

**APPLYING COGNITIVE LINGUISTICS TO SECOND LANGUAGE
IDIOM LEARNING**

BY

XINQING WANG

A thesis

submitted to Victoria University of Wellington
in fulfilment of the requirements for the degree of
Doctor of Philosophy

Victoria University of Wellington

2020

ABSTRACT

Idioms are known to cause great difficulty for second language (L2) learners, who may understand the literal meanings of the constituent words of idioms like *(be) waiting in the wings*, but often fail to interpret the idiomatic, figurative meaning of the expression. Proponents of Cognitive Linguistics (CL) claim that CL provides a pathway to more systematic and insightful learning of figurative expressions like idioms. They advocate that learners should be informed of the literal underpinning of idiomatic expressions and their relationship to the figurative meaning. This is supported by the results of several experimental studies employing ‘etymological elaboration’. However, little is known about how learners actually experience the CL-style explanations, or about how the learning is affected by other factors such as learners’ perceived transparency of the connection between the literal underpinnings and the idiomatic meanings, and their L1. The research reported in this thesis therefore (1) investigates the effectiveness of etymological elaboration in facilitating idiom comprehension and retention; (2) examines the problems that L2 learners, i.e., native-Chinese EFL learners in this study, experience when they encounter English figurative idioms, and identifies the factors influencing success in learning the meanings of idioms.

To achieve these objectives, a mixed methods design was employed. Etymological elaboration was implemented in a teaching experiment involving one-on-one interviews, in which 25 Chinese learners of English were presented with idioms whose meaning they were asked to guess first without and then with the aid of information about their literal underpinnings. After the correct figurative meaning was established, participants rated the transparency of the connection between the literal underpinning and the figurative meaning. One week later, the learners were presented with the same idioms and asked to recall their meaning. Follow-up interviews were also conducted to survey the learners’ experience with and awareness of idioms, and their general attitudes and strategies towards idiom learning. Participants’ responses and their recall of idiomatic meanings were scored by three raters. A combination of quantitative and qualitative analyses of the interview data investigated the learning process and the outcomes of the teaching experiment.

The major findings are: (1) Etymological elaboration can facilitate the interpretation and meaning retention of L2 idioms to a substantial degree; and the L2 idiom learning involves the interplay of multiple factors, including the transparency of the idioms, L1 transfer and cross-cultural differences, learners' prior L2 lexical knowledge, and their proficiency levels. (2) The degree of transparency of the literal-figurative connection influences meaning retention, especially for the low proficiency learners. However, the mnemonic effect is not confined to idioms that learners find most transparent, but also affects those that are "far-fetched". (3) The accuracy of meaning inference during the learning phase has a significant impact on memory for the idioms; many errors can be traced back to wrong guesses made in the prior learning phase, and some relate to false equivalents and partial equivalents in the L1. This suggests that trial-and-error learning potentially induces wrong memory traces and that teaching practices should therefore promote more accurate comprehension from the start, in order to facilitate better long-term memory for idioms. (4) More exposure to and better awareness of idioms help EFL learners foster positive attitudes towards idiom learning, which may facilitate the integration and automatization of figurative multiword expressions like idioms in their bilingual mental lexicon, and in turn lead to higher L2 proficiency. The findings of this study have implications for second language teaching and learning. The innovative research design and advanced statistical analyses contribute to the development of language teaching research methodology.

ACKNOWLEDGEMENTS

This PhD journey has been one of the greatest adventures I have ever undertaken. Even though this thesis documents my own research exploration during the journey, it could not have been completed without the support of many people, who facilitated my progress in their own different ways.

First of all, I wish to express my deepest gratitude to my supervisors Professor Frank Boers, Professor Paul Warren, and Associate Professor Peter Gu for their guidance and support. Frank and his work have been my source of inspiration since I embarked on this journey. He guided and motivated me through every stage of my study, even when he was far across the globe. Paul joined my supervision team when this project was under way. He taught and supported me through critical phases of my study. His timely and always thorough feedback on my writing and pertinent input was absolutely crucial for the progress of my thesis. My secondary supervisor Peter Gu was always available when I needed concrete suggestions about issues involved in doing a PhD. He has inspired me with his wisdom and his own experience as a diligent scholar and inquisitive learner. I feel truly blessed to be an apprentice of these three great scholars. I am immensely grateful for their time and input, patience and support.

I am also thankful to all my research participants, as well as those who facilitated my data collection in Shandong University in 2016. In particular, I would like to thank my Master's research supervisor, Professor Zhenqian Liu, who generously offered his own office space for my interviews, and my friend Yueming Sun, who assisted me in contacting every participant.

I would also like to thank my two raters and research teammates, Dr Mark Toomer and Dr Brian Strong. Our debates and discussions for my data processing were one of my fondest memories at the School of Linguistics and Applied Language Studies (LALS).

Dr Lisa Woods and Dr Rolando Coto Solano, thank you for helping me with statistical analysis. I saw the light through my analysis in my consultation sessions with each of them.

I'm also grateful to my examiners Dr Anna Siyanova, Dr Lynn Grant, and Professor Jeannette Littlemore for their comments and recommendations to further improve the quality of my thesis.

My heartfelt thanks to my officemates and great friends Diep Tran, Dr Jeremy Koay, and Associate Professor Kevin Parent, Setor Donne Novieto, and Yan Sun for their company and encouragement during my ups and downs. My sincere thanks to my PhD student peers and friends: Sonali Pawaskar, Ruojin Yan, Lingyu Li, Lin He, Douglas Meyer, George Aberi, Pakjira Leelertphan, Deborah Chua, Chi Duc Nguyen, Thuy Bui, Mengzhu Yan, Wei Wei, Elvenna Majuddin, Irene Fioravanti and many others from the Thesis Group and the Vocab Group, for the comradeship and unforgettable memories. I am also thankful to many staff at LALS and Victoria who were also part of my support system and who offered help and insights at different stages of my study: Dr Ha Hoang, Dr Patrick Coelho, Associate Professor Averil Coxhead, Dr Irina Elgort, Dr Victoria Chen, Dr Derek Wallace, Dr Sara Cotterall, Dr Lizzie Towl, and others whose name I may have missed here.

Special thanks to Huyen Le's (Hayley) family who offered me shelter from time to time. To Carol Mitchell, my New Zealand mum, for her positive psychology. To my fellow Toastmasters in Wellington, and many other intelligent minds, kind hearts and beautiful souls who have enriched my life during my studies.

Last and most importantly, I thank my mum and dad Jianmei Qu and Kaiwen Wang for their unconditional love and support throughout the years, and for always offering me the best education they can.

I am truly blessed to have you all, along this intellectual as well as spiritual, adventurous journey of my life. Thank you.

TABLE OF CONTENTS

ABSTRACT.....	i
ACKNOWLEDGEMENTS.....	iii
TABLE OF CONTENTS.....	v
LIST OF TABLES	x
LIST OF FIGURES	xii
ABBREVIATIONS.....	xiii
CHAPTER ONE: INTRODUCTION	1
1.1 Background to the study	1
1.2 Overview of the study and its aim	3
1.3 Organisation of the thesis.....	5
CHAPTER TWO: LITERATURE REVIEW	7
Introduction.....	7
2.1 What are idioms and what kind of idioms does this thesis focus on?.....	7
2.1.1 Idiom and Idiomaticity.....	9
2.1.2 Definitions and identifying features of idioms	10
2.1.3 Classifications of idioms and their relevance to L2 learning and teaching	15
2.1.3.1 Structural approaches.....	16
2.1.3.2 Semantic approaches.....	18
2.1.3.3 Functional approaches	21
2.1.3.4 A synthetic approach	22
2.1.4 Idioms in this study.....	23
2.2 Idiom processing and comprehension in first and second languages	24
2.2.1 L1 idiom processing and comprehension	25
2.2.1.1 Non-compositional models	25
2.2.1.2 Compositional models	28
2.2.1.3 The Hybrid Model.....	30
2.2.1.4 More general models of (figurative) language processing	32
2.2.1.5 Summary	36

2.2.2 L2 idiom processing and comprehension	37
2.2.2.1 How does L2 idiom processing differ from L1?.....	37
2.2.2.2 How do L2 learners comprehend idioms?	43
2.2.2.3 Summary	47
2.3 Enhancing second language idiom learning	48
2.3.1 Idioms in L2 teaching and learning	49
2.3.1.1 Why teach idioms?.....	49
2.3.1.2 General trends and approaches to L2 idiom teaching	55
2.3.2 Cognitive-linguistics-informed approach to L2 idiom learning	60
2.3.2.1 Introduction.....	60
2.3.2.2 Theoretical rationale for the CL-informed approach	61
2.3.2.3 From theories to practice—tenets and techniques	65
2.3.2.4 Etymological elaboration and rationale for this study	70
2.3.2.5 Research Objective One.....	76
2.3.3 Factors likely to affect L2 idiom learning under etymological elaboration.....	77
2.3.3.1 Introduction.....	77
2.3.3.2 Semantic analysability and transparency of idioms.....	80
2.3.3.3 L1 transfer and cross-cultural interference	85
2.3.3.4 Prior L2 lexical knowledge	86
2.3.3.5 L2 proficiency	87
2.3.3.6 Teaching procedure—trial-and-error and errorless learning	88
2.3.3.7 Other factors and considerations for this study.....	89
2.3.3.8 Research Objective Two	91
2.4 Summary of research objectives and questions	93
CHAPTER THREE: METHODOLOGY	95
3.1 Overview of the research design.....	95
3.2 Participants.....	98
3.3 Materials	102
3.3.1 Idiom selection.....	102
3.3.2 Etymology of idioms.....	108

3.3.3 Example sentences and context control	108
3.3.4 Format of material presentation	110
3.4 Data collection	111
3.4.1 Learning phase	112
3.4.2 Recall phase	116
3.5 Data processing	118
3.5.1 Transcription and translation.....	118
3.5.2 Rating the guesses and recall answers	119
3.5.2.1 Rating criteria and rating schemes	122
3.5.2.2 Rating procedure	125
3.5.2.3 Inter-rater reliability and accounting for disagreements	128
3.5.2.4 Strengths and limitations.....	129
3.6 Data analysis	131
CHAPTER FOUR: OVERVIEW OF THE LEARNING PROCESS AND OUTCOME	
.....	135
Introduction.....	135
4.1 Responses counted in the analysis	135
4.2 Overall learning gains	138
4.2.1 When initial idiom interpretation was successful (GuessBP = 1).....	138
4.2.2 When initial idiom interpretation was unsuccessful (GuessBP = 0 or 0.5)	140
4.2.3 The relationship between the correctness of interpretation and that of recall (GuessAP & Recall).....	142
4.3 Learning gains by learners and by idioms	143
4.3.1 Learning gains by individual learners.....	144
4.3.2 Learning gains by individual idioms.....	147
4.4 Summary and discussion.....	150
CHAPTER FIVE: INVESTIGATING L2 IDIOM INTERPRETATION	
Introduction.....	153
5.1 Factors affecting L2 idiom interpretation	154
5.1.1 Prior L2 lexical knowledge.....	154

5.1.2 L1 transfer and cross-cultural differences.....	158
5.2 Evaluating etymological elaboration via etymological semantic transparency	164
5.2.1 Results of etymological semantic transparency ratings	165
5.2.2 Regression analysis.....	166
5.2.2.1 Etymological semantic transparency and the correctness of aided interpretation	168
5.2.2.2 Etymological semantic transparency and L2 proficiency	170
5.2.2.3 The correctness of aided interpretation and L2 proficiency	171
5.3 Summary and discussion.....	172
CHAPTER SIX: INVESTIGATING L2 IDIOM MEANING RETENTION	175
Introduction.....	175
6.1 Factors affecting L2 idiom meaning retention.....	176
6.1.1 Model selection.....	176
6.1.2 Etymological semantic transparency and L2 proficiency	178
6.1.3 The correctness of guesses with the aid of etymological elaboration.....	183
6.1.3.1 Inferential analysis of Recall and GuessAP.....	184
6.1.3.2 Analysis of the recall errors	185
6.2 Qualitative analysis of other factors involved in L2 idiom learning	188
6.2.1 The EFL Learners' personal experience with and awareness of L2 idioms....	188
6.2.2 The EFL learners' strategies for L2 idiom comprehension and memorisation	190
6.2.3 The EFL learners' attitudes towards using and learning L2 idioms.....	193
6.2.4 Case analysis of outstanding individual learners	196
6.3 Summary and discussion.....	198
CHAPTER SEVEN: GENERAL DISCUSSION AND CONCLUSIONS	201
7.1 Major findings and pedagogical implications.....	201
7.1.1 Etymological elaboration as an inferencing prompt	202
7.1.2 Etymological elaboration as a mnemonic	204
7.1.3 Optimising the teaching procedure of etymological elaboration.....	206
7.1.4 L2 idiom learning in an EFL context	208

7.1.5 Other pedagogical aspects.....	210
7.2 Significance of the study and methodological contributions.....	211
7.3 Limitations of the study and suggestions for future research	212
7.4 Concluding remarks	214
REFERENCES	215
Appendix 1: Information Sheet for Participants	237
Appendix 2: Participant Consent Form.....	241
Appendix 3: Human Ethics Approval Confirmation	243
Appendix 4: Frequencies of the 84 idioms in COCA and Collins COBUILD	244
Appendix 5: Semantic transparency by three native English speakers.....	247
Appendix 6: Dictionary meanings and example sentences.....	251
Appendix 7: Etymological prompts of the idioms	259
Appendix 8: Directions for Interview One	264
Appendix 9: Semantic Transparency Rating Form (Set A)	265
Appendix 10: Meaning Recall Test Sheet (set A)	267
Appendix 11: General questions for the second interview (record sheet).....	268
Appendix 12: Excluded items by participants	269
Appendix 13: Breakdown figures of learning gains by participants.....	270
Appendix 14: The codes of idioms	274
Appendix 15: Breakdown figures of learning gains by idioms	275
Appendix 16: Descriptive data of etymological transparency ratings & correctness of responses of all the idioms (Total N = 470).....	277
Appendix 17: CLMM output for Chapter 5	280
Appendix 18: CLMM output for Chapter 6	287

LIST OF TABLES

Table 2.1 Moon's (1998) categories of FEIs	22
Table 2.2 Moon's (1998) alternative grouping of FEIs.....	23
Table 2.3 Teaching techniques proposed under the CL approach.....	66
Table 2.4 Factors investigated in learning-oriented studies of L2 idioms	79
Table 3.1 Set vs. Group design	96
Table 3.2 Data collection	98
Table 3.3 Descriptive statistics of TEM-4 scores	99
Table 3.4 Distribution of TEM-4 grades	99
Table 3.5 TEM-4 scores	100
Table 3.6 Target idioms in four sets	103
Table 3.7 Frequency distribution of the 84 idioms	104
Table 3.8 Distribution of the source domains	105
Table 3.9 Distribution of the degrees of semantic transparency	107
Table 3.10 Distribution of the phrase types	107
Table 3.11 Procedure of data collection.....	111
Table 3.12 Distribution of the datapoints.....	120
Table 3.13 Rating scheme	124
Table 3.14 Inter-rater reliability	128
Table 3.15 Codes and data types of factors and variables	132
Table 4.1 Summary of excluded cases (N)	136
Table 4.2 Cases of correct meaning guesses before the prompt (Total N of cases = 56)	139
Table 4.3 Overview of learning gain by cases	141
Table 4.4 Recall of literal underpinnings for each meaning recall score (Total N = 414)	
.....	142
Table 4.5 Comparison between GuessAP vs. Recall	143
Table 5.1 Insufficient L2 lexical knowledge by case and by idiom.....	155
Table 5.2 Unknown words by participant	156

Table 5.3 Idioms containing words with an unfamiliar homonymous/polysemous meaning.....	157
Table 5.4 L1-L2 degrees of similarity and cross-linguistic influences	160
Table 5.5 Counts of etymological transparency ratings (Total N = 414).....	165
Table 6.1 Factors & analysis in investigating L2 idiom meaning retention	176
Table 6.2 Meaning recall of idioms rated, by low and high transparency (total N = 414)	179
Table 6.3 Tracing recall errors from GuessBPs and GuessAPs (Total case N = 414)	186
Table 6.4 EFL learners' sources of exposure to L2 idioms.....	188
Table 6.5 Frequency of strategies used by the participants to memorise idioms.....	192
Table 6.6 Attitudes towards using L2 idioms by the EFL learners	193
Table 6.7 Characteristics of outstanding learners	196

LIST OF FIGURES

Figure 3.1 TEM-4 scores of 4 groups	101
Figure 3.2 Cue cards and the prompt	110
Figure 4.1 Change of mean scores between different phases for all the participants	145
Figure 4.2 Change of mean scores between different phases for all the idioms	148
Figure 5.1 Distribution of transparency ratings with different GuessAP scores	169
Figure 6.1 Probability of correct recall based on TEM-4 grade and Transparency	181

ABBREVIATIONS

Terms

CL	Cognitive Linguistics
EFL	English as a Foreign Language
L1	First language (i.e., native language)
L2	Second language
TEM-4	Test for English Majors-Band 4

Codes

GuessBP	Guess before the prompt
GuessAP	Guess after the prompt
Recall	Recall answers
RCOP	Recall of the prompt
Transparency	Etymological semantic transparency
UnknownW	Unknown words
PolyHomo	Polysemy/homonymy

CHAPTER ONE: INTRODUCTION

1.1 Background to the study

Idioms are a common type of multiword expression that are non-compositional. That is, the meaning of an idiom is not derivable by adding up the meanings of its component words. In addition, although some such expressions can also be used in a literal sense, their “idiomatic” meaning is figurative, e.g., *play by ear*, *apple of one’s eye*, *(be) waiting in the wings*. Dictionaries of English idioms defined along the above lines typically include about 6,000 entries (e.g., Ayto, 2009). However, according to an estimate from an idiom learning website (<https://www.theidioms.com>), there are more than 25,000 idioms in the English language. An issue with such estimates is that not all researchers consider the same phrases to be idioms (see further discussion in Chapter 2). In terms of the frequency of idioms in everyday language, Pollio, Barlow, Fine, and Pollio (1977) figured that most English speakers utter about 7,000 idioms per week, and Glucksberg (1989) estimated that on average about four idioms are used per minute of conversation. These researchers undoubtedly used a broader notion of idiomaticity than how idioms were defined above. It remains true, though, that these expressions are also figurative. Some estimates of the frequency of idioms as defined above have been based on their incidence in corpora. For instance, Stengers (2007) searched the online Word Banks corpus (the corpus which informed the Collins COBUILD book series) for examples of a random set of 500 expressions listed in a 5,000-entry idiom dictionary. She extrapolated from her counts that about 2,400 instances of idioms occur per million words of English discourse (i.e., one idiom per 500 words). However, this is likely to be an underestimation of the everyday incidence of idioms because the Word Banks corpus underrepresent spoken language. This is important to note, because idioms, as a class, are particularly frequent in conversation (McCarthy, 1998). Another reason why counts based on corpora are likely to underestimate the incidence of idioms is that idioms display greater variability than dictionary entries would lead us to assume, and so many examples of idioms in electronic corpora may escape word-string searches based on these dictionary entries.

Apart from the fact that idioms are pervasive in everyday speech, they also serve essential

discourse functions, such as providing evaluation and maintaining interpersonal relations and social bonding (O’Keeffe, McCarthy, & Carter, 2007). Moreover, they provide interesting insights into the use of words, languages and the thought processes of their speakers. Considering that idiomatic and figurative expressions embody metaphors and concepts peculiar to a group of people, it is no wonder that some have claimed that mastering a foreign language is like adopting a foreign mindset.

While idioms seldom pose comprehension problems to adult native speakers, they do pose a serious challenge to second language learners (Cieślicka, 2006; Irujo, 1993; Kövecses & Szabo, 1996). One source of comprehension problems is the semantic non-compositionality of idioms. A learner may understand *waiting* and *wing* as individual words but fail to work out the meaning of *(be) waiting in the wings*. The interpretation process is particularly likely to lead to inaccurate inferences when the first meaning of a constituent word that springs to mind is not congruent with the imagery behind the idiom, as would be the case with *waiting in the wings* if the learner interprets *wings* as birds’ wings (instead of part of a theatre where actors wait before they appear on the stage). Other sources of comprehension problems are L1 interference (if deceptively similar L1 expressions are available), cultural differences (e.g., Hu & Fong, 2010, on Chinese EFL learners’ difficulties with English idioms referring to the heart), and gaps in the learner’s knowledge of the culture-specific *source domains* or origins of particular L2 idioms (see, e.g., Boers, Demecheleer, & Eyckmans, 2004a, on Flemish EFL learners’ problems with idioms derived from cricket and baseball).

While using an L1 translation is a common strategy for explaining the meaning of single L2 words, translating idioms is often not an option when there is no equivalent in L1. A word-by-word literal translation of idioms is likewise unlikely to help learners overcome the kinds of obstacles to comprehension listed above. Explaining the meaning of idioms through paraphrase (in L2 or L1) remains an option, but for learners to fully grasp the meaning and function of the idiom such descriptions need to be sufficiently precise, because idioms are never just neutral alternatives to a more literal way of conveying the message (McCarthy, 1998, p. 145). Nonetheless, in order to be proficient in a second language, a learner needs to master a large number of multiword units such as idioms, collocations, and phrasal verbs. Therefore,

alternative methods that guide learners towards adequate comprehension and durable retention are well worth exploring.

A school of thought in linguistics which has been much concerned with figurative language, including idioms, is Cognitive Semantics, or more broadly, Cognitive Linguistics. Since Lakoff and Johnson's *Metaphors We Live By* (1980), the ubiquity of metaphor in natural language has attracted a lot of attention from scholars working in various fields of linguistics (e.g., Gibbs & Steen, 1999) including pedagogy-oriented applied linguistics (e.g., Boers & Littlemore, 2003). Research endeavours in applied linguistics have demonstrated that insights from Cognitive Linguistics (CL) have considerable potential to help L2 learners comprehend and remember figurative expressions such as idioms. Essentially, the CL-informed approach is to help learners appreciate how the established, figurative meaning of idioms is derived through metaphor (and metonymy) from their literal or original usage, and is thus far from arbitrary. Proponents of CL claim that this approach offers the prospect of more systematic and insightful learning of a wide array of figurative expressions, contrary to the 'traditional view', which held that idioms could only be learned through 'blind memorisation' (Boers, 2001, p. 35). However, the applicability and effectiveness of this CL-informed approach in the teaching and learning of idioms is yet to be determined. As we will see, there is a range of factors that affect L2 idiom learning and teachability through a CL-informed approach, such as the degree of semantic transparency of L2 idioms, and the impact of L1 transfer and cultural differences.

1.2 Overview of the study and its aim

Taking these factors into consideration, the research in this thesis explores CL-informed ways of aiding L2 learners in the comprehension and retention of idioms. While many previous research has focused on comparing the overall effects of different teaching approaches (e.g., Li, 2009; Skoufaki, 2008; Szczepaniak & Lew, 2011), this study aims to evaluate *etymological elaboration*, a particular CL-informed teaching technique proposed by Boers, Demecheleer, and Eyckmans (2004b), by exploring the factors that affect its effectiveness among learners in an EFL context. According to Boers et al. (2004a, 2004b, 2007), etymological elaboration involves informing learners about the origins or the literal underpinning of the idioms, in order

to facilitate the interpretation and memorisation of idioms. The learners/participants in the studies reported by Boers and colleagues were Dutch-speaking EFL students in Flanders. The benefits of the instructional intervention may have been relatively strong because of the proximity of the target language to their L1 and because of the familiarity of these students with Anglo-Saxon culture (thanks partly to aspects of a shared cultural heritage but also to abundant exposure through TV, films, etc.). In contrast to participants in the studies by Boers and colleagues, the participants in the project described in this thesis are EFL learners in mainland China.

This project has two broad but inter-related research objectives. Objective One is to investigate the effectiveness of etymological elaboration in facilitating idiom comprehension and retention. This includes an evaluation of what types of idioms and what type of learners are particularly amenable to such elucidations. Objective Two is to investigate the problems that L2 learners, that is, native-Chinese EFL learners in this study, experience when they encounter English figurative idioms, and to identify the factors influencing the learning success. Specific research questions under these two broad objectives will be discussed at the end of Chapter 2.

Overall, this is an explorative study that uses a mixed research design. To explore the thought processes of L2 learners, as they try to make sense of and recall idioms, a teaching experiment and a one-week-delayed post-test were embedded in one-on-one think-aloud interviews with 25 participants. During the interviews, different factors were investigated, including those factors involved in interpretation and memorisation, such as the degree of semantic transparency of selected idioms (as perceived by these L2 learners), L1 influence, and prior L2 lexical knowledge, as well as general learning factors, such as the learners' experience with, attitude towards, and learning strategies regarding English idioms. The learning focus of the teaching experiment is the receptive knowledge of idioms, that is, the participants' knowledge of the meaning of idioms. Both quantitative and qualitative methods were used in data analysis.

Findings from the study are expected to inform ways of L2 idiom teaching and learning in an EFL context. They may help teachers better understand students' idiom learning processes and needs, and fine-tune the implementation of CL-inspired instruction. They may also inform aspects of the design of textbooks and dictionaries that are intended to foster idiom learning.

1.3 Organisation of the thesis

Chapter 2 provides a literature review of relevant research topics that have informed the research questions of this study. First, by reviewing idiom research from linguistic, psycholinguistic, and applied linguistic fields, three major background questions will be answered: (1) what are idioms? (2) how does L2 idiom processing differ from L1 idiom processing? (3) what does prior research tell us about L2 idiom teaching and learning, especially regarding what applied cognitive linguists have proposed as methods for enhancing L2 idiom learning? Then, through critical reviews of CL-informed approaches to L2 idiom teaching and of factors involved in L2 idiom learning, ten research questions are proposed under the two research objectives mentioned above.

