and word order) were always kept in the reduced models to reflect the experimental design, even if their effects were not significant. The interactions with phrase type, word order, language background, vocabulary score and trial number were removed if they were not significant. Based on these tests, reduced models were reached to best fit the data. In the reduced models, variables with a *t/z* value above 1.96 are considered significant at the alpha=.05 level (Baayen et al., 2008). The significance of interactions was assessed by model comparisons with the anova () function. For all, vif () function from the *car* package (Fox et al., 2023) was used to test for collinearity. All variance inflation factors were around 1, indicating a very low degree of correlation among variables.

For models including continuous dependent variables, when there was a main effect of collocation, we used Bayes factors (Kass & Raftery, 1995) to compare the relative evidence for models with and without interactions between phrase type, word order and language background. The Bayes factors analysis was performed using the *ImBF* function from the *BayesFactor* package (Morey et al., 2022). When Bayes Factors (BFs) > 10, there is strong evidence for models including an interaction, while BFs > 1 provide evidence for models without an interaction. BFs between 3 and 10 provide weak to moderate evidence for models with an interaction.

5.3 Results

Reading time averages are shown in Table 5.1. L1 speakers' statistical results are reported in Tables 5.2 and 5.3. L2 speakers' statistical results are reported in Tables 5.4 and 5.5. The statistical results from the combined L1 and L2 models are displayed in Appendices 5.3 and 5.4.

Table 5.1

Summary of Measures in L1 and L2 Speakers

	L1 Speakers				L2 Speakers			
Measures (ms)	**Canonical collocations**	**Canonical novel word combinations**	**Reversed collocations**	**Reversed novel word combinations**	**Canonical collocations**	**Canonical novel word combinations**	**Reversed collocations**	**Reversed novel word combinations**
Whole phrase								
First run reading time	632 (307)	650 (305)	567 (284)	577 (282)	1093 (543)	1059 (569)	1046 (499)	1056 (539)
Regression path reading time	719 (358)	736 (335)	688 (317)	704 (318)	1272 (628)	1256 (575)	1256 (521)	1282 (588)
Selective regression path reading time	689 (334)	714 (313)	648 (273)	666 (276)	1218 (535)	1227 (550)	1211 (479)	1232 (554)
Total reading time	869 (448)	931 (455)	854 (446)	930 (457)	1797 (1201)	1945 (1186)	1890 (1249)	2025 (1307)
Total fixation counts	4.12 (1.99)	4.45 (2.10)	4.11 (2.05)	4.45 (2.07)	6.93 (4.28)	7.57 (4.34)	7.32 (4.38)	7.83 (4.65)
Final word								
First fixation duration	217 (62)	224 (67)	221(71)	222 (67)	286 (98)	292 (102)	279 (98)	279 (99)
First run reading time	243 (92)	265 (113)	255 (105)	263 (108)	388 (158)	404 (170)	418 (200)	427 (202)
Regression path reading time	314 (182)	357 (213)	324 (199)	327 (182)	553 (358)	575 (353)	518 (288)	539 (351)
Selective regression path reading time	276 (119)	295 (126)	283 (124)	288 (126)	457 (192)	473 (196)	475 (236)	493 (250)
Total reading time	355 (192)	382 (218)	383 (226)	400 (228)	714 (507)	767 (500)	767 (553)	867 (598)
Total fixation counts	1.67 (0.88)	1.75 (0.92)	1.80 (1.04)	1.88 (1.05)	2.67 (1.80)	2.91 (1.86)	2.85 (1.88)	3.25 (2.09)

5.3.1 L1 Speakers Results

5.3.1.1 Phrase Analysis

Table 5.2 shows that the phrase type effect was observed on total readings times (β = -34.84, t = -2.58, p = .012) and total fixation counts (β = -0.04, z = -2.84, p = .004): L1 speakers read collocations faster and fixated fewer times on collocations than novel word combinations. The main effect of word order was found on first run reading times (β = 35.15, t = 4.6, p < .001) and selective regression times (β = 21.82, t = -2.52, p = .014), with reversed constructions being processed more quickly than canonical ones. No significant main effects or interactions were observed for the other measures (all ps > .7). Furthermore, there was no interaction between phrase type and word order on total reading times ($\chi 2$ (1) = 0.15, p = .697; BFs = 0.09) and total fixation counts ($\chi 2$ (1) = 0.072, p = .788). The significant main effect of phrase type and the lack of interaction with word order indicates that L1 speaker consistently read collocations faster than novel word combinations, regardless of their word order.

5.3.1.2 Final Word Analysis

To avoid comparing final words from different syntactic constructions, reversed and canonical phrases were analysed separately (summarised in Table 5.3). For L1 speakers, the phrase type effect in canonical construction reached significance in the first run time model (β = -10.23, t = -2.40, p = .021) and in the regression path reading time model (β = -20.11, t = -2.46, p = .017).

5.3.2 L2 Speakers Results

5.3.2.1 Phrase Analysis

There was a phrase type effect on total reading time times (β = -70.60, t = -2.64, p = .009) and total fixation counts (β = -0.04, z = -4.24, p < .001), with collocations being processed more quickly than novel word combinations (see Table 5.4). The main effect of word order was not significant on any measure. There was no interaction between phrase type and word order on total reading times ($\chi 2$ (1) = 0.01, p = .996; BFs = 0.006) and total fixation counts ($\chi 2$ (1) = 0.35, p = .549). This means that the collocational processing advantage was independent of word order in L2

speakers as well. No other effects approached significance. LexTALE score, which measures participants' vocabulary level, was significant on all eye-tracking measures. The higher the score participants had, the faster they read. However, we did not detect any interactions between the LexTALE score and phrase type on any measures.

5.3.2.2 Final Word Analysis

Table 5.5 shows that there was a phrase type effect in the total reading time for verbs (β = -47.49, t = -2.04, p = .044) and in the total fixation count measures for nouns (β = -0.05, z = -2.25, p = .025) and verbs (β = -0.06, z = -2.45, p = .014). Participants were less likely to fixate on the final word in collocations than in novel word combinations. The interaction between LexTALE and phrase type was significant for the canonical constructions for the total fixation count measure (β = -0.05, z = -2.23, p = .026). L2 speakers with higher vocabulary scores were more likely to show fewer fixations on the final word of collocations relative to novel word combinations, while participants with lower vocabulary scores were more likely to show the opposite pattern (Figure 5.2).

Table 5.2

L1 Speakers: Summary Statistics for the Phrase Region

Factor	Statistics	FRRT	RPRT	SRPRT	TRT	TFC
Intercept [global mean]	*β*	608.49	714.22	681.99	899.98	1.42
	SE	23.69	33.62	30.84	40.8	0.04
	t/z	25.68	21.24	22.12	22.06	35.36
Phrase type [collocation]	*β*	-5.26	-7.94	-9.49	-34.84	-0.04
	SE	7.87	8.07	8.17	13.52	0.01
	t/z	-0.67	-0.93	-1.16	-2.58*	-2.84*
Word order [canonical]	*β*	35.15	14.25	21.82	4.5	0.002
	SE	7.65	8.07	8.66	12.66	0.01
	t/z	4.6*	1.77	2.52*	0.36	0.11

Note. FRRT: first run reading time, RPRT: regression path reading time, SRPRT: selective regression path reading time, TRT: total reading time, TFC: total fixation count.

* = $p < .05$.

Table 5.3

L1 Speakers: Summary Statistics for the Final Word Region

Factor	Statistics	FFRT	FRRT	RPRT	SRPRT	TRT	TFC
			Final words: nouns [canonical structure]				
Intercept [global mean]	*β*	219.53	252.27	332.72	283.15	368.3	0.53
	SE	4.37	6.97	15.07	9.28	16.06	0.03
	t/z	50.28	36.19	22.07	30.52	22.93	15.1
Phrase type [collocation]	*β*	-3.17	-10.23	-20.11	-8.55	-13.78	-0.02
	SE	2.64	4.27	8.19	5.11	8.25	0.03
	t/z	-1.2	-2.4*	-2.46*	-1.67	-1.67	-0.8
			Final words: verbs [reversed structure]				
Intercept [global mean]	*β*	221.22	258.48	324.8	285.23	389.7	0.58
	SE	5.14	7.49	12.55	9.36	18.34	0.04
	t/z	43.02	34.5	25.88	30.47	21.24	14.3
Phrase type [collocation]	*β*	-0.71	-3.86	-2.13	-2.94	-9.69	-0.02
	SE	2.58	3.7	7.2	4.66	9.29	0.03
	t/z	-0.27	-1.03	-0.3	-0.63	-1.04	-0.79

Note. FFRT: first fixation reading time, FRRT: first run reading time, RPRT: regression path reading time, SRPRT: selective regression path reading time, TRT: total reading time, TFC: total fixation count.

* = $p < .05$.

Table 5.4

L2 Speakers: Summary Statistics for the Phrase Region

Factor	Statistics	FRRT	RPRT	SRPRT	TRT	TFC
Intercept [global mean]	*β*	1066.61	1270.784	1225.63	1920.56	1.92
	SE	39.77	53.97	51.32	127.22	0.06
	t/z	26.82	23.55	23.88	15.1	32.34
Phrase type [collocation]	*β*	5.2	-5.51	-9.78	-70.6	-0.04
	SE	13.89	14.76	13	26.74	0.01
	t/z	0.38	-0.37	-0.75	-2.64*	-4.24*
Word order [canonical]	*β*	10.88	-4.51	-0.45	-43.92	-0.02
	SE	15.32	15.96	13.67	31.35	0.12
	t/z	0.71	-0.28	-0.033	-1.4	-1.31
LexTALE	*β*	-134.48	-185.82	-172.35	-343.85	-0.17
	SE	38.36	50.83	50.51	120.04	0.06
	t/z	-3.51*	-3.66*	-3.41*	-2.86*	-2.93*

Note. FRRT: first run reading time, RPRT: regression path reading time, SRPRT: selective regression path reading time, TRT: total reading time, TFC: total fixation count.

* = $p < .05$.

Table 5.5

L2 Speakers: Summary Statistics for the Final Word Region

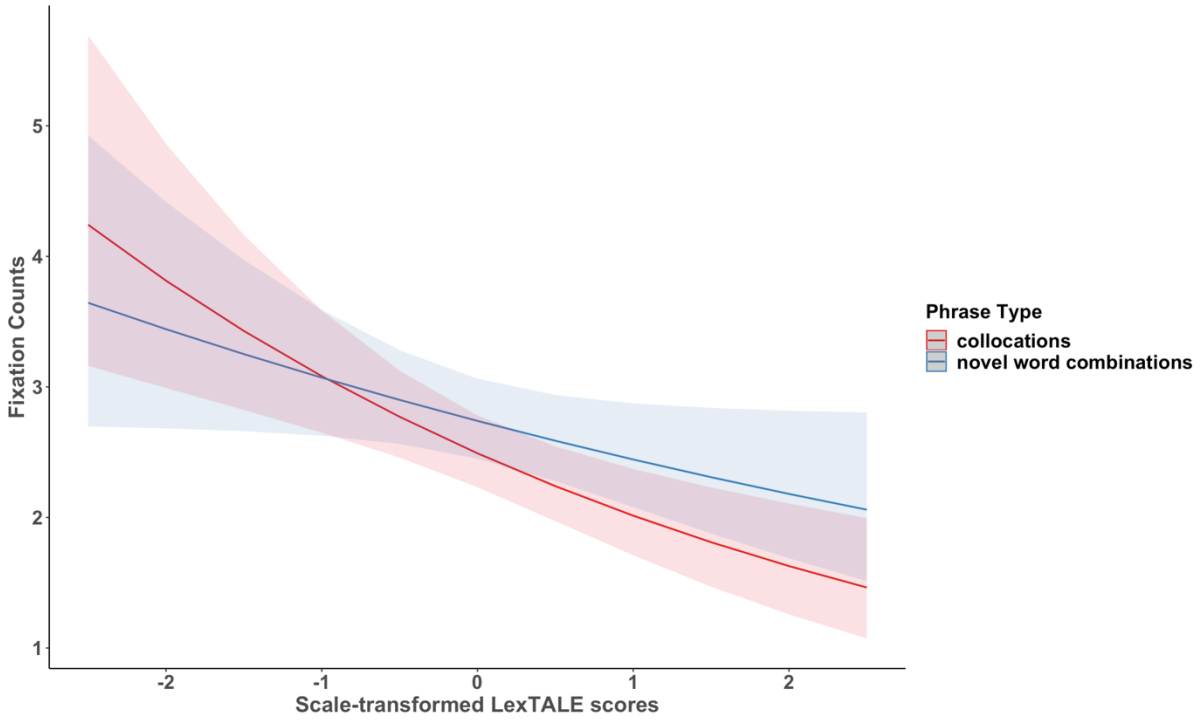
Factor	Statistics	FFRT	FRRT	RPRT	SRPRT	TRT	TFC
		Final words: nouns [canonical structure]					
Intercept [global mean]	*β*	287.57	393.11	558.09	461.38	736.57	0.96
	SE	5.97	10.77	26.32	15.07	46.63	0.05
	t/z	48.2	36.49	21.2	30.62	15.8	18.33
Phrase type [collocation]	*β*	2.73	-9.32	-10.95	-8.31	-24.58	-0.05
	SE	3.51	6.1	13.88	7.92	20.76	0.02
	t/z	0.78	-1.53	-0.79	-1.05	-1.18	-2.25*
LexTALE	*β*	-18.83	-52.52	-84.67	-61.23	-129.89	-0.16
	SE	5.89	10.3	25.1	14.13	44.97	0.05
	t/z	-3.2*	-5.01*	-3.37*	-4.33*	-2.89*	-3.07*
LexTALE* Phrase type	*β*						-0.05
	SE						0.02
	t/z						-2.23*
		Final words: verbs [reversed structure]					
Intercept [global mean]	*β*	278.4	423.66	529.28	484.11	814.5	1.03
	SE	6.23	15.16	20.71	17.82	56.39	0.06
	t/z	44.7	27.95	25.56	28.01	14.45	17.47
Phrase type [collocation]	*β*	-0.04	-4.87	-9.67	-9.31	-47.49	-0.06
	SE	4.01	8.94	13.51	9.82	23.25	0.03
	t/z	-0.01	-0.54	-0.72	-0.95	-2.04*	-2.45*
LexTALE	*β*	-17.7	-47.54	-76.61	-63.43	-174.4	-0.16
	SE	6.07	13.86	18.68	16.05	53.74	0.06
	t/z	-2.92*	-3.43*	-4.1*	-3.95*	-3.25*	-2.88*

Note. FFRT: first fixation reading time, FRRT: first run reading time, RPRT: regression path reading time, SRPRT: selective regression path reading time, TRT: total reading time, TFC: total fixation count.

* = $p < .05$.

Figure 5.2

L2 Speakers: Interaction between LexTALE Score and Phrase Type in Total Fixation Count for Canonical Constructions



5.3.3 Comparison of L1 and L2 Speakers

After looking at the L1 and L2 models separately, we built models with L1 and L2 data to investigate whether there are differences in collocation processing between the two groups of speakers other than the expected longer reading times for L2 speakers. Figure 3 shows the comparison between L1 and L2 speakers in reading collocations and novel word combinations on the phrase-level total reading time measure.

5.3.3.1 Phrase Analysis

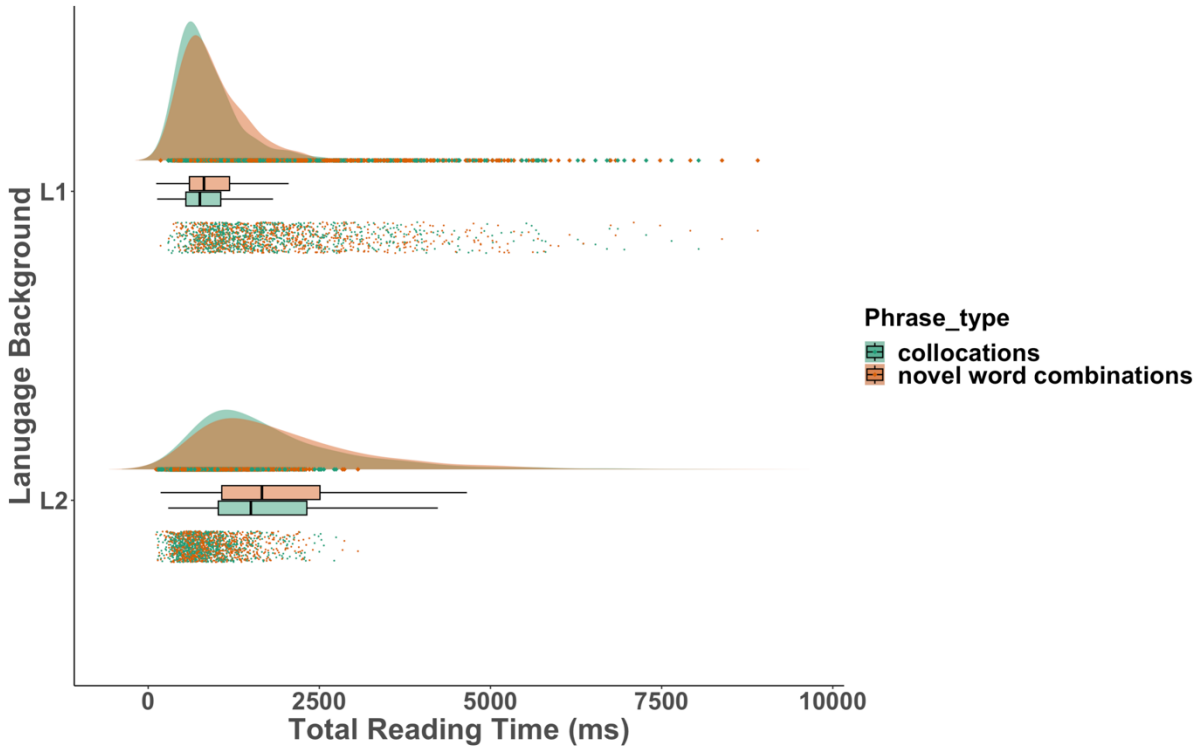
The L1 and L2 data were combined into one model. The L1 and L2 speakers differed on all measures (see Appendix 5.3). As expected, L2 speakers took longer to read than L1 speakers. The phrase type effect was observed in total reading times (β = -53.03, t = -3.28, p = .001) and total fixation counts (β = -0.04, z = -3.51, p < .001), with collocations being read faster than novel word combinations. There was also an effect of word order on first run reading times (β = 23. 59, t = 2.37, p = .019), with reversed constructions processed faster than canonical ones. The interaction between language background and phrase type was not significant on total reading times ($\chi 2$ (1) = 2.80, p = .089; BFs = 0.12) or fixation count ($\chi 2$ (1) = 0.002, p = .958). This indicates that L1 and L2 participants read collocations qualitatively similarly. There was no three-way interaction between language background, word order and phrase type in any measures.

5.3.3.2 Final Word Analysis

The language background variable was a consistent significant predictor on all measures (see Appendix 5.4), with L2 speakers reading final words more slowly, and fixating them more often than L1 speakers in both canonical and reversed structures. The phrase type effect, with collocations being read faster than novel word combinations, was observed in first run reading times for canonical structures (β = -9.95, t = -2.37, p = .02) and in the total reading time and fixation count (β = -29.03, t = -2, p = .048; β = -0.04, z = -2.02, p = .043) for reversed structures. No interaction between language background and phrase type was found for these measures. We also did not observe a three-way interaction between language background, phrase type, and word order.

BLE

Total Phrase Reading Time for Collocations and Novel Word Combinations by Language Background



5.4 Discussion

The current study tested the effects of changing word order on the processing of collocations in L1 and L2 speakers, while controlling for contextual predictability. The results showed that collocations were processed faster than novel word combinations, even when contextual predictability was matched. This result is in line with previous studies (Li et al., 2021; Öksüz et al., 2021; Sonbul, 2015; Vilkaitė & Schmitt, 2019; Wolter & Gyllstad, 2013; Wolter & Yamashita, 2015). Crucially, there was no interaction between phrase type and word order on phrase-level processing, suggesting that reversed collocations (in non-canonical word order) still show a collocation advantage. This result is in line with our hypothesis for the L1 speakers. Contrary to the hypothesis for the L2 speakers, this study showed that advanced L2 speakers do not differ qualitatively from L1 speakers regardless of the change in collocations' word order. While the collocation effect mainly showed up in later reading measures, L1 speakers also demonstrated a processing advantage for canonical collocations on the first run and the regression path reading measures.

5.4.1 Collocation Facilitation Effect

This study investigated whether the collocation effect is independent from the context effect and when the collocation effect emerges in reading, while controlling contextual predictability. The current results show that the collocation effect indeed exists in both early (first-run reading times and selective regression reading times, for L1 speakers only) and late reading measures (total reading times and fixations for both L1 and L2 speakers), though late eye-tracking measures demonstrated a more consistent effect. Such a finding is broadly in accord with MWS processing, which shows that language users are not only sensitive to word-level but also phrase-level frequency, such as lexical bundles, idioms, and collocations. Moreover, previous studies observed the facilitation effect from reading verb-noun collocations (McDonald & Shillcock, 2003a, 2003b; Vilkaitė, 2016) with base verbs. The present study found that collocations in their past tense can also elicit faster processing. This fits in with the study of Vilkaitė-Lozdienė (2022) showing that

different morphological forms of collocations are related and the overall exposure to a collocation contributes to the faster processing. Thus, the finding supports the usage-based approach of L1 acquisition (Bybee, 1998; A. Goldberg, 1995; Tomasello, 2003, 2009) and L2 acquisition (Ellis & Wulff, 2020), which proposes that language users acquire a set of linguistic constructions varying in form, size, and complexity and all these constructions are similarly affected by frequency.

When looking at the phrase level analyses, there was a consistent advantage for collocations over novel word combinations on late eye-tracking measures for both L1 and L2 speakers. However, final word analyses demonstrated a rather different pattern: only canonical collocations showed an effect on an early measure (first-run reading time and regression path reading time) amongst L1 speakers. Taken together, these findings are consistent with more recent eye-tracking studies (Li et al., 2021; Sonbul, 2015; Vilkaitė, 2016; Vilkaitė-Lozdienė, 2022) and seem to suggest that collocational status can affect an early stage of phrase processing. However, the present study generally observed a less robust and less consistent collocation effect on early measures, probably because contextual predictability was kept minimal and matched.

This study also sheds light on previous findings by Frisson et al. (2005), who found that the transitional probability effect no longer influenced verb-noun pairs processing once contextual predictability was controlled. However, it is worth noting that the verb-noun pairs in their study had a much lower average phrase-level frequency in comparison to recent collocation studies (outlined in the Introduction). This indicates that the association between verbs and nouns in their study was weaker than in other studies, which might have led to the second words becoming only weakly activated upon encountering the first words of those pairs. Exposure to highly frequently co-occurring words results in deeper entrenchment in memory. Consequently, these co-occurring words become more “lexicalised units” and “templated” in the mental lexicon. It is possible that lower phrase-level frequency verb-noun pairs are less entrenched in memory and encountering their collocational form takes longer to match to their collocational meaning. This also suggests that an early collocation effect might be restricted to high-frequency combinations only, and future research

will need to establish whether this is a monotonous effect, with an increasingly stronger collocation effect the higher the frequency of the combination becomes, or whether there is cut-off where an MWS becomes a "true" collocation. For this, it is essential that the contextual predictability of all parts of a collocation is controlled. To summarise, the lower collocational frequency examined in different studies might explain the lack of a collocation effect observed in the early measures of Frisson et al.'s study.

5.4.2 Passivized Collocations Facilitation Effect

Importantly, the present study revealed that collocations still maintain their processing advantage when word order is changed. There were consistent results in the phrase-level processing in L1 and L2 speakers: phrase type was a robust predictor and there was no interaction between phrase type and word order. It provides initial evidence that collocations are processed faster than novel word combinations regardless of variation in word order, suggesting that reversed collocations still retain their processing advantage.

However, the advantage in reversed collocations was limited to the later measures, since the final word processing pointed out a different pattern: only final words from canonical collocations elicited the processing advantage on early measures from L1 speakers. This seems to suggest that canonical collocations have an earlier effect than reversed collocations. Given that the collocation frequency of the reversed construction is much lower than that of the canonical construction (*they promoted development* vs. *development was promoted*), this might suggest that the earlier effect is frequency-related while the general advantage is more meaning-driven.

The results of the present study are in contrast with the findings of Vilkaitė-Lozdienė and Conklin (Vilkaitė-Lozdienė & Conklin, 2021), who showed that word [n] (e.g., *make*) was able to prime word [n+1] (*mistake*), while word [n+1] failed to facilitate word [n] in English collocations (i.e. no backward priming). The discrepancy can be related to the priming task not considering morphological change. As the authors themselves explained, the reversed collocations were

presented without the appropriate past participle forms (e.g., *mistake make* instead of *mistake that was made*), resulting in implausible phrases and unnatural forms.

As noted in the introduction, two previous studies (Vilkaitė, 2016; Vilkaitė-Lozdienė, 2022) demonstrated that certain collocations are flexible and can be modified by inserting words or different morphological forms without disrupting the online processing in L1 speakers. The present study extended this finding to reversed word order, which did not reduce the collocation facilitation effect at the phrase processing level when compared to corresponding controls. In a similar vein, Kyriacou et al. (2020) observed that in L1 speakers, reversed idioms were read more quickly and elicited fewer regressions than reversed controls, but reversed idioms were still read more slowly than active idioms, indicating that the faster reading found in reversed idioms can be attributed to the fact that the representation of the figurative meaning was activated. Thus, reversed idioms still retained their idiomatic meaning even in non-canonical configurations in L1 speakers. The current study extended Kyriacou et al.'s finding (2020) from idioms to collocations and from L1 speakers to L2 speakers. In contrast to Kyriacou et al.'s finding, the current study observed that reversed constructions were, in fact, read more quickly than canonical ones. Though it is often believed that parsing is more effortful with passive than active sentences, several studies have cast doubts on this generalisation (Carrithers, 1989; Traxler et al., 2014), and a recent series of self-paced reading studies found evidence for the reverse effect (Paolazzi et al., 2019). Additionally, as discussed in the introduction, contextual predictability was higher for canonical idioms than reversed idioms in Kyriacou et al.'s study, which could explain the processing advantage. Combined with the results of Kyriacou et al. (2020), the present findings of a collocation effect for reversed word order support the idea that formulaic language constructions are open to variation and modification.

5.4.3 The Question of Holistic Storage

Our finding suggests that the processing advantage in reversed collocations is most likely due to the activation of the collocational meaning rather than the actual collocational form.

Therefore, this finding provides evidence against holistic storage whereby form as well as meaning are stored as a whole (Goldinger, 1996; Jiang & Nekrasova, 2007; Wray, 2002). First, if a collocation's form (e.g., *improved quality*) is stored and retrieved as a whole chunk in the mental lexicon, the change in word order (*quality was improved*) should disrupt processing. On the contrary, the present study found the reverse, namely faster processing with reversed collocations. Secondly, it might be argued that reversed collocations can have their own lexical entries in memory, leading to an advantage in processing. However, the reversed collocations used in our study were rare (all MI scores of collocations in different morphological forms were lower than 1.5). In addition, Vilkaitė-Lozdienė (2022) also observed that the collocation advantage is maintained regardless of a change in morphological form. Combining the present study and Vilkaitė-Lozdienė 's findings, we argue that it is less likely that the collocational form is stored as one complete unit. Furthermore, our finding also suggests that the unitary semantic representation contributes to the faster processing of reversed collocations. It might be argued that the inserted words within non-adjacent collocations disrupted the collocational meaning, but the processing advantage was still found in the non-adjacent collocations (Vilkaitė, 2016). However, the majority of inserted words in the non-adjacent collocations in Vilkaitė's study provided extra information that modified the nouns of the collocations (e.g., *the degree of*, *at least some*, *a bit of*, *some of the*, *the best possible*). Hence, the canonical collocational meaning was retained, contributing to the processing advantage in non-adjacent collocations

5.4.4 Similarities between L1 and L2 Speakers

The present study showed that changing the word order of a collocation has similar effects on L1 and L2 speakers. L2 speakers showed a processing advantage for collocations in later processing measures at the phrase level. This is in line with previous findings, which showed that proficient L2 speakers are sensitive to phrase-level frequency or association strength between words. In line with the discussion above, the lack of an early effect on the final word can be related

to differences in collocational frequency between L1 and L2 speakers. It is safe to assume that L2 speakers received less exposure to the collocations of the target language, compared with L1 speakers. If indeed the early effect mainly reflects frequency, then it is not surprising that this effect will exerts itself much more strongly in L1 speakers.

Importantly, the phrase-level facilitation effect was observed for both canonical and reversed collocations. This provides initial evidence that experienced L2 speakers process variation collocations in a qualitatively similar manner to L1 speakers. This also casts doubt on the claim that L1 and L2 speakers process MWSs, such as collocations, in a fundamentally different manner, with L1 speakers assigning meaning to largely linguistic chunks and L2 speakers heavily relying on words and rules (Öksüz et al., 2021; Sonbul, 2015; Vilkaitė & Schmitt, 2019; Wray, 2002, 2008). Instead, the findings of our study suggest that L1 speakers and L2 speakers with high proficiency process large linguistic constructions in a similar manner, whereby they activate the unitary semantic meaning

The present results stand in contrast to Vilkaitė and Schmitt (2019), who found that the processing advantage in non-adjacent collocations disappeared in L2 speakers. In comparison to participants' Vocabulary Levels Test scores in their study, the average VLT score of the L2 speakers in our study was lower, making it unlikely that the divergent findings can be solely attributed to differences in proficiency. Rather, this difference in processing might reflect the difference in the constructions used. It is likely that the L2 speakers in Vilkaitė and Schmitt's study initially activated the second word of the collocation after encountering the first word, but the intervening words between the node and collocate complicated the integration of semantic information, resulting in a weakening of the activated collocate. Taken together, this shows that the facilitation effect of variation collocations depends on how collocations are modified and how closely together the words are presented.

5.4.5 The Proficiency Effect on Collocation Processing in L2 Speakers

We detected an interaction in the later measures of final word processing between vocabulary level and phrase type. More proficient L2 speakers fixated on the final word of collocations for a shorter time than on the final word of novel word combinations, while L2 speakers with low vocabulary levels spent more time reading the final word of collocations instead. On the one hand, less proficient L2 speakers might prefer to operate on the open-choice principle by selecting and combining single words (Wray, 2002, 2008). The greater dependence on the analysis and computation of individual components of the collocation slows the processing down. On the other hand, highly proficient L2 speakers have more exposure to collocations, likely resulting in frequently encountered collocations becoming lexicalised and forming their own unitary semantic representations in the mental lexicon. Hence, more proficient L2 speakers can either compute the familiar collocational meaning more quickly or activate its unitary semantic representation more straightforwardly. However, it is difficult to distinguish between these two explanations in the current study. The faster processing of collocations in proficient L2 speakers can be related to L2 vocabulary acquisition. Vocabulary knowledge consists not only of the size of the vocabulary but also its depth and breadth (i.e., how well those words are known, Anderson & Freebody, 1981). According to Nation (2001), knowing a word involves three main scopes: knowing the form, the meaning, and its use. If L2 speakers know the use of a word, it means they might know its collocational features (e.g., what words or types of words occur together), constraints on the word's usage (e.g., where and when the word can be used), and its grammatical functions (Nation, 2001). Collocations are thus seen as part of L2 word knowledge depth. In addition, vocabulary size in L2 is closely related to vocabulary depth (Schmitt, 2014). Therefore, increased vocabulary size in L2 speakers is associated with knowing which words co-occur and where these words are used, leading to faster processing of collocations. In addition, our study also observed a main effect of vocabulary knowledge in all L2 models, with L2 speakers with better vocabulary knowledge processing both collocations and novel word combinations faster than L2 speakers with smaller vocabularies. This

main effect of vocabulary level is expected since vocabulary size is closely related to general L2 proficiency, at least in English (Milton, 2013), and more proficient speakers tend to be faster readers (Berzak et al., 2018; Godfroid et al., 2020; Muñoz, 2017).