Chapter 3 introduces the methodology adopted for this study. It describes and explains the overall research design, participants, materials, and the procedure of data collection, data processing, and data analysis.

Chapters 4 to 6 report the results of the study and subsequent discussion of the results. Chapter 4 provides an overview of the learning process and outcomes under the examined teaching approach. Chapter 5 investigates the factors involved in L2 idiom interpretation during the learning phase. These factors include L2 lexical knowledge, L1 transfer, and the semantic transparency of idioms. Chapter 6 investigates further the factors involved in L2 idiom meaning retention that are identified in the preceding analyses. It also investigates factors related to more general aspects of idiom learning, including learners' experience with and awareness of L2 idioms, and their attitudes to and strategies for idiom learning. Together, the three chapters address the research questions regarding L2 idiom teaching and learning from different perspectives.

Chapter 7 concludes the whole thesis with a synthesis of the major research findings and a discussion of pedagogical implications. It also evaluates the significance and methodological contributions as well as the limitations of the study, and makes suggestions for future research.

CHAPTER TWO: LITERATURE REVIEW

Introduction

This chapter reviews previous research that has informed the research questions of this PhD study. First, by reviewing idiom research from three related but different fields—theoretical linguistics, psycholinguistics, and applied linguistics, three major background questions will be answered: (1) what are idioms? (2) how does L2 idiom processing differ from L1 idiom processing? (3) what does prior research tell us about L2 idiom teaching and learning, especially what have applied cognitive linguists proposed for enhancing L2 idiom learning? By answering the first background question, section 2.1 discusses the problems inherent in finding clear and universally accepted definitions of idioms and clarifies what kinds of idioms this thesis focuses on. By addressing the second background question, section 2.2 serves three purposes. First, this section explains why etymological elaboration was targeted as a teaching technique based on the particular features of L2 idiom processing in comparison with L1 processing; second, this section provides a rationale for the think-aloud approach of the research design (see section 2.2.2.2); and third, the section provides a discussion of some important factors that are likely to affect L2 idiom learning, especially L2 comprehension, which together with section 2.3, explains why certain factors were chosen to be investigated in this study. The focus of the literature review lies in addressing the third question, through a critical review of the CL-informed approach to L2 idiom teaching and learning. This is followed by a discussion of specific issues that leads to the current research objectives and questions.

2.1 What are idioms and what kind of idioms does this thesis focus on?

Defining *idiom* is one of the thorniest issues in the studies of idioms (Grant & Bauer, 2004; Liu, 2008; Moon, 1998). From the perspective of various studies in theoretical linguistics, psycholinguistics, and language teaching and learning, the term *idiom* is “ambiguous” and “used in conflicting ways” (Moon, 1998, p. 3), and refers to a variety of word combinations known by different names. For instance, some scholars use the term to cover a wide range of

conventionalised multiword items (e.g., Alexander, 1987; Makkai, 1972; McCarthy, 1992), including *tournures*, *phrasal verbs*, *irreversible binomials*, *phrasal/opaque compounds*, *similes*, *proverbs*, *familiar quotations*, etc.; some use it to refer to “fixed and semantically opaque or metaphorical” expressions (Moon, 1998, p. 4); and a few restrict it to ‘core idioms’, an even narrower group of multiword units that are both non-compositional and non-figurative (Grant & Bauer, 2004). Different views on what constitute idioms are also reflected in variation in the coverage of idiom dictionaries.

The ambiguity of the term is intrinsically due to the “multifaceted”, “complex” and “in many ways elusive” nature of idioms (Tabossi & Zardon, 1993, p.145). Different properties of idioms have been prioritized for different research purposes; consequently, definitions and classifications of idioms vary from researcher to researcher. Given this, it is essential to clarify the category or categories of expressions labelled in the literature as ‘idioms’ that are the focus of this thesis.

In this section, I will review ways in which ‘idioms’ as a language phenomenon have been delineated in previous research, as a way of clarifying the class of expressions I will be focusing on in my project. This class is often labelled ‘figurative idioms’ (Boers & Lindstromberg, 2012; Boers et al., 2004a) or ‘conventionalised figurative expressions’ (Boers & Webb, 2015). Therefore, the question of what idioms are will be answered by addressing four sub-questions:

- (1) **What does the term *idiom* mean in general use?** – The term *idiom* will be examined from its etymology and its general meanings in relation to the concept of *idiomaticity* (see 2.1.1).
- (2) **How are idioms defined by dictionaries of idioms and in previous research?** – Definitions of idioms from various sources will be considered, and identifying features of idioms underlying these definitions are highlighted and explained, as part of the grounding of the notion of idioms in the current study (see 2.1.2).
- (3) **How are idioms classified or categorised from different linguistic approaches?** – Classifications of idioms (including broader categories of idiomatic expressions) will be discussed with particular regard to language learning and teaching (see 2.1.3).

(4) **What are idioms in the current study?** – Answers to the above questions lead to a further clarification of idioms in the scope of this study. (see 2.1.4).

2.1.1 Idiom and Idiomaticity

The word **idiom** originates from the Latin word *idioma*, which was derived from the ancient Greek word *idios*, meaning ‘private, own, peculiar to oneself’. In relation to this, the *Oxford English Dictionary* (OED) (2019) listed three senses of idioms pertaining to language¹.

Sense 1 The specific character or individuality of a language; the manner of expression considered natural to or distinctive of a language; a language's distinctive phraseology.

Sense 2 a. A language, especially a person or people's own language; the distinctive form of speech of a particular people or country.

b. In narrower sense: a dialect or variety of a language; a form of a language limited to or distinctive of a particular area, category of people, period of time, or context.

Sense 3 A form of expression, grammatical construction, phrase, etc., used in a distinctive way in a particular language, dialect, or language variety; spec. a group of words established by usage as having a meaning not deducible from the meanings of the individual words.

Here, we can see that the senses of *idiom* range from a broad reference to the distinctive character of a language, through a dialect or variation of a language, to the more specific reference to conventionalised expressions.

In general use, the three senses of *idiom* in OED are reflected in the two levels of use, i.e., broad and narrow, discussed by Moon (1998). In broader uses, *idiom* refers to a particular manner of expressing things in a language which characterises a person or a group of people (cf. *Sense 1* and *Sense 2a*). In narrower uses, *idiom* represents those word combinations in a language that are “fixed and semantically opaque or metaphorical” (Moon, 1998, p. 4; cf. *Sense*

¹ There are also two other senses of *idiom* in OED, which were non-linguistic. E.g., *Sense 4* is “a distinctive style or convention in music, art, architecture, writing, etc.” (Retrieved May 11, 2019 from <https://www.oed.com/>).

3). The broad uses of idiom are often related to terms such as *idiomatic* and *idiomaticity* in a broad sense and to Sinclair's *idiom principle* (1987).

Idiomaticity, or “the quality of being marked by idioms” (Grant & Bauer, 2004, p. 39), is broadly used to refer to a manner of language use natural to native speakers, which covers a wide range of ‘nativelike’ ways of expression including idioms, collocations (see comparison between the two in 2.1.3.1), formulae like *What's up?* and *How are you doing?*, as well as other grammatical constructions (not to mention more regional or dialectal variations within a language, cf. *Sense 2b*). From a language learning perspective, idiomaticity is also described as “the extent to which a learner's language resembles that of a native speaker” (Idiomaticity, n.d.). The narrow and broad uses of the terms *idiom* and *idiomaticity*, therefore, render idiom “both a superordinate and a hyponymic term for lexical combinations” (Moon, 1998, p. 4), which to some extent contributes to the diverse views on and confusions of ‘idioms’ as mentioned earlier.

The focus in this thesis is on *idiom* in a narrow sense as ***Sense 3*** of OED. However, reflecting the mixed use of the term *idiom* in the literature, in the following review, I will use both narrow and broad senses. I will use *idiom* mostly in the narrow sense in discussions relating to the definitions and identifying features of idioms in 2.1.2. When discussing more general typological issues in 2.1.3, I will use it in a relatively broader sense as idiomatic expressions or idiomatic ‘multiword units’ (MWUs) (Grant & Bauer, 2004) and make a clear distinction in due course. To find out what properties identify idioms and what actually cause difficulties for L2 learners, we now turn to more practical and descriptive definitions of idioms.

2.1.2 Definitions and identifying features of idioms

Different dictionaries of English idioms provide various definitions of idioms. Here are some typical examples:

- 1 An idiom is a combination of two or more words which function as a unit of meaning.
(*Oxford Dictionary of Current Idiomatic English*, Cowie & Mackin, 1975, p. ix)
- 2 An idiom is a fixed group of words with a special different meaning from the meanings of the separate words.

(*Longman Dictionary of English Idioms*, 1979, front cover)

- 3 An idiom is a phrase whose meaning is different from the meaning of each word considered separately. These phrases have a fixed form—they usually cannot be changed.

(*Cambridge Dictionary of American Idioms* 2003, p, ix)

- 4 In practical terms this [i.e. ‘idiom’] includes a wide range of expressions that have become in a sense fossilized within the language and used in a fixed or semi-fixed way without reference to the literal meaning of their components.

(*Oxford Dictionary of Idioms*, 1999, Preface)

- 5 An idiom is a special kind of phrase. It is a group of words which have a different meaning when used together from the one it would have if the meaning of each words were taken individually. [...] Idioms are typically metaphorical: they are effectively metaphors which have become ‘fixed’ or ‘fossilized’.

(*Collins COBUILD Dictionary of Idioms*, 1999, p. v)

- 6 An idiom is a phrase, saying or a group of words that has a metaphorical (not literal) meaning, which has become accepted in common usage.

(*The Idioms*, Retrieved April 2019 from <https://www.theidioms.com/>)

Moreover, in previous studies of idiom learning and teaching, idioms are most commonly defined as a type of multiword expressions whose overall meanings cannot be predicted from adding up the meanings of their constituents.

The definitions listed above contain some recurring properties of idioms:

First, idioms are **multiword expressions** or **multiword units (MWUs)**. This means that idioms consist of two or more words (as described in *Definition 1*). However, the claim that idioms comprise minimally two words is disputed by some linguists (e.g. Hockett, 1958; Katz & Postal, 1963; Makkai, 1972) who consider idioms to include opaque polymorphemic words or compounds, such as *greenhouse* and *blackmail*, as examples of non-compositional linguistic units whose meaning cannot be derived from the sum of their components. However, this view is not widely accepted among linguists nor is it applied in language teaching practices. The reason for excluding single words as idioms for L2 learners was, as Liu (2008, p.15) put it, “simple and obvious”, since however difficult it is, learning a single word is basically the same

as grasping any other regular individual word, while learning idioms (comprising at least two words) is a different matter and more challenging, involving dealing with idiomatic meaning and invariant or restricted form. According to Grant (2003), *multiword units* are “fixed and recurrent patterns of lexical materials sanctioned by usage” (p.15). In other words, idioms have been conventionalised or institutionalised through usage by a large number of people in a speech community (Fernando & Flavell, 1981; Moon, 1998); and lexicogrammatical “fixed and recurrent patterns” (Grant, 2003, p. 15) are also recognised from the definitions of idioms.

Second, idioms are **relatively fixed in form**, as explicitly noted in *Definitions 3* and *4*. This (semi-)fixedness in form is identified as transformational deficiency by formal linguists such as Chafe (1968) and implies some degree of lexical and/or structural invariance with “preferred lexical realizations and often restrictions on aspect, mood or voice” (Moon, 1998, p. 7). According to Moon’s corpus-based study of ‘fixed expressions and idioms’ (FEIs, another term that includes idiomatic expressions), corpus evidence proves fixedness as a key feature of FEIs. Nevertheless, fixedness is clearly relative, since Moon also reports that “around 40% of database FEIs have lexical variations or strongly institutionalized transformations, and around 14% have two or more variations on their canonical forms” (p. 120). In addition, dictionaries like the *Collins COBUILD Dictionary of idioms* (1999, edited by Moon et al.), also report lexical variations within individual expressions (e.g., *a bag/box of tricks*) and variant forms across varieties of English (e.g., *have a green thumb* in American English/*have green fingers* in British English).

Third, **non-compositionality** is a key criterion unanimously agreed on among linguists and a recurrent theme in various definitions of idioms. The Principle of Compositionality (also known as *Frege’s principle*) is an influential principle in formal semantics, which states that “the meaning of a (syntactically complex) whole is a function only of the meanings of its (syntactic) parts together with the manner in which these parts were combined” (Pelletier, 1994, p.11). The meaning of an idiom is however not deducible from the sum of its components, and therefore idioms are non-compositional. A typical example frequently cited is *kick the bucket*, whose meaning ‘to die’ cannot be derived from combining the meaning of *kick* and *bucket*. In another sense, the meaning of an idiom exceeds the combined meanings of its components, and

this renders the form-meaning connection of the expression non-transparent to various degrees (Boers & Webb, 2015, referring to Cacciari & Glucksberg, 1991; Gibbs, Nayak & Cutting, 1989). The concept of non-compositionality as applied to idioms is especially complex if we take a historical perspective. This is because current non-compositional idioms are likely to have been compositional when they were coined. In addition, non-compositionality is a matter of degree “in relation to both syntactic structure and metaphoricality” (Moon, 1998, p. 8). On all accounts, non-compositional does not mean inexplicable. With figurativeness or metaphoricity involved, “idioms may require a more etymological/ historical approach” (Grant & Bauer, 2004, p. 51)—which is the core theme of this study.

Considering that idioms do not pose any problem for mature native speakers but are problematic for L2 learners to understand, it is no wonder that some researchers would argue that whether an expression is an idiom cannot be truly answered without raising the question of ‘idioms for whom’ (Liu, 2008, p. 13, referring to Makkai, 1993; Ruhl, 1989). This non-transparent form-meaning connection is often identified as a key comprehension obstacle for L2 learners, and this is often related to another feature of idioms—figurativeness, to which I now turn.

Figurativeness is sometimes but not always recognized in the definitions of idioms. The words ‘figurative’ and ‘metaphorical’ are often used interchangeably in previous publications, and the same is true in this thesis. According to *Collins COBUILD Dictionary of Idioms* (1999), idioms are “typically metaphorical” and “effectively metaphors which have become ‘fixed’ or ‘fossilized’” (p. v). Similarly, *Definition 6*, from a public website, *The Idioms*² (Theidioms.com), defines an idiom as “a phrase, saying or a group of words that has a metaphorical (not literal) meaning, which has become accepted in common usage”. While this definition is not from an academic source, it nevertheless reflects a lay view of what identifies idioms, and one that is important when we consider L2 learners’ perspective. Some applied linguists (e.g., Boers

² *The Idioms* (Theidioms.com) introduces itself as an educational website and ‘dictionary’ based in the UK since 1998 with multinational writers and linguists that provides users (language learners) with the world’s most famous English-language idioms, sayings, slang, and phrases. Its main objective is to reveal the meaning and origin of an idiom with original details for site visitors.

&Webb, 2015) also use the term *conventionalised figurative expressions* to refer to idioms in their idiom teaching research. Moreover, figurativeness is often used as a criterion in the classification and categorisation of idioms and idiomatic expressions (see further discussion in 2.1.3.2). The fact that figurativeness is not always recognized as an identifying feature of idioms is partly due to the diverse views on what constitute idioms, and partly due to the concept of figurativeness being used ambiguously in the literature, which will be further discussed in the context of *conceptual metaphors* in 2.3. Although not all non-compositional fixed MWUs are figurative, non-compositionality is often attributed to being figurative or metaphorical. It seems reasonable to assume that at some transitional stage between literal and figurative use, most MWUs had both a literal and a figurative meaning. The more abstract figurative meanings were shared and used in a speech community, and as language evolved with time, while some literal meanings or usage may have remained in use, many others were used less frequently and became outdated or even completely forgotten. In addition, some literal usage belongs to a special field unfamiliar to the ordinary language users (Mäntylä, 2004). As a result, idiomatic expressions became more commonly recognized by their figurative meaning. The connection between meaning and form became less clear, rendering the meaning unit non-compositional—thus ‘frozen’ and ‘fixed’. Such expressions are likely to be stored and retrieved from memory as a whole (see further discussion in 2.2) and tend to facilitate usage of the multiword expression as a chunk.

So far, we have noted the identifying features of idioms based on different definitions in idiom dictionaries and in some previous studies. That is, they consist minimally of two words, they show relative fixedness in form and non-compositionality in meaning, and they are often figurative or metaphorical. (Semi-)fixedness and non-compositionality are related to each other in the process of conventionalization or “institutionalization” (Grant, 2003, p. 36; Moon, 1998, p. 7); and to some extent, non-compositionality “entails certain barriers to structural changes” (Strässler, 1982, p. 24). All the definitions and identifying features mentioned above are compatible with the notion of idioms as the focus of the current project, i.e., *conventionalised figurative (multiword) expressions*. We will return to this notion again, after an overview of

classifications of idiomatic expressions, which will help further clarify the scope of idioms in this study.

2.1.3 Classifications of idioms and their relevance to L2 learning and teaching

In order to systematically account for different kinds of idiomatic expressions (idioms in a broad sense), theoretical linguists, psycholinguistics, corpus linguists, and lexicographers have proposed numerous classifications with different foci foregrounding different properties of idioms as well as what best represents those properties (Cowie, Mackin & McCaig, 1983; Fernando, 1996; Fernando & Flavell, 1981; Fraser, 1970; Grant & Bauer, 2004; Makkai, 1972; Moon, 1998; Strässler, 1982; Weinreich, 1969; Wood, 1986; etc.). While semanticists and syntacticians focus more on the internal features of idioms, functional linguists and sociolinguists provide pragmatic and functional accounts of idioms, investigating the real-world usage of idioms in texts. Drawing on one or more of these classifications, lexicographers take a more practical synthetic approach to the identification, categorisation and collection of idioms, in order to give guidance to language users and learners. For many of these studies, the availability of large corpora of authentic English have enabled in-depth research into the frequency and variability of idioms. As for language teaching, it is often claimed that “only if the set of idioms is divided into separate and distinguishable categories does the teaching of them in an ESL/EFL classroom become a more realistic task” (Grant, 2003, p.25). Furthermore, in this respect, pedagogy-oriented researchers propose metaphor-based classifications according to the sources or cognitive mechanisms of idioms, in order to facilitate L2 idiom learning, which will be explained in section 2.3, where the CL approach to idiom teaching is introduced.

For the purpose of this study, I will briefly review classifications of idioms or idiomatic expressions with particular regard to what matters for language learners and to the knowledge types that can inform L2 teaching. This section will in turn review structural, semantic, and functional approaches. For each approach, I will discuss specific issues relating to language teaching and learning. Some researchers combine different approaches, and the decision about which approach to discuss them under is mine and does not necessarily imply that the researchers would associate themselves primarily with that approach. Finally, I will introduce

a synthetic approach by corpus lexicographer Moon (1998) as a summary to this section. More detailed reviews on classifications of idioms and idiomatic expressions can be found in Fernando (1996), Moon (1998), Grant (2003), Liu (2008), and Dąbrowska (2018).

2.1.3.1 Structural approaches

Structural and syntactic classifications prioritize lexicogrammatical patterns of idiomatic expressions, reflecting their lexical and syntactic variabilities or irregularities. For the purpose of this study, I will review some seminal classifications that take a structural or form-related approach from two perspectives: (1) how idioms can be structurally distinguished from other types of MWUs; and (2) what kind of formal properties language learners need to know in order to be able to use idioms appropriately.

Early grammar-based studies examined the **transformational deficiencies** or potentials of idioms based on transformational-generative grammar (e.g., Fraser, 1970; Katz & Postal, 1963; Weinreich, 1969). Fraser (1970) proposes a 6-level *Frozenness Hierarchy* for assessing idioms based on their transformational flexibility, ranging from *L6 Unrestricted* (i.e., all transformations are possible), through *L5 Reconstitution*, *L4 Extraction*, *L3 Permutation*, *L2 Insertion*, and *L1 Adjunction*, to *L0 Completely Frozen* (i.e., no transformations are possible). The hierarchy captures the varying degrees of transformability of idioms from L5 to L0, as the relative unanalysability of idioms excludes them from L6. Such rule-based classifications are effective in describing and characterizing the structural ‘fixedness’ and ‘frozenness’ of idioms. However, because they are highly theoretical, technical and not very important to the purpose of this study (i.e., examining ways of helping L2 learners comprehend and remember the *meaning* of idioms), I will not go further into this approach (see detailed reviews in Dąbrowska, 2018; Moon, 1998).

Collocational restriction is another a key criterion that linguists and lexicographers consider as the basis of idiomaticity and which is used for both classifying idiomatic expressions and distinguishing idioms from non-idioms. Influenced by Firthian linguistics, Cowie, Mackin, and McCaig (1975, 1983) use **collocation** as a key criterion for their compilation of *Oxford Dictionary of Current Idiomatic English*, one of the earliest dictionaries of idiomatic expressions targeted at advanced foreign learners. Taking into account both collocational

restrictions (i.e., “the possibility of internal variation, or substitution of part for part” [1983, p. xii]) and semantic opaqueness, Cowie et al. use a complex categorisation and group idiomatic expressions into: (1) *pure idioms* (i.e., idioms in a strict sense, e.g., *kick the bucket*) (2) *figurative idioms* (which show very little variation in collocation) (3) *restricted collocation* (or *semi-idioms*, i.e., one word of which has a figurative meaning) and (4) *open collocations* (i.e., *free/loose collocations*) (pp. xii-xiii). Such a categorisation puts **collocations** and **idioms** onto a single continuum, on which the last category (open collocations) is marked as the most distinguishable from idioms, since the latter show no or very little variance. Similarly, many studies present idioms and habitual collocations as two related but different lexical types (see a critical review in this respect in Fernando, 1996).

Internal structure is also commonly used in classifying idiomatic expressions. An early influential study was conducted by Makkai³ (1972) from a stratificational structuralist view. He first distinguishes two types of idioms: (1) *idioms of encoding* or *phraseological idioms* (e.g., *drive at 60 m.p.h.* instead of *drive with*), which are similar to habitual collocations, and (2) *idioms of decoding* or *semantic idioms* (e.g., *red herring*) which “force the hearer to decode in a certain way” (pp. 56-57) and are potentially subject to misunderstanding. This second type seems to represent a narrower sense of idioms. Makkai also assumes that all the idioms of decoding are idioms of encoding but not vice versa (p. 57). He further divides idioms of decoding into two strata, *lexemic idioms* and *sememic idioms* (i.e., ‘sentence-long utterances’), according to their internal structure. Within lexemic idioms, he includes *tourneures* (e.g., *bite the bullet*, *be well-off*), *irreversible binomials* (e.g., *bread and butter*, *kith and kin*), *phrasal compounds* (e.g., *hot dog*, *bookworm*), *phrasal verbs* (e.g., *come across*, *put up with*), *incorporating verbs* (e.g., *babysit*, *whitewash*) and *pseudo-idioms* (i.e., one of the constituents of the expression is a “cranberry morph”, p.169). Within sememic idioms, he includes proverbs, familiar quotations, ‘first-base’ idioms (i.e., culturally bound expressions, e.g., *have two strikes again one*, *hit a home run*) and institutionalised expressions of greeting, politeness, indirectness,

³ Some scholars also identify Makkai’s classifications as a semantic (e.g., Moon, 1998) or functional approach (Fernando, 1996), as his classifications also touched upon non-compositionality (cf. *idioms of encoding*) and discursal functions (cf. *sememic idioms*) of idioms. I classify it as a syntactical approach for this thesis, based on his overall structural approach to language.

understatement and hyperbole. The latter stratum takes more cultural, pragmatic properties into account (for further sub-classifications, see Makkai, 1972, pp. 135-172). Makkai's classification uses a broad sense of idioms and provides structural accounts for a wide variety of idiomatic expressions, subsuming what this study labels as idioms. His approach clearly differentiates 'semantic idioms' from 'phraseological idioms'/habitual collocations, claiming that despite overlap, the latter are less likely to cause misunderstandings for a L2 speaker.

Moreover, lexicographers Cowie et al. (1983) highlight for learners the structural variety of idioms, grouping them into two general types based on **grammatical patterns**: *phrasal idioms*, including noun phrases, adjective phrases, prepositional phrases and adverbial phrases, and *clausal idioms*, which consist of 'verb + other component(s)' (p. xi).

In conclusion, structural approaches provide detailed lexical and syntactic descriptions and classifications of idioms. As noted by Cowie et al. (1983), "[t]he accurate and appropriate use of English expressions which are in the broadest sense idiomatic is one distinguishing mark of a native command of the language and a reliable measure of the proficiency of foreign learners" (p. x). The structural fixedness of idioms along with semantic opaqueness pose great challenges for L2 learners, especially those who endeavour to fully grasp idioms or use the language in a more nativelike manner. However, learning the formal properties is often not of immediate interest to learners, as they tend to shy away from using idioms themselves, and what concern them most are the semantic properties which are associated with cultural, conceptual aspects of languages. As fostering productive knowledge was not my ambition in this project, I will not go further into the form-oriented research. I will now turn to semantic approaches to idiom classification.

2.1.3.2 Semantic approaches

Although much attention has been given to the grammatical and structural properties of idioms, "idiomaticity is largely a semantic matter" (Cowie et al., 1983, p. xi). *Semantic transparency/opacity*, *literality/figurativeness*, and *compositionality* (in other words, *analysability* or *decomposability*) are commonly used to identify and classify idioms on a semantic basis (Alexander, 1987; Cacciari & Glucksberg, 1991; Fernando, 1996; Fernando &

Flavell, 1981; Moon, 1998; Wood, 1986). I will review two different approaches to the semantic classification of idioms, i.e., using scalar and dichotomous analyses.

In terms of **transparency**, a scale is often presented that is made up of categories such as *transparent*, *semi-transparent*, (*semi-opaque*) and *opaque idioms*. In terms of **figurativeness**, there is a scale that stretches from *literal*, *semi-literal/semi-idiom*, *metaphorical*, to *opaque* (e.g., Fernando & Flavell, 1981, p. 28). These two scales are sometimes mixed with each other, e.g., ranging from *literal phrases*, *semi-idioms*, *metaphorical idioms*, *figurative idioms* to *opaque/pure idioms* (Alexander, 1987, p. 110). Terms used at the very opaque end always represents a strict sense of idioms, i.e., ‘true idioms’ (Yorio, 1980), ‘pure idioms’ (Alexander, 1987; Fernando, 1996), or ‘full idioms’ (Fernando & Flavell, 1981). In addition, **decompositionality/analysability** is another criterion used to classify idioms in relation to how idioms are processed by L1 speakers (Gibbs & Nayak, 1989; see more discussion in section 2.2.1).

The scales of semantic properties provide a useful tool for describing and characterizing idiomaticity. These properties are also commonly used as variables in psycholinguistic research which examines idiom processing and acquisition, as will be further discussed in 2.2. While most researchers use similar descriptive scales, they sometimes do not agree with each other on which category some of the expressions are placed into. This is because transparency, figurativeness and analysability are inevitably subject to individual judgements.

To simplify their classification of MWUs for the purposes of language teaching and learning, Grant and Bauer (2004) proposed a more restricted definition of *idiom* based on a test using the two criteria of *non-compositionality* and *non-figurativeness*. To make the test easier to use by teachers and learners, they adopted a so-called ‘naïve approach’ to compositionality and figurativeness. They test whether an MWU is ‘non-compositional’ by seeing whether replacing each word in an MWU with its dictionary definition would change the meaning of the whole expression. They equate figurativeness with ‘untruth’ and judge whether an MWU is ‘figurative’ or ‘non-figurative’ by (a) the existence of an ‘untruth’ and (b) the possibility of reinterpretation of that ‘untruth’ “by use of an image or other means” (p. 51). For example, *He’s gone off the rails* entails an untruth, meaning *He’s no longer on the rails*, which can be reinterpreted as *He’s*

not moving forward in a controlled, guided way. In addition, they test whether only one word in the MWU is either non-literal or non-compositional. As a result of the test, they divide MWUs (a broad sense of *idioms*) into three groups: *core idioms* (non-compositional and non-figurative), *figuratives* (non-compositional but figurative), and *ONCEs*, i.e. MWUs containing one non-compositional element. They examined random samples of 19 idioms from the *Cambridge International Dictionary of Idioms* and the *Collins COBUILD Dictionary of Idioms*, and categorised them according to the proposed criteria. Their results reveal that the majority of idioms in the two idiom dictionaries are ‘figuratives’ or ‘ONCEs’, and that the set of ‘core idioms’ constitutes a small number of idiomatic MWUs. Crucially, Grant and Bauer have pointed out that ‘figuratives’ can be made sense of, but ‘core idioms’ cannot, and suggest that figuratives “may require a more etymological/historical approach” to learning (p. 52). This gives a strong motivation for the current study, which also draws support from cognitive linguistics (see further discussion in 2.1.4).

Grant and Bauer (2004) made a valuable attempt at distinguishing different types of MWUs for the purposes of teaching and learning, although their approach seems to rely largely on L1 speakers’ intuition and is not necessarily practical for L2 learners. Moreover, their classification yields a similar grouping to Cowie et al.’s (1983) *pure idioms*, *figurative idioms*, and *semi-idioms* (and excluding *open collocations*, which are clearly not idioms); and the boundaries of Grant and Bauer’s three categories are still inevitably fuzzy, though they claim the fuzziness is much reduced compared to the previous classifications.

The semantic criteria discussed above, whether they use a continuous scale or a dichotomous test, reflect not only the properties of idioms themselves but also the perceptions of people who are using or judging the expressions. In language teaching and learning, they can be considered as factors affecting the uptake of MWUs in L2, and can therefore be used to gauge the level of difficulty in learning, which is of great importance to the current study. A measurement of **L2-learners’ perceived transparency** of idioms is proposed and investigated in this study within an etymological approach. From an L2 perspective, this also involves the learners’ own cultural and linguistic knowledge (see further explanation in section 2.3.3).

2.1.3.3 Functional approaches

Functional approaches draw attention to the pragmatic and discursual functions of idiomatic expressions (e.g., Fernando, 1996; Moon, 1998; Strässler, 1982). Such approaches describe how and why idioms are used and help predict when and where idioms are likely to pose problem to language learners.

Fernando (1996) applied Halliday's (1973, 1985) systemic functional grammar and identified three major **discursual functions** of idiomatic expressions: *ideational*, *interpersonal*, and *relational*. In classifications based on these functions, most figurative idioms seem to belong to the ideational expressions, as a resource for communicating information about the world and the experiences of a speaker/writer. By contrast, institutionalised expressions such as greetings, politeness, etc., belong to interpersonal expressions, while multiword connectives such as *for example* and *as a result* constitute relational expressions. Ideational idioms are clearly the most relevant to this thesis. Amongst these ideational idioms, seven types of message content were identified by Fernando: actions (e.g., *wave the olive branch*), events (e.g., *paint the town red*), situations (e.g., *in the doldrums*), people and things (e.g., *a red herring*), attributes (e.g., *good as gold*), evaluations (e.g., *a Trojan horse*, *turn back the clock*), and emotions (e.g., *have one's heart in one's mouth*).

Moon (1998) categorises idiomatic expressions, which she called 'fixed expressions and idioms' (FEIs), according to five **text functions**: *informational*, *evaluative*, *situational*, *modalizing* and *organizational* (p. 217). Based on her database, Moon claimed that the majority of the FEIs are informational and/or evaluative (p. 219). In terms of attitude, "evaluation orientations are more strongly associated with metaphors [i.e., figurative idioms] than other types of FEI" (p. 247), with a higher proportion of negative than positive evaluations; however, a few FEIs can be either approbatory or deprecatory depending on the context.

Moreover, idiom use is often **register-sensitive** (Liu, 2008; Moon, 1998). Generally, idioms are more common in informal discourse types such as fiction or spoken discourse than in formal discourse. Amongst types of written discourse, existing corpus evidence shows that figurative idioms are commonly used in journalism, and we will see that this is supported by learners' observations in the current project.

2.1.3.4 A synthetic approach

Owing to the complex interrelated features of idioms, no matter which approach one takes and whether the measure is scalar or stratificational, the resulting classes or categories are inevitably fuzzy and sometimes overlap with each other. No single model can fully capture idiomaticity or classify various types of idiomatic expressions. Therefore, Moon (1998), as a corpus-based lexicographer, adopts an eclectic typology for idiomatic expressions or ‘FEIs’ (cf. idiomatic ‘MWUs’). Based on the possibility of accounting for an FEI in terms of lexicogrammar, pragmatics, or semantics, Moon uses three macro-categories for grouping: *anomalous collocations*, *formulae*, and *metaphors*, with each macro-category having its own finer categories (see Table 2.1). Along with the macro-categories, she presents an alternative classification with more descriptive and discrete criteria (see Table 2.2). This combination of two ways of classifying FEIs offers a more comprehensive account for the complex phenomenon of idiomaticity, incorporating various approaches to idiomatic expressions, which also makes a useful and accessible typological framework for language teaching and learning.

Based on her corpus analysis, Moon found that “the majority of metaphors [i.e., individual items] have frequencies of less than 1 per million” (p. 61). However, as a group, *metaphors* or figurative idioms constitute about one third of FEIs, showing that *metaphors* are a very common type of formulaic language.

Table 2.1 Moon’s (1998) categories of FEIs

problems of lexicogrammar	Anomalous collocations	ill-formed collocations cranberry collocations defective collocations phraseological collocations
problems of pragmatics	Formulae	simple formulae sayings proverbs (literal/metaphorical) similes
problems of semantics	Metaphors	transparent metaphors semi-transparent metaphors opaque metaphors

Table 2.2 Moon's (1998) alternative grouping of FEIs

phraseological collocations defective collocations	paradigmatically restricted
ill-formed collocations cranberry collocations	syntagmatically restricted
simple formulae sayings non-metaphorical proverbs	fixed, literal, discoursally meaningful
metaphorical proverbs similes transparent metaphors	non-literal, transparent
semi-transparent metaphors opaque metaphors	non-literal, non-transparent

2.1.4 Idioms in this study

For any line of idiom research, how one defines and classifies idioms is affected by one's research aims and interests. Importantly for this study, a pedagogical perspective will highlight that idioms may be grouped into different types, depending on the kind of learning required (Grant & Bauer, 2004). For this study, I chose to focus on idioms that are defined as *conventionalised figurative multiword expressions*, with a CL-informed approach. In CL publications, they have sometimes been labelled 'imageable idioms' (e.g., Boers & Demecheleer, 2001) by virtue of a possible link between the abstract meaning and a literal reading of the expression, i.e., where the literal reading conjures up a mental image.

With reference to previous classifications of idiomatic expressions, idioms in this study are treated as 'metaphors' in Moon's (1998) tripartite typology, and as 'figuratives' in Grant and Bauer's classification (2004), and the figuratives form the largest group of what most scholars consider idioms. It is also worth noting that idioms in this study will also include some of what Grant and Bauer call 'borderline figuratives', 'ONCEs' (for their overall figurative meaning), and 'core idioms' (e.g., *red herring*), which are interpretable with available source of origins under my etymological approach.

Furthermore, although many are figurative, phrasal verbs were excluded from this study, because they are easily distinguishable on the basis of form (Moon, 1998) and are large in number, and Grant (2003) argues that they deserve separate and thorough research. In addition, similes such as *(as) cool as a cucumber*, formulaic sequences such as *what's up*, and proverbs such as *the grass is always greener on the other side of the fence* were not of direct relevance to this study.

2.2 Idiom processing and comprehension in first and second languages

This section answers the second review question: how does L2 idiom processing differ from L1 idiom processing. It does this by providing an overview of psycholinguistic accounts for L1 and L2 idiom processing and by highlighting the differences between them (see subsections 2.2.1 to 2.2.2.1). As part of this, I will explain the rationale of the CL-inspired L2 idiom teaching approach by drawing from native speakers' idiom processing features and discussing what kind of teaching approach suits L2 learners' possessing features. On the one hand, people might find an 'analytic' teaching approach such as etymological elaboration 'unnatural' if they are convinced that native speakers process idioms holistically, i.e., "without recourse to any form-meaning matching of any sub-parts it may have" (Wray, 2008, p 12). However, if idioms are also (at least some of the time) analysed by native speakers, then using an 'analytic' teaching approach with L2 learners is perhaps not so 'unnatural', after all. On the other hand, we will also see how much L2 learners rely on the literal meanings of the component words of an idiom to process its idiomatic meaning.

After the review of online processing studies, I turn in subsection 2.2.2.2 to offline comprehension studies and explain how L2 learners comprehend idioms which are not yet familiar to them. I will also further explain the choice of etymological elaboration as a target of investigation. Moreover, I aim to provide a rationale for the think-aloud aspect of the research design with reference to an earlier study by Cooper (1999). Some important factors that are likely to affect L2 idiom learning, especially L2 comprehension, are also introduced in this subsection and are discussed in more detail in section 2.3.

2.2.1 L1 idiom processing and comprehension

As explained in 2.1, idioms are non-compositional, mostly figurative in meaning, and fixed or semi-fixed in form. How such expressions are represented and processed in the human mind has been of great interest and importance, and has also presented a great challenge to researchers. From the 1970s to the 1990s, a series of hypotheses and models were proposed supported by findings from experimental studies, and debates have formed around how idioms are represented and comprehended. Two main foci of these debates are whether idioms are stored and accessed as a single unit in the mental lexicon, and whether/how literal and figurative meanings are activated during comprehension. This subsection reviews major hypotheses and models of L1 idiom processing. These are based on different views of idioms and can accordingly be roughly grouped into four types: non-compositional models, compositional models, hybrid models, and more general models of figurative language processing.

2.2.1.1 Non-compositional models

Some of the assumptions found in early processing models emerged from linguistic accounts of idioms via the transformational and semantic analyses that were current at the time (e.g., Chafe, 1968; Fraser, 1974; Weinreich, 1969). As Swinney and Cutler (1979) stated, from those linguistic accounts two opposing views were identified, one treating idioms as “a separate, specialized list of items, which have special insertion privileges in the grammar” (p. 524) and the other treating idioms as “single lexical items which have much the same characteristics in the grammar as do any other lexical items” (ibid.). As an example of the first approach, they cited the work of Weinreich (1969), and as an example of the second approach, they cited Fraser (1974). Based on these approaches, Swinney and Cutler suggested two general processing models, i.e., the *Idiom List Hypothesis* and the *Lexical Representational Hypothesis*.

The **Idiom List Hypothesis**, as described by Swinney and Cutler (1979) and based on Bobrow and Bell (1973), posits that idioms are stored in and accessed from a separate list, or a special idiom lexicon that is not part of the general mental lexicon. In their study, Bobrow and Bell had two conditions—literal and idiomatic. In the literal condition, participants read four sentences with literal ambiguities (such as *Mary fed her dog biscuits*, i.e. ‘Mary fed biscuits to

her dog' or 'Mary fed dog biscuits to her'), and then saw a test sentence that had both literal and idiomatic meanings (e.g., *John gave Mary the slip* or *John and Mary buried the hatchet*). In the idiomatic condition, they read four sentences with both literal and idiomatic meanings and then saw a test sentence of the same type. Then the participants were asked to mark on paper whether they perceived the literal or idiomatic meaning of the test sentence first. Bobrow and Bell found that their participants gave different interpretations of the ambiguous test sentences depending on whether the prime contained literal or idiomatic ambiguities, and they took this as evidence for the existence of distinct idiomatic and literal modes of processing. Subsequent researchers interpreted this as indicating that the Idiom List Hypothesis prioritizes a literal analysis before the processing of the figurative meaning. Whether Bobrow and Bell's (1973) study supports such a claim or not, the hypothesis was soon rejected by subsequent studies, and Swinney and Cutler (1979) argue that post-perceptual tasks such as getting people to recall whether the idiomatic or literal meaning first appeared in their mind after reading the sentences do not provide an appropriate measure of real-time processing.

The **Lexical Representation Hypothesis** (Swinney & Cutler, 1979) posits that idioms are stored and retrieved from the lexicon as a long word, just like any ordinary word, and that "computation of both meanings—idiomatic and literal—is simultaneously initiated upon occurrence of the first word in the idiom string" (p.525). Swinney and Cutler substantiated the hypothesis by two experiments examining the speed of meaning retrieval via *Phrase Classification Tasks*, i.e., asking subjects to judge as quickly as possible whether or not a string of visually presented words formed a meaningful natural English phrase. The Lexical Representation Hypothesis predicts that if idiomatic phrases (e.g., *break the ice*) are single lexical entries then they should be processed more quickly than non-idiomatic expressions of similar length and syntactic complexity (e.g., *break the cup*). To check that their initial results had not been affected by variation in the types of idioms, Swinney and Cutler added linguistic frozenness⁴ as a control variable in their second experiment. The results of the two experiments

⁴ Swinney and Cutler employed Fraser's (1974) categories of 6-level frozenness as an indicator of transformational probabilities of idioms (see previous explanation of this categorisation in section 2.1.3.1), as the lexical status of an idiom may be affected by its linguistic frozenness (Swinney & Cutler, 1979, p. 531).

showed that idioms were judged to be acceptable English phrases much faster than their matched controls, and that there was no significant effect of frozenness (i.e., idiom types) on processing speed. Therefore, Swinney and Cutler conclude that idioms are stored and accessed as single lexical items. Although the Lexical Representation Hypothesis rejects the existence of a special ‘idiom list’ or any special processing mode, it still treats literal processing as an integral part of idiom processing. However, this latter view was countered by Gibbs in his *Direct Access Hypothesis*.

The **Direct Access Hypothesis** (Gibbs, 1980, 1985, 1986) asserts that given sufficient context, native speakers can access the idiomatic meaning of a familiar idiom directly, bypassing any literal processing. This view is supported by a series of studies that Gibbs conducted on people’s understanding of idioms (Gibbs, 1980, 1986), as well as other nonliteral utterances such as indirect requests and expressions of irony and sarcasm (Gibbs, 1983, 1984, 2002). In his 1980 study, three experiments examined native speakers’ comprehension and memory of idioms with both literal and idiomatic meanings in different conditions—with or without a conversational context. In one experiment, participants read the conversations one line at a time. The last line of the conversation contained an idiom used either figuratively or literally (e.g., *He’s singing a different tune*). Participants’ reaction time was measured when they made a True-or-False *Paraphrase Judgment* of this last sentence. They responded to the idiomatic sentences significantly faster than the literal ones in both with- and without-context conditions, although in the without-context condition the error rates were much higher for the literal sentences compared with the idiomatic ones. Based on the findings of this and the other two experiments in his study, Gibbs claims that people tend to automatically analyse the conventional idiomatic meaning before deciding if the literal meaning is appropriate. In his argument, Gibbs (1980) emphasizes the importance of sufficient contextual information in comprehension, and *conventionality* as the key to both idiom comprehension and memory. The Direct Access Hypothesis is also supported by Gibbs’ subsequent research (1986), as well as by the existence of those idioms (e.g., *by and large*) which do not have a well-defined literal meaning but are still processed and understood (Gibbs, 1985).

The three early models discussed above take it as a given that idioms are non-compositional, and so these models are often categorised as non-compositional or “direct look up” models (Glucksberg, 1993, p. 4), which view idiom processing as retrieval of the meaning of an idiom as a whole. This is in contrast with compositional models, which consider non-compositionality as a continuum and take heterogeneity of idioms into account. Major compositional models are the *Idiom Decomposition Hypothesis* and the *Configuration Hypothesis*.

2.2.1.2 Compositional models

The **Idiom Decomposition Hypothesis** (Gibbs & Nayak, 1989; Gibbs, Nayak, & Cutting, 1989) proposes that idioms exist on a continuum of semantic composition (or analysability; see also Nunberg, 1978), and holds that idioms are first processed in a compositional manner, i.e., via compositional analysis. However, it does not contend that readers/listeners necessarily analyse the literal meanings of idioms during their understanding of these figurative expressions. By ‘compositional analysis’, Gibbs et al. (1989) mean that “people recognize the individual parts of decomposable idioms as having independent, figurative meanings that contribute to an idiom’s overall figurative interpretation”, but that people do not necessarily “combine the literal meanings of words during idiom processing” (p. 589), which seems to be suggested by a *literal analysis* (e.g., Cacciari & Tabossi, 1988). To comprehend non-decomposable idioms, L1 speakers generally have to learn the conventional idiomatic meanings. L1 speakers’ intuition of the analysability of idioms was claimed to affect ease of comprehension (Gibbs et al., 1989) and acquisition (Gibbs, 1987, 1991). Gibbs and Nayak (1989) found that L1 speakers were able to classify idioms as being *normally decomposable* (e.g., *pop the question*, *miss the boat*), *abnormally decomposable* (e.g., *spill the beans*, *bury the hatchet*), and *semantically non-decomposable* (e.g., *kick the bucket*, *chew the fat*). Since then, these three categories (originally proposed by Nunberg, 1978) have been widely used as an indicator of decomposability/compositionality of idioms in subsequent research.

It is important to note that *decomposability* (or ‘analysability’) should not be confused with *transparency*, as these terms characterise idioms from different dimensions: the former is the degree to which individual meanings of the idiom’s word components contribute to the meaning of the whole phrase, and the latter is the extent to which the metaphorical meaning

can be deduced from its literal analysis (Cieślicka, 2015; Nunberg, Sag, & Wasow, 1994). Generally, transparent idioms are more decomposable and opaque idioms are often viewed as non-decomposable, but there are idioms that are both transparent and non-decomposable (e.g., *jump the gun*) and vice versa (see Nunberg et al., 2004; and more discussion on semantic decomposability in section 2.3.3.2).

The **Configuration Hypothesis** (Cacciari & Glucksberg, 1991; Cacciari & Tabossi, 1988; Tabossi & Zardon, 1993; Titone & Connine, 1994a) was first proposed by Cacciari and Tabossi (1988) with the support of three cross-modal priming experiments. This hypothesis treats idioms the same as any other familiar string of words; when such a string is encountered, the meanings of its constituents are activated while the string is recognized as a *configuration* (i.e., a unit representing the idiomatic meaning; Cacciari & Glucksberg, 1991). It posits that the configuration is recognized after a sufficient amount of the idiom is encountered (cf. word recognition, Marslen-Wilson, 1987), and such a recognition point was identified as the *idiomatic key* (Cacciari & Tabossi, 1988). More specifically, some parts of an idiom are more relevant than others for recognition of the idiom, and the speed of recognition depends on the location of the key and the context in which the idiom is encountered (ibid.). The Configuration Hypothesis emphasizes the predictability of an idiom as a major determinant in idiom comprehension/ recognition, which was further evidenced by a neurolinguistic study based on event related potentials (ERPs; Vespignani, Canal, Molinaro, Fonda & Cacciari, 2009).

Both the Decomposition Hypothesis and the Configuration Hypothesis suggest that idiom comprehension involves internal linguistic analysis of the idiom. While the former model differentiates the various degrees of analysability of idioms, the latter highlights the existence of the key recognition point(s) of an idiom configuration. Both hypotheses pointed out that people use different types of information simultaneously during idiom processing which affect both direct retrieval and compositional analysis.

As a test and extension of the Configuration Hypothesis, Titone and Connine (1994a) conducted three cross-modal experiments to examine the influence of *predictability* and *literality* (i.e., the degree to which an idiomatic expression has a plausible literal meaning) on idiom comprehension. They found that high-predictable idioms were accessed sooner than low-

predictable idioms and that the literal meanings were activated for all idiom types except high-predictable non-literal idioms. Overall, an idiomatic meaning is accessed very quickly in idiom comprehension but accessing this idiomatic meaning might not terminate ongoing literal processing, which suggests a competition between literal and figurative interpretations in the selection of the “best fit” (p. 1135). This view was further developed in the *Hybrid Model* (see below).

Importantly, in addition to the factors *predictability*, *literality* and *compositionality* which were acknowledged in the earlier models, *familiarity* was also identified as a key variable in (L1) idiom processing by Titone and Connine (1994b; see also Flores d'Arcais, 1993; Cronk & Schweigert, 1992; Cronk, Lima, & Schweigert, 1993; and Schweigert, 1986). In their exploration of the relationships between the above four dimensions, Titone and Connine (1994b) operationalised familiarity as frequency and meaningfulness, and found that predictability correlates significantly with familiarity. In line with this finding, Tabossi, Fanari, and Wolf (2009) attribute the overall fast recognition of idioms to the fact that they are expressions that are known to people, i.e., to familiarity.

2.2.1.3 The Hybrid Model

Taking into account features of both the compositional and non-compositional models, the **Hybrid Model** (Titone & Connine, 1999) proposes that idiomatic expressions function simultaneously as semantically non-compositional and compositional phrases. This means that idioms are highly automatised multiword expressions whose meanings can be directly retrieved from the mental lexicon. At the same time, literal processing of the component parts of an idiom (i.e., internal linguistic analysis) potentially allows the reader or listener to infer the original motivation of an idiom's figurative meaning (Cieślicka, 2015; Titone & Connine, 1999). According to this model, “activation of idiomatic meanings, and the activation and use of literal meanings during comprehension, will be a function of the degree to which idioms are conventional and compositional” (Titone & Connine, 1999, p. 1666).

Libben and Titone (2008) further develop the Hybrid Model within a *constraint-based* view, taking into account other constraint-satisfaction work in processing (e.g., MacDonald, Pearlmutter, Seidenberg, 1994; MacDonald & Seidenberg, 2006). By examining the effects of

familiarity, compositionality and word frequency on idiom processing, they find that idiom processing is determined by many factors, with different sources of information becoming available and interacting in a time-dependent manner and with decomposability playing a limited role at the beginning of comprehension.

From a functional perspective, Caillies and Butcher (2007) further develop the Hybrid Model within the *Construction-Integration* framework (Kintsch, 1998). They differentiate three levels of idiom representations, i.e., the surface, the propositional, and the situational levels (cf. the *superlemma* model in Sprenger, Levelt, & Kempen, 2006). In their argument, Caillies and Butcher (2007) postulate that people's world knowledge, including conceptual metaphors, may influence idiom comprehension at the situational level.

So far, the Hybrid Model has been supported by many experimental studies, which have examined idioms of different types (e.g., with different levels of compositionality and familiarity). Supporting evidence has also been found in studies of idiom production (Cutting & Bock, 1997; Sprenger, Levelt, & Kempen, 2006). For instance, Cutting and Bock (1997, p. 69) point out that "idioms are idioms because of how they are associated to conceptual representations, as well as how they are represented in the lexicon." One type of conceptual representation of idioms will be further discussed below under the *Conceptual Metaphor Hypothesis*.