In conclusion, the present findings add to the growing body of literature that supports the collocation advantage effect in both L1 and L2 speakers. Importantly, the change in collocations' word order, i.e., from the canonical construction to the reversed, passivized construction, did not disrupt the processing of collocations, and this was observed in both L1 and L2 speakers. This provides initial evidence that both L1 and proficient L2 speakers can maintain the processing advantages of variation collocations.

CHAPTER 6 GENERAL DISCUSSION

This chapter provides a general discussion of the results of the three experimental chapters (Chapters 3–5). Each experimental chapter has provided data that enriches our understanding of MWSs processing in both L1 and L2 speakers. The findings from each experimental chapter are summarised into different themes. Following this, I will discuss how those findings contribute to the current understanding of MWSs. I will end with some general conclusions and a discussion of limitations and highlight some directions for future research.

Prior research has convincingly established that the frequency at multiple linguistic levels influences both L1 and L2 speakers (e.g., Foster, 2001; Grainger, 1990; Li et al., 2021, 2022; Öksüz et al., 2021; Rayner & Raney, 1996; Sonbul, 2015; Vilkaitė, 2016; Vilkaitė & Schmitt, 2019). There is a great deal of evidence showing that more frequent words and word combinations are recognised and processed more quickly than less frequent counterparts. Most of the studies in MWSs processing to date have only focused on distributional information (e.g., Öksüz et al., 2021; Sonbul, 2015; Vilkaitė, 2016; Wolter & Gyllstad, 2013), i.e., phrase-level frequency and word-level frequency. However, additional properties of MWSs might influence the processing of MWSs, such as the degree of semantic decomposition and syntactic modification. Thus, the overall picture of how MWSs are processed is still not clear. This thesis has taken a step forward and shifted the focus away from distributional information and onto semantic and syntactic aspects. This thesis has assessed the role of semantic and syntactic properties in the processing of L1 and L2 MWSs in a sentence reading context in two ways: (1) by using the self-paced reading paradigm in combination with a stop-making-sense task and (2) by using the eye-tracking during the reading method.

6.1 Summary of Findings

6.1.1 Effects of Phrase-level Frequency

The two self-paced reading experiments (Chapter 3 and Chapter 4) and the eye-tracking study (Chapter 5) replicated the findings that both L1 and L2 speakers processed the corpus-defined collocations more quickly than novel word combinations. These findings are in line with previous

studies on the processing of L1 and L2 collocations (e.g., Gyllstad & Wolter, 2016; Li et al., 2022; Öksüz et al., 2021; Sonbul, 2015; Vilkaitė, 2016; Vilkaitė & Schmitt, 2019; Wolter & Yamashita, 2015) and contribute to the growing body of research showing that both L1 and L2 speakers are sensitive to the frequency distributional information at larger linguistic levels.

Findings from the prior and the present studies jointly suggest that both single words and MWSs are fundamental building blocks of language and highlight the parallels in processing words and larger linguistic units (Arnon et al., 2017; Arnon & Snider, 2010), which are difficult to accommodate in the word-and-rules language theories (e.g., Pinker & Ullman, 2002). The word-and-rules theories draw a clear distinction between the mental lexicon and mental grammar. The mental lexicon contains single words of which meanings and sounds are stored in memory. In contrast, mental grammar is a productive system and is used to combine stored single words and create larger linguistic units (e.g., complex words, phrases, and sentences). Under the word-and-rules view, MWSs are the product of mental grammar by assembling single words, and the frequency effect on MWS processing is not expected to be observed since MWSs are fully computed (Arnon & Snider, 2010).

Importantly, current findings support the usage-based models to first and second language acquisition, which proposes that language is built on constructions and is learnt by abstracting stored exemplars. Linguistic constructions and exemplars are not limited to single words and can be extended to multiple elements varying in size and complexity (Langacker, 1987; Tomasello, 2015). Under this view, MWSs are taken to be one type of construction. In addition, frequency is expected to exert an influence on linguistic constructions at different sizes and levels (e.g., syllables, morphemes, single words, phrases, and sentences), which stands contrary to the words-and-rules theories. Highly frequent constructions leave a stronger trace in memory (Bybee,1988, 2010), leading to faster recognition and encoding.

These findings also provide evidence to support the statistical learning theories of language acquisition (Romberg & Saffran, 2010; Saffran, Aslin, et al., 1996; Saffran, Newport, et al., 1996).

Corpus-extracted collocations are seen as recurrent occurring word strings and can be characterised by statistical information, including the higher co-occurring frequency, mutual information scores and transitional probabilities. In this thesis, collocations have demonstrated a processing advantage, suggesting that language users can make use of sequential statistical information and discover patterns in natural language. In addition, the faster processing of collocations was also observed in L2 speakers, which suggests that bilingual speakers can track the statistical information and find the linguistic patterns in their L2 as well (Weiss et al., 2009).

6.1.2 Beyond Distributional Information

Two self-paced reading experiments revealed the idioms' superiority effect in L1 speakers. L1 speakers read idioms in a figurative context more quickly than free word combinations (in the critical region, Chapter 3) and frequency-based collocations (in the spill-over region, Chapter 4). It might be argued that the idioms' superiority effect is driven by the higher subjective familiarity and objective frequency. Previous studies have compared idioms (e.g., *at the end of a day*; *break the ice*) to matched novel literal expressions (e.g., *at the end of war*; *break the glass*) (Cacciari & Tabossi, 1988; Gibbs et al., 1989; M. R. Libben & Titone, 2008; Siyanova-Chanturia, Conklin, & Schmitt, 2011; Swinney & Cutler, 1979; Tabossi et al., 2009). These literal expressions are relatively less familiar to language users, but idioms are well-known, leading to the faster speed of recognising idioms. In Chapter 4, our results indeed show that idioms were read faster than novel literal strings in L1 speakers, which is in line with those studies listed above. Furthermore, in contrast to prior research, the two studies in this thesis matched the print frequency and the familiarity rating between idioms and free word combinations and between idioms and collocations. Therefore, frequency and familiarity are unlikely to exert influence on the faster reading of idioms, compared to free word combinations and frequency-based collocations. Our findings, thus, suggest that the idioms' superiority effect is not only driven by the distributional information.

Apart from the distributional properties, conventionality and decomposability can be attributed to the observed idioms' superiority effect in this thesis. What sets idioms apart from free word combinations and frequency-defined collocations is the higher degree of conventionality and lower degree of decomposability rather than frequency and familiarity. One of the most prominent features of idioms is that the idiomatic meaning cannot be derived from assembling the meanings of the lexical parts. Therefore, idioms are highly conventionalised and less decomposable. In addition, comprehension of idioms can be accessed via individual word levels and via the overall conceptual level simultaneously and interactively (Kuiper et al., 2007; Sprenger et al., 2006). The comparison between complex words and idioms can be made. Complex words can be accessed and encoded via multiple routes (Kuperman et al., 2009). It is believed that an extra processing cost is associated with the decomposition route for complex words. In contrast, the direct access route is preferred since this route is considered to be more efficient and is linked to faster encoding and recognition (Bertram & Hyönä, 2003). Likewise, there are multiple routes available to activate and recognise an idiom interactively. An idiom can be recognised via the activation and computation of the lexical parts of idioms simultaneously and interactively. However, the activation of the conceptual level of an idiom is preferred due to the reduced processing cost. When encountering idioms, language processors are less likely to take the decomposition-based route and less likely to undergo full online linguistic computation. Hence, the processing of idioms is sped up. On the contrary, the decomposition route is preferred for frequency-based collocations and free word combinations, which is more time-consuming. Of note, it does not deny that the direct conceptual route is not available for collocations and free word combinations. Due to the fact that the meaning is more decomposable and less conventionalised for collocations and free word combinations, the direct access route is less likely to be taken compared to idioms.

6.1.3 Resolving Idiomatic Ambiguity

In Chapter 4, idioms processing partly demonstrated a pattern of resolving ambiguity in a similar way as single unbalanced ambiguous words in L1 speakers. When L1 speakers encountered ambiguous idioms after the biasing contexts (either towards a literal or a figurative interpretation of an idiom), the idiom used in the literal interpretation was read more slowly than the one used figuratively, in line with previous idiom studies (Gibbs, 1980; Ortony et al., 1978). For idioms in Chapter 4, the figurative interpretations were more frequent and salient than the literal interpretations. In a similar vein, when a disambiguating context precedes an unbalanced ambiguous target word, reading time is longer when the prior context supports the less frequent meaning of the ambiguous word, compared to those when the prior context supports the dominant meaning of that word, or either meaning of a balanced ambiguous word (Rayner & Frazier, 1989). Both cases of ambiguities (idioms and single words) demonstrate that the competition between alternate meanings in the subordinate context results in processing cost for the ambiguous idioms and single words.

Contrary to ambiguous single words, idioms in literal use did not elicit a significant processing cost compared to the frequency-matched word combinations (i.e., collocations). For unbalanced, ambiguous words, the pattern of results is different. Reading times for the ambiguous word are also longer than a frequency-matched control word when the context preceding the ambiguous word biases towards the subordinate meaning (Colbert-Getz & Cook, 2013; Rayner et al., 1994; Sereno et al., 1992). The inconsistent findings between ambiguous idioms and single words could be due to the frequency of use of subordinate interpretations. Though ambiguous idioms, on average, in Chapter 4 had more frequent and salient figurative interpretations related to the literal interpretations, the literal interpretations were not extremely infrequent (the average frequency of literal interpretations was 31%). In the study of ambiguous single words (Wiley & Rayner, 2000), the processing cost could be reduced between ambiguous words following the subordinate meanings

and the controls when ambiguous words have more frequent subordinate meanings (i.e., meaning frequencies ranging from 8%-30%).

6.1.4 Semantic Decomposability

In Chapter 4, we detected the effect of semantic decomposability, suggesting that access to components of an idiom facilitates the processing of idiomatic interpretation due to the involvement of facilitative links between the semantic representations of the components of an idiom and the whole idiom. Our findings add more evidence against the hypothesis that the idiomatic meaning of an idiom is accessed directly (the Direct Access Model, Gibbs, 1980; the Idiom List Hypothesis, Bobrow & Bell, 1973; the Lexical Representation Hypothesis, Swinney & Culter, 1979). If an idiom is processed by only looking up the idiom as a whole unit, one does not expect to observe any influence of a difference in semantic decomposability. On the other hand, current findings fit the Hybrid Model (Cacciari & Tabossi, 1988; Cutting & Bock, 1997; Sprenger et al., 2006) and the idiom Decomposition Hypothesis (Gibbs & Nayak, 1989; Gibbs et al., 1989). If language users take two routes simultaneously, the direct access route and the decomposition route, whereby the components of an idiom contribute to processing the idiomatic meaning, one would expect that less decomposable idioms will be more difficult to be parsed as the idiomatic meaning is more in conflict with the components' assembled meaning, and the processing of more decomposable idioms will be eased as the assembled meaning is more congruent with the idiomatic meaning. Indeed, the current results show that increasing decomposability of the first and the second component of an idiom sped up reading times for idioms used figuratively. Furthermore, the findings are in line with the decomposition effect found in studies of compound words. A processing advantage is associated with semantic decomposition for the semantically transparent compounds, while a processing cost is elicited for semantically opaque compounds (Frisson et al., 2008; Ji et al., 2011; Sandra, 1990; Zwitserlood, 1994). In the case of opaque compounds, when the direct access and decomposition

routes are taken, the assembled meaning by the decomposition route conflicts with the meaning derived by the direct access route, which would cause the increased cost.

In addition, semantic decomposability interacts with meaning dominance to affect accessing the figurative meaning of an idiom. These findings show that there is a shift between taking the decomposition and direct access route in the processing of an idiom's figurative sense. The shift of taking either route is dependent on the meaning dominance of the idiom. When the literal interpretation of an idiom is more dominant than the idiomatic interpretation, the increased decomposability has a bigger facilitative effect. However, the decomposability effect decreases when idioms tend to be interpreted in their conventionally idiomatic meanings. During the processing of an idiom that has a predominant idiomatic sense, the direct access route tends to outperform the decomposition route, leading to less online computation and decomposition of components of the idioms. In short, the decomposability effect is not always strong and apparent on all types of idioms or on both routes of accessing idiomatic interpretations. These findings might provide one reason why the results of the decomposability effect are mixed in the literature. Prior research looking at the effect of semantic decomposability has either focused on the idioms that are highly conventionalised and have predominant idiomatic interpretations or have not taken meaning dominance into account. The results have thus provided a more interactive explanation of these two factors.

6.1.5 The Processing Advantage in Reversed Collocations

The eye-tracking study (Chapter 5) revealed that collocations were read faster than their non-collocational counterparts by both L1 and L2 speakers, regardless of the word order or syntactic change, when the contextual predictability was matched. This finding extended the collocation advantage to reversed collocations. However, phrase level and final word level analyses generated different patterns in the processing of reversed collocations. The faster processing in reversed collocations was limited to late eye-tracking measures. On the contrary, canonical collocations not

only demonstrated the advantage from the late eye-tracking measures on the phrase level but also demonstrated the advantage from the earlier eye-tracking measures on the word level. Taken together, these results suggest that the canonical collocations have an earlier effect than the reversed collocations. In addition, canonical collocations have a much higher frequency than the reversed ones. Hence, the earlier collocation advantage is likely associated with frequency, but the activation of the collocational meaning drives the faster processing of reversed collocations.

The findings of the existing processing advantage in reversed collocations provide evidence against the holistic storage of MWSs whereby both form and meaning are stored as a whole unit (Jiang & Nekrasova, 2007; Wray 2002, 2012). Instead, we argue that the holistic activation of the collocational meaning contributes to the faster processing of reversed collocations rather than the holistic activation of collocational forms. This is based on the finding that reversed collocations still showed an advantage even though they are not presented in canonical forms. We would not expect to observe faster processing in reversed collocations if collocations were stored and represented in a holistic form. Additionally, one might argue that reversed collocations can be stored in their own forms. However, most reversed collocations in our study were not registered in corpora and, thus, are unlikely to have an entry in the mental dictionary.

In addition, recent research using the phrasal acceptability judgement task also provides evidence against the holistic storage of MWSs (Öksüz et al., 2021; Wolter & Yamashita, 2018). It has been found that there are parallel or interactive influences of the word-level and phrase-level frequencies of collocations on collocation processing. If the holistic view of collocation processing holds true, we would not expect to detect a facilitatory effect of word-level frequency on collocation processing. However, these two acceptability judgement experiments detected that word-level frequency still contributed to the processing of collocations even when the collocations were highly frequent.

The activation of the holistic collocational meaning does not mean that language users do not access the components of collocations and does not mean that we can draw a distinction between the direct access and decomposition routes of collocation processing. Instead, the direct access to the collocation interacts with the online computation of its lexical parts, which is similar to other findings in psycholinguistics where the dual routes are available for the processing of idioms (Cutting & Bock, 1997; Sprenger et al., 2006) and compound words (Kuperman et al., 2009; Pollatsek et al., 2000). Furthermore, the interactive relationship between direct access and computation is on a continuum scale. Since the dual representation and processing of the lexical parts of a collocation and the whole collocation are redundant, the direct access route might lead to a faster outcome than the decomposition route when language users are highly familiar with the collocation. This argument is supported by the finding that the larger the phrase-level frequency effect, the smaller the word-level frequency effect (Öksüz et al., 2021).

6.1.6 Processing L2 MWSs

The L2 speakers' performance in processing MWSs depends on the specific type of MWSs. On the one hand, in two self-paced reading experiments (Chapter 4 and Chapter 5), L2 speakers consistently demonstrated that collocations were read more quickly than novel word combinations. Similar to L1 speakers, these findings are taken as evidence for supporting the sensitivity to the frequency of larger linguistic constructions (Arnon et al., 2017; Arnon & Snider, 2010), the usage-based models of language acquisition (e.g., Bybee, 1998, 2006; Ellis, 2008; Langacker, 1987; Tomasello, 2009), and statistical learning of language (Saffran, Aslin, et al., 1996; Saffran, Newport, et al., 1996).

On the other hand, we did not detect the idioms' superiority effect in L2 speakers in two self-paced reading experiments. This finding stands in opposition to the findings of L1 speakers, whereby idioms were read more quickly than corpus-defined collocations and novel word combinations, respectively. However, no superiority effect of idioms in L2 speakers is generally in

line with the findings of much previous research (Carrol et al., 2016; Carrol & Conklin, 2014b; Siyanova-Chanturia, Conklin, & Schmitt, 2011). Additionally, we observed that L2 proficiency interacted with the type of MWSs. Less proficient L2 language users read free word combinations more quickly than idioms (Chapter 3) and read novel word combinations more quickly collocations (Chapter 6). The increasing L2 proficiency reduced the processing preference of free word combinations and novel word combinations. One might argue that L2 speakers did not know the idioms and had difficulties in comprehending idioms in these studies, leading to the lack of the superiority effect. However, the two studies from this thesis ensured that L2 speakers knew those idioms by pairing the self-paced reading paradigm with a concurrent stop-making-sense task. This paradigm implicitly excluded items that L2 participants experienced difficulties in understanding the chosen idioms.

The lack of an idiom superiority effect suggests that L2 speakers prefer to take the decomposition route whereby the lexical parts of idioms are computed online, though in theory there are two routes available to comprehending idioms (Kuiper et al., 2007; Sprenger et al., 2006). The direct access is generally assumed to be associated with faster processing, while the decomposition route is more time-consuming. Furthermore, the priority of the decomposition route is more likely to result in the computation and activation of the literal sense of an idiom. As the Literal Salience model (Cieślicka, 2006) hypothesises, the literal analysis of an idiom's components tends to be compulsory for L2 speakers. Therefore, it is more costly for L2 speakers to suppress and reject the literal sense of an idiom when processing idiomatic senses.

6.2 Limitations

The findings presented in this thesis should be discussed within the context of a number of limitations, together with possibilities of further investigation. First, similar to prior research in collocation processing, this thesis only focuses on lexical collocations. However, there is another type of collocation, namely grammatical collocations that typically consist of verb/noun/adjective +

preposition (e.g., *depend on*, *independent of*). The findings of how collocations are processed can thus not be generalised to all collocations.

Second, although recent years have witnessed a growing interest in investigating the representation, processing, and storage of MWSs, there is far less theoretical guidance and processing models available compared to single words. Therefore, this thesis borrows the recognition and processing hypotheses and models from complex words (e.g., compound words, inflected words). It is interesting to see that a similar processing mechanism seems to be used at different levels: morphological, lexical and MWS levels. From the theoretical perspective, future research could work on incorporating MWSs into existing reading or eye-tracking models. One of the findings from the current thesis is that both idioms and collocations are processed faster than their corresponding controls (in Chapter 4 and 5). This indicates the frequency effect can be extended to multi-word combinations. However, current eye-tracking reading models, such as the E-Z reader Model (Reichel et al., 2003), assumes that frequency of usage is computed for single words instead of larger linguistic constructions. Our finding suggests that current eye-tracking models may need to consider frequency at phrase-level and word-level (Zang, 2019).

Moreover, reading models, such as the Lexical Quality Hypothesis (LQH) (Perfetti & Hart, 2002; Perfetti & Stafura, 2014), assume that high-quality representations of words are precise in phonological and orthographic representations, which are connected to each other and are connected to the higher level of semantic representation. Thus, a high-quality representation of a word improves the ease of access to the word due to different sources of representations being activated. Nevertheless, the LQH is focused on the representation of single words and has not yet considered MWSs. Thus, the Lexical Quality Hypothesis will need to be modified to include the representation of multiple words that are statistically associated. Moreover, the LQH has not yet been extended to L2 reading, and it might be premature to assume that L2 readers are merely less proficient than L1 readers.

The third concern is more related to the methods used and the nature of this thesis. This thesis only focused on the visual presentation of MWSs processing. However, other modalities (e.g., auditory) and production have not been accounted for. Additional properties of MWSs might exert influence on recognising and processing MWSs when presented in different modalities, as single words research have shown some factors might be modality specific (Connine et al., 1990, 1993; Ziegler et al., 2008).

The fourth limitation is related to the experimental stimuli used in the studies. Although the first and second experiments were designed to control many item-related variables, such as phrase-level frequency, familiarity and contextual predictability, there were some inevitable shortcomings in the experimental sentences. For example, the use of a dummy pronoun (*it*) or an unresolved pronoun (e.g., *he* without an antecedent) in the disambiguating region that preceded critical items could potentially have resulted in ambiguity and slowed down reading for the following MWSs. To exclude this confounding factor, two item sets, including unresolved pronouns, were excluded, and the analysis was run again. The results of the reanalysis showed the same patterns in direction and similar levels of significance. While this indicates that these potentially problematic items did not impact the results, future research should employ only non-ambiguous experimental sentences.

Furthermore, the first experiment did not control language congruency, while previous studies found that congruent collocations (i.e., collocations in L2 that have a semantically equivalent collocation in L1) were responded to more quickly than incongruent collocations (i.e., not having a semantically equivalent L1 constructions) (Wolter & Gyllstad, 2011, 2013; Wolter & Yamashita, 2015, 2018). To address this limitation, the second and third experiment only used congruent collocations. With regards to the eye-tracking experiment, the length of the phrase-level region could only be matched when including of form of the verb *to be* and by excluding the subject in the reversed passive conditions. However, it should be noted that the analyses of the individual final words were split by voice, nullifying the concern of different grammatical structures. Nonetheless, by splitting by

voice, the number of observations is halved, resulting in decreased statistical power for the final word region analysis. In future, it would be ideal to have higher numbers of experimental items and/or participants for these sub-analyses.

The sixth limitation is related to the online experimental paradigm. The choice of the paradigm was necessitated by the Covid pandemic restrictions. Online experiments offer great advantages in participant diversity i.e., reaching participants outside universities and recruiting larger samples that increase the statistical power of the findings. However, one of the major concerns in running online experiments is data quality. For example, participants are less committed when they sit at home and are not observed by the experimenter. To address the concern, a stop-making-sense task was used, which is believed to be more sensitive to sentence processing (Boland et al., 1990; Gibson et al., 1996; Singh et al., 2016). Importantly, using this paradigm, the second experiment (in Chapter 4) reproduced the expected effects found in previous literature, i.e. the collocation processing advantage in both L1 and L2 speakers as well as the idiom processing advantage in L1 speakers, compared to control literal expressions. Furthermore, one specific advantage relevant to the current thesis is that participants were allowed to stop reading the sentence when they faced difficulties understanding. However, the paradigm caused a somewhat higher proportion of trial removal than standard self-paced reading experiments. We believe we had valid reasons for excluding those critical trials where participants pressed "NO." First, when pressing the "No" key, participants (especially L2 speakers) judged the sentence semantically implausible (to stop making sense), indicating difficulties in comprehending the sentence. Second, switching from a "Yes" response to a "No" response requires extra time, making it questionable whether these responses should be combined. To reduce the amount of data loss, future research using the same paradigm could include either a larger number of semantically implausible filler sentences or filler sentences that are more easily identifiable as being semantically implausible. Another way of addressing the data loss is to test participants in the laboratory with the traditional self-paced reading paradigm or eye-tracking-while-reading paradigm.

Another limitation is that the L2 participants in this thesis were proficient in English. Participants in the two self-paced reading experiments (Chapters 3 and 4) were recruited online. The descriptive results showed that many L2 participants had lived in English-speaking countries for a relatively long time and were skilled users of English due to the nature of the participant recruitment website (i.e., the website is mainly accessible to people living in the countries that join the organisation for economic co-operation and development). Although this thesis observed that L2 proficiency modulates the processing of MWSs, these findings may not apply or may even more noticeable, when testing a representative sample of L2 speakers or L2 speakers at lower proficiency levels.

An additional limitation is related to possible disparities between native English participants participating in the norming studies/pre-tests and those participating in the main experiments. The pre-test of the first experiment recruited participants in person, while the first main experiment was moved online due to the COVID-19 pandemic. The third main experiment recruited participants in the eye-tracking laboratory, whereas its corresponding contextual predictability norming study was conducted online, to ensure that the contexts could be normed by a sufficiently large number of participants. However, participants recruited online (aged around 30) were overall older than those recruited offline from the university (aged around 20). It would be beneficial for future research to consider the disparities of sampling between offline and online testing, and whether slightly different age groups normed the stimuli differently. Furthermore, these norming studies (e.g., contextual predictabilities, and semantic decomposability) were run with L1 speakers only, similar to previous studies examining MWS processing in L2 speakers (e.g., Carrol & Conklin, 2017; Vilkaitė & Schmitt, 2019). Though L2 speakers are less exposed to English than L1 speakers, it is reasonable to assume that language exposure for L2 speakers is proportional to L1 speakers. For example, though L1 speakers encounter both *provide information* and *control information* more often than L2 speakers, the relative frequency of the phrases for L2 speakers will very likely be in proportion to L1 speakers. Therefore, similar to L1 speakers, L2 speakers are also expected to encounter *provide*

information more often, due to the generally higher frequency of *provide information* than *control information*. Although this assumption is reasonable, we agree that in future work it would be beneficial to assess whether there are disparities between the norming results from native English participants and L2 participants' norming results.

Finally, there may be added variability across the three experiments due to the online and off-line recruiting and testing of participants. During the COVID-19 pandemic, remote testing was the only way to conduct our research. While this was not ideal, remote testing does have a few advantages. First, online testing is an efficient way to recruit many participants in a short amount of time, which would have taken much longer if done in-person. However, researchers need to take extra care to ensure that the online experiments are run in a well-constrained manner (e.g., by extensive piloting). Researchers need to be prepared for potential problems with the Internet connection, software and hardware performance, and timing issues associated with different browsers. Compared to in-person experiments, online experiments usually involve no direct experimenter supervision. It is, therefore, necessary to provide very clear instructions and experimental procedures for the participants and to include periodic attention and instruction checks so that distracted or less-motivated participants can be identified and removed later. Second, collecting data online via crowdsourcing platforms can reach a much wider range of participants than the typical (Psychology) undergraduate students. Participants recruited online represent greater diversity in age range, ethnicity, years of education and proficiency in L2. Our descriptive data show that participants in the two online studies were, on average, somewhat older and from a greater age range than the in-person participants. However, the L2 speakers' English proficiency did not vary much between online (M = 56.7; M = 53.8) and in-person (M = 53.82) experiments, according to the Vocabulary Level Test scores even though the L2 participants tested online (M = 7.52; M = 7.62) spent more time living in English-speaking countries than those tested in person (M = 2.5).

6.3 Future Research and Implications

We are still far from fully understanding how MWSs are recognised, represented, and stored in the mental lexicon. However, looking at the temporal activation of the full forms of MWSs and the components of MWSs will move us forward. Future research could address this issue by exploring how soon and in what order the MWSs' full form and the properties of the MWSs' first and second components emerge in the timeline of MWSs recognition. Previous research used acceptability judgement tasks to explore the interaction between the frequency of MWS components and the frequency of the whole MWS. However, employing the eye-tracking and ERP techniques will allow us to better understand the temporal resolution of MWs' components and the whole MWSs. For example, if the activation of the MWSs' full form is seen as the initial step, we might expect to see that the frequency of the whole MWSs plays a role in early eye-tracking measures. If the activation of the components of the MWS happens before the whole form MWS, we might expect to see that the frequencies of the MWS components contribute to the recognition in early measures, while the frequency of the whole MWS plays a role in late measures. By looking at the different time course for the different frequencies would help us to better understand the interaction between the decomposition and direct access of MWSs.

The current thesis investigated MWS processing from both L1 and L2 speakers' perspectives, which allows us to assess to what extent L1 and L2 speakers overlapped in their processing. However, this study did not tap into the bilingual lexicon. Prior research (Carrol & Conklin, 2014, 2017; Wolter & Yamashita, 2015) has shown inconsistent results in how the activation of L1 influences the processing of L2 MWSs. Research in L2 collocation (Wolter & Yamashita, 2015) found that L2 speakers (i.e., Japanese-English bilinguals) did not judge English translation of Japanese collocations faster than non-collocational control, which potentially suggests that L2 speakers' L1 was not activated at least when processing L2 collocations. In contrast to L2 collocation processing, L2 translations of L1 idioms were read more quickly than controls (Carrol et al., 2016; Carrol & Conklin, 2014b, 2017). The different findings between idioms and collocations might suggest that

idioms are more unitary at the semantic and conceptual level than collocations. Investigation of the activation of L1 on the processing of L2 MWSs may enrich our general understanding of the bilingual lexicon.

Our results suggest that frequency might not explain everything in the learning and processing of MWSs and future research should also take the semantic aspects of MWSs into consideration. On the one hand, both L1 and L2 speakers processed collocations more quickly than literal controls. On the other hand, when the distributional information (phrase-level frequency, familiarity, contextual predictability) was controlled for, there was a processing advantage for idioms over collocations in L1 speakers but not in L2 speakers. This finding indicates that idiomatic language remains challenging for L2 learners. Research has shown that words can be gradually and incidentally learned through repeated encounters in context (Chen & Truscott, 2010; Webb, 2007; Webb, 2013). However, our finding suggests that repeated exposure to idioms might not be an efficient way to acquire L2 idioms, and the idiomatic meaning might not be learned sufficiently well unless special attention is given to their semantic aspect. In general, whether L2 idioms can be learned gradually from repeated contexts, remains to be investigated from a pedagogical and empirical perspective. A conceivable focus in future research should be on the nature of the contexts the MWSs appear in, e.g., clearly literal or figurative. Such studies can investigate whether and how L2 speakers incidentally learn the figurative meaning of an idiom when exposed to varying strengths of contextual support for that figurative meaning. To illustrate, one group of L2 speakers is exposed to a context strongly biased towards the figurative meaning of an idiom, while the other group reads the idiom with minimal contextual support for the figurative meaning.

Therefore, acquiring L2 idioms would benefit from explicit attention in learning and teaching. L2 learners need to be made aware that some MWSs carry idiomatic meanings and are encouraged to interpret and use the idiomatic meanings of idioms without much reference to their alternative literal meanings. Though explicit learning might be beneficial in this case, the following

problem needs to be considered in teaching. Traditionally, in the teaching of L2 vocabulary, frequency is a well-accepted guidance for determining the importance of single words (Schmitt & Schmitt, 2014). However, individual idioms are not frequent, so frequency may not be a good guidance, and explicit instruction might therefore increase the relative frequency of these idioms.

6.4 Conclusion

In conclusion, the current thesis provides an important contribution to the processing of L1 and L2 MWSs. The findings have furthered our understanding of MWS processing beyond the distributional information and extended to semantic and syntactic features of MWSs. Idioms demonstrated a similar pattern to ambiguous words in resolving their ambiguity in L1 speakers whereby there is a competition between two interpretations when ambiguous targets following the disambiguating contexts towards the less frequent interpretation. Both L1 and L2 speakers demonstrated that the semantic decomposability of an idiom facilitated accessing the idiomatic interpretation and the contribution of semantic decomposability depended on the meaning dominance of the idiom. The current findings also revealed that the change of syntactic structure of MWSs, specifically, reversing the word order of collocations, did not eliminate the collocation processing advantage in both L1 and L2 speakers. Additionally, L2 proficiency mediated processing of MWSs, with less proficient L2 participants processing MWSs compositionally.