Such multi-level processing models as developed by Caillies and Butcher (2007) and Sprenger et al. (2006), as well as Peterson and Burgess' (1993) *Semantic-syntactic Autonomy Model*⁵ (supported by Peterson, Burgess, Dell, & Eberhard, 2001) are among more recent attempts to account for idiom representations at multiple levels, ranging from the phonetic, semantic, syntactic, to discourse levels, and from the lexical to more abstract conceptual levels. "Distinguishing between these different levels of representation [...] are [sic] extremely

⁵ The semantic-syntactic Autonomy Model (Peterson & Burgess, 1993) claims that there is dissociation between syntactic and semantic analyses in idiom comprehension. It proposes that a full syntactic analysis always proceeds during idiom processing but the semantic analysis of the literal meaning of an idiom is terminated once the figurative meaning is retrieved. According to Peterson and Burgess (1993), this model aligns with neurolinguistic findings on hemispheric asymmetries in the processing of literal and figurative language.

important because idioms, or MWE [i.e., multiword expressions] generally, can vary dramatically and independently at all of these levels” (Titone, Columbus, Whitford, Mercier, & Libben, 2015, p. 194) and this is true for both L1 and L2 speakers (ibid.). In addition, more recent priming studies of idiomatic expressions have found evidence of internal linguistic analysis not only at the syntactic level (Konopka & Bock, 2009; Snider & Arnon, 2012), but also at the phonological and semantic levels (Sprenger et al. 2006) (for a more detailed review, see Siyanova-Chanturia, 2015b).

2.2.1.4 More general models of (figurative) language processing

Another important hypothesis about idiom processing is the **Graded Salience Hypothesis** (Giora, 1997, 2003; Giora & Fein, 1999). By synthesizing earlier studies on the processing of literal and figurative language, Giora (1997, 2003) theorises that the comprehension of literal and figurative language is governed by the *salience* of meanings, which depends on factors such as conventionality, frequency, familiarity, and prototypicality. The Graded Salience Hypothesis proposes that salient meanings are processed first before less salient meanings are activated. Based on this assumption, it is argued that the figurative meaning of highly conventional idioms whose salient meaning is figurative tends to be accessed directly without activation of the less salient, literal meaning (Giora, 1997; cf. the Direct Access Hypothesis). Among the aforementioned factors that affect salience, familiarity is deemed the most crucial (Giora, 2003, p. 17). For example, language users are likely to have encountered an expression such as *pull someone’s leg* far more often in its figurative sense than in a literal sense, and so the former will be the default interpretation. In the case of *a free ride*, however, it is likely that language users will encounter literal as well as figurative uses, and so interpretation may need to rely more on context. As a test of the Graded Salience Hypothesis, Giora and Fein (1999) examined the comprehension of idioms with different degrees of familiarity in biasing contexts, by using *word fragment completion tests*—a so-called indirect semantic priming measure. The participants were 13 to 14-year-old students, who familiarised themselves with L1 Hebrew idioms through their studies at school for a year. At the end of the year, their knowledge of the idioms was tested, based on which the idioms in the experiment were divided into familiar and less familiar idioms. The researchers then asked participants first to read familiar and less

familiar idioms in contexts that biased towards idiomatic and literal interpretations, and then complete a word fragment which was related to either the literal or the figurative meaning of the target idiom. Their findings show that in comprehending familiar idioms, the salient idiomatic meaning was activated in both types of context, whereas the literal meaning was activated in the context biased towards the literal meaning but “hardly activated” in the context that was biased towards the idiomatic meaning (Giora & Fein, 1999, p. 1614). On the other hand, in the cases of less familiar idioms, where “the idiomatic meaning was no more salient than its literal interpretation”, the literal meaning was “highly activated” in both types of contexts, whereas the less salient idiomatic meaning was activated in the context biased towards the idiomatic meaning but only “marginally evoked” in the context that was biased towards the literal meaning (ibid.). Regardless of the reliability of the measurement of salience in this study, the findings seem to suggest that “familiar idioms always avail their salient idiomatic meanings, even when contextually incompatible, as in contexts biasing their interpretation toward the literal meaning” (Giora, 2003, p. 19), which was taken further as support for priority of salience of meaning over context effects in the Graded Salience Hypothesis (Giora, 2003; Giora & Fein, 1999).

In addition to the psycholinguistic approach, theoretical accounts of idiom processing have also been generated from pragmatic and cognitive linguistic studies. This includes Vega-Moreno’s *Relevance-theoretic Hypothesis* (2001, 2003, 2005) based on the *Relevance Theory* (Sperber & Wilson, 1986/1995) and Gibbs’ (1994, 1997) ‘*Conceptual Metaphor*’ Hypothesis based on the *Conceptual Metaphor Theory* (Lakoff & Johnson, 1980). Similar to the Graded Salience Hypothesis, these models treat idiom processing as part of more general figurative language processing.

The **Relevance-theoretic Hypothesis** (Vega-Moreno, 2001, 2003, 2005) holds a *path of least effort* view that idiom comprehension is neither literal nor figurative but relevance-driven, like utterance comprehension in general. According to this hypothesis, idioms of all types, regardless of degree of compositionality, are processed by the same relevance-driven inferential mechanisms (Vega-Moreno, 2001). These mechanisms always search for a best relevant interpretation by processing input in the form of “a set of highly accessible

encyclopaedic assumptions from the concepts encoded by the idiom string and its constituent words” (Vega-Moreno, 2003, p. 321). This procedure happens at both word and phrase levels, which explains why idioms often switch between literal and figurative meanings and why they often display different degrees of flexibility and productivity in the conversation flow (ibid.). Because retrieving the idiomatic meaning from memory is often faster and easier, compositional analysis of the constituents is likely to be avoided wherever possible (Vega-Moreno, 2001). Furthermore, Vega-Moreno (2001) differentiates between the processing of familiar and unfamiliar idioms. In processing familiar idioms, the hearer decodes an idiom and retrieves its idiomatic meaning without going through compositional analysis of the expression, while in processing unfamiliar idioms and idiom variants, “the meaning communicated by the speaker cannot be directly retrieved from the hearer’s lexicon. Thus, understanding these expressions often requires the exploitation of encyclopaedic entries of encoded concepts and of other conceptual assumptions in order to assign some content to the ad hoc concept” (p. 105). Overall, the Relevance-theoretic Hypothesis seems compatible with many existing psycholinguistic models, e.g., the Direct Access Hypothesis and the Hybrid Model (especially the constraint-based view).

The **Conceptual Metaphor Theory** (CMT) offers a view that metaphor is not just a figure of speech but a significant part of human cognition, and many concepts, especially abstract ones, are structured via metaphorical mappings of information from more concrete source domains to more abstract target ones (Lakoff & Johnson, 1980). For instance, people often conceptualize love by mapping the knowledge of physical journeys onto that love. The conceptual metaphor LOVE IS A JOURNEY⁶ is reflected in common expressions, such as *Their relationship has come to a dead end*, and *His marriage is on the rocks* (see more discussion of CMT in section 2.3). Based on this theory, Gibbs (1994) proposes a list of hypotheses about how metaphorical thoughts influence language use and comprehension (presented as below by Gibbs, Bogdanovich, Sykes, & Barr, 1997, p. 141):

⁶ Note that in cognitive linguistics, capital letters are used to indicate concepts (e.g. conceptual metaphors) rather than words.

H1 Metaphorical thought plays some role in the historical evolution of what words and expressions mean.

H2 Metaphorical thought motivates the linguistic meanings that have currency within linguistic communities, or is presumed to have some role in people's understanding of language.

H3 Metaphorical thought motivates an individual speaker's use and understanding of why various words and expressions mean what they do.

H4 Metaphorical thought functions in people's immediate on-line use and understanding of linguistic meaning.

While the first three hypotheses are well supported by linguistic and psycholinguistic evidence, the fourth hypothesis was relatively under-researched. To test this hypothesis, Gibbs et al. (1997) examined whether people's tacit knowledge of 'pre-existing' conceptual metaphors (e.g., ANGER IS HEATED FLUID IN A CONTAINER and ENCOURAGEMENT IS GIVING SOMEONE A DRUG) affects online processing of idioms (e.g., *blow your stack*, *a shot in the arm*). By using a self-paced reading task in combination with a visual lexical judgement task on words that are related or unrelated to conceptual metaphors (e.g., 'heat' related to ANGER IS HEATED FLUID IN A CONTAINER), Gibbs et al. found conceptual metaphors *can* be accessed quickly during online idiom processing, although this does not mean people have to access conceptual metaphors to understand idioms. In a reading-time experiment, Glucksberg, Brown, and McGlone (1993) found that conceptual metaphors were not automatically accessed during idiom comprehension. Moreover, some researchers argue that even though conceptual metaphors may be available, they may not always be accessed or used in any given context (e.g., Glucksberg, Keysar & McGlone, 1992). Overall, mixed results have been found in this line of research. As a consequence, it is unclear how metaphoric mappings affect online (L1) idiom processing. However, many studies have shown that "conceptual metaphors clearly help people make sense of why idioms mean what they do" (Gibbs et al., 1997, p. 149). Despite insufficient supporting evidence in L1 idiom processing research, Conceptual Metaphor Theory has been well acknowledged in the field of second language acquisition and has had significant impacts on L2 pedagogy, as will be discussed section 2.3.

2.2.1.5 Summary

To sum up, the hypotheses on L1 idiom processing reviewed in this section have offered various accounts regarding how idioms are processed, e.g., holistically as a single lexical unit or analytically as literal expressions, and regarding whether and how literal and figurative meanings are activated during comprehension. The differences among these views might partly be due to the heterogeneity of idioms, and this is reflected in the kinds of idioms that researchers focus on and how they interpret notions such as ‘compositionality’ (see Nunberg et al., 1994). The differences may also be partly due to the complex mechanism of idiom processing involving multiple types of information (e.g., the structure and types of an idiom, whether the idiom is encountered in a figuratively biased context). Therefore, it is not surprising to see that different experiments with manipulations of different conditions produce different results. Most recent models develop the earlier hypotheses by differentiating different types of idioms while accommodating the different factors or conditions involved in idiom comprehension.

Although processing *can* involve internal linguistic analysis, most research on L1 idiom processing suggests that the comprehension of familiar and non-decomposable idioms primarily involves direct retrieval of the idiomatic meaning, which is the most salient or default meaning for the L1 user. For idioms that are relatively less familiar or more decomposable, whether or not an L1 user will ‘analyse’ a given idiom depends on the nature of the idiom (e.g., compositionality, literality, predictability), the individual user’s familiarity with the expression, the salience of its literal/figurative meaning to the user, and the circumstances or context in which it is encountered.

From the review, we can see that relatively decomposable idioms (such as the ones in this study) are (at least sometimes) processed analytically also by L1 users. This suggests that the knowledge components that might emerge through the use of an analytic procedure (such as etymological elaboration) in teaching L2 idioms may not be too different from how idioms are represented in L1 users’ minds. If this is the case, then how far does L2 idiom processing deviate from L1 processing? What are the specific challenges that idioms pose to L2 learners? And what kind of teaching approach suits L2 learners? Now we will turn to idiom processing by L2 speakers.

2.2.2 L2 idiom processing and comprehension

As theoretical accounts of L1 idiom processing and monolingual research develop, research into crosslinguistic and L2 idiom processing has started to attract attention. Research into L2 idiom processing and comprehension can be classified into the following overlapping themes: (1) developing hypotheses and models for L2 idiom processing based on existing L1 models (Abel, 2003; Cieřlicka, 2006) by comparing idiom processing in native and non-native speakers (e.g., Carrol, Littlemore, & Dowens, 2018; Siyanova-Chanturia, Conklin, & Schmitt, 2011), (2) identifying strategies that L2 learners use during idiom comprehension (Cooper, 1999; Liontas, 2002), and (3) investigating factors that affect L2 idiom comprehension and acquisition (e.g., Carrol, Conklin, & Gyllstad, 2016; Carrol, Littlemore, & Dowens, 2018; Charteris-Black, 2002; Irujo, 1986a). These studies not only help us understand L2 idiom processing and comprehension, but also shed light on L2 idiom learning and teaching. Some of them, particularly Cieřlicka (2006) and Cooper (1999), provide important motivation for the teaching approach under examination in this thesis, and Cooper's (1999) study has inspired the design of the current study.

This section focuses mainly on the first two themes mentioned above. First, how L2 idiom processing differs from L1 idiom processing will be answered by reviewing relevant online processing studies. Then how L2 learners comprehend unfamiliar idioms will be explained by reviewing offline comprehension studies. Factors affecting L2 idiom comprehension, such as learners' L1 and crosslinguistic similarity, will be briefly discussed in the later part of this section, with further discussion to be carried out in section 2.3.3 under the topic of L2 idiom learning.

2.2.2.1 How does L2 idiom processing differ from L1?

With the abundance of theories and models, L1 idiom processing research has offered a starting point for investigating L2 idiom processing. By comparing idiom processing between native and non-native speakers of English, a few models have been proposed so far, in order to capture the 'psychological' reality of L2 idiom processing, e.g., the *Model of Dual Idiom Representations* (Abel, 2003) and the *Literal Salience Model* (Cieřlicka, 2006). In addition to online research that investigates the real-time processing of idioms that have been acquired but

which are not yet highly familiar to learners, there are comprehension studies that investigate L2 learners' use of strategies during the comprehension of less familiar or unknown English idioms, e.g., the *Heuristic Approach* (Cooper, 1999). Models also emerge from studies of learners of various second languages, such as the *Idiom Diffusion Model of Second Languages* (Liontas, 2002).

The **Model of Dual Idiom Representations** (DIR Model) proposed by Abel (2003) was probably the first model to integrate the L2 lexicon into idiom processing, or more precisely, idiom representation⁷. This model was empirically supported by two studies on non-native (L1 German) speakers' judgements of English idioms (Abel, 2003), in comparison with native speakers' judgements presented by Titone and Connine (1994b). The DIR Model considers idiom representation at both lexical and conceptual levels. It posits that at the lexical level, idioms exist as both *constituent entries*, i.e., lexical entries of the individual constituents of an idiom, and *idiom entries*, i.e., separate lexical entries that specify the idiomatic meaning; and in the absence of an idiom entry, conceptual representations of the literal constituents are activated during comprehension. By 'conceptual representations', Abel (2003) means "nonlinguistic entities that organize world knowledge and are represented at a general cognitive level" (p. 347). Much of the research in bilingual memory suggests that words in each language are stored in separate lexical systems, but that concepts are stored in a representation common to both languages (see e.g., De Groot, 1992; Kroll 1993; Kroll & Sholl, 1992). In this sense, the DIR model assumes that conceptual representations are autonomous with regard to the language system, and thus separates conceptual knowledge of idioms from their lexical meaning. Although there may be links between conceptual and lexical representations, the link is not on a one-to-one basis (Abel, 2003, p. 347). Conceptual representations can be wholly or partially retrieved depending on their relevance in a specific context (ibid.). According to Abel (2003), at least for some idioms, *conceptual metaphors* (Lakoff & Johnson, 1980) are activated, e.g., *smoke was coming out of his ear* (i.e., 'ANGER IS HEATED FLUID IN A CONTAINER');

⁷ Abel (2003) made a distinction between representation and processing, in that mental representation of language is at a higher and more abstract level than processing, although the mode of representation has direct influences on the processing (p. 342).

some idioms have one constituent that has a figurative meaning which activates the conceptual knowledge, e.g. *miss the boat* with the ‘boat’ referring to the opportunity to do something; and some are not conceptually motivated at all, e.g., *kick the bucket*, whose conceptual motivation has been lost/forgotten in the history of the English language.

According to this model, the nature of the lexical representation of an idiom is determined by its degree of decomposability and the frequency with which it is encountered. First, “[n]ondecomposable, i.e., noncompositional idioms require an idiom entry; decomposable idioms, which have compositional aspects, can be represented via constituent entries and can additionally develop an idiom entry” (Abel, 2003, pp. 342-343). This assumption is consistent with the Hybrid Model (Titone & Connine, 1999; see section 2.2.1.3) in that idioms are represented both compositionally and non-compositionally. However, the DIR Model also proposes that for the processing of decomposable idioms, the information provided by idiom entries is “supplementary” rather than “absolutely necessary” for successful comprehension, provided “conceptual representations are activated” (Abel, 2003, p. 343). This seems to allow for a more flexible, developmental mode for the L2 lexicon, i.e. that some idioms in the L2 may start off without idiom entries, which then develop as the learner becomes more familiar with the idiom. Second, the more frequently the idiom appears as an idiomatic configuration (see the Configuration Hypothesis, Cacciari & Tabossi, 1988; see section 2.2.1.2), the more likely it is that an idiom entry is formed. This second assumption is based on *frequency effects* (Abel, 2003, referring to other psycholinguistic research on word frequency effects; see also Arnon & Snider, 2010, and Siyanova-Chanturia, Conklin, & van Heuven, 2011 for evidence of phrase frequency effects). Frequency is stressed as a determining factor underlying the differences between native and non-native idiom processing. Because non-native speakers will have encountered idioms less frequently than native speakers, the development of idiom entries in the L2 lexicon is less well supported than in the L1 lexicon, and thus non-native speakers do not develop as many idiom entries as native speakers. This prediction in the number of idiom entries was supported by the findings that non-native speakers tend to decompose idioms while native speakers tend to judge them as non-decomposable (Abel, 2003), even though inter-rater reliability on whether or not an idiom is decomposable is not high. For instance, using a 75%

agreement criterion, it turns out that the judgements in both groups agreed on only half of the idioms (Abel, 2003; see also Titone & Connine, 1994b). In addition, the non-native speakers who reported to have had more exposure to English texts and thus to idioms tend to resemble native speakers in decomposability judgements (Abel, 2003).

With a paucity of idiom entries, non-native speakers more often “have to rely on the constituent entries and their corresponding conceptual representations during idiom processing” (Abel, 2003, p. 348). This links to the third assumption of the model, namely that if there is no idiom entry at the lexical level, conceptual knowledge of the idiom has to be accessed for successful comprehension. The third assumption of the DIR Model forms another important rationale for the CL-inspired teaching approach investigated in the current study, as it is suggested that an analytic approach that provides the conceptual knowledge of idioms would help L2 learners fully understand such expressions.

It is also worth noting that Abel (2003) argues that frequency—an objective measure of occurrence as found in a representative corpus of the language (p. 345, referring to e.g., Baayen & Lieber, 1996)—determines the development of ‘idiom entries’ at the *lexical* level, while familiarity—a subjective measure—is more relevant at the *conceptual* level. While a frequently occurring idiom is likely to be judged as familiar by an individual who has frequently seen or heard the idiom, an idiom that is familiar to an individual is not necessarily a frequent one in natural native-speaker discourse (see also Giora, 2003). As a proxy for the frequency with which her non-native-speaker participants encountered idioms, Abel (2003) used the participants’ biographical reports of how often they read English texts. As for decomposability, Abel (2003) employed a dichotomous division of decomposability (i.e., nondecomposable versus decomposable, rather than three categories in Gibbs & Nayak, 1989), and her interpretation of decomposability was not distinct from transparency (Abel, 2003, p. 344), even though she referred to the Idiom Decomposition Hypothesis (Gibbs et al., 1989) which differentiates decomposability from transparency. I will come back to the measurements of transparency in my discussion of the methodological concerns of the current study (see section 2.3.3.2). In general, the DIR Model is compatible with recent L1 processing models, e.g., the

Hybrid Model (Titone & Connine, 1999) and the Graded Salience Hypothesis (Giora, 1997; see section 2.2.1.4).

According to the Graded Salience Hypothesis, as I discussed in section 2.2.1.4, if the figurative (idiomatic) meaning of an idiom is much more frequent than its literal meaning, then L1 users are likely to activate the figurative meaning by default (because they will have been exposed to the figurative use far more often than the literal use). It will take contextual cues for them to activate the literal meaning instead. For L2 users, however, exposure to idioms is bound to be very limited and so these patterns of greater familiarity with the figurative use of a given idiom than with its literal use will often not apply. If so, a literal reading of the expression (i.e., a re-composition of the ‘constituent entries’, as in the DIR Model), is more likely to be arrived at. This leads to the development of the *Literal Salience Model* for L2 learners.

Based on the Graded Salience Hypothesis (Giora, 1997), the **Literal Salience Model** (Cieřlicka, 2006) ascribes salient status to a literal meaning in L2 idiom processing over a figurative meaning (which is generally more salient in L1 idiom processing). The Literal Salience Model proposes that L2 idiom comprehension requires an obligatory computation of the literal meanings of the idiom’s constituent words, regardless of its familiarity to an L2 learner and regardless of context bias. In other words, the literal meaning of an idiom remains the salient meaning (which is activated first and most strongly during online processing) even if the figurative meaning has been acquired and the L2 idiom has been automatized and incorporated into the L2 lexicon, and even if the idiom is presented in a context that biases towards the figurative meaning. This should be explained by the fact that L2 learners are most likely to be familiar with the literal meanings of the constituent words before they acquire the figurative use of an idiom, and that the literal meanings of the constituent words are likely to remain more frequently used than the idiomatic ones in an L2 learner’s language performance (Cieřlicka, 2006). This claim seems to be a bit radical, considering that the salience status might shift if a learner encounters an idiom repeatedly and becomes more familiar with the idiomatic meaning. However, Cieřlicka (2015) argues that a complete shift is unlikely to happen for foreign language learners whose exposure to idioms is limited by their learning environment. It is also claimed that “[l]iteral salience in L2 idiom processing is directly connected with the

process of building a lexical representation of an idiom entry in the course of its acquisition by an L2 learner” (op. cit., p. 121), which seems to fit the DIR Model well.

The Literal Salience Model was supported by Cieřlicka’s (2006) cross-modal priming experiment with Polish advanced learners of English (see additional neurolinguistic evidence in Cieřlicka & Heredia, 2011). Her findings prove that, unlike native speakers, L2 learners do tend to activate their knowledge of the literal meaning of the content words in idioms. However, the extent to which this spontaneous activation of literal meanings aids accurate idiom comprehension is not clear from that type of experiment—and it is one of the questions addressed in the current research project (see later in section 2.3.2.5).

Moreover, by making comparisons with existing L1 idiom processing models, Cieřlicka (2006) argues that her L2 results are compatible with the general tenets of the compositional models. However, in comparison to the Idiom Decomposition Hypothesis (Gibbs & Nayak, 1989; Gibbs et al., 1989) which “remains vague with respect to whether compositional meanings of idiom elements are literal in nature” (Cieřlicka, 2006, p. 135), the Literal Salience Model clearly emphasizes the role of the literal meaning in computing the figurative meanings of L2 idioms. On the other hand, despite agreeing with the Graded Salience Hypothesis on the priority of meaning salience (over context effects), the Literal Salience Model proposes that, in contrast to native speakers whose activation of literal and figurative meaning is a function of familiarity (Giora, 1997, 2003), the literal salience holds true whether or not an idiom is familiar to L2 learners—more specifically foreign language learners, who learn their L2 through formal instruction and live in the country where L2 is not currently used outside the classroom. In a later study, Cieřlicka (2011) further tested the suppression of literal meaning in biased/unbiased contexts by using the *self-paced reading paradigm* and found that a context that is biased towards the figurative meaning is helpful in suppressing irrelevant meanings, but its effects are mediated by the salience of idioms’ literal meanings (see also the *Retention Hypothesis*, Giora, 2002). The role of context in L2 idiom comprehension will be further discussed under two other models (namely the *Heuristic Approach* and the *Idiom Diffusion Model of Second Languages*).

By acknowledging the obligatory activation of literal meanings in L2 idiom processing, the Literal Salience Model suggests that “such techniques capitalizing on a literal analysis of idiomatic phrases might potentially serve as a useful mnemonic (i.e. mental imagery), promoting the meaningfulness of the vocabulary learning activity and enhancing the depth of processing which is essential in ensuring successful vocabulary retention” (Cieślicka, 2011, p. 30, referring to Nation, 1990, 2001). This suggestion aligns well with the cognitive linguistic view that etymological elaboration can harness this natural inclination of learners to seek a link between the literal and the figurative—provided the literal reading is congruent with the idiomatic, figurative meaning (see further discussion in section 2.3).

The two models discussed above provide sound theoretical accounts of online idiom processing in L2 speakers by drawing on the L1 idiom processing models. Setting aside crosslinguistic factors, frequency and familiarity seem to be the main factors underlying the differences between native and non-native speakers’ processing of idioms. So far, many empirical studies have supported the two L2 models with evidence of an overall processing advantage of native speakers over non-native speakers. For example, an eye-tracking study by Siyanova-Chanturia, Conklin, and Schmitt (2011) examined native and non-native speakers’ on-line processing of idioms (used literally or figuratively) and novel phrases in meaningful/supportive contexts. Besides L2 speakers’ overall slower reading compared to native speakers, this study showed that native speakers processed idioms much faster than novel phrases, and that literal and figurative uses of idioms were processed at a similar speed. By contrast, L2 speakers processed idioms and novel phrases at a similar speed, and their processing speed of figurative uses of idioms was slower than that of the literal ones. This suggests that for L2 speakers, the figurative reading of idioms incurred more processing cost than the literal one. To investigate how exactly L2 learners comprehend the figurative meaning of idioms, the next two studies to be reviewed look into the strategies that learners use to make sense of L2 idioms, particularly unfamiliar ones.

2.2.2.2 How do L2 learners comprehend idioms?

The preceding sections were about processing by L2 users of idioms that they had already acquired (but were not necessarily highly familiar with). These online processing experiments

tell us about the differences and similarities between L1 and L2 users in how knowledge of idioms is represented in the mind. We now turn to two comprehension studies and see how L2 learners try to make sense of unfamiliar or unknown idioms.