REFERENCES

Abel, B. (2003). English idioms in the first language and second language lexicon: A dual representation approach. *Second Language Research*, *19*(4), 329–358. https://doi.org/10.1191/0267658303sr226oa

Altenberg, B. (1998). *On the phraseology of spoken English: The evidence of recurrent word-combinations*. Oxford University Press.

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. *Comprehension and Teaching: Research Reviews*, 77–117.

Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology*, *16*(1–2), 285–311. https://doi.org/10.1080/09541440340000123

Arnon, I., & Christiansen, M. H. (2017). The role of multiword building blocks in explaining L1–L2 differences. *Topics in Cognitive Science*, *9*(3), 621–636. https://doi.org/10.1111/tops.12271

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

Ayto, J. (2009). *Oxford dictionary of English idioms*. Oxford University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, *37*(1), 94–117. https://doi.org/10.1006/jmla.1997.2509

Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations? *System*, *21*(1), 101–114. https://doi.org/10.1016/0346-251X(93)90010-E

Balota, D. A., Pilotti, M., & Cortese, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition*, *29*(4), 639–647. https://doi.org/10.3758/BF03200465

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, *17*(3), 364–390. https://doi.org/10.1016/0010-0285(85)90013-1

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

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, *60*(1), 92–111. https://doi.org/10.1016/j.jml.2008.06.003

Bertram, R., & Hyönä, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from eye movements when reading short and long Finnish compounds. *Journal of Memory and Language*, *48*(3), 615–634. https://doi.org/10.1016/S0749-596X(02)00539-9

Bertram, R., Schreuder, R., & Baayen, R. H. (2000). The balance of storage and computation in morphological processing: The role of word formation type, affixal homonymy, and productivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 489–511. https://doi.org/10.1037/0278-7393.26.2.489

Berzak, Y., Katz, B., & Levy, R. (2018). *Assessing language proficiency from eye movements in reading* (arXiv:1804.07329). arXiv. https://doi.org/10.48550/arXiv.1804.07329

Biber, D., & Conrad, S. (1999). Lexical bundles in conversation and academic prose. *Language and Computers*, *26*, 181–190.

Biber, D., Conrad, S., & Cortes, V. (2003). Lexical bundles in speech and writing: An initial taxonomy. In A. Wilson, P. Rayson, & T. McEnery (Eds.), *Corpus linguistics by the lune: A festschrift for Geoffrey Leech*. Peter Lang.

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

Bishop, H. (2004). The effect of typographic salience on the look up and comprehension of unknown formulaic sequences. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing and use* (pp. 227–248). John Benjamins.

Blasko, D. G., & Connine, C. M. (1993). Effects of familiarity and aptness on metaphor processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 295–308. https://doi.org/10.1037/0278-7393.19.2.295

Blumenthal-Dramé, A. (2018). Entrenchment from a psycholinguistic and neurolinguistic perspective. In S. Hans-Jörg (Ed.), *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge* (pp. 129–152). Berlin: Mouton de Gruyter. https://doi.org/10.1037/15969-007

Bobrow, S. A., & Bell, S. M. (1973). On catching on to idiomatic expressions. *Memory & Cognition*, *1*(3), 343–346. https://doi.org/10.3758/BF03198118

Boland, J. E., Tanenhaus, M. K., & Garnsey, S. M. (1990). Evidence for the immediate use of verb control information in sentence processing. *Journal of Memory and Language*, *29*(4), 413–432. https://doi.org/10.1016/0749-596X(90)90064-7

Brenner, G. (2003). *Webster's New World: American Idioms Handbook*. Houghton Mifflin Harcourt.

Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: Comparing a range of experiment generators, both lab-based and online. *PeerJ*, *8*, e9414. https://doi.org/10.7717/peerj.9414

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

Brysbaert, M., & Vitu, F. (1998). Word skipping: Implications for theories of eye movement control in reading. In *Eye guidance in reading and scene perception* (pp. 125–147). Elsevier.

Burgoyne, K., Kelly, J. M., Whiteley, H. E., & Spooner, A. (2009). The comprehension skills of children learning English as an additional language. *British Journal of Educational Psychology*, *79*(4), 735–747. https://doi.org/10.1348/000709909X422530

Butterworth, B. (1983). Lexical representation. In *Language production*. Academic Press.

Bybee, J. (1998). The emergent lexicon. In M. C. Gruber, D. Higgins, K. S. Olson, & T. Wysocki (Eds.), *The panels: Vol. CLS 34* (Issue 2). University of Chicago: Chicago Linguistic Society.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, *82*(4), 711–733.

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., & Corradini, P. (2015). Literal analysis and idiom retrieval in ambiguous idioms processing: A reading-time study. *Journal of Cognitive Psychology*, *27*(7), 797–811. https://doi.org/10.1080/20445911.2015.1049178

Cacciari, C., & Glucksberg, S. (1991). Understanding idiomatic expressions: The contribution of word meanings. In *Advances in psychology* (Vol. 77, pp. 217–240). Elsevier.

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.org/Oxford dictionary of English idioms

Carrithers, C. (1989). Syntactic complexity does not necessarily make sentences harder to understand. *Journal of Psycholinguistic Research*, *18*(1), 75–88. https://doi.org/10.1007/BF01069048

Carrol, G., & Conklin, K. (2014a). Eye-tracking multi-word units: Some methodological questions. *Journal of Eye Movement Research*, *7*(5). https://doi.org/10.16910/jemr.7.5.5

Carrol, G., & Conklin, K. (2014b). 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

Carroll, J. B., & White, M. N. (1973). Age-of-acquisition norms for 220 picturable nouns. *Journal of Verbal Learning and Verbal Behavior*, *12*(5), 563–576. https://doi.org/10.1016/S0022-5371(73)80036-2

Chafe, W. L. (1968). Idiomaticity as an anomaly in the Chomskyan paradigm. *Foundations of Language*, *4*(2), 109–127.

Chomsky, N. (1980). Rules and representations. *Behavioral and Brain Sciences*, *3*(1), 1–15. https://doi.org/10.1017/S0140525X00001515

Christiansen, M. H., & Arnon, I. (2017). More than words: The role of multiword sequences in language learning and use. *Topics in Cognitive Science*, *9*(3), 542–551. https://doi.org/10.1111/tops.12274

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

Clifton, C., Ferreira, F., Henderson, J. M., Inhoff, A. W., Liversedge, S. P., Reichle, E. D., & Schotter, E. R. (2016). Eye movements in reading and information processing: Keith Rayner's 40year legacy. *Journal of Memory and Language*, *86*, 1–19. https://doi.org/10.1016/j.jml.2015.07.004

Clifton, C., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. P. G. van Gompel, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 341–371). Elsevier.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge.

Colbert-Getz, J., & Cook, A. E. (2013). Revisiting effects of contextual strength on the subordinate bias effect: Evidence from eye movements. *Memory & Cognition*, *41*(8), 1172–1184. https://doi.org/10.3758/s13421-013-0328-3

Conklin, K., & Carrol, G. (2021). Words go together like 'bread and butter': The rapid, automatic acquisition of lexical patterns. *Applied Linguistics*, *42*(3), 492–513. https://doi.org/10.1093/applin/amaa034

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

Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics*, *32*, 45–61. https://doi.org/10.1017/S0267190512000074

Connine, C. M., Blasko, D. G., & Titone, D. A. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language*, *32*(2), 193–210.

Connine, C. M., Mullennix, J., Shernoff, E., & Yelen, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 1084–1096. https://doi.org/10.1037/0278-7393.16.6.1084

Cowie, A. P. (1998). *Phraseology: Theory, analysis, and applications*. Oxford University Press.

Cronk, B. C., Lima, S. D., & Schweigert, W. A. (1993). Idioms in sentences: Effects of frequency, literalness, and familiarity. *Journal of Psycholinguistic Research*, *22*(1), 59–82. https://doi.org/10.1007/BF01068157

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

Diessel, H. (2017). Usage-based linguistics. In M. Aronoff (Ed.), *Oxford research encyclopedia of linguistics*. New York, NY: Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.363

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, *27*(4), 429–446. https://doi.org/10.1016/0749-596X(88)90066-6

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., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *International Review of Applied Linguistics in Language Teaching*, *47*(2), 157–177. https://doi.org/10.1515/iral.2009.007

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior*, *20*(6), 641–655. https://doi.org/10.1016/S0022-5371(81)90220-6

Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition*, *18*(1), 91–126. https://doi.org/10.1017/S0272263100014698

Ellis, N. C. (2008). Usage-based and form-focused language acquisition: The associative learning of constructions, learned attention, and the limited L2 endstate. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 372–405). London: Routledge.

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

Ellis, N. C., & Wulff, S. (2020). Usage-based approaches to L2 acquisition: An introduction. In B. VanPatten, G. Keating, & S. Wulff (Eds.), *Theories in second language acquisition* (pp. 63–82). Routledge. https://doi.org/10.4324/9780429503986

Elsherif, M. M., Preece, E., & Catling, J. C. (2023). Age-of-acquisition effects: A literature review. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *49*(5), 812–847. https://doi.org/10.1037/xlm0001215

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, *112*, 777–813. https://doi.org/10.1037/0033-295X.112.4.777

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 idiom length and context in spoken idiom comprehension. *European Journal of Cognitive Psychology*, *22*(3), 321–334. https://doi.org/10.1080/09541440902843866

Fioravanti, I., Senaldi, M. S. G., Lenci, A., & Siyanova-Chanturia, A. (2021). Lexical fixedness and compositionality in L1 speakers' and L2 learners' intuitions about word combinations: Evidence from Italian. *Second Language Research*, *37*(2), 291–322. https://doi.org/10.1177/0267658320941560

Fodor, J. A., & Lepore, E. (2002). *The compositionality papers*. Oxford University Press.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, *12*(6), 627–635. https://doi.org/10.1016/S0022-5371(73)80042-8

Foster, P. (2001). Rules and routines: A consideration of their role in the task-based language production of native and non-native speakers. In *Researching Pedagogic Tasks*. Routledge.

Fox, J., Weisberg, S., Price, B., Adler, D., Bates, D., Baud-Bovy, G., Bolker, B., Ellison, S., Firth, D., Friendly, M., Gorjanc, G., Graves, S., Heiberger, R., Krivitsky, P., Laboissiere, R., Maechler, M., Monette, G., Murdoch, D., Nilsson, H., … R-Core. (2023). *car: Companion to Applied Regression* (3.1-2) [Computer software]. https://CRAN.R-project.org/package=car

Fraser, B. (1970). Idioms within a transformational grammar. *Foundations of Language*, *6*(1), 22–42.

Frisson, S. (2015). About bound and scary books: The processing of book polysemies. *Lingua*, *157*, 17–35. https://doi.org/10.1016/j.lingua.2014.07.017

Frisson, S., Niswander-Klement, E., & Pollatsek, A. (2008). The role of semantic transparency in the processing of English compound words. *British Journal of Psychology*, *99*(1), 87–107. https://doi.org/10.1348/000712607X181304

Frisson, S., & Pickering, M. J. (1999). The processing of metonymy: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1366–1383. https://doi.org/10.1037/0278-7393.25.6.1366

Frisson, S., & Pickering, M. J. (2001). Obtaining a figurative interpretation of a word: Support for underspecification. *Metaphor and Symbol*, *16*(3–4), 149–171. https://doi.org/10.1080/10926488.2001.9678893

Frisson, S., & Pickering, M. J. (2007). The processing of familiar and novel senses of a word: Why reading Dickens is easy but reading Needham can be hard. *Language and Cognitive Processes*, *22*(4), 595–613. https://doi.org/10.1080/01690960601017013

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*, 862–877. https://doi.org/10.1037/0278-7393.31.5.862

Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language Learning*, *67*(S1), 155–179. https://doi.org/10.1111/lang.12225

Gardner, D., & Davies, M. (2007). Pointing out frequent phrasal verbs: A corpus-based analysis. *TESOL Quarterly*, *41*(2), 339–359. https://doi.org/10.1002/j.1545-7249.2007.tb00062.x

Garner, J., & Crossley, S. (2018). A latent curve model approach to studying L2 N-Gram development. *The Modern Language Journal*, *102*(3), 494–511. https://doi.org/10.1111/modl.12494

Geeraert, K., Newman, J., & Baayen, R. H. (2017). Idiom variation: Experimental data and a blueprint of a computational model. *Topics in Cognitive Science*, *9*(3), 653–669. https://doi.org/10.1111/tops.12263

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

Gibbs, R. W. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory & Cognition*, *8*(2), 149–156. https://doi.org/10.3758/BF03213418

Gibbs, R. W., & Nayak, N. P. (1989). Psycholinguistic studies on the syntactic behavior of idioms. *Cognitive Psychology*, *21*(1), 100–138. https://doi.org/10.1016/0010-0285(89)90004-2

Gibbs, R. W., Nayak, N. P., & Cutting, C. (1989). How to kick the bucket and not decompose: Analyzability and idiom processing. *Journal of Memory and Language*, *28*(5), 576–593. https://doi.org/10.1016/0749-596X(89)90014-4

Gibson, E., Pearlmutter, N., Canseco-Gonzalez, E., & Hickok, G. (1996). Recency preference in the human sentence processing mechanism. *Cognition*, *59*(1), 23–59. https://doi.org/10.1016/0010-0277(95)00687-7

Gilhooly, K. J. (1984). Word age-of-acquisition and residence time in lexical memory as factors in word naming. *Current Psychological Research & Reviews*, *3*(2), 24–31. https://doi.org/10.1007/BF02686547

Giora, R. (2002). Literal vs. Figurative language: Different or equal? *Journal of Pragmatics*, *34*(4), 487–506. https://doi.org/10.1016/S0378-2166(01)00045-5

Giora, R., & Fein, O. (1999). Irony: Context and salience. *Metaphor and Symbol*, *14*(4), 241–257. https://doi.org/10.1207/S15327868MS1404_1

Giraudo, H., & Grainger, J. (2001). Priming complex words: Evidence for supralexical representation of morphology. *Psychonomic Bulletin & Review*, *8*(1), 127–131. https://doi.org/10.3758/BF03196148

Godfroid, A., Winke, P., & Conklin, K. (2020). Exploring the depths of second language processing with eye tracking: An introduction. *Second Language Research*, *36*(3), 243–255. https://doi.org/10.1177/0267658320922

Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.

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

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1166–1183. https://doi.org/10.1037/0278-7393.22.5.1166

Gonzalez-Gomez, N., & Nazzi, T. (2012). Phonotactic acquisition in healthy preterm infants. *Developmental Science*, *15*(6), 885–894. https://doi.org/10.1111/j.1467-7687.2012.01186.x

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language*, *29*(2), 228–244. https://doi.org/10.1016/0749-596X(90)90074-A

Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and lexical phrases. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis and applications* (pp. 145–160). Clarendon Press.

Granger, S., & Meunier, F. (2008). *Phraseology: An interdisciplinary perspective* (Vol. 139). John Benjamins Publishing.

Grant, L. E., & Nation, P. (2006). How many idioms are there in English? *ITL - International Journal of Applied Linguistics*, *151*(1), 1–14. https://doi.org/10.2143/ITL.151.0.2015219

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*. *35*, 151–166.

Grice, P. (1989). *Studies in the way of words*. Harvard University Press.

Gries, S. T. (2013). 50-something years of work on collocations: What is or should be next …. *International Journal of Corpus Linguistics*, *18*(1), 137–166. https://doi.org/10.1075/ijcl.18.1.09gri

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

Hamblin, J. L., & Gibbs, R. W. (1999). Why You Can't Kick the Bucket as You Slowly Die: Verbs in Idiom Comprehension. *Journal of Psycholinguistic Research*, *28*(1), 25–39. https://doi.org/10.1023/A:1023235403250

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

Holcomb, P. J. (1988). Automatic and attentional processing: An event-related brain potential analysis of semantic priming. *Brain and Language*, *35*(1), 66–85. https://doi.org/10.1016/0093-934X(88)90101-0

Holmes, V. M., & O'regan, J. K. (1992). Reading derivationally affixed French words. *Language and Cognitive Processes*, *7*(2), 163–192. https://doi.org/10.1080/01690969208409383

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.

Holsinger, E. (2013). Representing idioms: Syntactic and contextual effects on idiom processing. *Language and Speech*, *56*(3), 373–394. https://doi.org/10.1177/0023830913484899

Holsinger, E., & Kaiser, E. (2013). Processing (non)compositional expressions: Mistakes and recovery. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 866–878. https://doi.org/10.1037/a0030410

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

Hubers, F., Cucchiarini, C., & Strik, H. (2020). Second language learner intuitions of idiom properties: What do they tell us about L2 idiom knowledge and acquisition? *Lingua*, *246*, 102940. https://doi.org/10.1016/j.lingua.2020.102940

Hunston, S. (2022). *Corpora in applied linguistics* (2nd ed.). Cambridge University Press.

Hyönä, J., & Pollatsek, A. (1998). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1612–1627. https://doi.org/10.1037/0096-1523.24.6.1612

Inhoff, A. W., Lima, S. D., & Carroll, P. J. (1984). Contextual effects on metaphor comprehension in reading. *Memory & Cognition*, *12*(6), 558–567. https://doi.org/10.3758/BF03213344

Jarema, G., Busson, C., Nikolova, R., Tsapkini, K., & Libben, G. (1999). Processing compounds: A cross-linguistic study. *Brain and Language*, *68*(1), 362–369. https://doi.org/10.1006/brln.1999.2088

Ji, H., Gagné, C. L., & Spalding, T. L. (2011). Benefits and costs of lexical decomposition and semantic integration during the processing of transparent and opaque English compounds. *Journal of Memory and Language*, *65*(4), 406–430. https://doi.org/10.1016/j.jml.2011.07.003

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

Jolsvai, H., McCauley, S. M., & Christiansen, M. H. (2020). Meaningfulness beats frequency in multiword chunk processing. *Cognitive Science*, *44*(10), e12885. https://doi.org/10.1111/cogs.12885

Juhasz, B. J., Inhoff, A. W., & Rayner, K. (2005). The role of interword spaces in the processing of English compound words. *Language and Cognitive Processes*, *20*(1–2), 291–316. https://doi.org/10.1080/01690960444000133

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Katz, G., & Giesbrecht, E. (2006). Automatic identification of non-compositional multi-word expressions using latent semantic analysis. *Proceedings of the Workshop on Multiword Expressions: Identifying and Exploiting Underlying Properties*, 12–19. https://aclanthology.org/W06-1203

Kessler, R., Weber, A., & Friedrich, C. K. (2021). Activation of literal word meanings in idioms: Evidence from eye-tracking and ERP experiments. *Language and Speech*, *64*(3), 594–624. https://doi.org/10.1177/0023830920943625

Kilgarriff, A., Rychly, P., Smrz, P., & Tugwell, D. (2004). "The Sketch Engine". In G. Williams & S. Vessier (Eds.), *Proceedings of Euralex 2004* (pp. 105–116). Université de Bretagne-Sud.

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. https://doi.org/10.1002/tesq.66

Kiss, G. R., Armstrong, C., Milroy, R., & Piper, J. (1973). An associative thesaurus of English and its computer analysis. *The Computer and Literary Studies*, *153*.

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

Krishnamurthy, R., Sinclair, J., Daley, R., & Jones, S. (2004). *English collocation studies: The OSTI report* (Continuum). Continuum.

Kuiper, K., Egmond, M.-E. van, Kempen, G., & Sprenger, S. (2007). Slipping on superlemmas: Multi-word lexical items in speech production. *The Mental Lexicon*, *2*(3), 313–357. https://doi.org/10.1075/ml.2.3.03kui

Kuperman, V., Schreuder, R., Bertram, R., & Baayen, R. H. (2009). Reading polymorphemic Dutch compounds: Toward a multiple route model of lexical processing. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 876–895. https://doi.org/10.1037/a0013484

Kuperman, V., & Van Dyke, J. A. (2013). Reassessing word frequency as a determinant of word recognition for skilled and unskilled readers. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 802–823. https://doi.org/10.1037/a0030859

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463–470. https://doi.org/10.1016/S1364-6613(00)01560-6

Kutas, M., Lindamood, T. E., & Hillyard, S. A. (1984). Word expectancy and event-related brain potentials during sentence processing. In S. Kornblum & J. Requin (Eds.), *Preparatory states and processes* (pp. 217–237). Erlbaum, Hillsdale.

Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B., & Jensen, S. P. (2020). *Package 'lmertest.'* (3.1-3) [Computer software]. https://CRAN.R-project.org/package=lmerTest

Kyriacou, M., Conklin, K., & Thompson, D. (2020). Passivizability of idioms: Has the wrong tree been barked up? *Language and Speech*, *63*(2), 404–435. https://doi.org/10.1177/0023830919847691

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

Laufer, B. (1989). A factor of difficulty in vocabulary learning: Deceptive transparency. *The International Association of Applied Linguistics*, *6*(1), 10–20.

Leech, G., Rayson, P., & Wilson, A. (2001). *Word frequencies in written and spoken English: Based on the British National Corpus*. Longman.

Leinenger, M., & Rayner, K. (2013). Eye movements while reading biased homographs: Effects of prior encounter and biasing context on reducing the subordinate bias effect. *Journal of Cognitive Psychology*, *25*(6), 665–681. https://doi.org/10.1080/20445911.2013.806513

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, *44*(2), 325–343. https://doi.org/10.3758/s13428-011-0146-0

Li, H., Paterson, K. B., Warrington, K. L., & Wang, X. (2022). Insights into the processing of collocations during L2 English reading: Evidence from eye movements. *Frontiers in Psychology*, *13*, 845590. https://doi.org/10.3389/fpsyg.2022.845590

Li, H., Warrington, K. L., Pagán, A., Paterson, K. B., & Wang, X. (2021). Independent effects of collocation strength and contextual predictability on eye movements in reading. *Language, Cognition and Neuroscience*, *36*(8), 1001–1009. https://doi.org/10.1080/23273798.2021.1922726

Libben, G., Gibson, M., Yoon, Y. B., & Sandra, D. (2003). Compound fracture: The role of semantic transparency and morphological headedness. *Brain and Language*, *84*(1), 50–64. https://doi.org/10.1016/S0093-934X(02)00520-5

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., & 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

Lieven, E. V., Pine, J. M., & Barnes, H. D. (1992). Individual differences in early vocabulary development: Redefining the referential-expressive distinction. *Journal of Child Language*, *19*(2), 287–310. https://doi.org/10.1017/S0305000900011429

Lin, D. (1999). Automatic identification of non-compositional phrases. *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics*, 317–324.

Liu, D. (2003). The most frequently used spoken American English idioms: A corpus analysis and its implications. *Tesol Quarterly*, *37*(4), 671–700. https://doi.org/doi.org/10.2307/3588217

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*, 1171. https://doi.org/10.3389/FPSYG.2015.01171/BIBTEX

Lüdecke, D., Bartel, A., Schwemmer, C., Powell, C., Djalovski, A., & Titz, J. (2023). *sjPlot: Data visualization for statistics in social science* (2.8.14) [Computer software]. https://cran.r-project.org/web/packages/sjPlot/index.html

Macis, M., & Schmitt, N. (2017). The figurative and polysemous nature of collocations and their place in ELT. *ELT Journal*, *71*(1), 50–59. https://doi.org/10.1093/elt/ccw044

Malkiel, Y. (1959). Studies in irreversible binomials. *Lingua*, *8*, 113–160.

Martin, C., Vu, H., Kellas, G., & Metcalf, K. (1999). Strength of discourse context as a determinant of the subordinate bias effect. *Quarterly Journal of Experimental Psychology*, *52*(4), 813–839. https://doi.org/10.1080/713755861

Martinez, R., & Murphy, V. A. (2011). Effect of frequency and idiomaticity on second language reading comprehension. *TESOL Quarterly*, *45*(2), 267–290. https://doi.org/10.5054/tq.2011.247708

Masini, F. (2006). Binomial constructions: Inheritance, specification and subregularities. *Lingue e Linguaggio*, *2*, 207–232. https://doi.org/10.1418/23144

McCarthy, D., Keller, B., & Carroll, J. A. (2003). Detecting a continuum of compositionality in phrasal verbs. *Proceedings of the ACL Workshop on Multiword Expressions: Analysis, Acquisition and Treatment*, 73–80.

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

McGlone, M. S., Glucksberg, S., & Cacciari, C. (1994). Semantic productivity and idiom comprehension. *Discourse Processes*, *17*(2), 167–190. https://doi.org/10.1080/01638539409544865

Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first- and second-language learners. *Psychological Bulletin*, *140*, 409–433. https://doi.org/10.1037/a0033890

Milburn, E., Vulchanova, M., & Vulchanov, V. (2021). Collocational frequency and context effects on idiom processing in advanced L2 speakers. *Canadian Journal of Experimental Psychology / Revue Canadienne de Psychologie Expérimentale*, *75*, 169–174. https://doi.org/10.1037/cep0000214

Milburn, E., & Warren, T. (2019). Idioms show effects of meaning relatedness and dominance similar to those seen for ambiguous words. *Psychonomic Bulletin & Review*, *26*(2), 591–598. https://doi.org/10.3758/s13423-019-01589-7

Miller, G. A., & Selfridge, J. A. (1950). Verbal context and the recall of meaningful material. *The American Journal of Psychology*, *63*(2), 176–185. https://doi.org/10.2307/1418920

Milton, J. (2013). Measuring the contribution of vocabulary knowledge to proficiency in the four skills. In C. Bardel, C. Lindqvist, & B. Laufer (Eds.), *L2 vocabulary acquisition, knowledge and Use: New Perspectives on Assessment and Corpus Analysis* (pp. 57–78). Eurosla Monographs Series, 2.

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

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

Morey, R. D., Rouder, J. N., Jamil, T., Urbanek, S., Forner, K., & Ly, A. (2022). *Package 'BayesFactor': Computation of Bayes Factors for Common Designs* (0.9.12-4.4) [Computer software]. https://CRAN.R-project.org/package=BayesFactor

Muñoz, C. (2017). The role of age and proficiency in subtitle reading. An eye-tracking study. *System*, *67*, 77–86. https://doi.org/10.1016/j.system.2017.04.015

Nation, I. S. (2001). *Learning vocabulary in another language* (Vol. 10). Cambridge university press Cambridge.

Nazzi, T., Bertoncini, J., & Bijeljac-Babic, R. (2009). A perceptual equivalent of the labial-coronal effect in the first year of life. *The Journal of the Acoustical Society of America*, *126*(3), 1440–1446. https://doi.org/10.1121/1.3158931

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, *24*(2), 223–242. https://doi.org/10.1093/applin/24.2.223

Nesselhauf, N. (2004). Learner corpora: Learner corpora and their potential for language teaching. In J. M. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 125–152). John Benjamins.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*(2), 127–162. https://doi.org/10.1016/S0010-0285(03)00128-2

Niswander, E., Pollatsek, A., & Rayner, K. (2000). The processing of derived and inflected suffixed words during reading. *Language and Cognitive Processes*, *15*(4–5), 389–420. https://doi.org/10.1080/01690960050119643

Nunberg, G. (1978). *The pragmatics of reference.* Indiana: University of Indiana Linguistics Club.

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

Öksüz, D., Brezina, V., & Rebuschat, P. (2021). Collocational processing in L1 and L2: The effects of word frequency, collocational frequency, and association. *Language Learning*, *71*(1), 55–98. https://doi.org/10.1111/lang.12427

Onnis, L. (2012). The potential contribution of statistical learning to second language acquisition. In P. Rebuschat & J. N. Williams (Eds.), *Statistical learning and language acquisition* (pp. 203–235). Mouton de Gruyter.

Onnis, L., Monaghan, P., Richmond, K., & Chater, N. (2005). Phonology impacts segmentation in online speech processing. *Journal of Memory and Language*, *53*(2), 225–237. https://doi.org/10.1016/j.jml.2005.02.011

Ortony, A., Schallert, D. L., Reynolds, R. E., & Antos, S. J. (1978). Interpreting metaphors and idioms: Some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behavior*, *17*(4), 465–477. https://doi.org/10.1016/S0022-5371(78)90283-9

Pagin, P., & Westerståhl, D. (2010). Compositionality I: Definitions and variants. *Philosophy Compass*, *5*(3), 250–264. https://doi.org/10.1111/j.1747-9991.2009.00228.x

Paolazzi, C. L., Grillo, N., Alexiadou, A., & Santi, A. (2019). Passives are not hard to interpret but hard to remember: Evidence from online and offline studies. *Language, Cognition and Neuroscience*, *34*(8), 991–1015. https://doi.org/10.1080/23273798.2019.1602733

Paquot, M., & Granger, S. (2012). Formulaic language in learner corpora. *Annual Review of Applied Linguistics*, *32*, 130–149. https://doi.org/10.1017/S0267190512000098

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*. Routledge.

Pecina, P. (2010). Lexical association measures and collocation extraction. *Language Resources and Evaluation*, *44*(1), 137–158. https://doi.org/10.1007/s10579-009-9101-4

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Pelletier, F. J. (1994). The principle of semantic compositionality. *Topoi*, *13*(1), 11–24. https://doi.org/10.1007/BF00763644

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Vehoeven, C. Elbron, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–213). John Benjamins.

Perfetti, C. A., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*(1), 22–37. https://doi.org/10.1080/10888438.2013.827687

Phraseology: Theory, Analysis, and Applications—Google Books. (n.d.). Retrieved 5 April 2023, from https://books.google.co.uk/books?hl=en&lr=&id=Df-iQpNMLcgC&oi=fnd&pg=PR10&dq=Howarth,+P.+(1998).+The+phraseology+of+learners%E2%80%99+academic+writing.+In+A.+Cowie+(Ed.),+Phraseology:+Theory,+Analysis+and+Appli

cations+(pp.+161%E2%80%93186).+Oxford,+UK:+Oxford+University+Press.+Huey,&ots=fRYLITf5hd&sig=szcb54qaT1nP06DUOKDR9wwOOXo&redir_esc=y#v=onepage&q&f=false
Pickering, M. J., & Frisson, S. (2001). Processing ambiguous verbs: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 556–573. https://doi.org/10.1037/0278-7393.27.2.556
Pinker, S. (1991). Rules of language. *Science*, *253*(5019), 530–535. https://doi.org/10.1126/science.1857983
Pinker, S. (1999). *Words and rules: The ingredients of language*. Basic Books.
Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, *28*(1), 73–193. https://doi.org/10.1016/0010-0277(88)90032-7
Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, *6*(11), 456–463. https://doi.org/10.1016/S1364-6613(02)01990-3
Pollatsek, A., & Hyönä, J. (2005). The role of semantic transparency in the processing of Finnish compound words. *Language and Cognitive Processes*, *20*(1–2), 261–290. https://doi.org/10.1080/01690960444000098
Pollatsek, A., Hyönä, J., & Bertram, R. (2000). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 820–833. https://doi.org/10.1037/0096-1523.26.2.820
Powell, M. J. (2009). *The BOBYQA algorithm for bound constrained optimization without derivatives* (26; Cambridge NA Report NA2009/06). University of Cambridge.
R Core Team. (2022). *R: A language and environment for statistical* (4.2.1) [Computer software]. https://www.R-project.org/
Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422. https://doi.org/10.1037/0033-2909.124.3.372
Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, *14*(3), 191–201. https://doi.org/10.3758/BF03197692
Rayner, K., & Frazier, L. (1989). Selection mechanisms in reading lexically ambiguous words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 779–790. https://doi.org/10.1037/0278-7393.15.5.779
Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *Quarterly Journal of Experimental Psychology*, *53A*(4), 1061–1080. https://doi.org/10.1080/713755934
Rayner, K., Pacht, J. M., & Duffy, S. A. (1994). Effects of prior encounter and global discourse bias on the processing of lexically ambiguous words: Evidence from eye fixations. *Journal of Memory and Language*, *33*(4), 527–544. https://doi.org/10.1006/jmla.1994.1025
Rayner, K., & Raney, G. E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin & Review*, *3*(2), 245–248. https://doi.org/10.3758/BF03212426
Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*, *4*(3–4), SI21–SI49. https://doi.org/10.1080/01690968908406362
Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 514–528. https://doi.org/10.1037/a0020990

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

Rebuschat, P., & Williams, J. N. (2012). Implicit and explicit knowledge in second language acquisition. *Applied Psycholinguistics*, *33*(4), 829–856. https://doi.org/10.1017/S0142716411000580

Reichle, E. D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the E-Z Reader model. *Vision Research*, *39*(26), 4403–4411. https://doi.org/10.1016/S0042-6989(99)00152-2

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences*, *26*(4), 445–476. https://doi.org/10.1017/S0140525X03000104

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *WIREs Cognitive Science*, *1*(6), 906–914. https://doi.org/10.1002/wcs.78

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. https://doi.org/0.1126/science.274.5294.1926

Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology*, *69*(1), 181–203. https://doi.org/10.1146/annurev-psych-122216-011805

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*(4), 606–621. https://doi.org/10.1006/jmla.1996.0032

Saffran, J. R., & Wilson, D. P. (2003). From syllables to syntax: Multilevel statistical learning by 12-month-old infants. *Infancy*, *4*(2), 273–284. https://doi.org/10.1207/S15327078IN0402_07

Sag, I. A., Baldwin, T., Bond, F., Copestake, A., & Flickinger, D. (2002). Multiword expressions: A pain in the neck for NLP. *Computational Linguistics and Intelligent Text Processing*, 1–15.

Sandra, D. (1990). On the representation and processing of compound words: Automatic access to constituent morphemes does not occur. *The Quarterly Journal of Experimental Psychology Section A*, *42*(3), 529–567. https://doi.org/10.1080/14640749008401236

Santelmann, L. M., & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition*, *69*(2), 105–134. https://doi.org/10.1016/S0010-0277(98)00060-2

Schmitt, N. (2006). Formulaic language: Fixed and varied. *Estudios de Linguistica Inglesa Aplicada*. https://idus.us.es/handle/11441/34127

Schmitt, N. (2012). Formulaic language and collocation. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1–10). New York: Blackwell. https://doi.org/10.1002/9781405198431.wbeal0433.pub2

Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, *64*(4), 913–951. https://doi.org/10.1111/lang.12077

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, *18*(1), 55–88.

Schreuder, R., Burani, C., & Baayen, R. H. (2003). Parsing and semantic opacity. In E. M. H. Assink & D. Sandra (Eds.), *Reading complex words: Cross-language studies* (pp. 159–189). Springer US. https://doi.org/10.1007/978-1-4757-3720-2_8

Schweigert, W. A. (1986). The comprehension of familiar and less familiar idioms. *Journal of Psycholinguistic Research*, *15*(1), 33–45. https://doi.org/10.1007/BF01067390

Sereno, S. C., Pacht, J. M., & Rayner, K. (1992). The effect of meaning frequency on processing lexically ambiguous words: Evidence from eye fixations. *Psychological Science*, *3*(5), 296–300. https://doi.org/10.1111/j.1467-9280.1992.tb00676.x

Shi, J., Peng, G., & Li, D. (2023). Figurativeness matters in the second language processing of collocations: Evidence from a self-paced reading experiment. *Language Learning*, *73*(1), 47–83. https://doi.org/10.1111/lang.12516

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

Singh, R., Fedorenko, E., Mahowald, K., & Gibson, E. (2016). Accommodating presuppositions is inappropriate in implausible contexts. *Cognitive Science*, *40*(3), 607–634. https://doi.org/10.1111/cogs.12260

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

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., & 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. B. (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*, 776–784. https://doi.org/10.1037/a0022531

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

Smolka, E., Rabanus, S., & Rösler, F. (2007). Processing verbs in German idioms: Evidence against the configuration hypothesis. *Metaphor and Symbol*, *22*(3), 213–231. https://doi.org/10.1080/10926480701357638

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

Sprenger, S. A., Levelt, W. J. M., & 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

Stubbs, M. (2002). Two quantitative methods of studying phraseology in English. *International Journal of Corpus Linguistics*, *7*(2), 215–244. https://doi.org/10.1075/ijcl.7.2.04stu

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

Tabossi, P., Arduino, L., & Fanari, R. (2011). Descriptive norms for 245 Italian idiomatic expressions. *Behavior Research Methods*, *43*(1), 110–123. https://doi.org/10.3758/s13428-010-0018-z

Tabossi, P., Fanari, R., & Wolf, K. (2008). Processing idiomatic expressions: Effects of semantic compositionality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 313–327. https://doi.org/10.1037/0278-7393.34.2.313

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. (1979). Recognition of affixed words and the word frequency effect. *Memory & Cognition*, *7*(4), 263–272. https://doi.org/10.3758/BF03197599

Taft, M. (1981). Prefix stripping revisited. *Journal of Verbal Learning and Verbal Behavior*, *20*(3), 289–297. https://doi.org/10.1016/S0022-5371(81)90439-4

Taft, M. (1991). *Reading and the mental lexicon*. Erlbaum.

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *The Quarterly Journal of Experimental Psychology Section A*, *57*(4), 745–765. https://doi.org/10.1080/02724980343000477

Taylor, W. L. (1953). "Cloze procedure": A new tool for measuring readability. *Journalism & Mass Communication Quarterly*, *30*(4), 415–433. https://doi.org/10.1177/107769905303000401

Titone, D. A., & 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., & Libben, M. (2014). Time-dependent effects of decomposability, familiarity and literal plausibility on idiom priming: A cross-modal priming investigation. *The Mental Lexicon*, *9*(3), 473–496. https://doi.org/10.1075/ml.9.3.05tit

Titone, D. A., Lovseth, K., Kasparian, K., & Tiv, M. (2019). Are figurative interpretations of idioms directly retrieved, compositionally built, or both? Evidence from eye movement measures of reading. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *73*(4), 216. https://doi.org/10.1037/cep0000175

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

Tomasello, M. (2009). The usage-based theory of language acquisition. In E. L. Bavin (Ed.), *The Cambridge handbook of child language* (pp. 69–87). Cambridge University Press. https://doi.org/10.1017/CBO9780511576164.005

Traxler, M. J., Corina, D. P., Morford, J. P., Hafer, S., Hoversten, L. J., & NSF Science of Learning Center for Visual Language & Visual Learning (VL2). (2014). Deaf readers' response to syntactic complexity: Evidence from self-paced reading. *Memory & Cognition*, *42*, 97–111. https://doi.org/10.3758/s13421-013-0346-1

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). Bloomsbury Publishing.

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

Underwood, G., Schmitt, N., & Galpin, A. (2004). The eyes have it: An eye-movement study into the processing of formulaic sequences. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing, and use* (pp. 53–172). John Benjamins.

Van Lancker Sidtis, D. (2012). Formulaic language and language disorders. *Annual Review of Applied Linguistics*, *32*, 62–80. https://doi.org/10.1017/S0267190512000104

Vilkaitė, L. (2016). Are nonadjacent collocations processed faster? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 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

Vilkaitė-Lozdienė, L. (2022). Do different morphological forms of collocations show comparable processing facilitation? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*, 1328–1347. https://doi.org/10.1037/xlm0001130

Vilkaitė-Lozdienė, L., & Conklin, K. (2021). Word order effect in collocation processing. *The Mental Lexicon*, *16*(2–3), 362–396. https://doi.org/10.1075/ml.20022.vil

Vu, H., Kellas, G., & Paul, S. T. (1998). Sources of sentence constraint on lexical ambiguity resolution. *Memory & Cognition*, *26*(5), 979–1001. https://doi.org/10.3758/BF03201178

Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning*, *63*(1), 91–120. https://doi.org/10.1111/j.1467-9922.2012.00729.x

Weinreich, U. (1969). Problems in the analysis of idioms. In J. Puhvel (Ed.), *Substance and structure of language* (pp. 23–81). University of California Press.

Weiss, D. J., Gerfen, C., & Mitchel, A. D. (2009). Speech segmentation in a simulated bilingual environment: A challenge for statistical learning? *Language Learning and Development*, *5*(1), 30–49. https://doi.org/10.1080/15475440802340101

Wiechmann, D. (2008). On the computation of collostruction strength: Testing measures of association as expressions of lexical bias. *Corpus Linguistics and Linguistic Theory*. https://doi.org/10.1515/CLLT.2008.011

Wiley, J., & Rayner, K. (2000). Effects of titles on the processing of text and lexically ambiguous words: Evidence from eye movements. *Memory & Cognition*, *28*(6), 1011–1021. https://doi.org/10.3758/BF03209349

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 a comparison of congruent and incongruent collocation. *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: What accounts for L2 performance? *Studies in Second Language Acquisition*, *40*(2), 395–416. https://doi.org/10.1017/S0272263117000237

Wray, A. (2002). Formulaic language in computer-supported communication: Theory meets reality. *Language Awareness*, *11*(2), 114–131. https://doi.org/10.1080/09658410208667050

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

Wulff, S. (2008). *Rethinking idiomaticity: A usage-based approach*. Bloomsbury Publishing.

Zeno, S., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. Touchstone Applied Science Associates.

Ziegler, J. C., Petrova, A., & Ferrand, L. (2008). Feedback consistency effects in visual and auditory word recognition: Where do we stand after more than a decade? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 643–661. https://doi.org/10.1037/0278-7393.34.3.643

Zwitserlood, P. (1994). The role of semantic transparency in the processing and representation of Dutch compounds. *Language and Cognitive Processes*, *9*(3), 341–368. https://doi.org/10.1080/01690969408402123

APPENDICES

Appendix 3.1

Candidate items and filler items

Group	Items	Condition	Sentence
Group1	enters one's mind	restricted collocation	It enters my mind.
	blows one's mind	idiom	It blows my mind.
	affects one's mind	free word combination	It affects my mind.
	fries one's mind	filler	It fries my mind.
	rinses one's mind	filler	It rinses my mind.
	drops one's mind	filler	It drops my mind.
Group2	break the habit	restricted collocation	She won't break the habit.
	break the bank	idiom	She won't break the bank.
	break the glass	free word combination	She won't break the glass.
	break the paper	filler	She won't break the paper.
	break the cloud	filler	She wont' break the cloud.
	break the buffer	filler	She won't break the buffer.
Group3	kill time	restricted collocation	I might kill time.
	mark time	idiom	I might mark time.
	lose time	free word combination	I might lose time.
	rank time	filler	I might rank time.
	cook time	filler	I might cook time.
	push time	filler	I might push time.
Group4	lose one's faith	restricted collocation	They might lose their faith.
	lose one's edge	idiom	They might lose their edge.
	lose one's teeth	free word combination	They might lose their teeth.
	lose one's comma	filler	They might lose their comma.
	lose one's breeze	filler	They might lose their breeze.
	lose one's story	filler	They might lose their story.
Group5	meet one's goals	restricted collocation	I will meet my goals.
	meet one's match	idiom	I will meet my match.
	meet one's sister	free word combination	I will meet my sister.
	meet one's metal	filler	I will meet my metal.
	meet one's wall	filler	I will meet my wall.
	meet one's door	filler	I will meet my door.
Group6	burned one's throat	restricted collocation	He burned his throat.
	burned one's boats	idiom	He burned his boats.
	burned one's toast	free word combination	He burned his toast.
	bruned one's protein	filler	He burned his protein.
	burned one's idea	filler	He burned his idea.
	burned one's bubble	filler	He burned his bubble.
Group7	run a bath	restricted collocation	She will run a bath.

	run a mile	idiom	She will run a mile.
	run a race	free word combination	She will run a race.
	run a faith	filler	She will run a faith.
	run a pipe	filler	She will run a pipe.
	run a room	filler	She will run a room.
Group8	tie one's shoes	restricted collocation	No one will tie their shoes.
	fill one's shoes	idiom	No one will fill their shoes.
	dry one's shoes	free word combination	No one will dry their shoes.
	roast one's shoes	filler	No one will roast their shoes.
	steam one's shoes	filler	No one will steam their shoes.
	wake one's shoes	filler	No one will wake their shoes.
Group9	hold one's tongue	restricted collocation	Please, hold your tongue.
	hold one's horses	idiom	Please, hold your horses.
	hold one's plate	free word combination	Please, hold your plate.
	hold one's land	filler	Please, hold your land.
	hold one's shame	filler	Please, hold your shame.
	hold one's quality	filler	Please, hold your quality.
Group10	hold one's breath	restricted collocation	They have to hold their breath.
	save one's breath	idiom	They decided to save their breath.
	smell one's breath	free word combination	They have to smell their breath.
	paint one's breath	filler	They decided to paint their breath.
	make one's breath	filler	They have to make their breath.
	remain one's breath	filler	They have to remain their breath.
Group11	weak spot	restricted collocation	It is a weak spot.
	bright spot	idiom	It is a bright spot.
	quiet spot	free word combination	It is a quiet spot.
	curly spot	filler	It is a curly spot.
	late spot	filler	It is a late spot.
	dishonest spot	filler	It is a dishonest spot.
Group12	earn one's respect	restricted collocation	She is starting to earn his respect.
	earn one's wings	idiom	He is starting to earn his wings.
	earn one's wages	free word combination	He is starting to earn his wages.
	earn one's mirror	filler	She is starting to earn his mirror.
	earn one's sister	filler	He is starting to earn his sister.
	earn one's hesitation	filler	He is starting to earn his hesitation.
Group13	take ages	restricted collocation	It would take ages to do it.
	take pains	idiom	He would take pains to do it.
	take cash	free word combination	It would take cash to do it.
	take curves	filler	It would take curves to do it.
	take body	filler	It would take body to do it.
	take rumour	filler	He would take rumour to do it.
Group14	lucky guess	restricted collocation	It might be a lucky guess.

	lucky break	idiom	It might be a lucky break.
	lucky lady	free word combination	She might be a lucky lady.
	lucky bar	filler	It might be a lucky bar.
	lucky sand	filler	It might be a lucky sand.
	lucky hill	filler	She might be a lucky hill.
Group15	drop a hint	restricted collocation	I might drop a hint.
	drop a brick	idiom	I might drop a brick.
	drop a pen	free word combination	I might drop a pen.
	drop a house	filler	I might drop a house.
	drop a kitchen	filler	I might drop a kitchen.
	drop a lake	filler	I might drop a lake.
Group16	high demand	restricted collocation	There is a high demand.
	high office	idiom	He holds high office.
	high table	free word combination	He sits at the high table.
	high phone	filler	He sits at the high phone.
	high penny	filler	There is a high penny.
	high joint	filler	He holds high joint.
Group17	great hero	restricted collocation	It was done by the great hero.
	great minds	idiom	It was done by the great minds.
	great flood	free word combination	It was done by the great flood.
	great cell	filler	It was done by the great cell.
	great faint	filler	It was done by the great faint.
	great bacteria	filler	It was done by the great bacteria.
Group18	bad press	restricted collocation	They are not happy, because of bad press.
	bad blood	idiom	They are not happy, because of bad blood.
	bad music	free word combination	They are not happy, because of bad music.
	bad shadow	filler	They are not happy, because of bad shadow.
	bad babble	filler	They are not happy, because of bad babble.
	bad beauty.	filler	They are not happy, because of bad beauty.
Group19	big deal	restricted collocation	She is a big deal.
	big fish	idiom	She is a big fish.
	big girl	free word combination	She is a big girl.
	big flavour	filler	She is a big flavour.
	big nerve	filler	She is a big nerve.
	big cough	filler	She is a big cough.
Group20	hard news	restricted collocation	We hate the hard news.
	hard line	idiom	We take a hard line.
	hard rock	free word combination	We hate the hard rock.
	hard taxi	filler	We hate the hard taxi.

	hard bracket	filler	We take a hard bracket.
	hard bike	filler	We hate the hard bike.
Group21	cut prices	restricted collocation	We will cut prices.
	cut corners	idiom	We will cut corners.
	cut flowers	free word combination	We will cut flowers.
	cut coffee	filler	We will cut coffee.
	cut rains	filler	We will cut rains.
	cut surprise	filler	We will cut surprise.
Group22	full moon	restricted collocation	It is a full moon.
	full circle	idiom	It comes full circle.
	full page	free word combination	It is a full page.
	full stone	filler	It is a full stone.
	full hope	filler	It comes full hope.
	full boy	filler	It is a full boy.
Group23	easy target	restricted collocation	She is an easy target.
	easy prey	idiom	She is easy prey.
	easy task	free word combination	It is an easy task.
	easy socket	filler	She is an easy socket.
	easy roar	filler	She is easy roar.
	easy hazelnut	filler	It is an easy hazelnut.
Group24	last minute	restricted collocation	He did it at the last minute.
	last word	idiom	He has the last word.
	last week	free word combination	He did it last week.
	last blame	filler	He has the last blame.
	last jaw	filler	He did it at the last jaw.
	last charger	filler	He did it last charger.
Group25	run the club	restricted collocation	You will run the club.
	join the club	idiom	You will join the club.
	leave the club	free word combination	You will leave the club.
	field the club	filler	You will field the club.
	pull the club	filler	You will pull the club.
	hit the club	filler	You will hit the club.
Group26	black eye	restricted collocation	She has a black eye.
	black sheep	idiom	She is the black sheep.
	black dress	free word combination	She has a black dress.
	black wind	filler	She has a black wind.
	black plus	filler	She is the black plus.
	black trust	filler	She has a black trust.
Group27	take advice	restricted collocation	I don’t want to take advice.
	take sides	idiom	I don't want to take sides.
	take sugar	free word combination	I don't want to take sugar.
	take store	filler	I don't want to take store.

	take bell	filler	I don't want to take bell.
	take fever	filler	I don't want to take fever.
Group28	jump the queue	restricted collocation	I am going to jump the queue.
	jump the gun	idiom	I am going to jump the gun.
	jump the fence	free word combination	I am going to jump the fence.
	jump the soul	filler	I am going to jump the soul.
	jump the palm	filler	I am going to jump the palm.
	jump the smoke.	filler	I am going to jump the smoke.
Group29	short temper	restricted collocation	He has a short temper.
	short fuse	idiom	He has a short fuse.
	short tail	free word combination	He has a short tail.
	short ink	filler	He has a short ink.
	short mud	filler	He has a short mud.
	short diet	filler	He has a short diet.
Group30	clean record	restricted collocation	She doesn't have a clean record.
	clean hands	idiom	She doesn't have clean hands.
	clean glass	free word combination	She doesn't have a clean glass.
	clean refresh	filler	She doesn't have clean refresh.
	clean lesson	filler	She doesn't have a clean lesson.
	clean happiness	filler	She doesn't have a clean happiness.
Group31	free spirit	restricted collocation	She is a free spirit.
	free lunch	idiom	There is no free lunch.
	free milk	free word combination	There is no free milk.
	free lung	filler	She is a free lung.
	free flu	filler	There is no free flu.
	free safety	filler	There is no free safety.
Group32	firm action	restricted collocation	She takes firm action.
	firm hand	idiom	She needs a firm hand.
	firm base	free word combination	She needs a firm base.
	firm slot	filler	She needs a firm slot.
	firm dust	filler	She takes firm dust.
	firm blast	filler	She needs a firm blast.
Group33	give a talk	restricted collocation	I will give a talk.
	give a hand	idiom	I will give a hand.
	give a name	free word combination	I will give a name.
	give a trail	filler	I will give a trail.
	give a sage	filler	I will give a sage.
	give a mist	filler	I will give a mist.
Group34	cold facts	restricted collocation	These are such cold facts.
	cold fish	idiom	She is such a cold fish.
	cold drink	free word combination	It is such a cold drink.
	cold copies	filler	These are such cold copies.

	cold straw	filler	She is such a cold straw.
	cold lace	filler	It is such a cold lace.
Group35	raise hopes	restricted collocation	They raise hopes.
	raise hell	idiom	They raise hell.
	raise arms	free word combination	They raise arms.
	raise spell	filler	They raise spell.
	raise month	filler	They raise month.
	raise clock	filler	They raise clock.
Group36	raise one's voice	restricted collocation	There is no need to raise your voice.
	raise one's game	idiom	There is no need to raise your game.
	raise one's arms	free word combination	There is no need to raise your arms.
	raise one's sense	filler	There is no need to raise your sense.
	raise one's heart	filler	There is no need to raise your heart.
	raise one's right	filler	There is no need to raise your right.
Group37	light rain	restricted collocation	It is a light rain.
	light relief	idiom	It is light relief.
	light green	free word combination	It is a light green.
	light rule	filler	It is a light rule.
	light nature	filler	It is light nature.
	light scare	filler	It is a light scare.
Group38	catch one's eye	restricted collocation	They could catch your eye.
	catch one's death	idiom	You could catch your death.
	catch one's fish	free word combination	You could catch your fish.
	catch one's sky	filler	They could catch your sky.
	catch one's earth	filler	You could catch your earth.
	catch one's wound	filler	You could catch your wound.
Group39	bright smile	restricted collocation	That is a bright smile.
	bright spot	idiom	That is a bright spot.
	bright colour	free word combination	That is a bright colour.
	bright angle	filler	That is a bright angle.
	bright turn	filler	That is a bright turn.
	bright crime	filler	That is a bright crime.
Group40	fresh blood	restricted collocation	There is fresh blood.
	bad blood	idiom	There is bad blood.
	warm blood	free word combination	There is warm blood.
	itchy blood	filler	There is itchy blood.
	vice blood	filler	There is vice blood.
	spicy blood	filler	There is spicy blood.
Group41	closed mind	restricted collocation	He has a closed mind.
	closed book	idiom	He is a closed book.
	closed hand	free word combination	He opens her closed hand.
	closed nail	filler	He has a closed nail.

	closed foot	filler	He is a closed foot.
	closed spin	filler	He opens her closed spin.
Group42	game bird	restricted collocation	This is the game bird.
	early bird	idiom	She is the early bird.
	dead bird	free word combination	This is the dead bird.
	sporty bird	filler	This is the sporty bird.
	basic bird	filler	She is the basic bird.
	watery bird	filler	This is the watery bird.
Group43	catch the boat	restricted collocation	We might catch the boat.
	miss the boat	idiom	We might miss the boat.
	meet the boat	free word combination	We might meet the boat.
	swim the boat	filler	We might swim the boat.
	wear the boat	filler	We might wear the boat.
	bake the boat	filler	We might bake the boat.
Group44	out of order	restricted collocation	It is out of order.
	out of place	idiom	It is out of place.
	out of school	free word combination	It is out of school.
	out of wave	filler	It is out of wave.
	out of alarm	filler	It is out of alarm.
	out of stand	filler	It is out of stand.
Group45	in one's thoughts	restricted collocation	I am in his thoughts.
	in one's shoes	idiom	I am in his shoes.
	in one's flat	free word combination	I am in his flat.
	in one's rice	filler	I am in his rice.
	in one's cream	filler	I am in his cream.
	in one's tide	filler	I am in his tide.
Group46	plant a bomb	restricted collocation	He might plant a bomb.
	plant a seed	idiom	He might plant a seed.
	plant a tree	free word combination	He might plant a tree.
	plant a river	filler	He might plant a river.
	plant a dinner	filler	He might plant a dinner.
	plant a scarf	filler	He might plant a scarf.
Group47	fast track	restricted collocation	He is on the fast track.
	fast lane	idiom	He is in the fast lane.
	fast cars	free word combination	He is in the fast cars.
	fast rope	filler	He is in the fast rope.
	fast tray	filler	He is on the fast tray.
	fast foam	filler	He is in the fast foam.
Group48	miss the train	restricted collocation	You might miss the train.
	miss the boat	idiom	You might miss the boat.
	miss the meal	free word combination	You might miss the meal.
	miss the infinity	filler	You might miss the infinity.

	miss the force	filler	You might miss the force.
	miss the nostril	filler	You might miss the nostril.
Group49	sticky subject	restricted collocation	It is a sticky subject.
	sticky fingers	idiom	He has sticky fingers.
	sticky stuff	free word combination	He has got some sticky stuff.
	sticky profile	filler	He has sticky profile.
	sticky stock	filler	It is a sticky stock.
	sticky rise	filler	He has got some sticky rise.
Group50	in deep trouble	restricted collocation	I am in deep trouble.
	in deep water	idiom	I am in deep water.
	in deep snow	free word combination	I am in deep snow.
	in deep movie	filler	I am in deep movie.
	in deep bean	filler	I am in deep bean.
	in deep fruit	filler	I am in deep fruit.
Group51	hot topic	restricted collocation	This is a hot topic.
	hot potato	idiom	This is a hot potato.
	hot soup	free word combination	This is a hot soup.
	hot triangle	filler	This is a hot triangle.
	hot stop	filler	This is a hot stop.
	hot chance	filler	This is a hot chance.
Group52	face the truth	restricted collocation	You have to face the truth.
	face the music	idiom	You have to face the music.
	face the front	free word combination	You have to face the front.
	face the bark	filler	You have to face the bark.
	face the shout	filler	You have to face the shout.
	face the yell	filler	You have to face the yell.
Group53	key area	restricted collocation	It is a key area.
	grey area	idiom	It is a grey area.
	vast area	free word combination	It is a vast area.
	tasty area	filler	It is a tasty area.
	curtain area	filler	It is a curtain area.
	polite area	filler	It is a polite area.
Group54	gain weight	restricted collocation	She is starting to gain more weight.
	carry weight	idiom	She is starting to carry more weight.
	reduce weight	free word combination	She is starting to reduce more weight.
	digest weight	filler	She is starting to digest more weight.
	pour weight	filler	She is starting to pour more weight.
	chop weight	filler	She is starting to chop more weight.
Group55	follow one's advice	restricted collocation	They are going to follow your advice.
	follow one's nose	idiom	You are going to follow your nose.
	follow one's leader	free word combination	You are going to follow your leader.
	follow one's stripe	filler	You are going to follow your stripe.

	follow one's ham	filler	You are going to follow your ham.
	follow one's pork	filler	They are going to follow your pork.
Group56	find one's voice	restricted collocation	She might find her voice.
	find one's feet	idiom	She might find her feet.
	find one's book	free word combination	She might find her book.
	find one's hike	filler	She might find her hike.
	find one's sigh	filler	She might find her sigh.
	find one's floor	filler	She might find her floor.
Group57	save energy	restricted collocation	We have to save energy.
	save face	idiom	We have to save face.
	save water	free word combination	We have to save water.
	save mouth	filler	We have to save mouth.
	save storm	filler	We have to save storm.
	save court	filler	We have to save court.
Group58	soft drink	restricted collocation	I like the soft drink.
	soft touch	idiom	I like the soft touch.
	soft skin	free word combination	I like the soft skin.
	soft share	filler	I like the soft share.
	soft enemy	filler	I like the soft enemy.
	soft arrival	filler	I like the soft arrival.
Group59	sweet dreams	restricted collocation	She has sweet dreams.
	sweet tooth	idiom	She has a sweet tooth.
	sweet taste	free word combination	It has a sweet taste.
	sweet books	filler	She has sweet books.
	sweet chair	filler	She has a sweet chair.
	sweet knee	filler	It has a sweet knee.
Group60	read one's minds	restricted collocation	I can read their minds.
	read one's lips	idiom	I can read their lips.
	read one's list	free word combination	I can read their list.
	read one's drizzle	filler	I can read their drizzle.
	read one's thunder	filler	I can read their thunder.
	read one's lift	filler	I can read their lift.
Group61	poor response	restricted collocation	This is a poor response.
	poor relation	idiom	This is a poor relation.
	poor farmers	free word combination	They are poor farmers.
	poor giggle	filler	This is a poor giggle.
	poor scream	filler	This is a poor scream.
	poor blinks	filler	They are poor blinks.
Group62	change one's views	restricted collocation	They never change their views.
	change one's spots	idiom	They never change their spots.
	change one's boots	free word combination	They never change their boots.
	change one's terror	filler	They never change their terror.

	change one's horror	filler	They never change their horror.
	change one's feat	filler	They never change their feat.
Group63	wild guess	restricted collocation	This is a wild guess.
	wild cards	idiom	They are wild cards.
	wild horse	free word combination	This is a wild horse.
	wild year	filler	This is a wild year.
	wild glue	filler	They are wild glue.
	wild pencil	filler	This a wild pencil.
Group64	lose count	restricted collocation	I might lose count.
	lose sleep	idiom	I might lose sleep.
	lose heat	free word combination	I might lose heat.
	lose obesity	filler	I might lose obesity.
	lose trip	filler	I might lose trip.
	lose harm	filler	I might lose harm.
Group65	rough draft	restricted collocation	This is a rough draft.
	rough edges	idiom	He has rough edges.
	rough track	free word combination	This is a rough track.
	rough stress	filler	This is a rough stress.
	rough herbs	filler	He has rough herbs.
	rough shine	filler	This is a rough shine.