Cooper (1999) investigated the comprehension strategies used by L2 learners of English. The 18 participants in his study all lived in the US, spoke a range of L1s (i.e., Spanish, Japanese, Korean, Russian, or Portuguese), had different levels of English proficiency and diverse learning backgrounds. Cooper collected the data by using *think-aloud* interviews: each participant was asked to verbalise their thoughts as they attempted to interpret the meanings of 20 frequently used idioms, which were presented on note cards with a written context of one or two sentences. Based on the transcribed interview data that were analysed into T-units (i.e., the minimally terminable unit, or “one main clause plus any subordinate clause or nonclausal structure that is attached to or embedded in it”, Hunter, 1970, p. 4), Cooper identified eight strategies that L2 learners used in interpreting the meanings of the idioms. According to the order of frequency of use, they are *guessing from context* (GC, 28% of the time), *discussing and analysing the idiom* (DA; 24%), *using the literal meaning of the idiom* (LM; 19%), *requesting information* (RI, e.g., an unknown word in an idiom; 8%), *repeating or paraphrasing the idiom* (RP, 7%), *using background knowledge* (BK, e.g., where they learned or heard the idiom; 7%), *referring to an L1 idiom* (L1; 5%), and *other strategies* (OS; 2%). Among the eight strategies, the most often used ones were GC, DA and LM (71% in total). Cooper divided the eight strategies into two groups, preparatory strategies (i.e., RP, RI, and DA) and guessing strategies (i.e., GC, LM, BK, L1 and OS). Based on the explanation and the examples that Cooper (1999) gave for these strategies, *discussing and analysing the idiom or its context without guessing at the meaning* (DA), a so-called “preparatory strategy”, is closely related to the two “guessing strategies”, *guessing the meaning of the idiom from the context* (GC) and *using the literal meaning of the idiom as a key to its figurative meaning* (LM). Furthermore, among all the strategies, GC and LM are the ones that most often led to correct interpretations of the idioms (79 % in total). It should be noted however, that the overall success rate was 56% and this included cases where the participants were familiar with the idioms. In other words, participants’ familiarity with the idioms was not taken into account, which was

likely to affect the results to some extent. In addition, as Cooper (1999) pointed out, the contexts in which the idioms were presented in this study were explanatory (e.g. *By mistake, Kay let the cat out of the bag when she revealed the surprise*) and the results could be different with a non-supportive context. Nevertheless, the most frequently used sources of information turned out to be not just the context in which the idiom was presented, but also the literal meaning of the idiom. That said, according to the comments the participants made in the think-aloud data, they indicated that “a stumbling block in comprehension was often the lack of a clear and close relationship between the literal and figurative meanings of the idioms” (op. cit., p. 244). This seems to suggest that the transparency of the relationship between the literal and figurative meanings, in other words, *semantic transparency*, is a key factor that affects the success of L2 idiom comprehension. Overall, comprehending an unfamiliar idiom for L2 learners is more of a problem-solving process involving the evaluation of a series of possible meanings through trial and error. Therefore, Cooper (1999) summaries this as a **heuristic approach** to L2 idiom comprehension.

Cooper (1999) further suggests that the think-aloud method “can be adapted as a teaching tool” (p. 256), which can be applied to either an entire class or a single student (ibid., referring to Brown, 1996). On a single student basis, the teacher can have the student think aloud about how they interpret the meanings of idioms, and at the same time, guide the student towards a correct understanding by giving hints about the meanings of idioms or suggesting various comprehension strategies such as the ones mentioned above (ibid.). On the other hand, the focus of a think-aloud procedure is to get participants to report the content of their immediate awareness and ongoing thinking rather than to explain their behaviour or recall their thoughts (Cooper, 1999, referring to Olson, Duffy & Mack, 1984). Therefore, a think-aloud method is not only ideal for research that taps into the learners’ ongoing thinking process but also provides a useful mode for L2 idiom teaching practice. These two advantages of the method inspired the research design of the current project and fit very well with its research objectives (see a full account of research objectives in section 2.4 and research design in Chapter 3).

While Cooper (1999) investigated L2 idiom comprehension strategies with L2 learners of English, Liantas (2002) looked into the process of comprehension with L2 learners of multiple

languages, including English, Greek, Spanish, French and German, based on which he proposed the **Idiom Diffusion Model of Second Languages**. According to this model, L2 idiom comprehension consists of two stages: a *prediction* stage and a *confirmation, or replacement, and/or reconstruction* stage. At the first prediction stage, the learner constructs a number of hypotheses to predict the figurative meaning of an idiom. The construction of hypotheses depends on the availability of supporting context and the *conceptual-semantic image distance* of the L2 idiom from the L1 idiomatic expression, i.e., how close or distant an L2 idiom is from its equivalent L1 idiom (Liontas, 2015b, p. 305). In terms of conceptual-semantic image distance, Liontas (1999) differentiates three types of idioms: *lexical-level idioms*, i.e., L2 idioms with identical L1 idioms of the same lexical components and related image of the literal sense (e.g., *foindre les deux bouts* in French = *make ends meet*); *semi-level idioms*, i.e., L2 idioms with slightly different L1 equivalents (e.g., *in den sauren Apfel beißen* in German = *bite the bullet*, literal: ‘bite into the sour apple’); and *post-lexical-level idioms*, i.e., L2 idioms that do not match L1 corresponding idioms either lexically or pictorially (e.g., *echar toda la carne al asador* in Spanish = *put all your eggs in one basket*; literal: ‘to put all the meat on the spit’). Lexical-level idioms are the easiest to understand, for which context seems not necessary for comprehension. By contrast, comprehending post-lexical-level idioms requires much more cognitive effort, which will rely more on the context. On the other hand, L2 learners are limited by the amount of information that is available to them in a given task and by their own abilities to process the information (Liontas, 2002; 2015b). In the absence of context, they rely on the literal analysis of the constituent words of the idiom. At the confirmation, or replacement reconstructive stage, the earlier hypotheses and predictions are confirmed or replaced by the correct interpretation of the idiom, as the learner eliminates unlikely interpretations by focusing on the more relevant contextual constraints (ibid.).

In a later study, Liontas (2015b) summarises four major challenges L2 learners encounter in comprehending idioms, i.e., seeing an idiom without supporting context, not knowing vocabulary in the idiom, being unable to translate the idiom into L1 idiomatic expressions, and non-familiarity with the idiom. These challenges, together with the comprehension strategies identified by Cooper (1999), reveal that L2 learners make use of all kinds of information

available to them during the process of interpretation, i.e., the context where an idiom appears and their prior knowledge of L1 and L2, including lexical, conceptual, as well as experiential background knowledge.

Both the Heuristic Approach and the Idiom Diffusion Model of Second Languages discussed above capture the heuristic nature of L2 idiom comprehension, and identify some important factors affecting the ease and success of the process, namely the existence of supporting context, the degree of transparency of idioms, and crosslinguistic similarity between L1 and L2 (see further discussion of these factors in section 2.3.3).

2.2.2.3 Summary

While native speakers process and comprehend idioms effortlessly and even unconsciously, it often takes non-native speakers a considerable amount of time and effort to interpret idioms (even if they have seen them somewhere before). This difference between native and non-native speakers can be explained by the frequency of their exposure to and thus their familiarity with idioms, factors which are associated with the environment of language learning and use. In L1 acquisition, children are exposed to abundant opportunities to learn words and phrases in natural settings. Learning takes place naturally as the associations between language and external reality are built up, as well as the links between meaning and language form. Idiom acquisition enjoys the same status as learning other words and phrases. As native speakers are exposed to and use idioms frequently, many of these expressions are acquired in context as they grow up, thus creating representations of these idioms as whole units (i.e., ‘idiom entries’) in their mental lexicon, with the figurative meanings ‘entrenched’ in their memory and thus more salient to them. This in turn results in overall fast recognition and more direct retrieval of idioms during online processing compared to non-native speakers.

As the review of the relevant literature above has shown, while there are common patterns in L1 and L2 online processing (e.g., Titone et al., 2015), L2 learners, especially foreign language learners, tend to have less exposure to idioms and consequently develop many fewer idiom entries in their mental lexicon (Abel, 2003). The literal meanings of many idioms thus remain more salient than the figurative ones (Cieślicka, 2006). L2 learners naturally resort to the literal meaning of the idiom’s lexical components and relevant conceptual knowledge (if available)

to make sense of the figurative meaning of the idiom. This is where an analytic approach like etymological elaboration can come into play.

When encountering an unfamiliar idiom, L2 learners make use of a variety of strategies to decipher its meaning. They generally take a heuristic approach, which involves constructing, testing and reconstructing hypotheses by using contextual information, analysing the literal meanings of its lexical components, and referring to similar L1 expressions. Even so, the use of various strategies does not guarantee successful comprehension, which renders idioms a stumbling block for L2 learners to achieve full mastery of an L2 or foreign language. Challenges often come from the semantic opacity of the idioms, lack of supporting context, crosslinguistic differences between L1 and L2, and problems also related to learners' own linguistic competence and inferencing abilities. All of these leave a huge gap for L2 teaching to fill. In the next section, we will review studies about L2 teaching approaches to idioms and factors affecting L2 idiom learning, such as L1 transfer, semantic transparency of idioms and L2-proficiency-related factors.

2.3 Enhancing second language idiom learning

The foregoing sections of this chapter have clarified several background notions and issues about idioms and idiom processing. Section 2.1 has explained what idioms are, by reviewing various definitions and features of idioms, as well as by considering classifications of idiomatic expressions from different linguistic approaches. This section anchored our focus on figurative idioms, i.e., conventional figurative expressions. Section 2.2 has explained how L2 idiom processing differs from L1 idiom processing, by reviewing how idioms are processed in native speakers and non-native speakers, and has identified the gap for L2 idiom learning and teaching.

This section zooms in on this central topic—L2 idiom learning and teaching. I will review pedagogy-oriented studies of L2 idiom learning, in order to answer three questions:

- (1) What is the current status of idioms in language teaching and learning, regarding the motivations of and general approaches to L2 idiom teaching? (see 2.3.1)
- (2) What have applied cognitive linguists proposed as methods for enhancing L2 idiom learning? (see 2.3.2)

- (3) What factors have been identified in previous studies as affecting L2 idiom learning?
(see 2.3.3)

After presenting a brief overview of pedagogy-oriented research on L2 idioms learning in general, the focus of this section will be on the latter two questions, with a critical review of the CL-informed teaching approaches and a survey of factors that affect the learning of idioms in a second language. These considerations lead directly to the research objectives and questions of the current study (section 2.4).

2.3.1 Idioms in L2 teaching and learning

Why then should we teach idioms, and how can we teach idioms? The following two sections will provide brief answers to these two questions as an introduction to general instructional approaches to idioms.

2.3.1.1 Why teach idioms?

Considering their low frequency per item based on corpus counts, it makes sense that many scholars suppose that “idioms merit less teaching time” (Grant, 2007, p. 181). However, as discussed earlier in section 2.1.3.4, evidence from corpus research has shown that idioms, as a class, are ubiquitous in natural native-speakers’ discourse, and that they serve essential discourse functions, such as description, explanation and evaluation (see also O’Keeffe, McCarthy, & Carter, 2007; Simpson & Mendis, 2003). Moreover, idioms, as well as other types of idiomatic expressions such as collocations (in relation to the narrow and broad senses of idiomaticity as discussed in section 2.1.1), play an important role at different levels of language development and in various spheres of life. As part of figurative language, they are used as a valuable teaching medium in the classroom, to explain, clarify, summarise, evaluate, and so on (Littlemore et al., 2011). For instance, when evaluating a learner’s response to a proposed question, the teacher may say “(Correct)... Now we are *on the same wavelength!*” (example from the author’s own classroom experience). If the idiom *on the same wavelength* is not understood by the student, they may not fully appreciate the teacher’s evaluative feedback. As one type of formulaic language, idioms also play a role in “socializing novices to social dimensions such as politeness, hierarchy, and social identities including social roles and

statuses, and relationships” (Burdelski & Cook, 2012, p. 173; see also Wray, 2002).

The multifaceted nature of idioms is perhaps also reflected by the use of multiple labels in language research, e.g., figurative language, idiomatic expressions, and formulaic sequences (see the definition below). At the same time, these different labels highlight different aspects of the importance of idiom acquisition.

As an element of figurative language, some scholars view idiom acquisition as an integral part of the acquisition of *figurative competence*, i.e., the ability to comprehend and use figurative language (Cieślicka, 2006; 2015; Levorato, 1993), which is closely linked to cognitive development in general (Levorato, 1993). Similarly, some scholars include it in the development of *metaphoric competence*, i.e., the knowledge and ability to use metaphor (Littlemore & Low, 2006b; cf. *sociocultural competence*, in Bachman, 1990), as part of more general *communicative competence* (Hymes, 1971), as well as part of a complex problem-solving ability (Pollio & Smith, 1980). Lontas (2017), placing a special emphasis on the development of idiomaticity in second languages, introduces the term *idiomatic competence* as “the ability to understand and use idioms appropriately and accurately in a variety of sociocultural contexts, in a manner similar to that of native speakers, and with the least amount of mental effort” (p. 6; see also Lontas, 2015a), which “empowers learners to use language in socially responsible ways” (p. 8, Lontas, 2017).

A more common perspective is to classify idioms as a type of *formulaic sequence*, defined by Wray (2002) as “a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar” (p. 9). Such formulaic sequences have been argued to have an overall processing advantage in the native language relative to those in a second language and to novel expressions in the same language (e.g., Siyanova-Chanturia, Conklin, & Schmitt, 2011; Wray, 2002).

Some researchers also define formulaic sequences as “matching a single meaning or function to a form, although that form consists of multiple orthographic or phonological words” (Martinez & Schmitt, 2012, p. 299) and term them *multiword expressions* (MWEs; e.g.,

Kremmel, Brunfaut, & Alderson, 2017). These terms are all compatible with the current definitions of idioms (see previous section 2.1.2). From this point of view, it has been widely acknowledged that full mastery of formulaic sequences like idioms enhances proficiency, especially fluency in a second/foreign language (Ellis, 2012; Nattinger & DeCarrico, 1992; Weinert, 1995; Wray, 2002). For comprehension, the mastery of MWEs is likely to affect the accuracy and speed of L2 reading and listening. For production, the use of MWEs can increase the fluency and perceived lexical richness of learners' performance. In recent years, positive research evidence has been found for the role of MWEs in measures of L2 learners' oral proficiency (e.g., Boers, Eyckmans, Kappel, Stengers, & Demecheleer, 2006; Qi & Ding, 2011), reading comprehension (e.g., Kremmel et al., 2017), and writing performance (e.g., Li & Schmitt, 2009; Siyanova-Chanturia, 2015a). The importance of MWEs has started to be recognised in listening studies as well (e.g., Brunfaut & Révész, 2015).

Essentially, acquiring L2 idioms is a crucial step towards advanced-level or near-native language proficiency. As Cowie and Mackin (1975) state, “[f]amiliarity with a wide range of idiomatic expressions and the ability to use them appropriately in context are among the distinguishing marks of a native-like command of English” (p. vi). While this is true, learning idioms is infamously difficult. It is commonly observed that L2 learners have difficulty in understanding idioms. For instance, in a study of international students' comprehension of university lectures, Littlemore et al. (2011) asked 20 students to identify words or MWEs that they found difficult in four lecture transcripts. Their results showed that about 40% of the items reported to be difficult, despite consisting of familiar words, were metaphorical, among which are idioms, e.g., *melting pot*, *major shakeup*, and phrasal verbs, e.g., *lay down*, *stem from*. Moreover, when the students misinterpreted the metaphorical expressions, they were only aware of having a problem with about 4% of the cases. In a study of L2 reading comprehension, Martinez and Murphy (2011) tested 101 Brazilian learners of English with two texts that were composed of the same 2000 most frequent English words in the British National Corpus (BNC), with one text containing idiomatic MWEs (e.g., *over the hill*, *down to earth*, *it's about time*, *by and large*) and the other without. The comprehension test results showed that 98% of the learners scored significantly lower in the texts containing MWEs. Further, a comparison

between learners' test scores and their self-reported measures show that about 69% of the participants overestimated their comprehension of the texts with idiomatic MWEs. Such a marked gap between actual comprehension and self-reported comprehension of idioms was also found (in 63.56% of the participants) in a recent reading study with L2 middle-school Korean learners of English (Park & Chon, 2019). All these studies demonstrate the difficulty that metaphorical multiword expressions pose to non-beginner L2 learners and that learners are sometimes not even conscious of this difficulty. Besides problems of comprehension, it is also not uncommon to find that L2 learners tend to avoid using idioms (Yorio, 1989), although studies of highly advanced L2 learners, i.e., bilingual speakers of English, show that they do not seem to shy away from idioms, possibly due to the unavoidable exposure that results from living in an English-speaking country (e.g., Irujo, 1993; see also Laufer, 2000).

Comprehension issues often result from the semantic opacity and figurativeness of the idioms, from the lack of supporting context, from crosslinguistic differences between L1 and L2, as well as from learners' own linguistic competence and inferencing abilities (see previous discussion in section 2.2.2). When it comes to appropriate use of idioms, the learning task is even more demanding, since it includes acquiring knowledge of forms, which entails rules of transformation and patterns of collocation (see previous discussion in section 2.1.3.1), as well as grasping knowledge of appropriate pragmatic use (see section 2.1.3.3). On the whole, the factors underlying learning difficulties come both from the learners' backgrounds and from the very nature of idioms (see detailed discussion later in section 2.3.3). All these challenges and difficulties further suggest that L2 idiom learning, especially in a foreign language education setting, would benefit from a certain amount of explicit instruction.

Furthermore, leaving idioms to learners' own devices has its own risks, as research evidence mentioned above shows that L2 learners tend to overestimate how much they understand multiword expressions (Martinez & Murphy, 2011) and metaphorical expressions (Littlemore et al., 2011), including idioms (Park & Chon, 2019) that are made up of high-frequency words but where the overall meaning is non-compositional. Without being overtly aware of the actual existence of idioms, learners are more likely to overlook or misunderstand idioms in L2 reading and listening. To anticipate the results to be presented in later chapters, such phenomena were

also reflected in EFL learners' performance and comments in the current study. For example, one participant reflected on one common idiom that she had just learned by asking "Do native speakers really use such expressions [i.e., idioms]?" This also shows that the actual use of idioms in native-speakers' discourse could be underestimated by learners if they do not have sufficient exposure to L2 or have not received explicit teaching.

In addition, it has also been argued that "regardless of how broadly or narrowly one defines 'idiom', it is through the lexicalised phraseologies (that idioms are generally considered to be) that language expresses its resilience and color" (Liontas, 2017, p.7, referring to Correli, 2006). Along with this, many idioms have deep roots in the cultural and historical knowledge of a language (Smith, 1925). Tracing back to the origins of such expressions offers a good opportunity for learners to learn the culture of the language and of the people who speak it, e.g., sailors at sea (*showing the ropes*), hunters with their dogs (*barking up the wrong tree*), cooks in the kitchen (*putting things on the back burner*), boxers in the boxing ring (*throwing in the towel*), and so on. Revealing the underlying metaphors can also help them understand how language and thoughts are organised in the mind (see further discussion under the CL-informed approach). In this regard, avoiding the teaching of idioms does a great disservice to L2 learners.

So far, we have briefly reviewed the importance of idioms and idiom knowledge, and the challenges and difficulties of idiom learning. To sum up, there are four main reasons why we should teach idioms:

- 1) Idioms play an important role in language communication. They are ubiquitous in language and serve essential discourse functions.
- 2) Idiom acquisition is an integral part of second language acquisition, and mastering idioms bridges the gap between native speakers and non-native speakers.
- 3) Idioms can be difficult to learn in a second language. Mastery of comprehension and/or use of idioms requires some explicit instruction. Explicit teaching may also help raise awareness of idioms and facilitate learners' recognition of idioms when they encounter them, factors which in turn enhance the learning of idioms.
- 4) The etymology of idioms bears traces of the development of a language and the culture(s) in which it is used. Revealing the knowledge of idioms' origins invites cultural learning

into the classroom and may enrich language teaching and promote further independent learning.

Taking it further, many scholars (e.g., Irujo, 1986b; Littlemore & Low, 2006a; Lontas, 2015a, 2017; Simpson & Mendis, 2003) advocate that idioms should be incorporated into the L2 curriculum. It is argued that idiom learning should not be postponed until learners reach advanced levels of L2 proficiency (Park & Chon, 2019; see also Irujo, 1986a; Laufer, 2000). However, this does not mean that idioms should be included in classroom practices or textbooks without careful consideration. The amount of time that should be spent teaching different types of idiomatic expressions, the proficiency levels of students to whom they should be taught, and the methods that should be used all remain ongoing research topics in SLA (see more discussion in the following sections).

In support of the initiative of including idioms in the L2 curriculum, Lontas (2017) advances a strong rationale for teaching them in the classroom. He sets out five reasons very clearly, as follows:

Reason 1: “[I]dioms help learners to encounter and understand the workings of natural human language; that is, they help them to gain a deeper knowledge of the creative expression of human thought and language development over time.”

Reason 2: “[L]earners can go beyond the literal meaning of idioms and see the pivotal role that context plays in the understanding of idiomatic expressions.”

Reason 3: “[R]equiring learners to produce idioms in ways that native speakers use them enhances learners’ mastery of them, facilitating the building and mapping processes of idiom internalization.”

Reason 4: “[I]dioms afford learners the opportunity to examine their own mental images associated with idiomatic phrases and the conceptual metaphors mediating their figurative meanings.”

Reason 5: “[T]he study of idioms in the classroom can help the SLA profession to build a systematic program for the development of idiomatic competence in second language learners.”

(Lontas, 2017, pp. 9-16)

These reasons cover the comprehension (Reasons 1 & 2) and production of idioms (Reason 3) in association with how to teach them. As previously discussed in section 2.2.1, the first hurdle for L2 learners to overcome is difficulty in comprehension; the challenge of production follows later. In line with some L1 and L2 idiom processing models (Abel, 2003; Cutting & Bock, 1997; Gibbs et al., 1997; see sections 2.2.2.1, 2.2.1.3, and 2.2.1.4), Reasons 1 and 4 pinpoint the role of conceptual knowledge of the idiom's constituents in idiom comprehension and internalisation, and affirm the benefits of exploring etymological motivations of idioms in facilitating L2 idiom learning, which is at the core of the instructional approach investigated in the current project (see section 2.3.2). Ultimately, the SLA profession is challenged to propose "a systematic, theoretically informed program for developing idiomatic competence in L2 learners that is based on meaningful, authentic idiom use in the classroom and beyond" (Liontas, 2017, p. 5). This would involve not just including idiomatic expressions in language pedagogy but also in language testing. The five reasons listed here serve as a conclusion to the current discussion and a prelude to the next section on approaches to L2 idiom teaching.

2.3.1.2 General trends and approaches to L2 idiom teaching

While idiom acquisition is an integral part of the development of formulaic language, which makes a substantial contribution to second language development in general, L2 idiom teaching and learning has only started to receive considerable attention since the 1990s.

Various approaches to L2 idiom learning have emerged from different areas of interest in the field of second language acquisition. First, with the development of psycholinguistics and online measurement tools that are more sensitive than those that had previously been available, researchers have started to gain better understanding of the mechanisms involved in the processing of idioms and other types of formulaic sequences, and this may in turn inform L2 teachers of better ways of teaching (Siyanova-Chanturia & Martinez, 2015). Alongside this are advances in corpus linguistics and corpus tools, which have been found to facilitate the selection of items for teaching and for the development of materials. Most important is the emergence of cognitive linguistics and the Conceptual Metaphor Theory in the late 1980s, which has sparked a new wave of interest in idiom teaching research. This is because, according to this school of thought, idioms can be learned in a systematic way and are not 'dead metaphors'

that can only be learned by rote (Boers, 2011). In the meantime, the focus of vocabulary teaching research in general has gradually shifted from single words to multiword units, with idiomatic expressions such as idioms starting to receive more attention in L2 pedagogy.

While a large number of pedagogy-oriented researchers treat idioms as a marginal type of multiword expression (MWE)⁸, there are also many teachers and researchers who approach idioms, i.e., conventional figurative multiword expressions, as a type of linguistic unit worthy of study in their own right (e.g., Boers & Lindstromberg, 2008b; Boers, Demecheleer, & Eyckmans, 2004b; Boers, Piriz, Stengers, & Eyckmans, 2009; Cooper, 1998; Grant & Nation, 2006; Lennon, 1998; Liantas, 2017; Simpson & Mendis, 2003). A representative branch of this latter group are advocates of cognitive linguistics (CL), including Boers and his colleagues. Since this approach is of special interest to the current study, I will review the CL-informed/inspired approach separately in the next section. The present section briefly discusses some other strands of idiom teaching, regarding item selection and material design, teaching methods and strategies, as well as the general trends of practices as reflected in the teaching of MWEs or formulaic sequences. The aim here is to provide an overview of the status quo of L2 idiom teaching, which also helps contextualise both the current study and the targeted CL-inspired teaching approach with respect to the existing research literature.

Given the large number of idiomatic expressions in a language and the complex nature of idioms, the selection and organisation of idioms for teaching and learning, and the subsequent material design, are amongst major concerns for language teachers (Cooper, 1998; Irujo, 1986b; Liantas, 2017; Yorio, 1980). Corpus-based research is taking a lead in this area, where the frequency—a proxy for the usefulness of the items—and authenticity of the teaching materials are the most common criteria agreed upon by scholars (e.g., Grant, 2003; Liu, 2008; Simpson & Mendis, 2003). As introduced in section 2.1, Grant and Bauer (2004, based on Grant, 2003) made a valuable initial attempt at classifying idiomatic MWUs for different learning and teaching purposes (see a detailed review in section 2.1.3.2). Based on a classification of three

⁸ In this study, ‘multiword expression’ (MWE), ‘multiword unit’ (MWU, Grant & Bauer, 2004), ‘formulaic sequence’, as well as ‘idiomatic expression’ are terms used interchangeably, all in the broad sense of idiomaticity (see detailed discussion in section 2.1.1).

types of idiomatic expressions (i.e. ‘core idioms’, ‘figuratives’ and, ‘ONCEs’) according to their compositionality and figurativeness, and a corpus search, Grant (2003) made a comprehensive list of non-compositional and non-figurative ‘core idioms’. She also suggests that different types of MWUs deserve different amount of teaching time and require different types of instruction and learning strategies (see further below).

Other examples of corpus-driven studies in this area include Grant (2007), Liu (2003), and Simpson and Mendis (2003). All three studies give clear criteria for the identification of idioms, and provide suggestions for ensuring that instructional materials reflect how the idioms are used in natural discourse. It needs to be noted, though, that the variety and range of the selected idioms vary according to researchers’ prioritised registers and criteria for idiom identification, as well as the choice of corpora their search is based on. For example, Simpson and Mendis (2003) focused on idioms in academic speech, by using a specialised academic corpus (i.e., the Michigan Corpus of Academic Spoken English), and found that idioms exist among a wide range of academic speech events, performing a variety of significant pragmatic functions. Liu (2003) compiled lists of frequently used ‘idioms’ (including phrasal verbs and other fixed expressions) in contemporary spoken American English, by searching three spoken American corpora. Taking this approach further, Grant (2007) assessed idioms (i.e., ‘figuratives’, according to Grant & Bauer, 2004) in both British and American spoken discourse by comparing different corpora. It is noteworthy that although corpus research provides a powerful tool for selecting items, it has its own limitations, such as the size of a corpus and the rigour of criteria for idiom identification, as noted above. As Liu (2003) pointed out, many important idioms (e.g., *Let’s call it a day*) may not be represented or captured by a specific corpus search, and it is advised that other resources should be consulted for compiling textbooks and other teaching materials.

With regard to useful resources beyond corpora and specialised dictionaries, there have been a number of practical workbooks dedicated to idiom learning, such as *Working with English Idioms* (Peaty, 1983), *Idioms at Work* (McLay, 1987), *Idioms Organizer* (Wright, 1999), *English Idioms and their Cultural Origin* 英语习语及其文化源流 (李玉平, 2008), and *Words*,

words, words 咬文嚼字学英语 (李炜, 2008). These books introduce learners to sets of idioms, featuring different key aspects of idiom knowledge. Earlier books are organised by the grammatical structures (e.g., Peaty, 1983) and pragmatic functions of idioms (e.g., McLay, 1987), and more recent ones are organised by the metaphors, themes, keywords, or origins of idioms (e.g., 李炜, 2008; 李玉平, 2008; Wright, 1999; see more discussion on organising idioms for teaching in Liu, 2008 and those under the CL-informed approach in section 2.3.2). Moreover, with the development of the internet, a growing number of learning tools have emerged online and provide resource-rich information for idiom learning in multimedia modes such as videos with subtitles, e.g., *IdiomsTube*⁹.

Over the years, a variety of idiom teaching strategies, activities, and exercises have been suggested in a range of publications (e.g., Boers & Lindstromberg, 2008b; Cooper, 1998; Grant, 2003; Irujo, 1986b; Lennon, 1998; Liu, 2008; McPherron & Randolph, 2014; Simpson & Mendis, 2003). Common practices include identifying idioms in discourse contexts; guessing the meaning of idioms by using contextual information; taking quizzes and playing games with idioms, which involves guessing the meaning or the origins of idioms; comparing idioms with learners' L1 expressions; recalling idioms by filling blanks in a passage with parts or the whole of the idioms; making up sentences or writing passages by using idioms; practising idioms in communicative activities such as role play and storytelling; searching for idioms online, etc. These practices, combined with elaboration and feedback from the teacher and other resources, cater to the uptake of different aspects of idiom knowledge, i.e., their meaning, their forms, and their use. In general, there have been limited empirical studies regarding specific practices outside those of the CL-informed approach (e.g., Khonbi & Sadeghi, 2017, on teaching mode; Steinel et al., 2007 on learning and testing directions). This review will not go further into details here, but some of these studies will be discussed in section 2.3.3 in a review of the factors affecting L2 idiom learning. A comprehensive introduction of the macro-level pedagogical strategies and micro-level teaching activities of idioms can be found in Liu (2008)

⁹ IdiomsTube (www.idiomstube.com) is a web-based free app designed by Phoebe Lin at the Hong Kong Polytechnic University. It promotes idiom learning through watching authentic subtitled English videos accompanied by generated exercises and feedback, which caters to both self-directed learning and classroom teaching.

and a survey report of the usefulness of common teaching activities can be found in McPherron and Randolph (2014).

Besides the strategies and activities, some teaching principles have also emerged from various approaches and practices discussed above. They can be summarised as follows:

- 1) Regarding learners of different proficiency levels, teaching focus and activities should be adjusted accordingly (Cooper, 1998; Lennon, 1998). For beginners, the focus of learning should start from recognition (Yorio, 1980), which helps learners first identify and notice idioms. Initial steps of teaching can start from familiarising learners with this type of language phenomenon (McCarthy, 1998) by discussing and defining idioms and showing them in context (Cooper, 1998; Simpson & Mendis, 2003).
- 2) Regarding different types of idioms, a selective approach should be taken to teaching and learning (Grant, 2003; Simpson & Mendis, 2003). For a limited number of very opaque (difficult) idioms (cf. ‘core idioms’, Grant & Bauer, 2004) that have completely lost their origin and literal use (e.g., *by and large*), they can be learned as a long word, via a *holistic* approach. For the majority of idioms, whose conventional figurative meaning can be traced back to an original literal use, helping learners understand and internalise the expression by making use of etymological, conceptual, and cultural knowledge would be an effective strategy (Boers et al., 2004b), i.e., an *analytical* approach.
- 3) Regarding the development of instructional materials, these should reflect authentic (interactive) use in natural language (Liontas, 2017; Liu, 2003; McCarthy, 1998, etc) and should also be tailored to learners’ needs (Simpson & Mendis, 2003).

Overall, the major trends in L2 idiom teaching practices and principles have much in common with those of teaching formulaic sequences/MWEs (i.e., idioms in a broader sense) as reflected in experimental and empirical research for the past 20 years. Three major trends, as summarised by Boers and Lindstromberg (2012) are: (1) drawing learners’ attention to formulaic sequences as they are encountered, (2) stimulating lookups in dictionaries and the use of corpus tools, and (3) helping learners commit particular formulaic sequences to memory. More research attention has hitherto been paid to the receptive knowledge of MWEs compared to the productive knowledge. This is partly due to the general pattern of language acquisition whereby

comprehension runs ahead of production, so that learners' initial learning efforts mostly go to dealing with recognition and comprehension of the expressions. It has also been widely acknowledged that the development of formulaicity in an L2 is a slow and complex process (Boers & Lindstromberg, 2012; Kuiper, Columbus, & Schmitt, 2009; Qi & Ding, 2011). As part of this more general development regarding formulaic sequences, L2 idiom instruction has also been moving in the three directions outlined above, and has followed the same pattern with regard to the relationship of receptive and productive skills. However, given its particular focus on the uptake of meaning, it has largely oriented towards the third trend mentioned above. This is exemplified in the empirical studies of the prominent CL-informed instruction approach, which has been found particularly effective in fostering the comprehension and retention of figurative meanings of idioms. This approach is what we will now turn to.

2.3.2 Cognitive-linguistics-informed approach to L2 idiom learning

2.3.2.1 Introduction

Cognitive linguistics (CL) has become a popular branch of linguistic study since the 1980s, marked by the publication of the seminal book *Metaphors We Live By* (Lakoff & Johnson, 1980) and subsequent work (Johnson, 1987; Lakoff, 1987; Langacker, 1987, 1991). One of the central CL theories, Conceptual Metaphor Theory (CMT; Lakoff & Johnson, 1980) proposes an illuminating view that metaphors are not just a rhetorical device of language, but more profoundly an omnipresent cognitive mechanism that shapes our language, thought, and behaviour. Over the past 40 years, CL has generated a proliferation of research that has not only explored intra-linguistic structures and cross-linguistic differences, notably in semantics—hence ‘cognitive semantics’, but also inspired a new approach to second language learning and teaching.

With its exploration of the conceptual motivation of linguistic structures, cognitive linguistics, featuring CMT and Image Schema Theory (Gibbs & O'Brien, 1990; Johnson, 1987; Lakoff, 1987), shows great potential for second/foreign language pedagogy (Achard & Niemeier, 2004; Boers & Lindstromberg, 2008a; De Knop, Boers, & De Rycker, 2010; De Knop & De Rycker, 2008; Littlemore, 2009; Littlemore & Low, 2006a; Verspoor & Boers, 2013). This is because

the presentation of segments of language as motivated is likely to enhance language comprehension, retention, and cultural awareness (Boers & Lindstromberg, 2006). Evidence for this has been shown in a substantial number of studies on the comprehension and retention of L2 vocabulary, including polysemous single words, such as prepositions and verbs (e.g., Beréndi, Csábi, & Kövecses, 2008; Boers, 2000; Boers & Demecheleer, 1998; Cho, 2010; Li, 2009; Lindstromberg & Boers, 2005; Verspoor & Lowie, 2003), and metaphorical multiword expressions, such as idioms and phrasal verbs (e.g., Beréndi et al., 2008; Boers, 2000, 2001; Boers, Demecheleer & Eyckmans, 2004a, 2004b, Boers, Eyckmans, & Stengers, 2007; Condon, 2008; Gao & Meng, 2010; Guo, 2007; Li, 2009; Lindstromberg & Boers, 2005; Kövecses & Szabó, 1996; Yasuda, 2010).

Given the origins of CL, idioms have been a popular target of investigation from a CL perspective and have received a considerable amount of pedagogical interest in the last 30 years, since Kövecses and Szabó's (1996) first experiment. This is in contrast to earlier approaches when idioms were considered as 'dead' or fossilised expressions (Boers, 2001, 2011) and treated merely as an embellishment of language (Lennon, 1998).

A CL-informed teaching approach to idioms has been supported by three main theories. In the following section, I will first describe the three theories, i.e., the *Conceptual Metaphor Theory*, the *Dual Coding Theory*, and the *Levels of Processing Theory*, in terms of how they support the teaching approach. After that, I will review the development of idiom teaching within this approach based on experimental studies. Specific strategies and research methodologies will be evaluated, with a special focus on the more recently proposed strategy, i.e., etymological elaboration.

2.3.2.2 Theoretical rationale for the CL-informed approach

Conceptual Metaphor Theory (CMT) was first proposed by Lakoff and Johnson in *Metaphors We Live By* (1980). The main idea of the theory is that we conceptualise one mental domain by virtue of another, and that the concepts governing our thoughts structure our perception, action, and relationship with others in the world (Lakoff & Johnson, 1980; Lakoff, 1993). More specifically, people tend to comprehend and describe intangible domains of experience by seeking correspondences with concrete familiar domains; and the systematic

cross-domain mapping from concrete *source domains* onto abstract *target domains* is called conceptual metaphor (see a detailed elaboration of conceptual mapping in section 2.2.1.4).

Conceptual metaphors are grounded in our bodily experiences, and are manifested in language, notably in idioms, i.e., conventionalised figurative expressions. The conceptual mappings are “responsible for much of the meanings of idioms, the motivated nature of idioms, and the various kinds of cognitive mechanisms (like metaphor, metonymy, conventional knowledge) on which idioms are based” (Kövecses & Szabó, 1996, p. 327). For example, an often-cited conceptual metaphor ANGER IS HEATED FLUID IN A CONTAINER can be seen as conceptually motivating idioms such as *blow off steam*, *flip your lid*, and *hit the ceiling*. Within the conceptual metaphor, there exists a coherent organisation underlying such expressions (Lakoff, 1987). In the ANGER metaphor, features of the source domain HEATED FLUID IN A CONTAINER (cf. BODY IS A CONTAINER) can be systematically mapped to the target domain ANGER, as reflected in the entailments: “intense anger produces steam”, e.g., *blow off steam*; “intense anger produces pressure on the container”, e.g., *flip your lid*; and “when anger becomes too intense, the person explodes [...] and parts of container [i.e., the person] go up in the air”, e.g., *hit the ceiling* (Lakoff, 1987, pp. 381-389). Because of such systematic connections, the conceptual metaphors allow us to use expressions from one domain to talk about another. Supporting evidence of the psychological reality of conceptual metaphors and of the claim that they form the conceptual basis of idioms has been found in a series of studies by Gibbs in the 1990s (Gibbs, 1990, 1994; Gibbs & O’Brien, 1990).

Further, it is assumed that omnipresent cross-domain mappings may function as a tool in the comprehension of another language. This has inspired a new approach to L2 vocabulary learning, reviving ‘fossilized’ chunks of language like idioms, which were deemed to be unteachable but are actually explainable with reference to conceptual metaphors, linking the abstract figurative meaning of idioms to their concrete literal uses. For example, taking a back seat in a vehicle implies that someone else will be at the steering wheel and will thus likely be in control. By analogy, then, the meaning of the idiom *take a backseat* to refer to one’s non-determining role in a project or activity makes sense against the backdrop of a more generic Conceptual Metaphor LIFE IS A JOURNEY (Lakoff & Johnson, 1980), according to which

projects or activities are likened to journeys. In other words, the idiomatic meaning is motivated. If this indeed results in a clearer link between the form and the idiomatic meaning of the expression, it may be expected to facilitate acquisition (e.g., Steinel, Hulstijn, & Steinel, 2007). In addition, conceptual metonymy (see e.g., Littlemore, 2015)—systematic mappings within one conceptual system—and conventional knowledge are also involved in the motivating mechanism (Kövecses & Szabó, 1996).

As Boers (2011) points out, despite methodological concerns (see further discussion in the following section), the collective evidence from the empirical studies suggests rather strongly that making learners aware of the source domains or literal origins of idioms (in one way or another) is helpful for their retention of the expressions. This is because informing learners about the conceptual motivation not only renders the idiomatic meaning more transparent to them, but also stimulates a mental image associated with the literal use that makes the expression more durable in memory (e.g., Boers, 2000; Boers et al., 2004b). Behind this assumption lies the second supporting theory of the CL-informed approach, Dual Coding Theory.

Dual Coding Theory (DCT) is a theory of cognition and memory proposed by Paivio (1986, 1971). The general assumption of the theory is that there are two subsystems in human cognition: one is specialised for representing and processing verbal information, i.e., the language system, and the other is specialised for the representation and processing of nonverbal information, which is often referred to as the imagery system. The two systems are assumed to be distinct in nature and can function independently, and at the same time, their functions are interconnected, with activity in one system liable to activate that in the other (Paivio, 1986). Under this hypothesis, DCT consists of a wide range of assumptions and implications applying to many cognitive phenomena including: mnemonics, problem-solving, bilingual processing, metaphor comprehension, as well as second language learning (Paivio & Desrochers, 1981; also see Paivio, 2006, for further theoretical development). An important assumption about memory is that when both the language and imagery systems are involved in encoding, this benefits retrieval of information, as the information will have been stored in dual modes. That is why concrete words are easier to recall than abstract ones: Concrete concepts are represented

dually (word plus mental image), whereas abstract concepts are not. During processing and comprehension, the meaning of concrete words can be activated with both referential (i.e., via concrete object or entity) and verbal-associative networks, while abstract words rely largely on the verbal-associative network, particularly on verbal context cues (Paivio, 1986, p. 123). If this can be extended to idioms (though not a view held by the theory), presumably a long word with abstract meaning, then they can be rendered concrete or imageable by reviving their literal meaning, which in turn makes them more memorable (thanks to the dual coding). Likewise, with the more concrete sense of the literal underpinning, interpretation of the idiomatic meaning may not necessarily rely solely on the verbal contexts.

Overall, for second language learning, the DCT suggests that “language-learning strategies based on the systematic use of referent objects, pictures, activities, and mental imagery would be especially effective in promoting learning” (Paivio, 1986, p. 257). The DCT is compatible with the CMT with regard to their implications for L2 idiom learning, as they both highlight the role of mental imagery in conceptualization, i.e., concreteness and mental imagery are related, and both make learning easier. Thus, one major assumption of the research in this thesis is that the mental imagery associated with the literal underpinning of idioms facilitates L2 idiom interpretation and strengthens memorization. Moreover, of particular relevance to another interest of this study, the bilingual version of the DCT, i.e., the bilingual dual coding model (Paivio & Desrochers, 1980), may potentially account for interference effects of L1 on L2 idiom interpretation and learning via a shared imagery system.

Levels of Processing Theory (LoPT) (Cermak & Craik, 1979; Craik & Lockhart, 1972; Craik & Tulving, 1975) is another theory that the CL-informed teaching approach often draws on. According to this theory, we perceive stimuli (e.g., a word) via a series of processing stages (i.e., depth of processing) from analysis of sensory features (e.g., form and sound) to matching stored information (i.e., pattern recognition) and obtaining meaning; and deeper processing that involves a greater degree of semantic or cognitive analysis and creates memory traces that are “more elaborate, longer lasting, and stronger” (Craik & Lockhart, 1972, p. 578). It also suggests that if the stimulus matches our world knowledge and ‘semantic memory’ (Tulving, 1972)—in other words, the information is meaningful to us—then it yields better memory performance

(Craik & Tulving, 1975). Overall, “[r]etention is a function of depth, and various factors, such as the amount of attention devoted to a stimulus, its compatibility with the analyzing structures [determined by existing knowledge], and the processing time available, will determine the depth to which it is processed” (Craik & Lockhart, 1972, p. 676).

Mental operations—called elaborations—that build rich semantic associations around lexical items are considered ‘deep’ in this model. Connecting the meaning of an idiom to its literal underpinning qualifies as an example of this. The label coined by Boers et al. (2004b) for this particular type of elaboration is *etymological elaboration*. According to Boers et al., practising identification of “the source domains behind the figurative idioms involves a certain degree of cognitive effort [...] probably occurs at a ‘deeper’ level of processing than ‘shallow’ rote learning” (ibid., p. 58), and thus enhances memory.

Note that the term ‘etymological’ in Boers et al. (2004b) was used to refer to the ‘source domain’ (Lakoff & Johnson, 1980) of idioms. It could be interpreted broadly as the original context in which the expression was (and sometimes still is) used literally, and it contains etymological information about an idiom. In this thesis, such etymological information is also referred to as the *literal underpinning*. This is to differentiate the sense of its original use from a superficial summation of the literal meaning of each component word, which could be potentially misleading. For example, learners may misinterpret *wings* in *(be) waiting in the wings* as birds’ wings rather than part of a theatre where actors wait before they appear on the stage.

2.3.2.3 From theories to practice—tenets and techniques

I have given above a theoretical rationale for a CL-informed approach to teaching L2 figurative expressions. To sum up, three tenets can be synthesized as follows (see also Boers, 2011):

- 1) Conceptual Metaphor Theory, showing the relatedness of seemingly unrelated expressions through a shared conceptual metaphor or source domain, provides an organised way for vocabulary learning in general.
- 2) Seeking and recovering the cognitive motivation behind the meaning of L2 idioms facilitates comprehension, which provides a pathway for insightful learning.

- 3) Raising metaphor awareness involves creating mental images of the literal, concrete underpinning of the idioms and stimulates cognitive engagement, which will enhance the retention of idioms.

Under this approach, techniques employed in the instruction of L2 idioms include using conceptual metaphor (Beréñdi et al., 2008; Chen & Lai, 2013; Kövecses & Szabó, 1996; Li, 2009; Pan, 2019), source domains and etymology (Boers et al., 2004b, 2007), pictures and images (Boers, Píriz, Stengers, & Eyckmans, 2009; Ramonda, 2016), as well as directing learners' attention to phonological repetition, i.e. alliteration and rhyme (Lindstromberg & Boers, 2008a, 2008b). The majority of these studies focus on the comprehension and retention of the meaning of idioms, with a few focusing on the formal characteristics or lexical composition of the expressions (e.g. drawing attention to phonological patterns such as alliteration). Table 2.3 is an overview of specific techniques and tasks used in representative empirical studies.

Table 2.3 Teaching techniques proposed under the CL approach

Techniques	Tasks in experimental conditions	Representative studies
Using conceptual metaphors	Informing conceptual motivation by grouping idioms under conceptual metaphors; Identify conceptual metaphors of idioms	Beréñdi, Csábi, & Kövecses (2008); Boers (2000); Chen & Lai (2013); Gao & Meng (2010); Li (2009); Pan (2019); Skoufaki (2008)
Using source domains	Informing conceptual motivation by grouping idioms under source domains; Identifying/hypothesising source domains and idiomatic meanings, based on each other	Boers, Demecheleer, & Eyckmans (2004b); Boers, Eyckmans, & Stengers (2007); Skoufaki (2008)
Using etymological notes or pictures	Informing etymological information; Showing image relating to the etymology	Boers, Píriz, Stengers, & Eyckmans (2009); Guo (2008); Ramonda (2016); Szczepaniak & Lew (2011); Vasiljevic (2015)
Using phonological features	Noticing alliteration or rhyme	Boers & Lindstromberg, (2005); Lindstromberg & Boers (2008a, 2008b)

Experimental and quasi-experimental studies often test the effectiveness of the CL-informed techniques by comparing them with “traditional” methods, e.g., the grammar-translation

approach and rote learning. For example, some studies (e.g., Boers, 2000; Gao & Meng, 2010) compared the effects of grouping idioms under conceptual metaphors with those of grouping idioms under pragmatic and functional themes. Some studies (e.g., Pan, 2019) compared the learning condition of informing learners about the conceptual motivation of idioms grouped under conceptual metaphors with the control condition of informing learners about L1 translations of idioms grouped under functional themes. One study in Li (2009) compared three learning conditions: learning organised by semantic themes, learning organised by conceptual metaphors, and learning by conceptual metaphors plus images. While the general findings of these empirical studies show positive evidence favouring the CM technique, the range of target idioms used seems to be confined to the exemplar metaphors (and metonymies) concerning universal themes often associated with emotions like anger (Boers, 2000; Kövecses & Szabó, 1996; Gao & Meng, 2010; Li, 2009; Pan, 2019) and body parts like heart and liver (e.g., Hu & Fong, 2010; Kövecses & Szabó, 1996). Further, the general conceptual metaphor technique seems to have limited transferability, because of the influence of cross-cultural and cross linguistic interference from learners' L1 (e.g., Hu & Fong, 2010). This is not surprising in view of the nature of idioms and idiomaticity (see previous discussion in section 2.1) and the fact that idioms often originate from specific experiential domains that are peculiar to the culture surrounding a language, and particular domains may not be equally 'salient' across cultures (Boers & Demecheleer, 2001). For example, 'sailing' is a productive source domain in the English language, possibly due to early maritime exploration, generating idioms, e.g., *miss the boat*, *be left high and dry*, *learn the ropes*, while 'food' could be a more salient domain in French. In common domains such as 'sports', American English has idioms relating to baseball (e.g., *hit a home run*), while British English has idioms relating to cricket (e.g., *hit someone for six*) and horse-racing (e.g., *win hands down*).

Taking cultural variations into account, Boers and his colleagues (2001, 2004a, 2004b) explored various common source domains of a large number of English 'imageable idioms' (i.e., conventionalised figurative expressions in this study). Boers et al. (2004a, 2004b, 2007) proposed 'etymological elaboration' as a mnemonic technique (see a detailed review in the following section). This is a not a completely novel method for idiom teaching or learning, as

many idiom dictionaries (such as Oxford Dictionaries) and idiom workbooks provide etymological information for language users and learners to refer to. However, the usage of such information is particularly emphasised and exploited from a broad and insightful conceptual and cultural pedagogy-oriented perspective under the CL-informed approach, and the usefulness of such information is often overlooked.

As discussed in sections 2.2.2 and 2.3.1, etymological elaboration matches well the needs of L2 learners, who naturally tend to analyse the literal reading of idioms and to spontaneously activate images related to the literal meaning of constituent words (Abel, 2003; Cieřlicka, 2006; Cooper, 1999). This inclination is in contrast with most native speakers, for whom idiomatic meaning tends by default to be accessed directly, without resort to knowledge of the literal meaning (as previously discussed in section 2.2.1). In addition, the sense of metaphoricity or figurativeness may not be obvious to L1 speakers, as they learn the idiomatic meaning as it is. Personal knowledge of the literal underpinning of idioms may vary according to personal literacy and experiences, especially when an idiom originates from an invented literary image, a distant historical event, or a source domain that they have little personal experience with (e.g., someone who does not watch boxing matches may not know *on the ropes* originates from boxing). Besides, not all idioms are traceable to a specific source. This explains the facts that (1) native speakers, including ESOL teachers, may not be fully aware of the etymology or the literal underpinning of certain idioms and may therefore overlook the usage of etymology as teaching tool and (2) their perception of idioms, such as semantic transparency and ease of learning, might vary accordingly and differ from that of L2 learners. Because of these facts, it is necessary to raise ESOL teachers' consciousness of L2 learners' needs and show them that they can help learners appreciate idioms via etymological elaboration. As argued by Cooper (1998), "(t)hrough a teacher's conscious effort, students can receive effective practice in comprehending and producing idioms" (p. 259).