Appendix 3.2

Chapter 3 Experimental context sentences, items and filler sentences

Group	Condition	Items	Questions	Complete Sentences
1	free word combination	affects my mind		Walking down the street,/an idea/affects my mind/and I/record it/in my diary.
	restricted collocations	enters my mind		Walking down the street,/an idea/enters my mind/and I/record it/in my diary.
	idioms	blows my mind		Walking down the street,/an idea/blows my mind/and I/record it/in my diary.
	free word combination	affects my mind		This belief about/social distancing/affects my mind/and I/am not/sure what to do now.
	restricted collocations	enters my mind		This belief about/social distancing/enters my mind/and I/am not/sure what to do now.
	idioms	blows my mind		This belief about/social distancing/blows my mind/and I/am not/sure what to do now.
	free word combination	affects my mind		This quote,/"Racism is a scale",/affects my mind/and also/helps me/to understand this protest.
	restricted collocations	enters my mind		This quote,/"Racism is a scale",/enters my mind/and also/helps me/to understand this protest.
	idioms	blows my mind	1.Does the quote influence me?	This quote,/"Racism is a scale",/blows my mind/and also/helps me/to understand this protest.
2	free word combination	break the glass	2.Will something happen if you follow the advice?	I believe/you are not going to/break the glass/if you/follow the/advice.
	restricted collocations	break the habit		I believe/you are not going to/break the habit/if you/follow the/advice.
	idioms	break the bank		I believe/you are not going to/break the bank/if you/follow the/advice.
	free word combination	break the glass		Surely you know/we are not going to/break the glass/if we/only plan/to put a small weight on it.
	restricted collocations	break the habit		Surely you know/we are not going to/break the habit/if we/only plan/to do it just once.
	idioms	break the bank		Surely you know/we are not going to/break the bank/if we/only plan/to do it just once.
	free word combination	break the glass		Keep in mind,/it is possible to/break the glass/if you/are willing/to use force.
	restricted collocations	break the habit		Keep in mind,/it is possible to/break the habit/if you/are willing/to try harder.
	idioms	break the bank		Keep in mind,/it is possible to/break the bank/if you/are willing/to spend money.
3	free word combination	lose their teeth		Many men/are starting to/lose their teeth/by the/age of/40.

	restricted collocations	lose their faith		Many men/are starting to/lose their faith/by the/age of/40.
	idioms	lose their edge		Many men/are starting to/lose their edge/by the/age of/40.
	free word combination	lose their teeth		Soldiers exposed to the chemicals/started to/lose their teeth/in a/couple of/days.
	restricted collocations	lose their faith		Soldiers exposed to the chemicals/started to/lose their faith/in a/couple of/days.
	idioms	lose their edge		Soldiers exposed to the chemicals/started to/lose their edge/in a/couple of/days.
	free word combination	lose their teeth		This group of vulnerable patients/started to/lose their teeth/after the/unsuccessful operation/by the unqualified doctor.
	restricted collocations	lose their faith		This group of vulnerable patients/started to/lose their faith/after the/unsuccessful operation/by the unqualified doctor.
	idioms	lose their edge	3.Are patients getting better after the operation?	This group of vulnerable patients/started to/lose their edge/after the/unsuccessful operation/by the unqualified doctor.
4	free word combination	meet my sister		Do you think that/I will/meet my sister/in the/big competition/next week?
	restricted collocations	meet my goal		Do you think that/I will/meet my goals/in the/big competition/next week?
	idioms	meet my match		Do you think that/I will/meet my match/in the/big competition/next week?
	free word combination	meet my sister		From time to time,/I do/meet my sister/in the/football games/every month.
	restricted collocations	meet my goals		From time to time,/I do/meet my goals/in the/football games/every month.
	idioms	meet my match		From time to time,/I do/meet my match/in the/football games/every month.
	free word combination	meet my sister		I try to stay positive,/but it seems impossible to/meet my sister/during this/uncertain time/as the result of the pandemic.
	restricted collocations	meet my goal		I try to stay positive,/but it seems impossible to/meet my goals/during this/uncertain time/as the result of the pandemic.
	idioms	meet my match	4.Does it seem very promising at the present time?	I try to stay positive,/but it seems impossible to/meet my match/during this/uncertain time/as a result of the pandemic.
5	free word combination	run a race		I have heard/she would/run a race/to relieve/the stress/from the upcoming exam.

	restricted collocations	run a bath		I have heard/she would/run a bath/to relieve/the stress/from the upcoming exam.
	idioms	run a mile		I have heard/she would/run a mile/to relieve/the stress/from the upcoming exam.
	free word combination	run a race		Frankly speaking,/I would/run a race/if I/was the/one who received such news.
	restricted collocations	run a bath		Frankly speaking,/I would/run a bath/if I/was the/one who received such news.
	idioms	run a mile		Frankly speaking,/I would/run a mile/if I/was the/one who received such news.
	free word combination	run a race		Please keep in mind,/do not try to/run a race/especially when/you are/not in the right state of mind.
	restricted collocations	run a bath		Please keep in mind,/do not try to/run a bath/especially when/you are/not in the right state of mind.
	idioms	run a mile	5.Should you try doing this when you are not in the right state of mind?	Please keep in mind,/do not try to/run a mile/especially when/you are/not in the right state of mind.
6	free word combination	dry their shoes		Sad to say,/no-one was able to/dry their shoes/when the/power was/cut off.
	restricted collocations	tie their shoes		Sad to say,/no-one was able to/tie their shoes/when the/power was/cut off .
	idioms	fill their shoes		Sad to say,/no-one was able to/fill their shoes/when the/power was/cut off .
	free word combination	dry their shoes		I wonder/who is going to/dry their shoes/these days/with so/much going on.
	restricted collocations	tie their shoes		I wonder/who is going to/tie their shoes/these days/with so/much going on.
	idioms	fill their shoes		I wonder/who is going to/fill their shoes/these days/with so/much going on.
	free word combination	dry their shoes		Last week, they could not/find anyone to/dry their shoes/and now/they are/having difficulties.
	restricted collocations	tie their shoes		Last week, they could not/find anyone to/tie their shoes/and now/they are/having difficulties.
	idioms	fill their shoes	6.Did they find someone to perform the task?	Last week, they could not/find anyone to/fill their shoes/and now/they are/having difficulties.
7	free word combination	hold your plate		Could we/kindly ask you to/hold your plate/as many/people are/in line ahead of you?
	restricted collocations	hold your tongue		Could/we kindly ask you to/hold your tongue/as many/people are/in line ahead of you?
	idioms	hold your horses		Could/we kindly ask you to/hold your horses/as many/people are/in line ahead of you?

	free word combination	hold your plate		I have to remind you that/you should/hold your plate/because everyone/is doing/it here.
	restricted collocations	hold your tongue	7.Is everyone else doing it?	I have to remind you that/you should/hold your tongue/because everyone/is doing/it here.
	idioms	hold your horses		I have to remind you that/you should/hold your horses/because everyone/is doing/it here.
	free word combination	hold your plate		Just as we agreed from the start,/I want you to/hold your plate/right after/you arrive/here.
	restricted collocations	hold your tongue		Just as we agreed from the start,/I want you to/hold your tongue/right after/you arrive/here.
	idioms	hold your horses		Just as we agreed from the start,/I want you to/hold your horses/right after/you arrive/here.
8	free word combination	smell their breath		Please,/ask them to/smell their breath/and they/will find/there is a horrible smell.
	restricted collocations	hold their breath		Please,/ask them to/hold their breath/and they/will find/help will be coming soon.
	idioms	save their breath		Please,/ask them to/save their breath/and they/will find/there is no way of getting in.
	free word combination	smell their breath		They have made a plan/and are going to/smell their breath/when the/group of/people are approaching them.
	restricted collocations	hold their breath	8.Did they make a plan?	They have made a plan/and are going to/hold their breath/when the/group of/people are approaching them.
	idioms	save their breath		They have made a plan/and are going to/save their breath/when the/group of/people don't listen to them.
	free word combination	smell their breath		In this situation,/they would better/smell their breath/mainly because/there is/a policeman checking for drunkenness.
	restricted collocations	hold their breath		In this situation,/they would better/hold their breath/mainly because/there is/a promise that things will improve soon.
	idioms	save their breath		In this situation,/they would better/save their breath/mainly because/there is/no one agreeing with them.
9	free word combination	quiet spot	9.Are we certain that this is mentioned?	We are certain that/this is the/quiet spot/that had/been mentioned/in the local paper.
	restricted collocations	weak spot		We are certain that/this is the/weak spot/that had/been mentioned/in the local paper.
	idioms	bright spot		We are certain that/this is the/bright spot/that had/been mentioned/in the local paper.
	free word combination	quiet spot		Do you know/whether there is a/quiet spot/that we/all should/avoid?

	restricted collocations	weak spot		Do you know/whether there is a/weak spot/that we/all should/notice?
	idioms	bright spot		Do you know/whether there is a/bright spot/that we/all should/avoid?
	free word combination	quiet spot		Do not be surprised by/the fact that this/quiet spot/will be/used by/everyone who knows about it.
	restricted collocations	weak spot		Do not be surprised by/the fact that this/weak spot/will be/used by/everyone who knows about it.
	idioms	bright spot		Do not be surprised by/the fact that this/bright spot/will be/known by/everyone who watches the show.
10	free word combination	lucky lady		Nothing can help,/all we need here is a/lucky lady/that can/deal with/the situation.
	restricted collocations	lucky guess		Nothing can help,/all we need here is a/lucky guess/that can/deal with/the situation.
	idioms	lucky break		Nothing can help,/all we need here is a/lucky break/that can/deal with/the situation.
	free word combination	lucky lady		To be honest with you,/we would better have a/lucky lady/that can/cope with/the problem.
	restricted collocations	lucky guess		To be honest with you,/we would better have a/lucky guess/that can/cope with/the problem.
	idioms	lucky break		To be honest with you,/we would better have a/lucky break/that can/cope with/the problem.
	free word combination	lucky lady		I wish/there was a/lucky lady/that could/help me/out of this trap.
	restricted collocations	lucky guess		I wish/there was a/lucky guess/that could/help me/out of this trap.
	idioms	lucky break	10.Do I need help to get out of trouble?	I wish/there was a/lucky break/that could/help me/out of this trap.
11	free word combination	drop a pen		You would better be careful/and do not/drop a pen/during the/important meeting/tomorrow.
	restricted collocations	drop a hint		You would better be careful/and do not/drop a hint/during the/important meeting/tomorrow.
	idioms	drop a brick		You would better be careful/and do not/drop a brick/during the/important meeting/tomorrow.
	free word combination	drop a pen		Dave suggests to me/to keep calm and to not/drop a pen/when they/are discussing/the issue.
	restricted collocations	drop a hint	11.Is it Dave who suggests this?	Dave suggests to me/to keep calm and to not/drop a hint/when they/are discussing/the issue.
	idioms	drop a brick		Dave suggests to me/to keep calm and to not/drop a brick/when they/are discussing/the issue.

	free word combination	drop a pen		Tell your brother/to be careful and not to/drop a pen/in the/middle of/his romantic dinner.
	restricted collocations	drop a hint		Tell your brother/to be careful and not to/drop a hint/in the/middle of/his romantic dinner.
	idioms	drop a brick		Tell your brother/to be careful and not to/drop a brick/in the/middle of/his romantic dinner.
12	free word combination	great flood		We should get fully prepared,/since/the great flood/will come/at any/time of the day.
	restricted collocations	great hero		We should get fully prepared,/since/the great hero/will come/at any/time of the day.
	idioms	great minds		We should get fully prepared,/since/the great minds/will come/at any/time of the day.
	free word combination	great flood		From a news broadcast,/I have heard that/the great flood/will be/coming to/our village.
	restricted collocations	great hero		From a news broadcast,/I have heard that/the great hero/will be/coming to/our village.
	idioms	great minds		From a news broadcast,/I have heard that/the great minds/will be/coming to/our village.
	free word combination	great flood		Everything should be arranged/before/the great flood/is going/to arrive/in our town/this afternoon.
	restricted collocations	great hero		Everything should be arranged/before/the great hero/is going/to arrive/in our town/this afternoon.
	idioms	great minds	12.Should we get prepared in advance?	Everything should be arranged/before/the great minds/is going/to arrive/in our town/this afternoon.
13	free word combination	bad music	13.Did the band members get along well?	There certainly was/a lot of/bad music/simply because/band members/hated each other.
	restricted collocations	bad press		There certainly was/a lot of/bad press/simply because/band members/hated each other.
	idioms	bad blood		There certainly was/a lot of/bad blood/simply because/band members/hated each other.
	free word combination	bad music		Have you noticed/there were claims of/bad music/that had/appeared in/the tabloids?
	restricted collocations	bad press		Have you noticed/there were claims of/bad press/that had/appeared in/the tabloids?

	idioms	bad blood		Have you noticed/there were claims of/bad blood/that had/appeared in/the tabloids?
	free word combination	bad music		The record shows/there is a history of/bad music/within the/production company/in North London.
	restricted collocations	bad press		The record shows/there is a history of/bad press/within the/production company/in North London.
	idioms	bad blood		The record shows/there is a history of/bad blood/within the/production company/in North London.
14	free word combination	cut flowers		It may be possible that/they are going to/cut flowers/if they/cannot find/a better solution.
	restricted collocations	cut prices		It may be possible that/they are going to/cut prices/if they/cannot find/a better solution.
	idioms	cut corners		It may be possible that/they are going to/cut corners/if they/cannot find/a better solution.
	free word combination	cut flowers		If I were you,/I would not/cut flowers/under the/current circumstances/and wait for several days.
	restricted collocations	cut prices		If I were you,/I would not/cut prices/under the/current circumstances/and wait for several days.
	idioms	cut corners		If I were you,/I would not/cut corners/under the/current circumstances/and wait for several days.
	free word combination	cut flowers		In the end,/they did not/cut flowers/even though/the temptation/had been high.
	restricted collocations	cut prices		In the end,/they did not/cut prices/even though/the temptation/had been high.
	idioms	cut corners	14.Did they give in to their temptation?	In the end,/they did not/cut corners/even though/the temptation/had been high.
15	free word combination	full page		Pete was happy that/it had reached a/full page,/although it/took longer/than expected.
	restricted collocations	full moon		Pete was happy that/it had reached a/full moon,/although it/took longer/than expected.
	idioms	full circle		Pete was happy that/it had reached/full circle,/although it/took longer/than expected.
	free word combination	full page		I expect/it to become a/full page,/but it/will take/a few more days.
	restricted collocations	full moon		I expect/it to become a/full moon,/but it/will take/a few more days.
	idioms	full circle		I expect/it to come/full circle,/but it/will take/a few more days.
	free word combination	full page		Jim assured me/that it will be a/full page/but I/have no/confidence in him.
	restricted collocations	full moon		Jim assured me/that it will be a/full moon/but I/have no/confidence in him.

	idioms	full circle	15.Do I trust Jim?	Jim assured me/that it will come/full circle,/but I/have no/confidence in him.
16	free word combination	easy task		I am convinced that/Joe has an/easy task/because he/can easily/finish it in 10 minutes.
	restricted collocations	easy target		I am convinced that/Joe is an/easy target/because he/can easily/be deceived.
	idioms	easy prey		I am convinced that/Joe is/easy prey/because he/can easily/be deceived.
	free word combination	easy task		There is no point assuming that/it will be an/easy task/and everyone/knows that/it will require a lot of components.
	restricted collocations	easy target		There is no point assuming that/she will be an/easy target/and everyone/knows that/she will fight against it.
	idioms	easy prey		There is no point assuming that/she will be/easy prey/and everyone/knows that/she will fight her corner.
	free word combination	easy task		Lucy told us that/Tom has an/easy task/because he/almost always/gets preferential treatment.
	restricted collocations	easy target		Lucy told us that/Tom is an/easy target/because he/almost always/believes everything.
	idioms	easy prey	16.Did Lucy tell us Tom is very clever?	Lucy told us that/Tom is/easy prey/because he/almost always/believes everything.
17	free word combination	last week		It is said that/Gregory wanted to have it/last week,/but I/heard that/he failed.
	restricted collocations	last minute		It is said that/Gregory wanted to have it/last minute,/but I/heard that/he failed.
	idioms	last word		It is said that/Gregory wanted to have the/last word,/but I/heard that/he failed.
	free word combination	last week		Anna managed to/have it/last week,/but Boris/did not,/which explains why he was so sad.
	restricted collocations	last minute		Anna managed to/have it/last minute,/but Boris/did not,/which explains why he was so sad.
	idioms	last word		Anna managed to/have the/last word,/but Boris/did not,/which explains why he was so sad.
	free word combination	last week		It turns out/she was not able to get it in/last week,/so she/needed to/have someone else handle it.
	restricted collocations	last minute		It turns out/she was not able to get it/last minute,/so she/needed to/have someone else handle it.
	idioms	last word	17.Did she need help?	It turns out/she was not able to get in the/last word,/so she/needed to/have someone else handle it.
18	free word combination	black dress		You might not know/who has the/black dress,/it is/better not/to ask about it.
	restricted collocations	black eye		You might not know/who has the/black eye,/it is/better not/to ask about it.

	idioms	black sheep		You might not know/who is the/black sheep,/it is/better not/to ask about it.
	free word combination	black dress		We never knew that/Joan had the/black dress,/which explains/why she/seems a bit tense.
	restricted collocations	black eye		We never knew that/Joan had the/black eye,/which explains/why she/seems a bit tense.
	idioms	black sheep		We never knew that/Joan was the/black sheep,/which explains/why she/seems a bit tense.
	free word combination	black dress		I was unaware/who had the/black dress,/but then/my aunt/told me.
	restricted collocations	black eye		I was unaware/who had the/black eye,/but then/my aunt/told me.
	idioms	black sheep	18.Did I know it before my aunt told me?	I was unaware/who was the/black sheep,/but then/my aunt/told me.
19	free word combination	take sugar		In the end,/I decided not to/take sugar,/but everything/turned out/fine.
	restricted collocations	take advice		In the end,/I decided not to/take advice,/but everything/turned out/fine.
	idioms	take sides		In the end,/I decided not to/take sides,/but everything/turned out/fine.
	free word combination	take sugar		Donald had been told repeatedly/not to/take sugar,/but he/ignored it/completely.
	restricted collocations	take advice		Donald had been told repeatedly/not to/take advice,/but he/ignored it/completely.
	idioms	take sides		Donald had been told repeatedly/not to/take sides,/but he/ignored it/completely.
	free word combination	take sugar		The woman told us that/she was going to/take sugar,/but it/is unclear/whether she did in the end.
	restricted collocations	take advice		The woman told us that/she was going to/take advice,/but it/is unclear/whether she did in the end.
	idioms	take sides	19.Do we know what she did in the end?	The woman told us that/she was going to/take sides,/but it/is unclear/whether she did in the end.
20	free word combination	jump the fence		I can assure you that/I am not going to/jump the fence,/no matter/what Linda/told you!
	restricted collocations	jump the queue		I can assure you that/I am not going to/jump the queue,/no matter/what Linda/told you!
	idioms	jump the gun		I can assure you that/I am not going to/jump the gun,/no matter/what Linda/told you!
	free word combination	jump the fence		Let's not be hasty,/I guess you do not want to/jump the fence,/which would/make things/worse.

	restricted collocations	jump the queue	20.Should you be patient?	Let's not be hasty,/I guess you do not want to/jump the queue,/which would/make things/worse.
	idioms	jump the gun		Let's not be hasty,/I guess you do not want to/jump the gun,/which would/make things/worse.
	free word combination	jump the fence		Seeing how rushed he is,/I am sure he is going to/jump the fence,/even though/it might/end up badly.
	restricted collocations	jump the queue		Seeing how rushed he is,/I am sure he is going to/jump the queue,/even though/it might/end up badly.
	idioms	jump the gun		Seeing how rushed he is,/I am sure he is going to/jump the gun,/even though/it might/end up badly.
21	free word combination	short tail		I was surprised about/the monkey's/short tail/and the/amount of/trouble it could cause.
	restricted collocations	short temper		I was surprised about/the monkey's/short temper/and the/amount of/trouble it could cause.
	idioms	short fuse		I was surprised about/the monkey's/short fuse/and the/amount of/trouble it could cause.
	free word combination	short tail		The little creature/is known for its unusually/short tail/and it/should be/handled with extra care.
	restricted collocations	short temper	21.Should the animal be handled with extra care?	The little creature/is known for its unusually/short temper/and it/should be/handled with extra care.
	idioms	short fuse		The little creature/is known for its unusually/short fuse/and it/should be/handled with extra care.
	free word combination	short tail		You've been warned,/do not laugh at its/short tail/when you/see the/neighbour's dog!
	restricted collocations	short temper		You've been warned,/do not laugh at his/short temper/when you/see the/neighbour's son!
	idioms	short fuse		You have been warned,/do not laugh at his/short fuse/when you/see the/neighbour's son!
22	free word combination	clean glass	22.Should we blame her for the mess?	Since/she has a/clean glass,/she cannot/be blamed/for the mess.
	restricted collocations	clean record		Since/she has a/clean record,/she cannot/be blamed/for the mess.
	idioms	clean hands		Since/she has/clean hands,/she cannot/be blamed/for the mess.
	free word combination	clean glass		It is said that/the guy has a/clean glass,/but I/think that/they are mistaken.
	restricted collocations	clean record		It is said that/the guy has a/clean record,/but I/think that/they are mistaken.

	idioms	clean hands		It is said that/the guy has/clean hands,/but I/think that/they are mistaken.
	free word combination	clean glass		I am surprised that/the respectable-looking woman does not have a/clean glass,/which is/why she/has asked for a different one.
	restricted collocations	clean record		I am surprised that/the respectable-looking woman does not have a/clean record,/which is/why she/has a hard time finding a job.
	idioms	clean hands		I am surprised that/the respectable-looking woman does not have/clean hands,/which is/why she/has a hard time finding a job.
23	free word combination	free milk		Please trust me,/there is no such thing as/free milk,/no matter/what you/have been told.
	restricted collocations	free spirit		Please trust me,/there is no such thing as a/free spirit,/no matter/what you/have been told.
	idioms	free lunch		Please trust me,/there is no such thing as a/free lunch,/no matter/what you/have been told.
	free word combination	free milk		Although they described it/as some kind of/free milk,/it quickly/became clear/that they were wrong.
	restricted collocations	free spirit	23.Is their description wrong?	Although they described him/as some kind of/free spirit,/it quickly/became clear/that they were wrong.
	idioms	free lunch		Although they described it/as some kind of/free lunch,/it quickly/became clear/that they were wrong.
	free word combination	free milk		It is not true that/there is no such thing as/free milk,/I know/several people/who claim they had one.
	restricted collocations	free spirit		It is not true that/there is no such thing as a/free spirit,/I know/several people/who claim they had one.
	idioms	free lunch		It is not true that/there is no such thing as a/free lunch,/I know/several people/who claim they had one.
24	free word combination	firm base	24.Are these adolescents finishing their holidays?	It is obvious that/these adolescents need a/firm base/to finish/their schooling/during these uncertain times.
	restricted collocations	firm action		It is obvious that/these adolescents need/firm action/to finish/their schooling/during these uncertain times.
	idioms	firm hand		It is obvious that/these adolescents need a/firm hand/to finish/their schooling/during these uncertain times.
	free word combination	firm base		At this stage,/Lizzy would need a/firm base/to be/well-prepared for/what will come down the line.

	restricted collocations	firm action		At this stage,/Lizzy would need/firm action/to be/well-prepared for/what will come down the line.
	idioms	firm hand		At this stage,/Lizzy would need a/firm hand/to be/well-prepared for/what will come down the line.
	free word combination	firm base		The politician said that/children nowadays need a/firm base/in order/to become/model citizens.
	restricted collocations	firm action		The politician said that/children nowadays need/firm action/in order/to become/model citizens.
	idioms	firm hand		The politician said that/children nowadays need a/firm hand/in order/to become/model citizens.
25	free word combination	give a name		Laura said/she was going to/give a name/after delivering/the baby/at home.
	restricted collocations	give a talk		Laura said/she was going to/give a talk/after delivering/the baby/at home.
	idioms	give a hand		Laura said/she was going to/give a hand/after delivering/the baby/at home.
	free word combination	give a name		In the end,/I decided to/give a name/because they/kept asking/me to do so.
	restricted collocations	give a talk		In the end,/I decided to/give a talk/because they/kept asking/me to do so.
	idioms	give a hand		In the end,/I decided to/give a hand/because they/kept asking/me to do so.
	free word combination	give a name	25.Will I do it for money?	To be honest,/I do not want to/give a name/even though/you offer/lots of money.
	restricted collocations	give a talk		To be honest,/I do not want to/give a talk/even though/you offer/lots of money.
	idioms	give a hand	25.Will I do it for money?	To be honest,/I do not want to/give a hand/even though/you offer/lots of money.
26	free word combination	raise arms		The instructor taught us that/we had to/raise arms/and be/an inspiration/to others.
	restricted collocations	raise hopes		The instructor taught us that/we had to/raise hopes/and be/an inspiration/to others.
	idioms	raise hell		The instructor taught us that/we had to/raise hell/and be/an inspiration/to others.
	free word combination	raise arms		Do not say you have not been warned,/you would better not/raise arms/if you/are not/prepared for what is going to happen next.
	restricted collocations	raise hopes		Do not say you have not been warned,/you would better not/raise hopes/if you/are not/prepared for what is going to happen next.

	idioms	raise hell		Do not say you have not been warned,/you would better not/raise hell/if you/are not/prepared for what is going to happen next.
	free word combination	raise arms	26.Are they going to ignore opposition?	It is madness,/they are going to/raise arms/regardless of/all opposition/that they might face.
	restricted collocations	raise hopes		It is madness,/they are going to/raise hopes/regardless of/all opposition/that they might face.
	idioms	raise hell	26.Are they going to ignore opposition?	It is madness,/they are going to/raise hell/regardless of/all opposition/that they might face.
27	free word combination	raise your arms		Here is the trick,/you have to/raise your arms/if you/want to/be noticed.
	restricted collocations	raise your voice		Here is the trick,/you have to/raise your voice/if you/want to/be noticed.
	idioms	raise your game		Here is the trick,/you have to/raise your game/if you/want to/be noticed.
	free word combination	raise your arms	27.Is there no way of getting noticed?	The only way of getting recognised by them/is to/raise your arms/up to/a very/high level in this fierce competition.
	restricted collocations	raise your voice	27.Is there a way of getting noticed?	The only way of getting recognised by them/is to/raise your voice/up to/a very/high level in this fierce competition.
	idioms	raise your game		The only way of getting recognised by them/is to/raise your game/up to/a very/high level in this fierce competition.
	free word combination	raise your arms		We think/it would be better for you to/raise your arms/because no-one/is paying/any attention.
	restricted collocations	raise your voice		We think/it would be better for you to/raise your voice/because no-one/is paying/any attention.
	idioms	raise your game		We think/it would be better for you to/raise your game/because no-one/is paying/any attention.
28	free word combination	bright colour		We were so glad,/there was a/bright colour/that could/possibly lift/her spirits.
	restricted collocations	bright smile		We were so glad,/there was a/bright smile/that could/possibly lift/her spirits.
	idioms	bright spot		We were so glad,/there was a/bright spot/that could/possibly lift/her spirits.
	free word combination	bright colour	28. Do we all agree on this?	We all agree that/Ron needs to find a/bright colour/that can/make him/feel better.
	restricted collocations	bright smile	28. Do we all agree on this?	We all agree that/Ron needs to find a/bright smile/that can/make him/feel better.
	idioms	bright spot		We all agree that/Ron needs to find a/bright spot/that can/make him/feel better.
	free word combination	bright colour		It is said that/Yolanda is looking for a/bright colour/in her/very dull/life.
	restricted collocations	bright smile		It is said that/Yolanda is looking for a/bright smile/in her/very dull/life.

	idioms	bright spot		It is said that/Yolanda was looking for a/bright spot/in her/very dull/life.
29	free word combination	closed hand		It is clear,/Frank shows a/closed hand,/so it/is hard/to have him open it up.
	restricted collocations	closed mind		It is clear,/Frank has/a closed mind,/so it/is hard/to have him open up.
	idioms	closed book		It is clear,/Frank is/a closed book,/so it/is hard/to have him open up.
	free word combination	closed hand	29.Is it Tina who we are talking about?	Sometimes,/Tina can show/a closed hand,/so it/is impossible/to know what she is holding.
	restricted collocations	closed mind	29.Is it Tina who we are talking about?	Sometimes,/Tina can be a/closed mind,/so it/is impossible/to talk to her.
	idioms	closed book		Sometimes,/Tina can be/a closed book,/so it/is impossible/to talk to her.
	free word combination	closed hand		It is said that/Michael punches with/a closed hand,/which is/not a/pleasant thing to experience.
	restricted collocations	closed mind		It is said that/Michael is a bit of/a closed mind,/which is/not a/pleasant thing to experience.
	idioms	closed book		It is said that/Michael is a bit of a/closed book,/which is/not a/pleasant thing to experience.
30	free word combination	out of school		I believe that/the student is/out of school,/therefore this/needs to/be communicated to his parents.
	restricted collocations	out of order		I believe that/the student is/out of order,/therefore this/needs to/be communicated to his parents.
	idioms	out of place		I believe that/the student feels/out of place,/therefore this/needs to/be communicated to his parents.
	free word combination	out of school		The teacher informs me that/the young man is/out of school,/but I/do not/know what needs to be done.
	restricted collocations	out of order		The teacher informs me that/the young man is/out of order,/but I/do not/know what needs to be done.
	idioms	out of place		The teacher informs me that/the young man looks/out of place,/but I/do not/know what needs to be done.
	free word combination	out of school	30.Should we ignore the bully?	It is worrying that/this bully is/out of school,/so we/would better/keep an eye on him.
	restricted collocations	out of order		It is worrying that/this bully is/out of order,/so we/would better/keep an eye on him.
	idioms	out of place	30.Should we ignore the bully?	It is worrying that/this bully seems/out of place,/so we/would better/keep an eye on him.
31	free word combination	in his flat		If his girlfriend/had been/in his flat,/the tragedy/would never/have happened.

	restricted collocations	in his thoughts		If his girlfriend/had been/in his thoughts,/the tragedy/would never/have happened.
	idioms	in his shoes		If his girlfriend/had been/in his shoes,/the tragedy/would never/have happened.
	free word combination	in his flat		If the workers/had been/in his flat,/then everything/would have/turned out differently.
	restricted collocations	in his thoughts		If the workers/had been/in his thoughts,/then everything/would have/turned out differently.
	idioms	in his shoes		If the workers/had been/in his shoes,/then everything/would have/turned out differently.
	free word combination	in his flat	31.Did the accident already happen?	If his daughter/had been/in his flat,/the accident/could have/been avoided.
	restricted collocations	in his thoughts		If his daughter/had been/in his thoughts,/the accident/could have/been avoided.
	idioms	in his shoes	31.Did the accident already happen?	If his daughter/had been/in his shoes,/the accident/could have/been avoided.
32	free word combination	planted a tree	32.Do I want to attract attention for the event?	With the aim of attracting attention,/I have/planted a tree/for the/big event/later in the week.
	restricted collocations	planted a bomb		With the aim of attracting attention,/I have/planted a bomb/for the/big event/later in the week.
	idioms	planted a seed		With the aim of attracting attention,/I have/planted a seed/for the/big event/later in the week.
	free word combination	planted a tree		One of the most popular speakers/must have/planted a tree/simply because/the area/had nothing before his arrival.
	restricted collocations	planted a bomb		One of the most popular speakers/must have/planted a bomb/simply because/the area/is full of remains.
	idioms	planted a seed		One of the most popular speakers/must have/planted a seed/simply because/the area/is full of motivated students.
	free word combination	planted a tree		The researcher from that lab/must have/planted a tree/because later/the committee/questioned him.
	restricted collocations	planted a bomb		The researcher from that lab/must have/planted a bomb/because later/the committee/questioned him.
	idioms	planted a seed		The researcher from that lab/must have/planted a seed/because later/the committee/changed their mind.
33	free word combination	miss the meal	33.Do you need to get prepared?	Stay prepared,/otherwise you might/miss the meal,/which would/be a/real shame.
	restricted collocations	miss the train		Stay prepared,/otherwise you might/miss the train,/which would/be a/real shame.

	idioms	miss the boat		Stay prepared,/otherwise you might/miss the boat,/which would/be a/real shame.
	free word combination	miss the meal		If he is not well-organised,/he is going to/miss the meal,/then his/parents will/become very upset.
	restricted collocations	miss the train		If he is not well-organised,/he is going to/miss the train,/then his/parents will/become very upset.
	idioms	miss the boat		If he is not well-organised,/he is going to/miss the boat,/then his/parents will/become very upset.
	free word combination	miss the meal		Unfortunately,/they will/miss the meal,/which means/that they/will have no place to stay for tonight.
	restricted collocations	miss the train		Unfortunately,/they will/miss the train,/which means/that they/will have no place to stay for tonight.
	idioms	miss the boat		Unfortunately,/they will/miss the boat,/which means/that they/will have no place to stay for tonight.
34	free word combination	in deep snow		Your brother does not know what to do,/he is/in deep snow/and needs/urgent help/right now.
	restricted collocations	in deep trouble		Your brother does not know what to do,/he is/in deep trouble/and needs/urgent help/right now.
	idioms	in deep water		Your brother does not know what to do,/he is/in deep water/and needs/urgent help/right now.
	free word combination	in deep snow	34.Did Lewis need help?	Lewis was aware that/he was stuck/in deep snow/and he/called his/best friend to help him out.
	restricted collocations	in deep trouble	34.Did Lewis need help?	Lewis was aware that/he was stuck/in deep trouble/and he/called his/best friend to help him out.
	idioms	in deep water		Lewis was aware that/he was stuck/in deep water/and he/called his/best friend to help him out.
	free word combination	in deep snow		Due to his own carelessness,/Bill found himself/in deep snow/and he/did not/know what to do anymore.
	restricted collocations	in deep trouble		Due to his own carelessness,/Bill found himself/in deep trouble/and he/did not/know what to do anymore.
	idioms	in deep water		Due to his own carelessness,/Bill found himself/in deep water/and he/did not/know what to do anymore.
35	free word combination	hot soup		Be careful,/this is a/hot soup/and you/do not/want to do anything stupid.
	restricted collocations	hot topic		Be careful,/this is a/hot topic/and you/do not/want to do anything stupid.
	idioms	hot potato		Be careful,/this is a/hot potato/and you/do not/want to do anything stupid.
	free word combination	hot soup		Pete did not want to/deal with the/hot soup/and he/decided/to go for a walk.