Further, etymological elaboration, or the source domain technique, has also been combined with the use of pictures as a mode of teaching. Inspired by the Dual Coding Theory, relevant studies (Boers et al., 2009, Ramonda, 2016; Szczepaniak & Lew, 2011; and Vasiljevic, 2015) explored the potential of using images of the etymological information to enhance

comprehension and retention of idioms' meaning and form. So far, mixed results have been found in comparing the mnemonic effects of verbal notes with those of pictorial support. As the use of pictures was not of interest in this project, I will not go further into this strand. Instead, I will further investigate the efficacy of etymological elaboration as a means of stimulating *mental* imagery.

The studies listed in Table 2.3 and discussed above suggest that CL-informed approaches to idiom teaching are relatively effective in comparison with other, 'mainstream' approaches. Nevertheless, as noted by Boers (2013), many of these quasi-experimental studies have obvious shortcomings in design and statistical analysis: many are small-scale; some show limited statistical rigor (especially earlier ones) or small effect sizes, and the early studies did not even report effect sizes or even all the information usually needed for the estimation of effect sizes (e.g., standard deviations or other measures of variance); and some results are hard to interpret due to confounding variables. Therefore, replication studies are still necessary to consolidate previous claims. More importantly, much of the research to date has focused on broad issues, such as comparing teaching techniques under different learning conditions and their effects on retention. However, many issues have been overlooked and remain under-researched, including the following:

Issue 1: Due to the quantitative, pre-test—treatment—post-test nature of the quasi-experimental studies that have so far yielded evidence in support of a CL-informed approach, it is not clear to what extent learners spontaneously resort to the kind of figurative reasoning which CL interventions aim to promote. Neither is it clear how learners actually experience the usefulness of CL-style explanations about the links between literal and figurative meaning.

Issue 2: Only minimal attention has so far been given to evaluating the effectiveness of CL-style explanations at the levels of individual items and learners. There are various factors that are likely to influence the effectiveness of a particular instructional approach to L2 idioms, such as the degree of semantic transparency of the idioms, and obstacles to idiom interpretation such as L1 transfer and cultural differences. It is likely that some (types of) idioms lend themselves better to a CL-inspired approach

than others. On the other hand, learner traits such as proficiency level are often neglected in previous studies, even though these are likely to influence the efficacy of teaching.

Issue 3: With very few exceptions, the quasi-experimental studies to date have pitted one or the other CL-inspired ways of presenting idioms against a ‘mainstream’ presentation. Given that the bulk of these studies has furnished evidence in support of CL, the time is now ripe for fine-tuning and optimizing the approach.

In the next section, I will further discuss these critical issues through reviewing the core studies of etymological elaboration.

2.3.2.4 Etymological elaboration and rationale for this study

In the literature review above, I have introduced etymological elaboration as an important technique for enhancing L2 idiom learning, with supporting evidence from L2 idiom processing in contrast to L1 processing, and I have shown how its origins have been informed by CL theories. In this section, I will explain the rationale for this project by reviewing the core study by Boers et al. (2004b), where they first proposed etymological elaboration.

In this study, Boers et al. reported two experiments conducted with a computer-aided program in a self-study setting, in order to gauge the effectiveness of etymological elaboration. The participants were English majors at a university college. The total program included exercises on 400 idioms, and each idiom was tackled in three exercise components. (Note that the two experiments employed these exercises differently for different purposes).

In one component, learners were presented with the idiom in isolation and asked to choose the most likely domain of origin from multiple choices. For example, when presented with *jump the gun*, they could choose between source domains such as sports, war, jurisdiction, etc. Following this, the literal underpinning was displayed for learners through words (e.g., an athlete who jumps the gun in a racing contest sets off before the starting pistol has been fired). The task of hypothesising about the source domain is supposed to prompt a cognitive effort, i.e., deep processing, and together with recovering the literal underpinning in the feedback, to stimulate mental imagery that facilitates dual coding.

The second component asked learners to choose the correct figurative meaning of the expressions from another multiple-choice exercise. For *jump the gun*, for example, they were given the choice between (a) defend someone at your own risk, (b) do something before the appropriate time, and (c) be startled by an unexpected event. The correct choice was subsequently pointed out to them. The first and second tasks are expected to help learners associate the figurative meaning of an idiom with its literal underpinning, i.e., providing the conceptual motivation.

The third component of the exercise programme presented learners with a gap-fill task, where the idiom was incorporated in a meaningful context and one missing content word needed to be supplied. For example, “Although we had agreed not to tell anyone about my pregnancy yet until we were absolutely sure, my husband jumped the _____ and told his parents straightaway.” This exercise was used as a post-test to evaluate the effectiveness of the preceding steps.

The first experiment reported in Boers et al. (2004b) aimed to assess the mnemonic effect of presenting the literal underpinning of the idioms by comparing two conditions—each using one of the two multiple-choice components. In the experimental condition, 51 learners tackled the multiple-choice exercise on source domains, and one week later they were asked to do the completion exercise, i.e., they skipped the exercise on the figurative meaning of the idioms. In the control condition, 70 learners did the multiple-choice exercise on the figurative meaning of the idioms, and then the completion exercise as a delayed post-test. The 121 participants included 80 first-year students and 41 third-year student, who, within each year level, were randomly assigned to experimental and control groups. Each participant tackled several sets of idioms (with each set targeting 20 to 30 idioms) (total N of idioms =175), with the participants in the corresponding control group and experimental group receiving the same sets. One week later, the participants took gap-fill tests where they could obtain a score of 1 or 0 per test item. The success rate in the experimental group (39%) was higher than in the control group (28%), although no inferential statistics were given. Boers et al. took this as positive evidence for the mnemonic effects of etymological elaboration. Moreover, in this experiment, learners in the experimental group (Exp-group for short) were allowed to click on each of five possible source

domains on the identify-the-origin exercise multiple times until they found the correct origin (in the program, students needed to choose the right option before they could advance to the next item). Boers et al. counted the clicks that learners made, i.e., relative ease of identification—as a proxy of ‘transparency’. 68% of the idioms, receiving one or two clicks, were deemed as ‘etymologically transparent’ and the other 32%, receiving more than two clicks, as ‘etymologically opaque’. They found the mnemonic effect based on the one-week delayed post-test was rather weak for the ‘etymologically opaque idioms’ (Exp-group 31% vs. control-group 29% successful recall), while the effect was slightly more obvious on the ‘etymologically transparent idioms’ (Exp-group 40% vs. control-group 32%).

In the second experiment, Boers et al. (ibid.) adjusted the design. A total of 274 idioms were used. Participants (English majors at the same university college) ranged from year one to year four. All participants learned sets of 20 to 30 idioms by tackling the three exercise components of the program within one session in the following order: (1) the multiple-choice exercise on the figurative meaning of the idioms, which served as a pre-test (with no corrective feedback this time), (2) the multiple-choice exercise on the idioms’ origins (three origins for each item, rather than five as in the previous experiment) plus corrective feedback, and (3) the completion exercise, which served as a post-test again. In both multiple-choice exercises, learners were allowed only one chance to click on an option (in contrast to unlimited chances in the previous experiment). Correct choice of origin was taken as indicating that the idiom was ‘etymologically transparent’. The results of the first two exercises show that in 70.04% of the cases, the meaning was chosen correctly in the pre-test. Of these, 66.08% also had correct recognition of the origin, which was significantly more than prediction by chance. The authors took this result as strong evidence for etymological elaboration being a ‘realistic pathway for idiom interpretation.’ (p. 72). They further predicted that it could be used as tool to encourage learners to try and infer the idiomatic meaning via its literal underpinning. Whether this is a realistic and effective way of teaching idioms is one of the principal questions investigated in this study.

Further, based on the immediate gap-fill test, the recall rate of the idioms whose meaning was successfully chosen at the pre-test was unsurprisingly the highest (80.14%). The test scores

were slightly lower when the learners failed to identify the origin (74.63%, no *p* value provided). When learners failed to choose the correct meaning in the pre-test, the average post-test scores for idioms whose source domains were hard to guess were very similar to those for idioms with better guessable source domains (67.96% vs. 68.68%). Although the measures of transparency (i.e., predictability of the origin) and the interval between the ‘treatment’ and the post-test differ between the two experiments, the first experiment seems to show that the mnemonic effects are discounted with relatively ‘opaque’ idioms after one week, while the second experiment seems to show no effects of transparency on short-term memory.

Overall, the findings of the two experiments in Boers et al. (ibid.) provide indirect yet positive evidence for the potential of etymological elaboration in facilitating idiom interpretation and retention. These positive findings were later reinforced by Boers et al.’s (2007) study via an improved design by using the same online program plus a separate pen-and-paper experiment. In the computer-assisted experiment, the learners in the experimental group (who tackled the identify-the-source exercise before the identify-the-meaning exercise) significantly outperformed those in the control group (who did the exercises in the opposite direction) both in the identifying-the-meaning exercise and in the gap-fill test, i.e., the retention of idioms. To further assess the contribution of etymological elaboration in idiom comprehension, Boers et al. (2007) also conducted a pen-and-paper experiment, where the learners were asked to explain the figurative meanings of idioms with written materials in the order of (1) idioms in isolation, (2) idioms in verbal contexts borrowed from a dictionary, and (3) idioms in context plus etymological information. The researchers found that in 29.5% of the cases where a learner could not interpret an idiom with the contextual clues, the etymological information played a key role in solving the comprehension problem; however, the contribution of etymological information varied across idioms to a great extent.

Following these findings, numerous subsequent studies have applied and tested etymological elaboration in various learning contexts (with ESL or EFL learners from China, Japan, Iran, etc.), by comparing it with other pedagogical approaches, such as rote learning (Guo, 2008; Noroozi & Salehi, 2013) and pictorial elucidations (Szczepaniak & Lew, 2011; Vasiljevic, 2015) as mentioned above. Rather surprisingly, however, there have been few other studies that

examine the effectiveness of etymological elaboration, and those which have examined it are either unpublished theses (e.g. Coryell, 2012; Zhang, 2009) or articles in local or non-prestigious venues (e.g., Bagheri & Fazel, 2010). While most of these studies found that etymological elaboration had a facilitative effect on retention, especially for the meaning of idioms (e.g., Vasiljevic, 2015), it should be noted that not all studies so far have yielded positive evidence for etymological elaboration. Szczepaniak and Lew (2011), for example, found etymological notes have no positive effects on the retention of idiom form and meaning in comparison with pictorial illustrations. That study was conducted in independent learning conditions where participants were asked to read a booklet containing the definition of an idiomatic meaning plus an example sentence with or without an etymological note and/or a picture, which seems comparable to Boers et al.'s (2004b) online self-study program. With no direct interaction with a teacher, the low post-test scores could possibly result from 'inattentive reading' (ibid., p. 341). However, the effects of etymological elaboration could potentially be reinforced through explicit teaching, where interactions and corrective feedback might stimulate more cognitive engagement from learners (i.e., deeper processing). Therefore, the current study aims to explore the effects of etymological elaboration via a think-aloud procedure in a teaching experiment (see previous discussion of think-aloud protocols in section 2.2.2; cf. Cooper, 1998). A think-aloud interview is adopted as a means of explicit teaching, also with consideration of the critical **Issue 1** (see above), as it provides a channel into learners' thinking process as they tackle the meaning of idioms. Most of the relevant studies so far have been quasi-experimental studies (e.g., Guo, 2008 with 70 Chinese EFL learners; Noroozi & Salehi, 2013 with 60 Iranian EFL learners) which compared the overall effects of the approach with rote learning, but provide hardly any details of the teaching procedures, which limits their replicability. More importantly, few studies (except Coryell, 2012, unpublished ESL classroom-based action research) look into how learners actually experience the usefulness of CL-style explanations about the links between literal and figurative meaning. The current study was designed to fill in this gap.

Second, a question underlying the unclear effects of etymological elaboration in Boers et al. (2004b) is to what extent the semantic transparency of an idiom would mediate the mnemonic

effects of elaboration. Leaving aside the inconsistent nature of the findings reported in Boers et al.'s (2004b) two experiments, what these findings do not tell us is whether it matters if learners find the proposed motivation of an idiom's meaning transparent *after* it is presented to them. For example, learners may understandably fail to guess that *jump the gun* originates from track sports, but they might nonetheless find that the latter explanation makes good sense when it is subsequently given to them. By contrast, in a case such as *throw in the towel* they may not only fail to guess that the expression originates from boxing, but perhaps also find the information about this origin non-illuminating if they happen to be unfamiliar with boxing. In the case of a fair number of idioms, such as *a red herring*, the origins proposed in dictionaries (e.g., that the distracting scent of smoked herring was used to train hunting dogs) may come across as rather speculative and even far-fetched. The question these observations raise is whether information about a literal underpinning can always be expected to help learners remember the meaning of an idiom. Therefore, in this study, I will ask learners to evaluate the connection between the proposed literal underpinning of idioms and their figurative meaning after these are both explained to them (see further discussion of the transparency in section 2.3.3.2). Moreover, as pointed out in **Issue 2**, L1 transfer and cultural differences, as well as learners' prior lexical knowledge, such as unknown or misinterpreted keywords (Boers et al., 2004b) are also likely to affect the effectiveness of etymological elaboration (see further discussion in 2.3.3), which will be explored in the think-aloud procedure.

Third, in response to **Issue 3**, regarding fine-tuning and optimizing the CL-informed approach, a question that remains to be answered is whether it is best to engage learners in an inferencing procedure, where the literal underpinning is given to them as a basis for their inferencing, or whether it is safer to first explain the idiomatic meaning and then use the literal underpinning to make the idiomatic meaning more memorable. Ideally, a between-group design would be used to compare the two procedures. However, because of concerns involving the scope of factors to be investigated and the duration of the project, I will explore this issue by examining the error rates in the learning and its impact on the retention within the former trial-and-error condition, where the literal underpinning is given as an inference prompt.

Due to the fact that most EFL learners struggle in comprehending idioms while they may not feel the need to actually use the expressions themselves, this study is primarily concerned with the comprehension and retention of the meanings of idioms, although learners' attitudes towards using idioms themselves will also be explored as part of the investigation. Moreover, given that etymological elaboration is assumed to involve deeper processing and that learners have limited ability to allocate processing resources towards different aspects of lexical acquisition in learning tasks (Barcroft, 2002), I will use meaning recall as the post-test for the learning outcome of the teaching experiment. This follows Barcroft's (2002) *type of processing-resource allocation (ToPRA) model*, which suggests that "[w]hen processing demands are sufficiently high, [...] semantic elaboration can increase learning rates for the semantic (conceptual) properties of words while simultaneously decreasing learning rates for the structural properties of words" (p. 325).

2.3.2.5 Research Objective One

To sum up, in response to the critical issues identified in previous studies under the CL-informed approach and, more centrally, to the under-researched issues about etymological elaboration, this project intends to achieve two objectives.

The first objective is to further investigate the effectiveness of etymological elaboration in fostering idiom comprehension and retention. This concerns the following questions:

- (1) To what extent do learners' interpretations of the meanings of idioms benefit from etymological elaboration?
- (2) To what extent does learners' retention of the meanings of idioms benefit from etymological elaboration?
- (3) Does the procedure of inferencing basing on etymological elaboration have an impact on error rates in retention?

The second objective will be discussed in section 2.3.3, and involves the examination of the impact of factors that are likely to affect the effectiveness of this teaching approach. This comprises (a) identifying the obstacles in the learning process and (b) exploring the mechanism of the interplay between factors on learning. In other words, the first objective is intertwined with the second objective, in revealing a full picture of L2 idiom learning with the help of

etymological elaboration. To further specify these factors, we will review relevant studies on the factors that affect L2 idiom learning in the following section.

2.3.3 Factors likely to affect L2 idiom learning under etymological elaboration

2.3.3.1 Introduction

As discussed in the previous section, etymological elaboration has been proposed both as a mnemonic and as a way of helping students make an ‘informed’ or ‘educated’ guess or inference about the idiomatic meaning of expressions (Boers et al., 2004b; 2007). Although knowledge about the literal underpinnings does not guarantee that the correct meanings of idioms can be easily guessed, Boers et al. assumed this would prompt more cognitive engagement on the part of the students. As I continue my examination of the effectiveness of etymological elaboration in both the interpretation and retention of idioms, certain factors have to be considered. During the initial stage of idiom interpretation, there are a few factors which are likely to prevent learners from arriving at a correct understanding of the literal underpinning, such as the presence in the idiom of unknown words or deceptively familiar words (Boers et al., 2004b; see also previous discussion of comprehension issues in section 2.3.1.1). Moreover, when learners are informed about the literal underpinning or the origin of the idiom, there are still factors that are likely to interfere with the interpretation. These include cultural differences, lack of familiarity with the source domain, competing meanings that learners may come up with (e.g., *a full plate* could be interpreted to mean an abundance of resources as well as the correct interpretation of having a full workload to deal with), and non-transparent connections between the literal underpinning and the idiomatic meaning. As a result, providing the literal underpinning will lead to more successful inferences for some idioms than for some others. One of my principal research questions concerns the relevance of such issues. We will come back to research questions in section 2.4.

These factors have been investigated in the wider domain of SLA, and not just in CL-informed approaches. While idioms have been well researched in linguistics and psycholinguistics from the perspective of monolingual and first language performance and acquisition, they have not received sufficient attention in the field of second language acquisition. “[L]ittle is known

about the linguistic and cognitive mechanisms underlying the understanding and acquisition of idioms in second languages, a field of research in its own right” (Liontas, 2017, p. 6). In addition, there have been more studies on idiom comprehension than acquisition. This is possibly because comprehension is a prerequisite of acquisition, and that many studies approached acquisition by investigating learners’ comprehension of idioms as part of learning (Liu, 2008, p. 93). While psycholinguistic studies of L2 idiom processing and comprehension can inform us about how learning can be better facilitated (Siyanova-Chanturia & Martinez, 2015), the focus and objectives of such investigations could be slightly different from those of learning. The more learning-oriented comprehension studies focus more on specific problems that learners encounter in interpreting idioms, and on how certain factors may affect learning outcomes including retention and production.

From the review of psycholinguistic studies of idiom processing and comprehension in section 2.2, we have seen that factors affecting both native and non-native speakers’ online processing of idioms include (1) familiarity with the idioms resulting from frequency of exposure to them (e.g., Abel, 2003), which is related to the environment of language learning and use (see further discussion in section 2.3.3.7); (2) analysability of the idioms, which is often measured by decomposability, literal plausibility, or transparency (e.g., Abel, 2003; Titone & Connine, 1994b); and (3) the specific linguistic context in which idioms are encountered (i.e., whether this context is supporting/revealing or uninformative/ambiguous with regard to the idiomatic meaning; e.g., Giora, 2003; Cieřlicka, 2011; see relevant offline studies, e.g., Liontas, 2002, 2003; see also Gibbs, 2017). Research also shows that factors particularly affecting idiom comprehension by L2 learners are the learners’ L1 (e.g., Cooper, 1999; Liontas, 2003), their level of proficiency in the L2, their inferencing abilities and their cognitive styles (see Table 2.4 and further discussion below). In addition, learning-oriented studies (Boers et al., 2004a; Charteris-Black, 2002; Irujo, 1986a; Laufer, 2000; Türker, 2016) have investigated how the learning of L2 idioms is affected by cross-linguistic differences, including lexical and conceptual knowledge, and by cultural differences. Table 2.4 is a summary of these learning-oriented studies, based on the factors investigated and their effects on particular outcomes of idiom learning, which is roughly divided into comprehension, retention and production.

Overall, it has been acknowledged that L2 idiom learning is a very complex process with various factors coming into play.

Table 2.4 Factors investigated in learning-oriented studies of L2 idioms

Factors investigated	Representative studies	Learning outcomes
Analysability/Decomposability	Bortfeld (2003); Carrol, Littlemore, & Dowens (2018)	Comprehension
*Semantic transparency (& Imageability)	Boers & Webb (2015); Carrol, Littlemore, & Dowens (2018); Karlsson (2013); Ramonda (2016); <u>Steinel et al. (2007)</u>	Comprehension & <u>Retention</u>
*L1 transfer (of lexical & conceptual knowledge)	Carrol, Conklin, & Gyllstad (2016); Charteris-Black (2002); Hu & Fong (2010); Irujo (1986a); Johnson (1989); Kellerman (1977, 1979); Liontas (1999, 2003); Suñer (2018); Türker (2016)	Comprehension
	Irujo (1986a; 1993); Laufer (2000)	Production
*Cross-cultural variation (of source domains)	Boers & Demecheleer (2001); <u>Boers et al. (2004a)</u> ; Hu & Fong (2010); Karlsson (2013)	Comprehension & <u>Retention</u>
*L2 proficiency	Johnson & Rosano (1993)	Comprehension (of metaphor)
*Prior L2 lexical knowledge	Park & Chon (2019); <u>Zyzik (2011)</u>	Comprehension & <u>Production</u>
Cognitive styles	Boers, Eyckmans & Stengers (2006); Boers & Littlemore (2000); Boers, Lindstromberg, Littlemore, Stengers & Eyckmans (2008); Johnson & Rosano (1993)	Comprehension (of MWEs & metaphor)
Context of use	Karlsson (2013); Liontas (2003); Suñer (2018)	Comprehension
Learning environment, frequency effects, & familiarity	Carrol, Littlemore, & Dowens (2018); Karlsson (2013); Suñer (2019)	Comprehension
*Learner attitudes	Laufer (2000)	Production

(Note. *Factors are to be reviewed and investigated in this study; studies which investigate retention or production of idioms are underlined to differentiate them from those focusing solely on comprehension in the same categories.)

In this study, while examining the effectiveness of etymological elaboration as a teaching approach, I am particularly interested in how some of the factors relating to idioms and to learners may affect the learning process and outcome. In the following sections, I will review relevant studies concerning the following factors: (1) analysability and semantic transparency

of idioms (2.3.3.2), (2) L1 transfer and cross-cultural influences (2.3.3.3), (3) prior L2 lexical knowledge (2.3.3.4), (4) learners' L2 proficiency (2.3.3.5), and (5) procedures involved in teaching idioms (2.3.3.6). I will also discuss the possible interactions between the first four factors in the learning process. Teaching procedure is a separate issue I will discuss within the CL-informed approach, with a focus on error treatment and its impact on memory. After that, I will discuss some other learning-related factors (2.3.3.7) that are relevant to the current study, as well as some pedagogical considerations that are of particular relevance to the design of my research. These factors include context of use, the frequency and familiarity of idioms in particular learning environments, and learners' attitudes towards learning and using idioms.

2.3.3.2 Semantic analysability and transparency of idioms

Analysability, or decomposability, and semantic transparency are the two most frequently examined factors of idioms in online processing and offline comprehension studies. Although closely related and sometimes mixed, they measure idioms from different dimensions (Cieślicka, 2015). As discussed in section 2.2.1.2, decomposability is the degree to which individual meanings of the idiom's word components contribute to the meaning of the whole phrase, while transparency is the extent to which the metaphorical meaning can be deduced from its literal analysis (Cieślicka, 2015; Nunberg, Sag, & Wasow, 1994). In the following, I will discuss the two measures individually and rationalise my choice of measure for this study, i.e., etymological semantic transparency.

1. Analysability, or decomposability, often operationalised by classifying idioms into three categories, i.e., normally decomposable idioms, abnormally decomposable idioms, and non-decomposable idioms, was first proposed by Nunberg (1978) and has become popular since Gibbs and Nayak (1989) adopted it in their investigation of idiom comprehension by native speakers of English. Although widely used in studies of online idiom processing (see previous section 2.2.1), the reliability of decomposability ratings has been challenged by researchers, because of inconsistent judgements between native speakers for most idioms which show that people may not have a shared intuition of compositionality except for a limited number of expressions (e.g., Titone & Connine, 1994b, Tabossi, Fanari, & Wolf, 2008). In terms of the impact of the factor, findings are inconsistent regarding the role of decomposability in idiom

processing. Earlier studies suggest that decomposability facilitates the early stages of idiom processing (e.g., Gibbs & Nayak, 1989; Gibbs et al., 1989), whereas more recent studies find that decomposability facilitates the later stages of comprehension which involve integrating particular meanings into a relevant context (Titone & Connine, 1999; Libben & Titone, 2008). Furthermore, the impact of decomposability on processing speed in native speakers could possibly be overridden by familiarity, as suggested by Tabossi, Fanari, and Wolf (2008).

Decomposability is also considered as a measure in L2 idiom comprehension studies, and many researchers treat decomposability as synonymous with semantic transparency (e.g., Abel, 2003; Liu, 2008). While the two notions are very closely related in that decomposable idioms, whose constituents individually contribute to the overall meaning, tend to be more transparent than non-decomposable ones (Gibbs, 1993), decomposability and transparency are actually two different dimensions of idioms (Cieślicka, 2015). Moreover, decomposability is normally rated by native speakers for idioms that are already known to them. However, judgements that involve overt metalinguistic analysis, such as a decomposability rating that requires raters to be able to identify individual components of an idiom, could be problematic for L2/foreign language learners who have limited knowledge of the target language. In this respect, compared to decomposability, semantic transparency seems a more approachable measure of idioms pertaining to L2 learners.

2. Semantic transparency is defined by Nunberg et al. (1994) as “the ease with which the motivation for the use (or some plausible motivation—it needn’t be etymologically correct) can be recovered” (p. 498). I will use this as a working definition of transparency for this study. However, it should be noted that in different studies semantic transparency might be operationalised differently, and therefore that one should be cautious about generalising assumptions in relation to this notion. For example, some researchers, e.g. Gyllstad & Wolter (2016, on collocation processing) operationalise semantic transparency as compositionality (based on Howarth’s, 1996, definitions of collocations and free combinations), which is similar to similar to Grant & Bauer’s, 2004, ‘naïve’ approach to compositionality, i.e., whether each word component can be replaced by another word without changing the meaning of whole expression. Some researchers such as Steinel et al. (2007) operationalise it as “the degree of

semantic overlap or similarity between the literal and the figurative meaning of an idiom” (p. 478). When semantic transparency is considered as a measure of analysability of idioms or a proxy for the ease of learning, some related issues about the transparency rating need to be kept in mind.