	restricted collocations	hot topic		Pete did not want to/deal with the/hot topic/and he/decided to/go for a walk.
	idioms	hot potato		Pete did not want to/deal with the/hot potato/and he/decided/to go for a walk.
	free word combination	hot soup	35.Would Jim do it immediately?	Jim realised that/it was a/hot soup/and he/decided to/wait for a while.
	restricted collocations	hot topic		Jim realised that/it was a/hot topic/and he/decided to/wait for a while.
	idioms	hot potato	35.Would Jim do it immediately?	Jim realised that/it was a/hot potato/and he/decided to/wait for a while.
36	free word combination	face the front	36.Did he have something to do with it?	I have no idea/why he has to/face the front/when this/has nothing/to do with him?
	restricted collocations	face the truth		I have no idea/why he has to/face the truth/when this/has nothing/to do with him?
	idioms	face the music		I have no idea/why he has to/face the music/when this/has nothing/to do with him?
	free word combination	face the front		You need to/sit down and/face the front/and do/not leave/until you understand what you have done.
	restricted collocations	face the truth		You need to/sit down and/face the truth/and do/not leave/until you understand what you have done.
	idioms	face the music		You need to/sit down and/face the music/and do/not leave/until you understand what you have done.
	free word combination	face the front		Because Sue had been so horrible,/she had to/face the front/and she/was told/to change her attitude.
	restricted collocations	face the truth		Because Sue had been so horrible,/she had to/face the truth/and she/was told/to change her attitude.
	idioms	face the music		Because Sue had been so horrible,/she had to/face the music/and she/was told/to change her attitude.
37	free word combination	vast area		We were told clearly that/this is a/vast area/and that/we should/stay alert.
	restricted collocations	key area		We were told clearly that/this is a/key area/and that/we should/stay alert.
	idioms	grey area		We were told clearly that/this is a/grey area/and that/we should/stay alert.
	free word combination	vast area		That forbidden forest/at the country's border remains a/vast area/that has/to be/regulated yet.
	restricted collocations	key area		That forbidden forest/at the country's border remains a/key area/that has/to be/regulated yet.
	idioms	grey area		That forbidden forest/at the country's border remains a/grey area/that has/to be/regulated yet.
	free word combination	vast area	37.Do we need to take notes?	Our boss told us that/we should explore this/vast area/and take/lots of/notes.
	restricted collocations	key area		Our boss told us that/we should explore this/key area/and take/lots of/notes.

	idioms	grey area	37.Do we need to take notes?	Our boss told us that/we should explore this/grey area/and take/lots of/notes.
38	free word combination	find her book		I am informed that/once she can/find her book,/she will/be starting/to deal with this problem.
	restricted collocations	find her voice		I am informed that/once she can/find her voice,/she will/be starting/to deal with this problem.
	idioms	find her feet		I am informed that/once she can/find her feet,/she will/be starting/to deal with this problem.
	free word combination	find her book	38.Has Mary already found a solution?	As soon as/Mary is able to/find her book,/she will/look for/a solution.
	restricted collocations	find her voice	38.Has Mary already found a solution?	As soon as/Mary is able to/find her voice,/she will/look for/a solution.
	idioms	find her feet		As soon as/Mary is able to/find her feet,/she will/look for/a solution.
	free word combination	find her book		Give her some time,/she will/find her book,/then she/will tell/the whole story.
	restricted collocations	find her voice		Give her some time,/she will/find her voice,/then she/will tell/the whole story.
	idioms	find her feet		Give her some time,/she will/find her feet,/then she/will tell/the whole story.
39	free word combination	save water	39.Is the crisis international?	People have realised that/the most urgent thing now is to/save water/because of/the big/international crisis.
	restricted collocations	save energy		People have realised that/the most urgent thing now is to/save energy/because of/the big/international crisis.
	idioms	save face		People have realised that/the most urgent thing now is to/save face/because of/the big/international crisis.
	free word combination	save water		The news reported that/the man tried to/save water/partly because/he had/to live within his means.
	restricted collocations	save energy		The news reported that/the man tried to/save energy/partly because/he had/to compete later in the day.
	idioms	save face		The news reported that/the man tried to/save face/partly because/he had/made an embarrassing mistake.
	free word combination	save water		It might be tricky,/but I believe we can/save water/if we/try to/help each other out.
	restricted collocations	save energy		It might be tricky,/but I believe we can/save energy/if we/try to/help each other out.
	idioms	save face		It might be tricky,/but I believe we can/save face/if we/try to/help each other out.

40	free word combination	soft skin		He does not like/the idea of having a/soft skin,/mainly because/he thinks/it means he is not masculine enough.
	restricted collocations	soft drink		He does not like/the idea of having a/soft drink,/mainly because/he thinks/it means he is not masculine enough.
	idioms	soft touch		He does not like/the idea of having a/soft touch,/mainly because/he thinks/it means he is not masculine enough.
	free word combination	soft skin		I do not understand/why you suddenly want a/soft skin,/would you/care to/tell me the reason?
	restricted collocations	soft drink		I do not understand/why you suddenly want a/soft drink,/would you/care to/tell me the reason?
	idioms	soft touch		I do not understand/why you suddenly want a/soft touch,/would you/care to/tell me the reason?
	free word combination	soft skin	40.Did I believe that I was responsible for telling you about it?	I was aware that/Rachel had a/soft skin,/but I/did not/think it was my responsibility to tell you about it.
	restricted collocations	soft drink		I was aware that/Rachel had a/soft drink,/but I/did not/think it was my responsibility to tell you about it.
	idioms	soft touch	40.Did I believe that I was responsible for telling you about it?	I was aware that/Rachel had a/soft touch,/but I/did not/think it was my responsibility to tell you about it.
41	free word combination	sweet taste		Fiona told me that/it had a/sweet taste,/but she/did not/want to be reminded of it.
	restricted collocations	sweet dreams		Fiona told me that/she had/sweet dreams,/but she/did not/want to be reminded of them.
	idioms	sweet tooth		Fiona told me that/she had a/sweet tooth,/but she/did not/want to be reminded of it.
	free word combination	sweet taste	41.Was I surprised about this?	I was not surprised that/the apple had a/sweet taste,/simply because/my friend/had told me it would.
	restricted collocations	sweet dreams	41.Was I surprised about this?	I was not surprised that/the child had/sweet dreams,/simply because/my friend/had told me he would.
	idioms	sweet tooth		I was not surprised that/the child had a/sweet tooth,/simply because/my friend/had told me he would.
	free word combination	sweet taste		It is true,/this cereal used to have a/sweet taste,/but that/is no/longer the case.
	restricted collocations	sweet dreams		It is true,/this girl used to have/sweet dreams,/but that/is no/longer the case.
	idioms	sweet tooth		It is true,/this girl used to have a/sweet tooth,/but that/is no/longer the case.

42	free word combination	read their list		There is no option left,/we need to/read their list/in order/to know/exactly what is going on.
	restricted collocations	read their minds		There is no option left,/we need to/read their minds/in order/to know/exactly what is going on.
	idioms	read their lips		There is no option left,/we need to/read their lips/in order/to know/exactly what is going on.
	free word combination	read their list		They might be surprised/if you manage to/read their list/under the/present circumstances.
	restricted collocations	read their minds		They might be surprised/if you manage to/read their minds/under the/rather unusual/present circumstances.
	idioms	read their lips		They might be surprised/if you manage to/read their lips/under the/present circumstances.
	free word combination	read their list		This is urgent,/I need someone who can/read their list/and also/report it/back to me.
	restricted collocations	read their minds		This is urgent,/I need someone who can/read their minds/and also/report it/back to me.
	idioms	read their lips	42.Can I do it by myself?	This is urgent,/I need someone who can/read their lips/and also/report it/back to me.
43	free word combination	wild horse		We had been warned,/this animal is a/wild horse/and you/have to/be careful with it.
	restricted collocations	wild guess		We had been warned,/this prediction is a/wild guess/and you/have to/be careful with it.
	idioms	wild cards		We had been warned,/these people are/wild cards/and you/have to/be careful with them.
	free word combination	wild horses	43.Do we already know the right way back?	No matter how bad the situation is,/we cannot rely on such a/wild horse/that might/help us/to search for the right way back home.
	restricted collocations	wild guess	43.Do we already know the right way back?	Given how bad the situation is,/we might have to rely on such a/wild guess/that can/help us/to search for the right way back home.
	idioms	wild cards		No matter how bad the situation is,/we cannot rely on such/wild cards/that might/help us/to search for the right way back home.
	free word combination	wild horses		Enough is enough,/we do not need a/wild horse/to make/the situation/even more chaotic.
	restricted collocations	wild guess		Enough is enough,/we do not need a/wild guess/to make/the situation/even more chaotic.

	idioms	wild cards		Enough is enough,/we do not need/wild cards/to make/the situation/even more chaotic.
44	free word combination	change their boots	44.Is it worth trying to persuade them?	No matter how hard you try,/they are not going to/change their boots/and it/is a/waste of time trying to persuade them.
	restricted collocations	change their views		No matter how hard you try,/they are not going to/change their views/and it/is a/waste of time trying to persuade them.
	idioms	change their spots		No matter how hard you try,/they are not going to/change their spots/and it/is a/waste of time trying to persuade them.
	free word combination	change their boots		They are very stubborn/and will not/change their boots/just because/you ask/them.
	restricted collocations	change their views		They are very stubborn/and will not/change their views/just because/you ask/them.
	idioms	change their spots		They are very stubborn/and will not/change their spots/just because/you ask/them.
	free word combination	change their boots		I know them very well,/they will never/change their boots/because they/always think/they know better.
	restricted collocations	change their views		I know them very well,/they will never/change their views/because they/always think/they know better.
	idioms	change their spots		I know them very well,/they will never/change their spots/because they/always think/they know better.
45	free word combination	rough track	45.If they want to proceed, do they need to make improvements?	If they want to proceed,/they need to fix the/rough track,/so that/everyone will/be able to trek along it.
	restricted collocations	rough draft		If they want to proceed,/they need to fix the/rough draft,/so that/everyone will/be able to understand it.
	idioms	rough edges		If they want to proceed,/they need to fix the/rough edges,/so that/everyone will/be able to agree with it.
	free word combination	rough track		The engineer will need to/spend time on improving the/rough track,/especially if/he wants/to be taken seriously.
	restricted collocations	rough draft		The engineer will need to/spend time on improving his/rough draft,/especially if/he wants/to be taken seriously.
	idioms	rough edges		The engineer will need to/spend time on improving his/rough edges,/especially if/he wants/to be taken seriously.
	free word combination	rough track		The supervisor was never/going to ignore the/rough track,/as he/was worried/that someone might end up seriously injured.
	restricted collocations	rough draft		The supervisor was never/going to ignore their/rough draft,/as he/was worried/that someone might end up seriously infuriated.

	idioms	rough edges		The supervisor was never/going to ignore their/rough edges,/as he/was worried/that someone might end up seriously injured.

Filler sentence

When she takes off/her chunky coat/off the sky,/her brother/comes out/from the living room, saying hello to her.
In London,/I live in a small hat./As soon as/I could/afford it,/I will buy a flat in Paris.
I am being breathed to/all kinds of events,/and as long as I/am broken,/I will/be doing anything.
Julian loves being himself,/being independent/and making his/own decision,/which is/why people think he is so needy.
Everything we observe/has a direct impact on our brains./Unfortunately through television and/social water,/we are/exposed to many bad things.
With her other hand,/she is holding/her mountain,/through which/Joanna is/telling a story about her job.
There has been an increasing number of people/who cheat on/their marriage partners,/so the/number of/divorces is much lower than ever in recent years.
A raindrop from the newspaper/tells a 70-year-old woman/children stories/about time travel/from her garden.
I decided to bomb my two balconies/and fill them/with plants,/so that/I could/sit out there and water them.
A piece of cake/will be interviewed soon,/so you can/write down/the number/and give us the information.
Now the children/have left home/for summer crab,/she does/not have/much time on her hands.
Since stores are closed/every day,/customers have/an unlimited/amount of time/ to examine the goods in stores.
There is no doubt that/good quality of sleep can/seriously harm/your health/and cause/many diseases in the long term.
Before I realised what was happening, the bag grabbed a pine tree and started to talk to it.
The hospital is/empty now/because it is full of people/in desperate/need of/medical attention.
In the last two weeks,/the situation has become/less indigo and worrying,/with more/strikes and/protests spreading across the country.
There was/a lot of pressure/on the team tonight/and that/is why/they were entirely relaxed.
To be unsuccessful/in a river,/producers must devote more time/than ever/to developing/relationships with clients.

Who would ever guess from the smiling expression on his face that just a moment ago he had been so happy?
He was not working very hard/and eventually/achieved her tower of/becoming a/well-respected professor/in her field.
For some reason/we do not fully walk,/she is the winner/as she/was doing/so poorly in the competition.
Once you have/passed your driving test,/your career/will be/destroyed and/you can start to build up your experience.
Newly born babies/like to/ride their bikes/no matter/how bad/the weather is.
Engineers were not able to find the source of the problem in the food, so they decided to call a desk.
Primary school teachers/suggest that/each kid/should visit a dentist/least twenty times/a day.
Remember to always/add a good dose/of bleach to/the machine/when you/wash your laptops.
Kate's hair barked at her laptop since for a while it was warm and dry outside.
Tim went to/a laboratory/on Tuesday morning/where they/served the/most delicious chemistry.
She is a/good wardrobe/because she always avoids/dealing with/the problems/of her staff.
Learning a new language/takes a lot of hair,/so it/sounds like/I can/easily achieve it.
As lockdown eases,/Sue went to/the garden center/in the village/and a/new prisoner cut her tea very short.
For your own/and other passengers' safety,/please smoke and/set fire/when the/plane takes off.
Rachel plans to/go to the post office/this afternoon/since her/sugar level is/running low.
Many 3-year-old babies send their grandparents to public school because the tuition fee is very low.
The movie that/was shown at/the local painting/attracted a/high number/of beer.
This article is/worth eating/as some of/the information/in the/article is not particularly relevant.
Fewer and fewer people/are starting to/work from their scooter/after the/whole country/went into lockdown.
As we all know,/children usually start/their primary school/between the/ages of/23 and 24.
Claire burned the vase,/so that/she could/listen to/her favourite/classical pieces.
I went to a hospital and a lorry who examined me told me to take some time off work and try to rest.

Since it was raining,/Ella hung the/clothes outside/so that/they would/dry faster.
The weather is/hot and sunny./You would better/wear more/layers to/keep yourself warm.
Can you speak/a little louder please?/My hearing is/very good/after the/successful operation.
The door is solid,/so he can have a peek/and know/what is/going on/inside.
They had/a safe journey:/there was/heavy snow and/the car crashed into/the mountain on the way back home.

Appendix 3.3

Summary of the Best-fitting Models of MWSs on the Spill-over Region for L1 Speakers

	L1 Spill-over Region			
Predictors	***β***	***SE***	***t***	***p***
(Intercept)	6.35	0.022	288.03	<0.001
idioms	-0.03	0.01	-1.75	0.081
restricted collocations	-0.02	0.01	-1.45	0.147
Random Effects				
σ^2		0.04		
$\tau_{00\ items}$		0		
$\tau_{00\ Subject}$		0.03		
ICC		0.46		
$N_{\ Subject}$		84		
$N_{\ items}$		133		
Observations		10377		
Marginal R^2 / Conditional R^2		0.002 / 0.462		

Note. The reference level is free word combinations, reading times are log-transformed

Appendix 3.4

Summary of the Best-fitting Models of MWSs on the Spill-over Region for L2 Speakers

	L2 Spill-over Region			
Predictors	***β***	***SE***	***t***	***p***
(Intercept)	6.42	0.02	266.51	<0.001
profiency	-0.0006	0.002	-0.32	.749
idioms	0.0009	0.02	0.05	.962
restricted collocations	-0.01	0.02	-0.82	.412
Random Effects				
σ^2		0.06		
$\tau_{00\ items}$		0.01		
$\tau_{00\ Subject}$		0.03		
ICC		0.38		
$N_{\ Subject}$		73		
$N_{\ items}$		133		
Observations		8783		
Marginal R^2 / Conditional R^2	0.001 / 0.380			

Note. The reference level is free word combinations, reading times are log-transformed

Appendix 3.5

Reports of Data Loss for Each Cleaning Step by Group and Condition

Removing trials of which comprehension questions answered incorrectly				
Group	**Total**	**Condition**		
		free word combinations	restricted collocations	idioms
L1 speakers	68 trials (0.6%)	21 trials (0.19%)	20 trials (0.18%)	27 trials (0.24%)
L2 speakers	82 trials (0.84%)	26 trials (0.26%)	38 trials (0.39%)	18 trials (0.18%)

Removing sentences judged as implausible				
Group	**Total**	**Condition**		
		free word combinations	restricted collocations	idioms
L1 speakers	155 trials (1.02%)	76 trials (0.68%)	37 trials (0.33%)	42 trials (0.37%)
L2 speakers	97 trials (0.99%)	38 trials (0.39%)	21 trials (0.21%)	38 trials (0.39%)

Removing trials of which the critical region based on the cut-off points				
Group	**Total**	**Condition**		
		free word combinations	restricted collocations	idioms
L1 speakers	45 trials (0.4%)	19 trials (0.17%)	13 trials (0.11%)	13 trials (0.11%)
L2 speakers	198 trials (2.03%)	63 trials (0.64%)	60 trials (0.61%)	75 trials (0.77%)

Removing trials of which the spill-over region based on the cut-off points				
Group	**Total**	**Condition**		
		free word combinations	restricted collocations	idioms
L1 speakers	31 trials (0.27%)	10 trials (0.08%)	6 trials (0.05%)	15 trials (0.13%)
L2 speakers	125 trials (1.28%)	40 trials (0.41%)	46 trials (0.47%)	39 trials (0.4%)

Trial removal on the critical region based on 2.5 SD				
Group	**Total**	**Condition**		
		free word combinations	restricted collocations	idioms
L1 speakers	334 trials (2.98%)	111 trials (0.99%)	117 trials (1.04%)	106 trials (0.94%)
L2 speakers	281 trials (2.89%)	84 trials (0.86%)	86 trials (0.88%)	111 trials (1.14%)

Trial removal on the spill-over region based on 2.5 SD				
Group	**Total**	**Condition**		
		free word combinations	restricted collocations	idioms
L1 speakers	326 trials (2.91%)	118 trials (1.05%)	110 trials (0.98%)	98 trials (0.87%)
L2 speakers	310 trials (3.19%)	109 trials (1.12%)	89 trials (0.91%)	112 trials (1.15%)

Appendix 4.1

73 sets of candidate items

set	condition	final word	items
1	idioms	air	hot air
1	collocations	air	clean air
1	novel combinations	air	rural air
2	idioms	apple	rotten apple
2	collocations	apple	fresh apple
2	novel combinations	apple	soft apple
3	idioms	area	grey area
3	collocations	area	vast area
3	novel combinations	area	rare area
4	idioms	arrow	straight arrow
4	collocations	arrow	curved arrow
4	novel combinations	arrow	metal arrow
5	idioms	bath	early bath
5	collocations	bath	warm bath
5	novel combinations	bath	white bath
6	idioms	beer	small beer
6	collocations	beer	strong beer
6	novel combinations	beer	French beer
7	idioms	bee	busy bee
7	collocations	bee	little bee
7	novel combinations	bee	tiny bee
8	idioms	bird	early bird
8	collocations	bird	small bird
8	novel combinations	bird	social bird
9	idioms	blanket	wet blanket
9	collocations	blanket	thick blanket
9	novel combinations	blanket	soft blanket
10	idioms	blood	fresh blood
10	collocations	blood	royal blood
10	novel combinations	blood	natural blood
11	idioms	blow	final blow
11	collocations	blow	single blow

11	novel combinations	blow	recent blow
12	idioms	book	closed book
12	collocations	book	classic book
12	novel combinations	book	Greek book
13	idioms	brush	broad brush
13	collocations	brush	soft brush
13	novel combinations	brush	cheap brush
14	idioms	bullet	silver bullet
14	collocations	bullet	plastic bullet
14	novel combinations	bullet	metal bullet
15	idioms	cat	fat cat
15	collocations	cat	dead cat
15	novel combinations	cat	happy cat
16	idioms	chapter	new chapter
16	collocations	chapter	main chapter
16	novel combinations	chapter	nice chapter
17	idioms	cheque	blank cheque
17	collocations	cheque	stolen cheque
17	novel combinations	cheque	missing cheque
18	idioms	cheese	big cheese
18	collocations	cheese	soft cheese
18	novel combinations	cheese	bad cheese
19	idioms	chestnut	old chestnut
19	collocations	chestnut	sweet chestnut
19	novel combinations	chestnut	bad chestnut
20	idioms	collar	white collar
20	collocations	collar	high collar
20	novel combinations	collar	clean collar
21	idioms	colours	true colours
21	collocations	colours	dark colours
21	novel combinations	colours	prime colours
22	idioms	corner	tight corner
22	collocations	corner	bottom corner
22	novel combinations	corner	closed corner
23	idioms	cow	silly cow
23	collocations	cow	dairy cow
23	novel combinations	cow	hungry cow
24	idioms	day	rainy day
24	collocations	day	sunny day

24	novel combinations	day	weird day
25	idioms	elephant	white elephant
25	collocations	elephant	African elephant
25	novel combinations	elephant	male elephant
26	idioms	drawer	top drawer
26	collocations	drawer	locked drawer
26	novel combinations	drawer	front drawer
27	idioms	face	long face
27	collocations	face	young face
27	novel combinations	face	blue fish
28	idioms	feet	cold feet
28	collocations	feet	front feet
28	novel combinations	feet	nice feet
29	idioms	feet	itchy feet
29	collocations	feet	sore feet
29	novel combinations	feet	injured feet
30	idioms	fingers	green fingers
30	collocations	fingers	long fingers
30	novel combinations	fingers	large fingers
31	idioms	fingers	sticky fingers
31	collocations	fingers	slender fingers
31	novel combinations	fingers	skinny fingers
32	idioms	fish	cold fish
32	collocations	fish	fresh fish
32	novel combinations	fish	blue fish
33	idioms	flag	white flag
33	collocations	flag	national flag
33	novel combinations	flag	common flag
34	idioms	fuse	short fuse
34	collocations	fuse	slow fuse
34	novel combinations	fuse	thin fuse
35	idioms	grapes	sour grapes
35	collocations	grapes	ripe grapes
35	novel combinations	grapes	toxic grapes
36	idioms	hand	old hand
36	collocations	hand	small hand
36	novel combinations	hand	wide hand
37	idioms	hand	firm hand
37	collocations	hand	steady hand

37	novel combinations	hand	pretty hand
38	idioms	hands	wrong hands
38	collocations	hands	strong hands
38	novel combinations	hands	weak hands
39	idioms	hat	black hat
39	collocations	hat	new hat
39	novel combinations	hat	great hat
40	idioms	horse	dark horse
40	collocations	horse	grey horse
40	novel combinations	horse	weak horse
41	idioms	lane	fast lane
41	collocations	lane	narrow lane
41	novel combinations	lane	flat lane
42	idioms	light	green light
42	collocations	light	white light
42	novel combinations	light	useful light
43	idioms	lining	silver lining
43	collocations	lining	inner lining
43	novel combinations	lining	hidden lining
44	idioms	lunch	free lunch
44	collocations	lunch	light lunch
44	novel combinations	lunch	huge lunch
45	idioms	mark	black mark
45	collocations	mark	red mark
45	novel combinations	mark	green mark
46	idioms	meat	dead meat
46	collocations	meat	fresh meat
46	novel combinations	meat	extra meat
47	idioms	mouth	big mouth
47	collocations	mouth	dry mouth
47	novel combinations	mouth	flat mouth
48	idioms	name	big name
48	collocations	name	full name
48	novel combinations	name	short name
49	idioms	nut	tough nut
49	collocations	nut	mixed nut
49	novel combinations	nut	smashed nut
50	idioms	pocket	deep pockets
50	collocations	pocket	small pockets

50	novel combinations	pocket	nice pockets
51	idioms	point	high point
51	collocations	point	main point
51	novel combinations	point	large point
52	idioms	potatoes	small potatoes
52	collocations	potatoes	roast potatoes
52	novel combinations	potatoes	nice potatoes
53	idioms	potato	hot potato
53	collocations	potato	baked potato
53	novel combinations	potato	fresh potato
54	idioms	ride	free ride
54	collocations	ride	long ride
54	novel combinations	ride	bad ride
55	idioms	sheep	black sheep
55	collocations	sheep	dead sheep
55	novel combinations	sheep	heavy sheep
56	idioms	sheet	clean sheet
56	collocations	sheet	single sheet
56	novel combinations	sheet	light sheet
57	idioms	shot	a long shot
57	collocations	shot	a single shot
57	novel combinations	shot	a hard shot
58	idioms	skin	thick skin
58	collocations	skin	dark skin
58	novel combinations	skin	pretty skin
59	idioms	slip	pink slip
59	collocations	slip	tiny slip
59	novel combinations	slip	thin slip
60	idioms	spark	bright spark
60	collocations	spark	tiny spark
60	novel combinations	spark	lovely spark
61	idioms	spot	sore spot
61	collocations	spot	ideal spot
61	novel combinations	spot	hidden spot
62	idioms	spot	black spot
62	collocations	spot	quiet spot
62	novel combinations	spot	huge spot
63	idioms	spoon	greasy spoon
63	collocations	spoon	wooden spoon

63	novel combinations	spoon	coffee spoon
64	idioms	stomach	strong stomach
64	collocations	stomach	flat stomach
64	novel combinations	stomach	huge stomach
65	idioms	straw	final straw
65	collocations	straw	fresh straw
65	novel combinations	straw	light straw
66	idioms	tape	red tape
66	collocations	tape	sticky tape
66	novel combinations	tape	fabric tape
67	idioms	water	deep water
67	collocations	water	warm water
67	novel combinations	water	mixed water
68	idioms	wall	brick wall
68	collocations	wall	stone wall
68	novel combinations	wall	vast wall
69	idioms	wire	live wire
69	collocations	wire	thin wire
69	novel combinations	wire	clear wire
70	idioms	wolf	a lone wolf
70	collocations	wolf	grey wolf
70	novel combinations	wolf	a weak wolf
71	idioms	wood	dead wood
71	collocations	wood	dark wood
71	novel combinations	wood	firm wood
72	idioms	word	have the final word
72	collocations	word	single word
72	novel combinations	word	normal word
73	idioms	work	dirty work
73	collocations	work	manual work
73	novel combinations	work	clever work

Appendix 4.2

Experimental items and sentences

set	condition	items	disambiguation region	control region	critical and spill-over region	question
1	idiom_figurative	hot air	We had the opportunity to meet the infamous Jack,	who is full of	hot air,/as Fiona/told us.	Did we miss the opportunity?
1	idiom_literal	hot air	We had the opportunity to work in the engine room,	which is full of	hot air,/as Fiona/told us.	Did we miss the opportunity?
1	collocation	clean air	We had the opportunity to hike in the Swiss Alps,	which is full of	clean air,/as Fiona/told us.	Did we miss the opportunity?
1	novel combination	rural air	We had the opportunity to visit the Welsh countryside,	which is full of	rural air,/as Fiona/told us.	Did we miss the opportunity?
2	idiom_figurative	rotten apple	Since he causes so many problems in the workplace,	there is no doubt that the	rotten apple/needs to/be taken out.	Does the person need to be taken out?
2	idiom_literal	rotten apple	Since our restaurant only uses the best ingredients,	there is no doubt that the	rotten apple/needs to/be taken out.	Does the apple need to be taken out?
2	collocation	fresh apple	Since we usually use overripe things to bake cakes,	there is no doubt that the	fresh apple/needs to/be taken out.	Does the apple need to be taken out?
2	novel combination	soft apple	Since our restaurant only uses the best ingredients,	there is no doubt that the	soft apple/needs to/be taken out.	Does the apple need to be taken out?
3	idiom_figurative	grey area	As the rules are not totally clear,	there is a	grey area/that we/have not noticed.	Have we noticed everything?
3	idiom_literal	grey area	As the wall is not painted properly,	there is a	grey area/that we/have not noticed.	Have we noticed everything?
3	collocation	vast area	As Russia is such a big country,	there is a	vast area/that we/have not noticed.	Have we noticed everything?

3	novel combination	rare area	Although most of them is known to us,	there is a	rare area/that we/have not noticed.	Have we noticed everything?
4	idiom_figurative	early bird	The best bargains will be sold out quickly,	so probably only the	early bird/will have/the chance of getting the product.	Can everyone get the product?
4	idiom_literal	early bird	Seeds are in short supply,	so probably only the	early bird/will have/the chance of getting the food.	Can every bird get the food?
4	collocation	small bird	Seeds are in short supply,	so probably only the	small bird/will have/the chance of getting the food.	Can every bird get the food?
4	novel combination	social bird	Building this type of nest requires cooperation,	so probably only the	social bird/will have/the chance of completing it before winter.	Can every bird complete building nest?
5	idiom_figurative	fresh blood	To create new ideas,	we need to bring in some	fresh blood/as soon/as possible.	Do we need to bring in someone new?
5	idiom_literal	fresh blood	To save the patient's life,	we need to bring in some	fresh blood/as soon/as possible.	Do we need to bring in something new?
5	collocation	royal blood	To earn international recognition,	we need to bring in some	royal blood/as soon/as possible.	Do we need to bring in someone new?
5	novel combination	natural blood	To save the animal's life,	we need to bring in some	natural blood/as soon/as possible.	Do we need to bring in something new?
6	idiom_figurative	final blow	Peter lost his job after a divorce,	unfortunately this is the	final blow/that crushed /him.	Was Peter single?
6	idiom_literal	final blow	Peter was hit by an uppercut in the contest,	unfortunately this is the	final blow/that crushed/him.	Was Peter hit in the contest?
6	collocation	single blow	Peter was only hit by an uppercut in the contest,	unfortunately this is the	single blow/that crushed/him.	Was Peter hit in the contest?

6	novel combination	recent blow	Peter was also hit by an uppercut in the contest,	unfortunately it is the	recent blow/that crushed/him.	Was Peter hit in the contest?
7	idiom_figurative	closed book	I cannot understand abstract art,	it is a	closed book/to me/if I am honest with you.	Do I enjoy abstract art?
7	idiom_literal	closed book	I cannot access the content inside,	it is a	closed book/to me/if I am honest with you.	Do I have access to the content?
7	collocation	classic book	I was forced to read "Great Expectations",	it is a	classic book/to me/if I am honest with you.	Have I never read "Great Expectations"?
7	novel combination	greek book	I cannot read the content inside,	it is a	Greek book/to me/if I am honest with you.	Can I read the content?
8	idiom_figurative	big cheese	After her promotion to vice president,	Anna needs to deal with the	big cheese/very carefully/to avoid mistakes.	Does Anna need to be careful?
8	idiom_literal	big cheese	After taking it from the shelf,	Anna needs to deal with the	big cheese/very carefully/to avoid mistakes.	Does Anna need to be careful?
8	collocation	soft cheese	After taking it from the shelf,	Anna needs to deal with the	soft cheese/very carefully/to avoid mistakes.	Does Anna need to be careful?
8	novel combination	bad cheese	After taking it from the shelf,	Anna needs to deal with the	bad cheese/very carefully/to avoid mistakes.	Does Anna need to be careful?
9	idiom_figurative	true colours	The kitten looked timid in her new home,	but then it displayed its	true colours/sooner than/expected.	Did the kitten reveal its personality?
9	idiom_literal	true colours	After stripping off layers of paint,	the cupboard displayed its	true colours/sooner than/expected.	Did the cupboard reveal its colour?
9	collocation	dark colours	After stripping off layers of paint,	the cupboard displayed its	dark colours/sooner than/expected.	Did the cupboard reveal its colour?

9	novel combination	prime colours	After stripping off layers of paint,	the cupboard displayed its	prime colours/sooner than/expected.	Did the cupboard reveal its colour?
10	idiom_figurative	tight corner	As he had hardly any options left,	he found himself in a	tight corner/and couldn't/get out of it.	Could he get out of trouble?
10	idiom_literal	tight corner	As the farm fences are built at a sharp angle,	the animal found itself in a	tight corner/and couldn't/get out of it.	Could the animal get out of the corner?
10	collocation	bottom corner	When the frightened fox ran down the garden,	the animal found itself in a	bottom corner/and couldn't/get out of it.	Could the animal get out of the corner?
10	novel combination	closed corner	When the frightened fox ran down the garden,	the animal found itself in a	closed corner/and couldn't/get out of it.	Could the animal get out of the corner?
11	idiom_figurative	silly cow	That girl really does not think before doing something;	there is little doubt that the	silly cow/was responsible/for it.	Was the girl at fault?
11	idiom_literal	silly cow	An animal had kicked over the pail of milk;	there is little doubt that the	silly cow/was responsible/for it.	Was the animal at fault?
11	collocation	dairy cow	An animal had kicked over the pail of milk;	there is little doubt that the	dairy cow/was responsible/for it.	Was the animal at fault?
11	novel combination	hungry cow	An animal had kicked over the pail of milk;	there is little doubt that the	hungry cow/was responsible/for it.	Was the animal at fault?
12	idiom_figurative	rainy day	I recommend saving money from now on,	which is good for a	rainy day,/and to/make a habit out of it.	Do I recommend making a habit out of it?
12	idiom_literal	rainy day	I recommend wearing a pair of boots,	which is good for a	rainy day,/and to/make a habit out of it.	Do I recommend making a habit out of it?
12	collocation	sunny day	I recommend wearing a pair of short pants,	which is good for a	sunny day,/and to/make a habit out of it.	Do I recommend making a habit out of it?

12	novel combination	weird day	I recommend remaining positive,	which is good for a	weird day,/and to/make a habit out of it.	Do I recommend making a habit out of it?
13	idiom_figurative	long face	She is disappointed with the current situation,	as I can tell from her	long face/with little/difficulty.	Is she disappointed with the current situation?
13	idiom_literal	long face	She looks almost identical to her mom,	as I can tell from her	long face/with little/difficulty.	Does she look like her mom?
13	collocation	young face	She still does not have enough experience,	as I can tell from her	young face/with little/difficulty.	Does she lack experience?
13	novel combination	wide face	She looks almost identical to her mom,	as I can tell from her	wide face/with little/difficulty.	Does she look like her mom?
14	idiom_figurative	itchy feet	Lauren loves travelling around the world every year,	it is certain she will get	itchy feet/sooner or/later.	Does Lauren hate travelling?
14	idiom_literal	itchy feet	If Lauren walks barefooted through these fields,	it is certain she will get	itchy feet/sooner or/later.	Will Lauren be comfortable walking barefooted?
14	collocation	sore feet	If Lauren walks barefooted through these fields,	it is certain she will get	sore feet/sooner or/later.	Will Lauren be comfortable walking barefooted?
14	novel combination	injured feet	If Lauren walks barefooted through these fields,	it is certain she will get	injured feet/sooner or/later.	Will Lauren be comfortable walking barefooted?
15	idiom_figurative	white flag	After a fierce fight,	our team displayed a	white flag/and it/worked.	Did our team's action work?
15	idiom_literal	white flag	To distinguish between the colours,	our team displayed a	white flag/and it/worked.	Did our team's action work?
15	collocation	national flag	During the opening ceremony,	our team displayed a	national flag/and it/worked.	Did our team's action work?

15	novel combination	common flag	To not distinguish us from others,	our team displayed a	common flag/and it/worked.	Did our team's action work?
16	idiom_figurative	old hand	Oliver has been working in this field for many years,	apparently he is an	old hand/and everyone/knows it.	Does Oliver work in the field?
16	idiom_literal	old hand	Oliver stumbled upon a horrible crime scene,	apparently he saw an	old hand/and everyone/knows it.	Did Oliver come across a crime scene?
16	collocation	small hand	Oliver inherited many features from his dad,	apparently he has a	small hand/and everyone/knows it.	Does Oliver look like his dad?
16	novel combination	wide hand	Oliver inherited many features from his dad,	apparently he has a	wide hand/and everyone/knows it.	Does Oliver look like his dad?
17	idiom_figurative	firm hand	To make those children behave well,	they need someone with a	firm hand/if at/all possible.	Do they need someone else to get involved?
17	idiom_literal	firm hand	To open the stuck closed door,	they need someone with a	firm hand/if at/all possible	Do they need someone else to get involved?
17	collocation	steady hand	To repair the complicated artwork	they need someone with a	steady hand/if at/all possible.	Do they need someone else to get involved?
17	novel combination	pretty hand	To find the best jewellery model for the show,	they need someone with a	pretty hand/if at/all possible.	Do they need someone else to get involved?
18	idiom_figurative	fast lane	He was having a successful career and exciting life,	he enjoyed being in the	fast lane/but often/people resented him.	Did he enjoy his career and life?
18	idiom_literal	fast lane	Driving at night when there were fewer cars,	he enjoyed being in the	fast lane/but often/people honked at him.	Did he enjoy driving at night?
18	collocation	narrow lane	While driving in the remote countryside,	he enjoyed being in the	narrow lane/but often/he had to slow down.	Did he enjoy driving in the countryside?

18	novel combination	flat lane	Compared to cycling uphill and downhill,	he enjoyed being in the	flat lane/but often/he got bored of it quickly	Did he enjoy cycling in the flat lane?
19	idiom_figurative	green light	Keep explanations simple in this project,	if you want to get the	green light/that you/are looking for.	Should you keep your explanations simple?
19	idiom_literal	green light	Press this small round button,	if you want to get the	green light/that you/are looking for.	Should you press the round button?
19	collocation	white light	Press this small round button,	if you want to get the	white light/that you/are looking for.	Should you press the round button?
19	novel combination	useful light	Press this small round button,	if you want to get the	useful light/that you/are looking for.	Should you press the round button?
20	idiom_figurative	free lunch	I need to be more motivated and work harder,	there is no	free lunch/for students/at this university.	Should I work less?
20	idiom_literal	free lunch	I had to pay regular price,	there is no	free lunch/for students/at this university.	Is there a free lunch at this university?
20	collocation	light lunch	They always provide fatty and oily food,	there is no	light lunch/for students/at this university.	Is there a light lunch at this university?
20	novel combination	huge lunch	They are very strict with food portions,	there is no	huge lunch/for students/at this university.	Is there a huge lunch at this university?
21	idiom_figurative	black mark	After John received a speeding ticket,	there was a	black mark/on his/file.	Is John's file clean?
21	idiom_literal	black mark	After John dropped the paint brush,	there was a	black mark/on his/file.	Is John's file clean?

21	collocation	red mark	After John dropped the paint brush,	there was a	red mark/on his/file.	Is John's file clean?
21	novel combination	green mark	After John dropped the paint brush,	there was a	green mark/on his/file.	Is John's file clean?
22	idiom_figurative	dead meat	John's girlfriend found him cheating on her,	so he is basically	dead meat/as far/as I know.	Is John having a good relationship with his girlfriend?
22	idiom_literal	dead meat	A chick was ran over in the dark,	so it is basically	dead meat/as far/as I know.	Was the chick run over when it was light?
22	collocation	fresh meat	I just bought the chicken from the butcher's	so it is basically	fresh meat/as far/as I know.	Did I buy the chicken from a supermarket?
22	novel combination	extra meat	A chicken was cooked after the big dinner,	so it is basically	extra meat/as far/as I know.	Was the chicken cooked before the big dinner?
23	idiom_figurative	big mouth	I never give confidential information to Emma	because she has a	big mouth,/as my/friend pointed out.	Was it my friend who pointed this out?
23	idiom_literal	big mouth	Emma is not happy with herself	because she has a	big mouth,/as my/friend pointed out.	Was it my friend who pointed this out?
23	collocation	dry mouth	Emma drinks plenty of cold water all day long	because she has a	dry mouth,/as my/friend pointed out.	Was it my friend who pointed this out?
23	novel combination	flat mouth	Emma is not happy with herself	because she has a	flat mouth,/as my/friend pointed out.	Was it my friend who pointed this out?
24	idiom_figurative	big name	When I looked at who had been arrested,	I was surprised to see a	big name/in the/middle of the front page.	Was I surprised by what I saw?
24	idiom_literal	big name	When I looked at the child's large writing,	I was able to see a	big name/in the/middle of the page.	Was I able to see the name?

24	collocation	full name	When I looked at who had been arrested,	I was surprised to see a	full name/in the/middle of the front page.	Was I surprised to see the name?
24	novel combination	short name	When I looked at the child's large writing,	I was able to see a	short name/in the/middle of the page.	Was I able to see the name?
25	idiom_figurative	tough nut	Inequality is an ever-growing problem.	It looks like the	tough nut/won't be/easily dealt with.	Can the problem be easily dealt with?
25	idiom_literal	tough nut	The shell is extremely big and thick.	It looks like the	tough nut/won't be/easily dealt with.	Can the nut be easily dealt with?
25	collocation	Brazil nut	The shell is extremely big and thick.	It looks like the	Brazil nut/won't be/easily dealt with.	Can the nut be easily dealt with?
25	novel combination	smashed nut	The bits of shell are all over the place.	It looks like the	smashed nut/won't be/easily dealt with.	Can the nut be easily dealt with?
26	idiom_figurative	high point	Being all together sharing drinks and food,	this is a	high point/and we/will remember it forever.	Do we dislike being together and sharing things?
26	idiom_literal	high point	We will soon be able to see the valley below,	this is a	high point/and we/will remember it forever.	Can we already see the valley?
26	collocation	main point	You should write this down carefully,	this is a	main point/and we/will discuss it afterwards.	Do we refuse to discuss the point?
26	novel combination	large point	We should use a different tool,	this has a	large point/and we/will be more accurate with a smaller one.	Should we use the same tool?
27	idiom_figurative	hot potato	Forcing everyone to take a vaccine,	this is certainly a	hot potato/and we/should be careful with it.	Should we be careful in dealing with vaccination?

27	idiom_literal	hot potato	Watch out and don't use your hands,	this is certainly a	hot potato/and we/should be careful with it.	Should we be careful with the potato?
27	collocation	baked potato	As you can see on his dinner plate,	this is certainly a	baked potato/and we/should get one for ourselves.	Do we want to get one for ourselves?
27	novel combination	fresh potato	As you can tell from its colour and firmness,	this is certainly a	fresh potato/and we/should get one for ourselves.	Do we want to get one for ourselves?
28	idiom_figurative	free ride	As Tom is the boss's son,	Tom will get a	free ride/at some/point.	Is Tom the boss's child?
28	idiom_literal	free ride	As many people drive across this area,	Tom will get a	free ride/at some/point.	Do many people drive across this area?
28	collocation	long ride	As many people drive across this area,	Tom will get a	long ride/at some/point.	Do many people drive across this area?
28	novel combination	bad ride	Countryside lanes are narrow and rugged,	Tom will get a	bad ride/at some/point.	Are countryside lanes rough?
29	idiom_figurative	clean sheet	To reach the men's football final,	we'd better keep a	clean sheet/for our/own good.	Is it good for us to lose the match?
29	idiom_literal	clean sheet	As Lewis is staying over,	we'd better keep a	clean sheet/for our/own good.	Is it good for us to have a dirty sheet?
29	collocation	single sheet	To take less stuff in our backpacks,	we'd better keep a	single sheet/for our/own good.	Is it good for us to take as much stuff as possible?
29	novel combination	light sheet	To take less stuff in our backpacks,	we'd better keep a	light sheet/for our/own good.	Is it good for us to take as much stuff as possible?

30	idiom_figurative	long shot	With gambling the only option left to make quick money,	he took his chance even though it is a	long shot/and he/might fail.	Does he have many options to make money?
30	idiom_literal	long shot	With the football match ending and him standing far from the goal,	he took his chance even though it is a	long shot/and he/might fail.	Did the team win the competition?
30	collocation	single shot	With the shooting competition ending and the team losing,	he took his chance even though it is a	single shot/and he/might fail.	Did the team win the competition?
30	novel combination	hard shot	With many defenders in front of the goal,	he took his chance with the free kick even though it is a	hard shot/and he/might fail.	Did the team win the competition?
31	idiom_figurative	thick skin	Mary was still in a good mood after receiving tons of criticism,	mainly because of the	thick skin/that she/had developed.	Was Mary in a bad mood?
31	idiom_literal	thick skin	Mary was not bruised at all after being hit by a falling stone,	mainly because of the	thick skin/that she/was born with.	Was Mary injured?
31	collocation	dark skin	Mary was in a good mood after getting a main role in the film,	mainly because of the	dark skin/that she/was born with.	Was Mary in a bad mood?
31	novel combination	pretty skin	Mary was in a good mood after getting a main role in the film,	mainly because of the	pretty skin/that she/was born with.	Was Mary in a bad mood?
32	idiom_figurative	sore spot	Kate's divorce was not pleasant	it remained a	sore spot/and she/thought about it often.	Was Kate divorced?
32	idiom_literal	sore spot	Kate's head was hit in a car accident,	it remained a	sore spot/and she/thought about it often.	Was Kate hit by a car?
32	collocation	ideal spot	Kate used to have a glass of wine on the rooftop,	it remained an	ideal spot/and she/thought about it often.	Did Kate like drinking wine?

32	novel combination	hidden spot	Kate secretly used to have a glass of wine on the rooftop,	it remained a	hidden spot/and she/thought about it often.	Did Kate like drinking wine?
33	idiom_figurative	greasy spoon	The couple only ate healthy plant-based food,	so they weren't happy about the	greasy spoon/and they/complained to us about it.	Did the couple complain?
33	idiom_literal	greasy spoon	The parents demanded that dishes should be done perfectly,	so they weren't happy about the	greasy spoon/and they/complained to us about it.	Did the parents complain?
33	collocation	wooden spoon	The couple were enthusiastic about protecting plants,	so they weren't happy about the	wooden spoon/and they/complained to us about it.	Did the couple complain?
33	novel combination	coffee spoon	The couple wanted cutlery to eat their soup,	so they weren't happy about the	coffee spoon/and they/complained to us about it.	Did the couple complain?
34	idiom_figurative	strong stomach	You should not watch this strange movie	unless you have a	strong stomach/like the/rest of us.	Do we recommend you watch the film?
34	idiom_literal	strong stomach	Don't buy meat from this store	unless you have a	strong stomach/like the/rest of us.	Do we recommend you buy meat from the store?
34	collocation	flat stomach	Go and finish the whole set of exercise,	unless you have a	flat stomach/like the/rest of us.	Do we recommend you laze around?
34	novel combination	huge stomach	Do not order the big burger here,	unless you have a	huge stomach/like the/rest of us.	Do we recommend you order the burger?
35	idiom_figurative	red tape	People are fed up with bureaucracy,	there is too much	red tape/and other/unnecessary things.	Are people happy with bureaucracy?

35	idiom_literal	red tape	It is not the recommended way of wrapping up a gift,	there is too much	red tape/and other/unnecessary things.	Is this the recommended way of wrapping a gift?
35	collocation	sticky tape	It is not the recommended way of wrapping up a gift,	there is too much	sticky tape/and other/unnecessary things.	Is this the recommended way of wrapping a gift?
35	novel combination	fabric tape	It is not the recommended way of wrapping up a gift,	there is too much	fabric tape/and other/unnecessary things.	Is this the recommended way of wrapping a gift?
36	idiom_figurative	deep water	The algebra class is too difficult for the twins and	they are now in	deep water;/they will/likely need more help.	Will the twins be fine without help?
36	idiom_literal	deep water	The whales were guided back to the open seas and	they are now in	deep water;/they will/likely need no further help.	Will the whales need more help?
36	collocation	warm water	The whales were guided back to the open seas and	they are now in	warm water;/they will/likely need no further help.	Will the whales need more help?
36	novel combination	mixed water	The whales were guided back to the open seas and	they are now in	mixed water;/they will/likely need no further help.	Will the whales need more help?
37	idiom_figurative	brick wall	As his funding was suddenly cut,	John ran into a	brick wall/all of/a sudden.	Was John's funding continued?
37	idiom_literal	brick wall	As he was careless at the construction site,	John ran into a	brick wall/all of/a sudden.	Did John run into a wall?
37	collocation	stone wall	While he was exploring the unfamiliar town,	John ran into a	stone wall/all of/a sudden.	Did John run into a wall?
37	novel combination	vast wall	While he was exploring the unfamiliar town,	John ran into a	vast wall/all of/a sudden.	Did John run into a wall?

38	idiom_figurative	live wire	Since Lucas is always so active and energetic,	we need to manage the	live wire/to the/best of our abilities.	Should we ignore Lucas?
38	idiom_literal	live wire	Since the machine has short-circuited but is still plugged in,	we need to manage the	live wire/to the/best of our abilities.	Should we ignore the machine?
38	collocation	thin wire	Since the machine has short-circuited but is still plugged in,	we need to manage the	thin wire/to the/best of our abilities.	Should we ignore the machine?
38	novel combination	clear wire	Since the machine has short-circuited but is still plugged in,	we need to manage the	clear wire/to the/best of our abilities.	Should we ignore the machine?
39	idiom_figurative	final word	As she is the boss of the company,	Mary was given the	final word/before they/wrapped up.	Was Mary given the final word?
39	idiom_literal	final word	As the charades game was ending,	Mary was given the	final word/before they/wrapped up	Was Mary given the last word?
39	collocation	single word	As she struggles with spelling,	Mary was given the	single word/before they/wrapped up	Does Mary struggle with spelling?
39	novel combination	normal word	As she struggles with spelling,	Mary was given the	normal word/before they/wrapped up.	Does Mary struggle with spelling?
40	idiom_figurative	dirty work	William is the right-hand man of a mafia boss,	his job involves a lot of	dirty work/as you/might imagine.	Is William out of work?
40	idiom_literal	dirty work	William works for a cleaning company,	his job involves a lot of	dirty work/as you/might imagine.	Is William out of work?
40	collocation	manual work	William works as a self-employed builder,	his job involves a lot of	manual work/as you/might imagine.	Is William out of work?
40	novel combination	clever work	William works for a design studio,	his job involves a lot of	clever work/as you/might imagine.	Is William out of work?

Appendix 4.3

Appendix 4.3.1 Norming meaning dominance

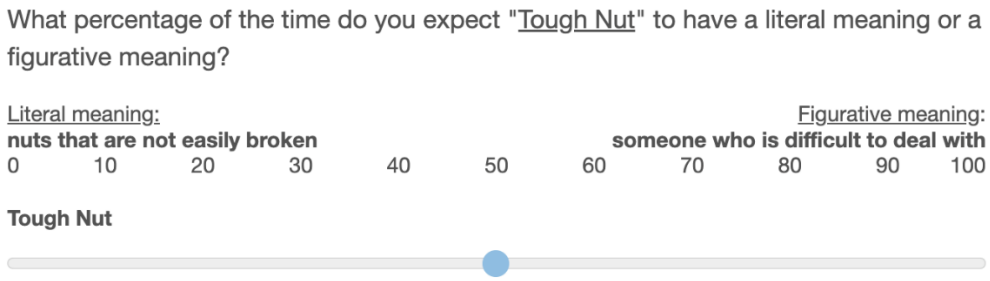
Section 2 - part B:
You are asked to **rate what percentage of the time you expect a phrase to have a literal or figurative meaning.**
Before you start, we ask you to familiarise yourself with **what figurative and literal meanings of a phrase are**.

Phrases sometimes have both literal and figurative meanings. **A literal meaning** refers to taking words in their usual or most basic sense. For example, "on Monday" just means on that day of the week. **A word or phrase** can also be used in a different way from its usual, straightforward meaning, to create a particular mental picture. This less straightforward, more unusual interpretation is known as **a figurative meaning.**

For example: "at the end of the day"
1) Literal usage: "At the end of the day, he went to bed and fell asleep".
Here, "at the end of the day" just means exactly that – once the day is finished.

2) Figurative usage: "At the end of the day, it's his decision". In this sentence, "at the end of the day" does not mean once the day has ended but means "Ultimately, it is his decision".

Example 1 of meaning dominance



Example 2 of meaning dominance

What percentage of the time do you expect "Sore Spot" to have a literal meaning or a figurative meaning?

Literal meaning:
a spot that is painful

Figurative meaning:
something that is likely to make someone upset/angry

0 10 20 30 40 50 60 70 80 90 100

Sore Spot

Appendix 4.3.2 Norming local semantic decomposability

Section 2 - part C:

You are asked to determine to what extent the literal meanings of components of a phrase individually contribute to the figurative meaning of the phrase as expressed in the paraphrase, on a scale of 1-7 with

1 = "Not contribute at all to the figurative meaning"

7 = "100% contribute to the figurative meaning".

The figurative meaning of a phrase is listed above the scale.

Example 1 of local semantic decomposability

To what extent do the literal meanings of components "Sore" and "Spot" individually contribute to the figurative meaning of the phrase Sore Spot on a scale of 1-7 with

1 = "Not contribute at all to the figurative meaning" and

7 = "100% contribute to figurative meaning"?

Sore Spot: something that is likely to make someone upset/angry

	1	2	3	4	5	6	7
Sore	○	○	○	○	○	○	○
Spot	○	○	○	○	○	○	○

Example 2 of local semantic decomposability

To what extent do the literal meanings of components
"Tough" and "Nut" individually contribute to the figurative meaning of the phrase
Tough Nut on a scale of 1-7 with
1 = "Not contribute at all to the figurative meaning" and
7 = "100% contribute to figurative meaning"?

Tough Nut: someone who is difficult to deal with

	1	2	3	4	5	6	7
Tough	○	○	○	○	○	○	○
Nut	○	○	○	○	○	○	○

Appendix 4.3.3 Norming global semantic decomposability

Example 1 of global semantic decomposability

You are asked to rate how the literal meaning of "**Tough Nut**" contributes to its figurative meaning on a scale of 1-7 with
1 = "Not contribute at all to the figurative meaning" and
7 = "100% contribute to figurative meaning"?

Literal meaning:
nuts that are not easily broken

Figurative meaning:
someone who is difficult to deal with

1 2 3 4 5 6 7

Tough Nut

Example 2 of global semantic decomposability

You are asked to rate how the literal meaning of "**Red Tape**" contributes to its figurative meaning on a scale of 1-7 with
1 = "Not contribute at all to the figurative meaning" and
7 = "100% contribute to figurative meaning"?

Literal meaning:
tape that is in red

Figurative meaning:
official rules and processes that seem unnecessary and delay results

1 2 3 4 5 6 7

Red Tape

Appendix 4.4

Examples of multiple choice test to check whether L2 speakers understood idioms in reading experiment

9. Please choose the most likely interpretation of the last two words in the sentence.
Sentence: "The couple only ate healthy plant-based food, so they weren't happy about the **greasy spoon** ..."

1. a small restaurant, especially one that sells a lot of delicious food

2. a spoon that is covered with oil

3. a small, cheap restaurant, especially one that sells a lot of fried food

3. Please choose the most likely interpretation of the last two words in the sentence.
Sentence: "Inequality is an ever-growing problem. It looks like the **tough nut** ..."

1. a nut that is very hard and strong

2. someone who is very strong and persistent

3. someone who is very unpleasant and difficult to deal with

Appendix 4.5

The Best-fitting Models of MWSs on Critical and Spill-over Regions across L1 and L2 Speakers

Critical Region				
Predictors	***β***	***SE***	***t***	***p***
(Intercept)	7.47	0.11	67.83	<0.001
collocation	-0.1	0.03	-3.3	0.001
[idiom-figurative]	-0.1	0.029	-3.62	<0.001
[idiom-literal]	-0.06	0.029	-2.18	0.029
Language-background [L2]	0.15	0.04	3.6	<0.001
word1 frequency	-0.02	0.009	-2.22	0.026
word2 frequency	-0.04	0.008	-4.13	<0.001
trial number	-0.05	0.007	-7.42	<0.001
collocation*Language-background [L2]	0.03	0.025	1.19	0.233
idiom-figurative*Language-background [L2]	0.05	0.026	2.09	0.037
idiom-literal*Language-background [L2]	0.02	0.025	0.72	0.474
Random Effects				
σ^2	0.08			
$\tau_{00\ ID}$	0.06			
$\tau_{00\ items.x}$	0.01			
$\tau_{11\ items.x.Language\text{-}background\ [L2]}$	0			
$\rho_{01\ items.x}$	-0.19			
ICC	0.48			
N_{ID}	204			
$N_{items.x}$	120			
Observations	5831			
Marginal R^2 / Conditional R^2	0.078 / 0.523			

Spill-over Region				
Predictors	***β***	***SE***	***t***	***p***
(Intercept)	6.6	0.03	243.15	**<0.001**
novel word combinations	-0.02	0.02	-0.76	0.445
idiom-figurative	-0.06	0.02	-2.79	**0.005**
idiom-literal	-0.04	0.02	-1.69	0.092
Language-background [L2]	0.01	0.03	0.31	0.754
trial number	-0.06	0.005	-11.87	**<0.001**
novel word combinations* Language-background [L2]	0.02	0.02	0.88	0.378
idiom-figurative*Language-background [L2]	0.05	0.02	2.54	**0.011**
idiom-literal *Language-background [L2]	0.006	0.02	0.27	0.789
Random Effects				
σ^2	0.05			
$\tau_{00\ ID}$	0.04			
$\tau_{00\ items.x}$	0.01			
$\tau_{11\ items.x.Language\text{-}background\ [L2]}$	0			
$\rho_{01\ items.x}$	0.41			
ICC	0.54			
N_{ID}	204			
$N_{items.x}$	120			
Observations	5997			
Marginal R^2 / Conditional R^2	0.042 / 0.558			

Note: reading times were log-transformed, idiom-literal refers to idioms in literal contexts and idiom-figurative refers to idioms in figurative contexts; novel word combinations were coded as 0 for the critical region and idiom-figurative was coded as 0 for the spill-over region; L1 speakers were coded 0

Appendix 4.6

The Best-fitting Models of Idioms on Critical and Spill-over Regions across L1 and L2 Speakers

Critical Region				
Predictors	***β***	***SE***	***t***	***p***
(Intercept)	7.36	0.14	53.29	<.001
trial number	-0.04	0.008	-5.3	<.001*
noun decomposability	-0.07	0.02	-3.68	<.001*
word1 frequency	-0.03	0.009	-3	.003*
word2 frequency	-0.03	0.01	-2.42	.015*
adjective decomposability	-0.05	0.02	-2.15	.032*
meaning dominance	0.001	0.0004	1.25	.212
context [literal context]	0.05	0.02	2.43	.015*
Language-background [L2]	0.18	0.04	4.63	<.001*
adjective decomposability* meaning dominance	0.001	0.0005	2.51	.012*
noun decomposability * context [literal context]	0.05	0.013	4.04	<.001*
adjective decomposability* context [literal context]	0.07	0.02	2.91	.004*
Random Effects				
σ^2	0.08			
$\tau_{00\ ID}$	0.06			
$\tau_{00\ items.x}$	0.01			
$\tau_{11\ items.x.Language\text{-}background\ [L2]}$	0			

Spill-over Region				
Predictors	***β***	***SE***	***t***	***p***
(Intercept)	6.54	0.03	239.72	<.001
trial number	-0.06	0.006	-9.06	<.001*
noun decomposability	-0.03	0.01	-2.31	.021
adjective decomposability	0.001	0.02	0.06	.955
meaning dominance	0.0001	0.0003	0.28	.779
context [literal context]	0.03	0.01	1.83	.067
Language-background [L2]	0.07	0.03	1.97	.048
adjective decomposability* meaning dominance	0.0006	0.0002	2.5	.012
context [literal context]* Language-background [L2]	-0.0501	0.02	-3.19	.001
Random Effects				
σ^2	0.04			
$\tau_{00\ ID}$	0.04			
$\tau_{00\ items.x}$	0			
$\tau_{11\ items.x.Language\text{-}background\ [L2]}$	0			

$\rho_{01\ items.x}$	-0.6	$\rho_{01\ items.x}$	0.86
ICC	0.46	ICC	0.55
N_{ID}	204	N_{ID}	204
$N_{items.x}$	40	$N_{items.x}$	40
Observations	2929	Observations	3025
Marginal R^2 / Conditional R^2	0.098 / 0.51	Marginal R^2 / Conditional R^2	0.051 / 0.570

Note. Reading times were log-transformed; decomposability and dominance were centered; Context has two levels: idiom-literal and idiom-figurative; idiom-literal refers to idioms in literal contexts; idioms in figurative contexts were coded as 0.

Appendix 4.7

Reports of Data Loss for Each Cleaning Step by Group and Condition

Removing sentences judged as implausible					
Group	**Total**	**Condition**			
		free word combination	collocations	idiom-literal	idiom-figurative
L1 speakers	523 trials (15%)	169 trials (4.86%)	155 trials (4.45%)	104 trials (2.99%)	96 trials (2.73%)
L1 speakers	948 trials (20.3%)	259 trials (5.55%)	205 trials (4.39%)	219 trials (4.69%)	265 trials (5.67%)

Removing trials of which comprehension questions answered incorrectly					
Group	**Total**	**Condition**			
		free word combination	collocations	idiom-literal	idiom-figurative
L1 speakers	103 trials (2.96%)	31 trials (0.89%)	18 trials (0.52%)	21 trials (0.6%)	33 trials (0.95%)
L1 speakers	202 trials (4.33%)	50 trials (1.07%)	52 trials (1.11%)	37 trials (0.79%)	63 trials (1.35%)

Removing trials of which the critical region based on the cut-off points					
Group	**Total**	**Condition**			
		free word combination	collocations	idiom-literal	idiom-figurative
L1 speakers	137 trials (3.94%)	45 trials (1.29%)	28 trials (0.81%)	32 trials (0.92%)	32 trials (0.92%)
L1 speakers	102 trials (2.18%)	32 trials (0.69%)	20 trials (0.43%)	21 trials (0.45%)	29 trials (0.62%)

Removing trials of which the spill-over region based on the cut-off points					
Group	**Total**	**Condition**			
		free word combination	collocations	idiom-literal	idiom-figurative
L1 speakers	66 trials (1.89%)	20 trials (0.58%)	17 trials (0.49%)	17 trials (0.49%)	12 trials (0.35%)
L1 speakers	96 trials (2.06%)	26 trials (0.56%)	24 trials (0.51%)	30 trials (0.64%)	16 trials (0.34%)

Trial removal on the critical region based on 2.5 SD					
Group	**Total**	**Condition**			
		free word combination	collocations	idiom-literal	idiom-figurative
L1 speakers	10 trials (0.29%)	1 trial (0.03%)	1 trial (0.03%)	6 trials (0.17%)	2 trials (0.06%)
L1 speakers	16 trials (0.34%)	2 trials (0.04%)	8 trials (0.17%)	4 trials (0.09%)	2 trials (0.04%)

Trial removal on the spill-over region based on 2.5 SD					
Group	**Total**	**Condition**			
		free word combination	collocations	idiom-literal	idiom-figurative
L1 speakers	14 trials (0.4%)	4 trials (0.11%)	1 trial (0.03%)	4 trials (0.11%)	5 trials (0.14%)
L1 speakers	13 trials (0.28%)	2 trials (0.04%)	4 trials (0.08%)	4 trials (0.08%)	3 trials (0.06%)

Appendix 5.1

Summary of Language Proficiency Data

	L1 speakers (*M*, *SD*)	**L2 speakers (*M*, *SD*)**
Age of English acquisition	1.08 (0.47)	7.42 (3.08)
Time lived in an English-speaking country (in years)	20.2 (5.3)	2.5 (2.5)
LexTALE score	91.96 (6.75)	66.21 (11.72)
VLT score	84.1 (4.49)	53.82 (17.82)
Self-reporting of English proficiency (a 6-point scale)		
Proficiency in writing	4.90 (0.37)	3.16 (0.72)
Proficiency in speaking	4.95 (0.31)	3.13 (0.96)
Proficiency in listening	4.92 (0.35)	3.45 (0.89)
Proficiency in reading	4.88 (0.40)	3.53 (0.60)
Self-reporting of time using English (per day)		
Writing (in hours)	5.5 (4.96)	2.45 (1.94)
Speaking (in hours)	9.75 (4.53)	1.71 (1.97)
Watching/listening (in hours)	5.2 (5.07)	2.57 (1.82)
Reading (in hours)	4.59 (4.90)	2.54 (1.90)

Appendix 5.2

Summary of Items Characteristics

Items characteristics	**Collocations (*M*, *SD*)**	**Novel Word Combinations (*M, SD*)**
Verb length	6.2 (1.4)	6.25 (1.4)
Noun length	7.6 (1.7)	7.7 (1.7)
Item length	20.8 (2.4)	20.9 (2.6)
Verb past form frequency	8254.9 (7049.1)	8104.1 (7897.9)
Verb root frequency	6734.9 (6517.4)	7022.3 (7048.3)
Verb lemma frequency	19604.9 (19401.8)	19475.2 (20305.6)
Noun form frequency	8615.3 (5279.3)	7984.2 (5567.1)
Noun lemma frequency	12249.9 (8596.4)	12481.7 (11388.8)
Phrasal past form frequency	59.5 (88.4)	5.3 (2.5)
Phrasal root form frequency	93.9 (111.1)	6.1 (4.4)
Phrasal lemma frequency	230.4 (298.0)	19.5 (21.4)
MI (past form)	3.51 (1.62)	1.43 (1.44)
MI (root form)	4.50 (1.96)	1.47 (1.13)
MI (lemma)	3.75 (3.24)	0.72 (2.41)
Predictability (active)	0.006 (.019)	0.001 (.025)
Predictability (passive)	0.011 (.025)	0.006 (.023)

Note. Length is in characters. Frequency information was taken from BNC with Sketch Engine and all frequency measures are expressed raw frequency.

Appendix 5.3

L1L2 speakers: Summary statistics for the Collocation Region

Factor	Statistics	FRRT	RPRT	SRPRT	TRT	TFC
Intercept [global mean]	β	838.45	993.58	954.83	1409.66	1.67
	SE	25.9	34.96	33.07	73.32	0.039
	t/z	32.34	28.42	28.87	19.25	42.98
Collocation [collocation]	β	0.29	-6.14	-8.97	-53.03	-0.04
	SE	8.93	9.52	8.66	16.18	0.01
	t/z	0.03	-0.65	-1.04	-3.28*	-3.51*
Voice [active]	β	23.59	4.8	10.67	-17.49	-0.01
	SE	9.94	10.01	9.28	17.77	0.01
	t/z	2.37*	0.48	1.15	-0.98	-0.62
Language-background [L2]	β	239.6	293.08	284.55	473.8	0.24
	SE	24.25	32.57	31.6	68.78	0.04
	t/z	9.88*	9.00*	9.00*	6.89*	6.55*

Note. FRRT: first run reading time, RPRT: regression path reading time, SRPRT: selective regression path reading time, TRT: total reading time, TFC: total fixation count.
* = $p < .05$.

Appendix 5.4

L1L2 speakers: Summary statistics for the Final Word Region

Factor	Statistics	FFRT	FRRT	RPRT	SRPRT	TRT	TFC
		Final words: nouns					
Intercept [global mean]	β	253.54	322.63	445.54	372.24	552.73	0.74
	SE	4.08	7.93	16.72	10.22	27.28	0.04
	t/z	62.1	40.69	26.64	36.44	20.26	20.55
Collocation [collocation]	β	-0.25	-9.95	-15.03	-8.19	-19.07	-0.03
	SE	2.35	4.2	8.44	4.95	12.04	0.02
	t/z	-0.1	-2.37*	-1.78	-1.65	-1.58	-1.81
Language-background [L2]	β	34.28	70.88	114.02	89.79	187.82	0.23
	SE	3.86	7.35	15.74	9.57	25.68	0.03
	t/z	8.88*	9.64*	7.24*	9.38*	7.31*	6.55*
		Final words: verbs					
Intercept [global mean]	β	249.37	341.15	426.74	384.53	601.46	0.8
	SE	4.42	9.39	13.92	11.43	33.12	0.04
	t/z	56.46	36.33	30.64	33.63	18.15	20.25
Collocation [collocation]	β	-0.21	-4.83	-6.09	-6.19	-29.03	-0.04
	SE	2.45	5.36	8.42	6.28	14.50	0.02
	t/z	-0.09	-0.9	-0.72	-0.99	-2.00*	-2.02*
Language-background [L2]	β	29.08	83.25	103.31	100.36	214.13	0.23
	SE	4.18	8.64	12.6	10.43	30.92	0.04
	t/z	6.96*	9.63*	8.2*	9.62*	6.93*	6.19*

Note. FFRT: first fixation reading time, FRRT: first run reading time, RPRT: regression path reading time, SRPRT: selective regression path reading time, TRT: total reading time, TFC: total fixation count.
* = $p < .05$.

Appendix 5.5

Chapter 5 Experimental items and sentences

set	condition	item	sentence	question
1	canonical collocation	they promoted development	After World War II,*they promoted development*in the*inner city.	NO_QUESTION
1	canonical novel word combinations	they controlled development	After World War II,*they controlled development*in the*inner city.	NO_QUESTION
1	reversed collocation	development was promoted	After World War II,*development was promoted*in the*inner city.	NO_QUESTION
1	reversed novel word combinations	development was controlled	After World War II,*development was controlled*in the*inner city.	NO_QUESTION
2	canonical collocation	he improved quality	We are pleased to hear that*he improved quality*to increase*sales.	NO_QUESTION
2	canonical novel word combinations	he discussed quality	We are pleased to hear that*he discussed quality*to increase*sales.	NO_QUESTION
2	reversed collocation	quality was improved	We are pleased to hear that*quality was improved*to increase*sales.	NO_QUESTION
2	reversed novel word combinations	quality was discussed	We are pleased to hear that*quality was discussed*to increase*sales.	NO_QUESTION
3	canonical collocation	they allowed access	Everyone thinks that*they allowed access*to the*building but the reality is the opposite.	Is the reality different from what everyone expects?
3	canonical novel word combinations	they offered access	Everyone thinks that*they offered access*to the*building but the reality is the opposite.	Is the reality different from what everyone expects?
3	reversed collocation	access was allowed	Everyone thinks that*access was allowed*to the*building but the reality is the opposite.	Is the reality different from what everyone expects?
3	reversed novel word combinations	access was offered	Everyone thinks that*access was offered*to the*building but the reality is the opposite.	Is the reality different from what everyone expects?
4	canonical collocation	they confirmed findings	After a few months,*they confirmed findings*from*the previous studies.	NO_QUESTION

4	canonical novel word combinations	they addressed findings	After a few months,*they addressed findings*from*the previous studies.	NO_QUESTION
4	reversed collocation	findings were confirmed	After a few months,*findings were confirmed*from*the previous studies.	NO_QUESTION
4	reversed novel word combinations	findings were addressed	After a few months,*findings were addressed*from*the previous studies.	NO_QUESTION
5	canonical collocation	she received treatment	Even though she was very busy,*she received treatment*every*few weeks.	Did she have the treatment?
5	canonical novel word combinations	she arranged treatment	Even though she was very busy,*she arranged treatment*every*few weeks.	Did she have the treatment?
5	reversed collocation	treatment was received	Even though she was very busy,*treatment was received*every*few weeks.	Did she have the treatment?
5	reversed novel word combinations	treatment was arranged	Even though she was very busy,*treatment was arranged*every*few weeks.	Did she have the treatment?
6	canonical collocation	they caused trouble	The woman said that*they caused trouble*in this*neighbourhood.	NO_QUESTION
6	canonical novel word combinations	they stopped trouble	The woman said that*they stopped trouble*in this*neighbourhood.	NO_QUESTION
6	reversed collocation	trouble was caused	The woman said that*trouble was caused*in this*neighbourhood.	NO_QUESTION
6	reversed novel word combinations	trouble was stopped	The woman said that*trouble was stopped*in this*neighbourhood.	NO_QUESTION
7	canonical collocation	they provided services	The council announced*they provided services*in this*area for poor people.	Did the council only care about the wealthy?
7	canonical novel word combinations	they considered services	The council announced*they considered services*in this*area for poor people.	Did the council only care about the wealthy?
7	reversed collocation	services were provided	The council announced*services were provided*in this*area for poor people.	Did the council only care about the wealthy?
7	reversed novel word combinations	services were considered	The council announced*services were considered*in this*area for poor people.	Did the council only care about the wealthy?
8	canonical collocation	she achieved results	As Jane's partner expected,*she achieved results*in the*workplace.	NO_QUESTION

8	canonical novel word combinations	she ignored results	As Jane's partner expected,*she ignored results*in the*workplace.	NO_QUESTION
8	reversed collocation	results were achieved	As Jane's partner expected,*results were achieved*in the*workplace.	NO_QUESTION
8	reversed novel word combinations	results were ignored	As Jane's partner expected,*results were ignored*in the*workplace.	NO_QUESTION
9	canonical collocation	they created conditions	The head teacher believed*they created conditions*to improve*the learning environment.	Did the head teacher try to improve the learning environment?
9	canonical novel word combinations	they produced conditions	The head teacher believed*they produced conditions*to improve*the learning environment.	Did the head teacher try to improve the learning environment?
9	reversed collocation	conditions were created	The head teacher believed*conditions were created*to improve*the learning environment.	Did the head teacher try to improve the learning environment?
9	reversed novel word combinations	conditions were produced	The head teacher believed*conditions were produced*to improve*the learning environment.	Did the head teacher try to improve the learning environment?
10	canonical collocation	they adopted policies	It has been argued that*they adopted policies*to control*inflation.	NO_QUESTION
10	canonical novel word combinations	they studied policies	It has been argued that*they studied policies*to control*inflation.	NO_QUESTION
10	reversed collocation	policies were adopted	It has been argued that*policies were adopted*to control*inflation.	NO_QUESTION
10	reversed novel word combinations	policies were studied	It has been argued that*policies were studied*to control*inflation.	NO_QUESTION
11	canonical collocation	he gained experience	The document showed that*he gained experience*in the*management of big projects.	NO_QUESTION
11	canonical novel word combinations	he expected experience	The document showed that*he expected experience*in the*management of big projects.	NO_QUESTION
11	reversed collocation	experience was gained	The document showed that*experience was gained*in the*management of big projects.	NO_QUESTION
11	reversed novel word combinations	experience was expected	The document showed that*experience was expected*in the*management of big projects.	NO_QUESTION
12	canonical collocation	they performed duties	The workers said that*they performed duties*which*were outlined in the contract.	Did workers break the contract?

12	canonical novel word combinations	they completed duties	The workers said that*they completed duties*which*were outlined in the contract.	Did workers break the contract?
12	reversed collocation	duties were performed	The workers said that*duties were performed*which*were outlined in the contract.	Did workers break the contract?
12	reversed novel word combinations	duties were completed	The workers said that*duties were completed*which*were outlined in the contract.	Did workers break the contract?
13	canonical collocation	he lost confidence	Tom told me that*he lost confidence*during*the difficult interview.	Did Tim have an easy interview?
13	canonical novel word combinations	he kept confidence	Tom told me that*he kept confidence*during*the difficult interview.	Did Tim have an easy interview?
13	reversed collocation	confidence was lost	Tom told me that*confidence was lost*during*the difficult interview.	Did Tim have an easy interview?
13	reversed novel word combinations	confidence was kept	Tom told me that*confidence was kept*during*the difficult interview.	Did Tim have an easy interview?
14	canonical collocation	they developed software	The annual report showed that*they developed software*for business*management by the end of 2015.	NO_QUESTION
14	canonical novel word combinations	they presented software	The annual report showed that*they presented software*for business*management by the end of 2015.	NO_QUESTION
14	reversed collocation	software was developed	The annual report showed that*software was developed*for business*management by the end of 2015.	NO_QUESTION
14	reversed novel word combinations	software was presented	The annual report showed that*software was presented*for business*management by the end of 2015.	NO_QUESTION
15	canonical collocation	she influenced behaviour	Kelly claimed that*she influenced behaviour*at work,*so that her workplace became more pleasant.	Did Kelly's workplace become less pleasant to work in?
15	canonical novel word combinations	she recommended behaviour	Kelly claimed that*she recommended behaviour*at work,*so that her workplace became more pleasant.	Did Kelly's workplace become less pleasant to work in?
15	reversed collocation	behaviour was influenced	Kelly claimed that*behaviour was influenced*at work,*so that her workplace became more pleasant.	Did Kelly's workplace become less pleasant to work in?
15	reversed novel word combinations	behaviour was recommended	Kelly claimed that*behaviour was recommended*at work,*so that her workplace became more pleasant.	Did Kelly's workplace become less pleasant to work in?

16	canonical collocation	they spread news	I was disappointed to hear that*they spread news*before*the information was confirmed.	NO_QUESTION
16	canonical novel word combinations	they accepted news	I was disappointed to hear that*they accepted news*before*the information was confirmed.	NO_QUESTION
16	reversed collocation	news was spread	I was disappointed to hear that*news was spread*before*the information was confirmed.	NO_QUESTION
16	reversed novel word combinations	news was accepted	I was disappointed to hear that*news was accepted*before*the information was confirmed.	NO_QUESTION
17	canonical collocation	she expressed concern	In the interview,*she expressed concern*about*the changing situation.	Did the changing situation raise concerns?
17	canonical novel word combinations	she revealed concern	In the interview,*she revealed concern*about*the changing situation.	Did the changing situation raise concerns?
17	reversed collocation	concern was expressed	In the interview,*concern was expressed*about*the changing situation.	Did the changing situation raise concerns?
17	reversed novel word combinations	concern was revealed	In the interview,*concern was revealed*about*the changing situation.	Did the changing situation raise concerns?
18	canonical collocation	they saved energy	I've heard that*they saved energy*more*effectively than before because they cared about the environment.	NO_QUESTION
18	canonical novel word combinations	they sold energy	I've heard that*they sold energy*more*effectively than before because they wanted to earn more money.	NO_QUESTION
18	reversed collocation	energy was saved	I've heard that*energy was saved*more*effectively than before because they cared about the environment.	NO_QUESTION
18	reversed novel word combinations	energy was sold	I've heard that*energy was sold*more*effectively than before because they wanted to earn more money.	NO_QUESTION
19	canonical collocation	he offered advice	I'm sure that*he offered advice*on all*aspects of the financial situation.	NO_QUESTION
19	canonical novel word combinations	he expected advice	I'm sure that*he expected advice*on all*aspects of the financial situation.	NO_QUESTION
19	reversed collocation	advice was offered	I'm sure that*advice was offered*on all*aspects of the financial situation.	NO_QUESTION
19	reversed novel word combinations	advice was expected	I'm sure that*advice was expected*on all*aspects of the financial situation.	NO_QUESTION

20	canonical collocation	they raised cash	Everyone was happy that*they raised cash*for the*environment protection fund.	Did the environment protection fund receive money?
20	canonical novel word combinations	they left cash	Everyone was happy that*they left cash*for the*environment protection fund.	Did the environment protection fund receive money?
20	reversed collocation	cash was raised	Everyone was happy that*cash was raised*for the*environment protection fund.	Did the environment protection fund receive money?
20	reversed novel word combinations	cash was left	Everyone was happy that*cash was left*for the*environment protection fund.	Did the environment protection fund receive money?
21	canonical collocation	they required effort	Students knew full well that*they required effort*to complete*a degree in rocket science.	Is it hard to complete a degree in rocket science?
21	canonical novel word combinations	they needed effort	Students knew full well that*they needed effort*to complete*a degree in rocket science.	Is it hard to complete a degree in rocket science?
21	reversed collocation	effort was required	Students knew full well that*effort was required*to complete*a degree in rocket science.	Is it hard to complete a degree in rocket science?
21	reversed novel word combinations	effort was needed	Students knew full well that*effort was needed*to complete*a degree in rocket science.	Is it hard to complete a degree in rocket science?
22	canonical collocation	they satisfied demand	The city council worked hard and*they satisfied demand*for new*primary schools.	Did the city council ignore the needs of primary schools?
22	canonical novel word combinations	they adjusted demand	The city council worked hard and*they adjusted demand*for new*primary schools.	Did the city council ignore the needs of primary schools?
22	reversed collocation	demand was satisfied	The city council worked hard and*demand was satisfied*for new*primary schools.	Did the city council ignore the needs of primary schools?
22	reversed novel word combinations	demand was adjusted	The city council worked hard and*demand was adjusted*for new*primary schools.	Did the city council ignore the needs of primary schools?
23	canonical collocation	they attracted investment	To boost local economy,*they attracted investment*to create*more jobs.	Were more jobs created to boost local economy?
23	canonical novel word combinations	they obtained investment	To boost local economy,*they obtained investment*to create*more jobs.	Were more jobs created to boost local economy?
23	reversed collocation	investment was attracted	To boost local economy,*investment was attracted*to create*more jobs.	Were more jobs created to boost local economy?
23	reversed novel word combinations	investment was obtained	To boost local economy,*investment was obtained*to create*more jobs.	Were more jobs created to boost local economy?

24	canonical collocation	they attracted attention	After a short while,*they attracted attention*to the*new recruiters.	NO_QUESTION
24	canonical novel word combinations	they shifted attention	After a short while,*they shifted attention*to the*new recruiters.	NO_QUESTION
24	reversed collocation	attention was attracted	After a short while,*attention was attracted*to the*new recruiters.	NO_QUESTION
24	reversed novel word combinations	attention was shifted	After a short while,*attention was shifted*to the*new recruiters.	NO_QUESTION
25	canonical collocation	they reduced pressure	In the end,*they reduced pressure*from*the principal at the school meetings.	NO_QUESTION
25	canonical novel word combinations	they accepted pressure	In the end,*they accepted pressure*from*the principal at the school meetings.	NO_QUESTION
25	reversed collocation	pressure was reduced	In the end,*pressure was reduced*from*the principal at the school meetings.	NO_QUESTION
25	reversed novel word combinations	pressure was accepted	In the end,*pressure was accepted*from*the principal at the school meetings.	NO_QUESTION
26	canonical collocation	he expressed opinions	It was not that surprising that*he expressed opinions*when*everyone was present.	NO_QUESTION
26	canonical novel word combinations	he expressed solutions	It was not that surprising that*he expressed solutions*when*everyone was present.	NO_QUESTION
26	reversed collocation	opinions were expressed	It was not that surprising that*opinions were expressed*when*everyone was present.	NO_QUESTION
26	reversed novel word combinations	solutions were expressed	It was not that surprising that*solutions were expressed*when*everyone was present.	NO_QUESTION
27	canonical collocation	it required planning	Given the state the project was in,*it required planning*beyond*what was already forecasted.	Could we complete the project with the original plan?
27	canonical novel word combinations	it required material	Given the state the project was in,*it required material*beyond*what was already forecasted.	Could we complete the project with the original plan?
27	reversed collocation	planning was required	Given the state the project was in,*planning was required*beyond*what was already forecasted.	Could we complete the project with the original plan?
27	reversed novel word combinations	material was required	Given the state the project was in,*material was required*beyond*what was already forecasted.	Could we complete the project with the original plan?

28	canonical collocation	they brought peace	Margaret is not sure that*they brought peace*to the*village.	Does Margaret know everything in the village?
28	canonical novel word combinations	they brought meat	Margaret is not sure that*they brought meat*to the*village.	Does Margaret know everything in the village?
28	reversed collocation	peace was brought	Margaret is not sure that*peace was brought*to the*village.	Does Margaret know everything in the village?
28	reversed novel word combinations	meat was brought	Margaret is not sure that*meat was brought*to the*village.	Does Margaret know everything in the village?
29	canonical collocation	he received payment	John is happy to hear that*he received payment*without*any more problems.	Is John facing a financial issue?
29	canonical novel word combinations	he received finance	John is happy to hear that*he received finance*without*any more problems.	Is John facing a financial issue?
29	reversed collocation	payment was received	John is happy to hear that*payment was received*without*any more problems.	Is John facing a financial issue?
29	reversed novel word combinations	finance was received	John is happy to hear that*finance was received*without*any more problems.	Is John facing a financial issue?
30	canonical collocation	they wasted energy	The local newspaper reporter wrote*they wasted energy*in an*irresponsible way.	NO_QUESTION
30	canonical novel word combinations	they wasted labour	The local newspaper reporter wrote*they wasted labour*in an*irresponsible way.	NO_QUESTION
30	reversed collocation	energy was wasted	The local newspaper reporter wrote*energy was wasted*in an*irresponsible way.	NO_QUESTION
30	reversed novel word combinations	labour was wasted	The local newspaper reporter wrote*labour was wasted*in an*irresponsible way.	NO_QUESTION
31	canonical collocation	they resolved disputes	While it was a difficult task,*they resolved disputes*amicably*in the end.	NO_QUESTION
31	canonical novel word combinations	they resolved mistakes	While it was a difficult task,*they resolved mistakes*amicably*in the end.	NO_QUESTION
31	reversed collocation	disputes were resolved	While it was a difficult task,*disputes were resolved*amicably*in the end.	NO_QUESTION
31	reversed novel word combinations	mistakes were resolved	While it was a difficult task,*mistakes were resolved*amicably*in the end.	NO_QUESTION

32	canonical collocation	they ensured success	You don't need to worry about it, as*they ensured success*after*a lot of hard work.	NO_QUESTION
32	canonical novel word combinations	they ensured clarity	You don't need to worry about it, as*they ensured clarity*after*a lot of hard work.	NO_QUESTION
32	reversed collocation	success was ensured	You don't need to worry about it, as*success was ensured*after*a lot of hard work.	NO_QUESTION
32	reversed novel word combinations	clarity was ensured	You don't need to worry about it, as*clarity was ensured*after*a lot of hard work.	NO_QUESTION
33	canonical collocation	she avoided confusion	Although it was not easy,*she avoided confusion*as much*as she could.	NO_QUESTION
33	canonical novel word combinations	she avoided judgement	Although it was not easy,*she avoided judgement*as much*as she could.	NO_QUESTION
33	reversed collocation	confusion was avoided	Although it was not easy,*confusion was avoided*as much*as she could.	NO_QUESTION
33	reversed novel word combinations	judgement was avoided	Although it was not easy,*judgement was avoided*as much*as she could.	NO_QUESTION
34	canonical collocation	we protected workers	After the accident happened,*we protected workers*as much*as possible.	Were actions taken after the accident?
34	canonical novel word combinations	we protected tourists	After the accident happened,*we protected tourists*as much*as possible.	Were actions taken after the accident?
34	reversed collocation	workers were protected	After the accident happened,*workers were protected*as much*as possible.	Were actions taken after the accident?
34	reversed novel word combinations	tourists were protected	After the accident happened,*tourists were protected*as much*as possible.	Were actions taken after the accident?
35	canonical collocation	it influenced behaviour	It is still not clear why*it influenced behaviour*and we*need more details.	Do we have a clear idea?
35	canonical novel word combinations	it influenced planning	It is still not clear why*it influenced planning*and we*need more details.	Do we have a clear idea?
35	reversed collocation	behaviour was influenced	It is still not clear why*behaviour was influenced*and we*need more details.	Do we have a clear idea?
35	reversed novel word combinations	planning was influenced	It is still not clear why*planning was influenced*and we*need more details.	Do we have a clear idea?

36	canonical collocation	he increased efficiency	It is hard to believe that*he increased efficiency*by making*a small change.	Did a small change make a difference?
36	canonical novel word combinations	he increased discipline	It is hard to believe that*he increased discipline*by making*a small change.	Did a small change make a difference?
36	reversed collocation	efficiency was increased	It is hard to believe that*efficiency was increased*by making*a small change.	Did a small change make a difference?
36	reversed novel word combinations	discipline was increased	It is hard to believe that*discipline was increased*by making*a small change.	Did a small change make a difference?
37	canonical collocation	he corrected errors	Lisa was upset to hear that*he corrected errors*without*even asking her.	NO_QUESTION
37	canonical novel word combinations	he corrected photos	Lisa was upset to hear that*he corrected photos*without*even asking her.	NO_QUESTION
37	reversed collocation	errors were corrected	Lisa was upset to hear that*errors were corrected*without*even asking her.	NO_QUESTION
37	reversed novel word combinations	photos were corrected	Lisa was upset to hear that*photos were corrected*without*even asking her.	NO_QUESTION
38	canonical collocation	they carried passengers	I did not know that*they carried passengers*even*though some people objected to it.	NO_QUESTION
38	canonical novel word combinations	they carried cigarettes	I did not know that*they carried cigarettes*even*though some people objected to it.	NO_QUESTION
38	reversed collocation	passengers were carried	I did not know that*passengers were carried*even*though some people objected to it.	NO_QUESTION
38	reversed novel word combinations	cigarettes were carried	I did not know that*cigarettes were carried*even*though some people objected to it.	NO_QUESTION
39	canonical collocation	they sought fortune	I am not at all surprised that*they sought fortune*no matter*what.	NO_QUESTION
39	canonical novel word combinations	they sought victory	I am not at all surprised that*they sought victory*no matter*what.	NO_QUESTION
39	reversed collocation	fortune was sought	I am not at all surprised that*fortune was sought*no matter*what.	NO_QUESTION
39	reversed novel word combinations	victory was sought	I am not at all surprised that*victory was sought*no matter*what.	NO_QUESTION

40	canonical collocation	they attracted visitors	It was fun to see how easily*they attracted visitors*to the*quirky house.	Did people like the house?
40	canonical novel word combinations	they attracted relatives	It was fun to see how easily*they attracted relatives*to the*quirky house.	Did people like the house?
40	reversed collocation	visitors were attracted	It was fun to see how easily*visitors were attracted*to the*quirky house.	Did people like the house?
40	reversed novel word combinations	relatives were attracted	It was fun to see how easily*relatives were attracted*to the*quirky house.	Did people like the house?
41	canonical collocation	he had prevented damage	Anonymous sources told us that*he had prevented damage*in the*last minute.	Did we know the sources of the information?
41	canonical novel word combinations	he had prevented murder	Anonymous sources told us that*he had prevented murder*in the*last minute.	Did we know the sources of the information?
41	reversed collocation	damage had been prevented	Anonymous sources told us that*damage had been prevented*in the*last minute.	Did we know the sources of the information?
41	reversed novel word combinations	murder had been prevented	Anonymous sources told us that*murder had been prevented*in the*last minute.	Did we know the sources of the information?
42	canonical collocation	she had written letters	I noticed that*she had written letters*with*the pen I gave her as her birthday gift.	Did I give her a pen as the birthday gift?
42	canonical novel word combinations	she had written numbers	I noticed that*she had written numbers*with*the pen I gave her as her birthday gift.	Did I give her a pen as the birthday gift?
42	reversed collocation	letters had been written	I noticed that*letters had been written*with*the pen I gave her as her birthday gift.	Did I give her a pen as the birthday gift?
42	reversed novel word combinations	numbers had been written	I noticed that*numbers had been written*with*the pen I gave her as her birthday gift.	Did I give her a pen as the birthday gift?
43	canonical collocation	we introduced measures	It was a rumour that*we introduced measures*as soon*as possible.	NO_QUESTION
43	canonical novel word combinations	we introduced pictures	It was a rumour that*we introduced pictures*as soon*as possible.	NO_QUESTION
43	reversed collocation	measures were introduced	It was a rumour that*measures were introduced*as soon*as possible.	NO_QUESTION

43	reversed novel word combinations	pictures were introduced	It was a rumour that*pictures were introduced*as soon*as possible.	NO_QUESTION
44	canonical collocation	she has given priority	We are puzzled to hear that*she has given priority*to people*whom she just met.	NO_QUESTION
44	canonical novel word combinations	she has given membership	We are puzzled to hear that*she has given membership*to people*whom she just met.	NO_QUESTION
44	reversed collocation	priority has been given	We are puzzled to hear that*priority has been given*to people*whom she just met.	NO_QUESTION
44	reversed novel word combinations	membership has been given	We are puzzled to hear that*membership has been given*to people*whom she just met.	NO_QUESTION
45	canonical collocation	she heard sounds	She complained that*she heard sounds*even*from the next room.	Could she hear things from the next room?
45	canonical novel word combinations	she heard records	She complained that*she heard records*even*from the next room.	Could she hear things from the next room?
45	reversed collocation	sounds were heard	She complained that*sounds were heard*even*from the next room.	Could she hear things from the next room?
45	reversed novel word combinations	records were heard	She complained that*records were heard*even*from the next room.	Could she hear things from the next room?
46	canonical collocation	it encouraged investment	Most people thought*it encouraged investment*but nothing*really happened.	Did most people think nothing would happen?
46	canonical novel word combinations	it encouraged opposition	Most people thought*it encouraged opposition*but nothing*really happened.	Did most people think nothing would happen?
46	reversed collocation	investment was encouraged	Most people thought*investment was encouraged*but nothing*really happened.	Did most people think nothing would happen?
46	reversed novel word combinations	opposition was encouraged	Most people thought*opposition was encouraged*but nothing*really happened.	Did most people think nothing would happen?
47	canonical collocation	he reversed direction	It has become clear that*he reversed direction*in the*end.	NO_QUESTION
47	canonical novel word combinations	he reversed attitude	It has become clear that*he reversed attitude*in the*end.	NO_QUESTION
47	reversed collocation	direction was reversed	It has become clear that*direction was reversed*in the*end.	NO_QUESTION

47	reversed novel word combinations	attitude was reversed	It has become clear that*attitude was reversed*in the*end.	NO_QUESTION
48	canonical collocation	she accepted defeat	Her parents were not surprised by the fact that*she accepted defeat*in the*end.	Were her parents shocked?
48	canonical novel word combinations	she accepted delay	Her parents were not surprised by the fact that*she accepted delay*in the*end.	Were her parents shocked?
48	reversed collocation	defeat was accepted	Her parents were not surprised by the fact that*defeat was accepted*in the*end.	Were her parents shocked?
48	reversed novel word combinations	delay was accepted	Her parents were not surprised by the fact that*delay was accepted*in the*end.	Were her parents shocked?