First, transparency of an idiom is a matter of degree, and the perception of degree is idiosyncratic and peculiar to individual raters, which “necessarily remains a subjective measure, biased by the linguistic and cultural background of the informants” (Steinel et al., 2007, p. 457). This idiosyncrasy is not only commonly shown by inconsistent ratings among native speakers (see an in-depth investigation in Ramonda, 2016; see also, Boers & Webb, 2015) but is also reflected in the marked disparity of ratings between native speakers and EFL learners (Boers & Webb, 2015). In a rating study with 7 native English teachers and 33 advanced Japanese EFL learners, Boers and Webb (2015) found a moderate level of agreement among native speakers (correlation coefficient ranging from .29 to .71, with an average of .54, $SD = .13$), and a higher agreement among EFL learners (.51 to .85, with an average of .69, $SD = .072$). When comparing the mean ranking by the seven teachers and the learners, they find discrepancies between the two groups, with correlations ranging from .28 to .85, although pooled together they produce a high mean at .79. This suggests that native speakers and L2 learners may not see eye to eye on the transparency of idioms and that it takes more than two raters to obtain an acceptable reliability. It was further argued that in order to understand learners’ sense of learning difficulty and to predict possible misinterpretations, teachers and designers of learning material have to consider learners’ L1 and cultural background and prior L2 lexical knowledge, i.e., factors learners bring to bear in tackling the meaning of MWEs like idioms. This leads to the next issue.

Second, semantic transparency is affected by many factors (Boers & Demecheleer, 2001; Boers & Webb, 2015; Carrol, Littlemore, & Dowens, 2018). Besides decomposability mentioned above, the degree of transparency is determined by the interplay of various factors including: learner-dependent factors such as familiarity with the idiom, prior vocabulary knowledge, L1 equivalence, and cross-cultural differences,; and item-dependent factors such as the underlying

conceptual metaphor, the cultural knowledge required for interpretation of a specific idiom, imageability, as well as the clarity of the etymology.

Third, the perception of transparency is not constant but liable to change due to learning effects or increased familiarity. As found by Malt and Eiter (2004) in an experimental study with non-native speakers of English, learners' perception of transparency increased significantly after the meaning of an idiom was learned and used, regardless of whether the learned meaning was true or false. This finding aligns with Keysar and Bly's finding (1995) with native speakers. In a rating study with native and non-native English speakers, Carrol, Littlemore, and Dowens (2018) found that familiarity has a clear effect on perceptions of transparency: highly familiar idioms seem to be more transparent than less familiar idioms, and this effect was observed in both native and non-native speakers.

In terms of the impact of transparency on idiom comprehension and retention, previous studies have yielded convergent evidence that more transparent idioms are comprehended better and faster than less transparent ones (e.g., Karlsson, 2013). As for retention, the effects are relatively weak with mixed results from experimental studies. In addition, different operationalisations of semantic transparency and the varied learning conditions of these studies render their findings hard to compare. Findings from one experiment conducted by Boers et al. (2004b) show that opaque idioms, whose source domains learners failed to identify before correct feedback was given, were remembered almost as well as transparent ones in an immediate gap-fill task (68% vs. 68.7%). With guessability of source domains (measured by mouse clicks on multiple choice options) used as a proxy of transparency, Boers et al. (ibid.) took this as positive evidence for the use of etymological elaboration. However, since this computer-aided experiment did not reveal learners' actual thoughts or their direct evaluation of the etymological motivation, the researchers also suggested that research containing think-aloud procedures is needed to test such a speculation.

Steinel et al. (2007) investigated effects of transparency and imageability on L2 idiom learning in paired-associate learning tasks (pairing L2-L1 or L1-L2)—a different procedure from the CL approach. They found that 'imageability'—rather than 'transparency', i.e., 'the degree of semantic overlap or similarity between the literal and figurative meanings'—was a better

predictor of meaning and form retention measured three weeks later. Transparency was only found to affect recognition, i.e., recalling the meaning of L2 idiom by typing in the L1 equivalent/paraphrase, but had no facilitative effects on production, i.e., recalling the form of L2 idiom based on the L1 equivalent/paraphrase. On the basis of these findings, Steinel et al. (2007) claimed that “transparency does not have any mnemonic effect and cannot be said to be a good predictor of learning after all” (p. 479).

This study will further investigate the effects of semantic transparency of idioms on the learning outcome under the etymological elaboration technique.

3. Etymological semantic transparency. Learners’ perceptions of semantic transparency will be measured during the implementation of the etymological elaboration technique. To differentiate this measure from general use of semantic transparency, I will call it *etymological semantic transparency*, or etymological transparency, the latter being a term first used in Boers et al.’s study (2007): “Etymological transparency is inevitably a subjective experience and should be conceived as a gradable concept with idioms occupying a position on a continuum between extremely transparent [...] and extremely opaque [...]” (p. 51).

In essence, it is the same as Nunberg et al.’s (1994) definition of semantic transparency (see above). However, etymological transparency used here is slightly different in the following aspects: (1) the transparency is operationalised as the clarity of the link between the literal and the figurative meaning of the expression, in terms of the ease of seeing how the literal use could have given rise to the figurative use of the expression; (2) it is implemented on the condition that the target idioms all have an identifiable etymology, as very opaque idioms whose etymology is untraceable are excluded from etymological elaboration; (3) the rating task is implemented after learners have learned both the idiomatic meanings and the literal underpinning (cf. Carrol, Littlemore, & Dowens, 2018, who collected semantic transparency ratings before non-native speakers got to know the meaning of the idioms); (4) L2 learners’ perception of etymological transparency comes inevitably through the lens of their own native language and culture, and is affected by their prior knowledge in general; and (5) the objective of etymological transparency is to measure how well learners appreciate the etymological

motivation, which will be used to gauge the efficacy of etymological elaboration as a mnemonic.

Concerning the methodology, besides the measure of etymological transparency obtained with EFL learners, a measure of semantic transparency from native speakers was also used in preparation of the teaching materials. The latter measure served as a rough estimation of the ease of learning of the idioms, for the purpose of creating balanced sets of target idioms (see further discussion in section 3.3.1).

2.3.3.3 L1 transfer and cross-cultural interference

L1 knowledge plays an indispensable role in L2 idiom acquisition. Research evidence of L1 transfer occurring in L2 idiom comprehension and production has been found among L2 English learners (even at higher proficiency levels) with different L1 backgrounds, including Spanish (Irujo, 1986a), Hebrew (Laufer, 2000), French (Boers & Demecheleer, 2001), Malay (Charteris-Black, 2002), Chinese (Hu & Fong, 2010), Swedish (Carrol et al., 2016), and also among L2 learners of other languages, e.g., L2 learners of Greek, Spanish, French and German (Liontas, 2002) and Korean (Türker, 2016). The general finding is that the effects of L1 transfer, positive or negative, and the degree of the impact often depend on the relationship between the L1 and L2, i.e., how different/similar L2 idioms are from/to L1 idioms. If there is an L1 equivalent whose lexical makeup and meaning is identical or similar to L2, L1 often facilitates L2 idiom comprehension; whereas if the L1 idiom is very different from the target L2 idiom, or there is no idiomatic counterpart in L1, learners often find difficulty in understanding the L2 idiom, especially in the absence of helpful contextual clues. On the other hand, partial formal similarity between the L1 and L2 is more likely to cause negative transfer and avoidance of use in L2 idiom production than complete difference, and generates more interference errors (Irujo, 1986a; Laufer, 2000).

Evidence has also been found that false cognates or ‘false friends’ are likely to impede idiom interpretation when cross-cultural differences are involved (Hu & Fong, 2010). For example, in Hu and Fong’s (2010) study, English idioms containing *heart* and *mind* were often confusing to L1 Chinese speakers, as the Chinese culture does not share the western duality of *mind* vs. *body* and of *mind* vs. *heart*. Their participants tend to interpret *speak one’s mind* (‘say

something without reservation’) as being honest and true to someone, which is related to the concept of *heart* (i.e., 真心, pinyin: *zhen xin*) in Chinese. Such example also demonstrates that L1 transfer and cross-cultural interferences involve not only formal and semantic differences but also conceptual differences (e.g., Pavlenko, 1999), which seems particularly true of languages with very different language roots or formal configurations, such as logographic languages e.g., Chinese, Korean, and Japanese, versus phonographic languages e.g., English. More recent studies (e.g., Charteris-Black, 2002, Türker, 2016) have examined the role of L1 conceptual knowledge (i.e., knowledge about source domains, target domains and conceptual processing) as well as linguistic knowledge in L2 idiom learning. They found that idioms with an equivalent conceptual basis and linguistic form are the easiest to learn, and the most difficult are those with a different conceptual basis and an equivalent linguistic form, as well as those which are culturally specific with a different conceptual basis and a different linguistic form. Moreover, the frequency of L1 and L2 idioms and the availability of situational context seem to moderate the impact of L1 transfer in L2 idiom acquisition (see Charteris-Black, 2002, Türker, 2016).

Of particular relevance to this study is how L2 learners spontaneously refer to their L1 linguistic and conceptual knowledge in the process of interpretation (without many contextual clues) under etymological elaboration and to what extent this may affect their retention of idioms, as reflected in the success and error rates (see further discussion below).

2.3.3.4 Prior L2 lexical knowledge

The effects of prior lexical knowledge are often overlooked in L2 idiom studies (except e.g., Boers et al., 2007; Park & Chon, 2019). Studies of L2 idiom processing generally assume that L2 learners are familiar with the literal meaning of the lexical components of idioms (e.g., Cieřlicka, 2006). Studies of MWE learning often claim that their selected idioms contain only high-frequency words that are familiar to their participants (e.g., Steinel et al., 2007). However, learners’ insufficient prior L2 lexical knowledge, such as the presence in the idioms of unknown words and deceptively familiar words, i.e., polysemes and homonymous words, has been observed to be an initial obstacle to L2 idiom interpretation in several CL-informed studies (Boers et al., 2004b; Boers & Webb, 2015; Hu & Fong, 2010). Moreover, the presence

of familiar words with an unfamiliar form-meaning link is one of the common causes of ‘deceptive transparency’ (Laufer, 1989, 1997) in L2 idiom interpretation.

Regarding the effects on learning, one study by Zyzik (2011) found significant effects of lexical knowledge on L2 (Spanish) idiom production (rather than idiom recognition). The idioms with known words received significantly better scores than those with unknown words in the active recall of form (although the effect size ($d = .27$) is small). Furthermore, as Boers and Webb (2015) point out, in order to understand L2 learners’ perception of idioms and to estimate L2 learners’ difficulties in learning, native-speaking teachers and textbook designers need to take into consideration learners’ L1 and cultural background as well as their L2 lexical knowledge. Therefore, prior lexical knowledge is one of the factors to be examined in this study.

2.3.3.5 L2 proficiency

Another variable that is largely missing from the available body of research on L2 idiom learning is the potential role of individual learner traits. Exceptions are explorations of so-called cognitive style differences among learners when they are taught idioms with the etymological elaboration (e.g., Boers, Eyckmans, & Stengers, 2006). It is conceivable that some learners experience etymological elaborations as more useful than other learners do and are thus more inclined to appreciate the mnemonic potential of such elaborations. By contrast, for learners who find it harder to appreciate the connection between an idiom’s meaning and its literal underpinning, the mnemonic benefits of etymological elaboration may be confined mostly to relatively transparent connections. What seems worth exploring, then, is whether such predispositions might be related to L2 student profiles more generally.

Although previous research shows that learners’ level of L2 proficiency does not seem to affect idiom interpretation in general (e.g., Wu, Chen, & Huang, 2006), it is nonetheless possible that this variable mediates the mnemonic effects of etymological elaboration. For example, it is possible that more proficient learners, i.e., high achievers within a given student population, reap the mnemonic benefits more readily than their comparatively low-achieving peers. This possibility appears compatible with findings from research on language-learning aptitude (Robinson, 2013; Skehan, 2015), where one of the known predictors of learning success is associative memory ability. High achievers are likely to have an advantage in this regard, and,

in theory, this could also apply to how well they remember the meaning of an idiom in association with additional information, such as its literal underpinning. In addition, the high achievers are also likely to have built a larger L2 vocabulary and to have developed greater familiarity with the L2 in general. If so, they will be more likely to be familiar with the lexical constituents of new idioms, and this could reduce the learning burden. It is now relatively well established that the ability to learn new L2 lexical items tends to increase as one becomes more proficient in the language (e.g., Elgort & Warren, 2014), which is sometimes referred to as the Matthew effect—the rich get richer faster, i.e., learners with a comparatively large vocabulary pick up additional lexical items faster than learners with a comparatively small vocabulary. Research evidence also shows that more advanced learners' idiomatic performance resembles that of native speakers, which is also reflected in a *dynamic model of multilingualism* (Herdina & Jessner, 2002). In sum, there are indeed grounds for hypothesizing that, within an otherwise homogenous population of L2 students, those who have relatively higher L2 proficiency may be more inclined to engage with and remember the etymological elaborations proposed by a teacher, regardless of whether some of these elaborations seem far-fetched. An evaluation of this hypothesis will therefore be part of the present study as well.

In addition, the possibility that more proficient learners might benefit the most from etymological elaborations should also be compatible with research on vocabulary learning strategies (Gu, 2003, 2013) which suggests that individuals who welcome a wide range of strategies tend to be relatively successful learners. This assumption will also be explored in the current study.

2.3.3.6 Teaching procedure—trial-and-error and errorless learning

Error treatment and assessment are important aspects of teaching and learning. Appropriate treatment and assessment should help improve learners' learning of idioms, and help teachers understand learner difficulties and optimise teaching procedures (Liu, 2008). It is one of the concerns for the implementation of etymological elaboration.

As described in section 2.3.2.4, the procedure of the teaching experiment in this study consists of the guessing of meanings with and without the aid of etymological elaboration. This is inspired by Cooper's (1999) study which proposed a heuristic model for L2 idiom processing

and suggested that a trial-and-error method (adapted from the think-aloud protocols) could be applied to L2 idiom pedagogy—a view shared by Irujo (1993) and Lennon (1998). However, there is ongoing debate concerning the benefits of meaning guessing, with some arguing that wrong guesses can interfere with the establishment in memory of the desired associations and that errorless learning might therefore be more advisable than learning through trial and error (e.g. Boers & Strong, 2019a, 2019b; Warmington & Hitch, 2014). While the guessing task in the current design will be followed by prompt corrective feedback and explicit explanation—an indispensable part in CL-guided interpretation (Skoufaki, 2008; Hu & Fong, 2010)—during the interaction between the teacher and a student, it inevitably engenders errors in learning. Therefore, the effects of the engendered errors (such as L1 inference errors mentioned above) on meaning retention will be examined as part of the investigation of this study (see also previous discussion in section 2.3.2.4 regarding Issue 3).

2.3.3.7 Other factors and considerations for this study

The context in which an idiom occurs is undoubtedly highly relevant in both L1 and L2 idiom comprehension (see previous discussion in section 2.2). Analysing the contextual information is one of the most commonly used strategies in L2 idiom comprehension (Cooper, 1999; Lontas, 2002, 2003; Karlsson, 2013). Seeing idioms in and out of a meaningful context makes a significant difference in the accuracy of idiom interpretation (Lontas, 2003). However, an experiment reported by Boers et al. (2007, pp. 52-53) suggests that to a large extent, learners fail to comprehend idioms even when the items are embedded in rich contexts such as those provided in a corpus-based idiom dictionary. Besides this, it is agreed among these studies that the ease and accuracy of comprehension also depend on the transparency or opacity of specific items. Thus, in the current study, which has a special focus on the extent to which unpacking the literal meaning can help L2 idiom interpretation (and retention), the context in which idioms were embedded was carefully treated as part of the learning materials, but not considered as a target of investigation. The context of use was manipulated in respect of richness in information, in order to facilitate the interpretation and comprehension of idioms before and after the etymological elaboration (see detailed discussion in section 3.3.3)

Item frequency and familiarity, in association with external learning environment, are also key factors affecting idiom processing and comprehension (as discussed in section 2.2). However, it seems of relatively less direct impact on L2 idiom learning (see e.g., Suñer, 2019 with a plea for increasing enhanced input and explicit knowledge of idioms for L2 learners even in the stay-abroad learning context). First, research evidence shows that item frequency based on the British National Corpus (BNC) and *Collins Cobuild Idioms Dictionary* (2002) does not seem to predict whether an L2 learner (and also a L1 speaker) will know the meaning of an idiom (Karlsson, 2013). While item frequency based on a native-speaker corpus is viewed as an objective measure of frequency for L1 idiom processing and comprehension, it does not apply to L2 learners and especially to foreign language learners whose personal exposure to idioms varies drastically on an individual basis. On the other hand, item frequency is also hard to measure based on language speakers' intuitions (Abel, 2003), and one may refer to familiarity ratings instead. While the degree of familiarity is likely to affect idiom comprehension and use by L2 learners in English-speaking countries where they have a substantial amount of exposure to idioms, this may not apply to foreign language learners, who have fewer chances to encounter English idioms outside the classroom. Therefore, in the current study with Chinese EFL learners, item frequency and familiarity were not treated as the target of investigation. However, idiom frequency and familiarity were considered and carefully treated in the idiom selection and data collection. Item frequency indicated by *Collins COBUILD Dictionary of Idioms* (1999) based on occurrences in a native-speaker corpus was used as a proxy of the usefulness of learning target (see detailed explanation in section 3.3.3). A norming study was conducted with more advanced English-major students to ensure that most of the selected idioms were unknown to the actual participants before the study. Additionally, during the teaching experiment, learners' prior knowledge of the idioms was also checked by asking them whether an item was known or unknown to them at the start of learning, which also served as a pre-test for the study (see detailed explanation in section 3.4.1).

To sum up the above, while they were not specific targets of the investigation, item frequency and familiarity and context of use were considered and carefully treated in the item selection, material design and data collection, so as to guarantee the pedagogical ecological validity.

Learner attitudes towards the learning and use of idioms is another under-researched area in second language idiom learning. Although most previous studies claim that L2 learners in general have positive attitudes towards learning idioms in another language (e.g., Cooper, 1998; Liu, 2008) and perceive idioms as an important part of L2 learning (McPherron & Randolph, 2014), very few studies delve into L2 learners' attitudes towards the learning and use of idioms, and even fewer explain how these attitudes may relate to other factors affecting learning, such as cross-linguistic differences and L2 proficiency (except e.g., Laufer, 2000). In order to get a fuller picture of Chinese EFL learners' idiom learning, a follow-up interview will be conducted, during which biographical information is collected from the learners regarding their exposure to idioms (cf. Abel, 2003), their attitudes towards learning and using idioms, as well as the learning strategies they use when encountering unknown multiword expressions like idioms (see details of data collection in section 3.4).

2.3.3.8 Research Objective Two

In conclusion, the factors discussed above influence different aspects of L2 idiom learning and they can be interrelated in various ways. So far, specific factors have often been investigated as single effects in most learning-oriented studies, and more recent studies start to investigate the interplay between different factors. On the other hand, although numerous studies have been carried out that explore effects on comprehension, less is known about the retention of L2 idioms. Against this backdrop, Research Objective Two of this study is to investigate the factors that influence L2 idiom interpretation and retention, with a specific focus on learners' perceived semantic transparency of idioms, L1 transfer and interference, prior L2 lexical knowledge, and L2 proficiency. Attempts will be made to identify the extent to which these factors influence idiom comprehension and retention, with exploration of potential interactions between these factors. This is intertwined with Research Objective One (see 2.3.2.5) in probing which interactions and stimuli affect the learning process and outcome under the etymological elaboration approach and in exploring the best procedure that leads to better learning performance. Specific research questions include the following.

Regarding both the process and outcome:

- (1) To what extent does prior lexical knowledge affect L2 idiom interpretation and retention?
- (2) To what extent do L1 transfer and cross-cultural interferences affect L2 idiom interpretation and retention?

Regarding the learning outcome:

- (3) Does learners' perceived transparency of idioms affect L2 idiom retention and how?
- (4) Does L2 proficiency affect L2 idiom retention and how?

Regarding learning in general:

- (5) What do the EFL learners' experiences with and awareness of idioms tell us about L2 idiom learning?
- (6) What do their use of strategies in learning vocabulary tell us about L2 idiom learning?
- (7) What are the learners' attitudes towards learning and using L2 idioms?

2.4 Summary of research objectives and questions

To sum up, the study aims to achieve two research objectives. Objective One is to investigate the effectiveness of etymological elaboration in fostering idiom comprehension and retention.

Questions addressed are:

- RQ 1. To what extent do learners' interpretations of the meanings of idioms benefit from etymological elaboration?
- RQ 2. To what extent does learners' retention of the meanings of idioms benefit from etymological elaboration?
- RQ 3. Does the procedure of inferencing based on etymological elaboration have an impact on error rates in retention?

Objective Two is to investigate the problems that L2 learners, native-Chinese EFL learners in particular, experience when they encounter English figurative idioms, and to identify the key factors of learning success. As mentioned above, questions involved are:

- RQ 4. To what extent does prior lexical knowledge affect L2 idiom interpretation and retention?
- RQ 5. To what extent do L1 transfer and cross-cultural differences affect L2 idiom interpretation and retention?
- RQ 6. Does learners' perceived transparency of idioms affect L2 idiom retention and how?
- RQ 7. Does L2 proficiency affect L2 idiom retention and how?
- RQ 8. What do the EFL learners' experiences with and awareness of idioms tell us about L2 idiom learning?
- RQ 9. What do their use of strategies in learning vocabulary tell us about L2 idiom learning?
- RQ 10. What are the learners' attitudes towards using and learning L2 idioms?

Ultimately, this study aims to optimize etymological elaboration as a CL-informed teaching approach, in order to enhance second language idiom learning.

CHAPTER THREE: METHODOLOGY

3.1 Overview of the research design

This chapter introduces the research design, the process of data collection and preparation for analysis, and explains the rationale behind each step. By nature, this is an exploratory study of L2 idiom learning, with the primary aim of investigating factors that may influence the effectiveness of the CL-inspired instructional approach, or in other words, ‘etymological elaboration’ (Boers et. al, 2004a, 2004b, 2007). It examines the problems that L2 learners, native-Chinese EFL learners in particular, experience when they encounter figurative idioms in their target language, and identifies the key factors of learning success, while taking individual item and learner traits into account. It also investigates how native-Chinese EFL learners’ comprehension and learning of English figurative idioms relate to their attitudes and strategies towards L2 vocabulary learning. All these objectives were achieved within one experiment, where data were collected and then analysed by means of both quantitative and qualitative procedures in response to research questions from different perspectives. Comprehension in the experiment was captured by learners’ *interpretation* of idioms with limited contextual support in the process of comprehension. *Interpretation* and *guessing* are used interchangeably, especially when I talk about guesses. Comprehension and retention are both part of learning. This study treats retention as the learning outcome after the learning stage in the first week which includes trials of interpretation.

As an attempt to gauge the scope of applicability of a teaching approach at the item level among a certain cohort of learners, rather than comparing different teaching conditions, the experiment in this study took the form of one-on-one interviews in a think-aloud mode. This allowed the researcher to observe the learning process of each individual learner on each item and to probe into the factors lying behind that process. In the experiment, each participant was individually interviewed twice with a one-week interval. In the first interview, the participants were presented with a set of figurative idioms whose meaning they were asked to guess, first without and then with the prompt of information about their etymology or **literal underpinning** (the two terms will be interchangeably used in this thesis). After the figurative meaning of each

idiom was explained to them, the participants were asked to rate the transparency of its connection to the proposed underpinning. One week later, the learners were presented with the same sets of idioms and asked to recall their meaning and to verbalise what they could remember of the explanations given the previous week. In this second interview, after the recall test, a series of follow-up questions were also asked about the participants' own English idiom learning experience (and their more general vocabulary learning experiences) and their impression of the preceding learning process. Details of the implementation will be fully explained in section 3.4.

This mixed-design approach generated rich data available for further exploration. The data took various forms, including responses from the participants in the teaching experiment, test results and answers to interview questions. Unsurprisingly, the implementation of this approach, as well as the processing, analysis and interpretation of the resulting data, were time-consuming and demanding, and required careful organisation. Taking into account the research aims, time available to the project, and the conceivable number of target items that participants would be able to learn in one learning session, a design was adopted that had **4 sets of items, with 21 items in each set, and a minimum of 6 participants** allocated to each set. This allowed the collection of learner responses to a relatively large set of stimuli—84 idioms—while at the same time keeping the time invested in the interviews manageable for the interviewees. The design is shown in Table 3.1. Each set of 21 idioms (Set A, B, C, and D) was presented to six people, with an additional seventh participant in one group. Later, Group A, B, C, and D will be used to refer to the groups of participants who were given the corresponding sets of idioms.

Table 3.1 Set vs. Group design

Set	Idioms	Participants	Group
Set A	21	7	Group A
Set B	21	6	Group B
Set C	21	6	Group C
Set D	21	6	Group D

Given the design outlined above, the sample size in terms of how many participants were exposed to each idiom was relatively small (six or seven participants). An alternative design would have involved presenting a single set of 21 idioms to all 25 participants. However, since the primary aim of the current study was to gauge the generalizability of the CL-approach to idiom teaching, it was felt that it was important to achieve a larger sample size in terms of the number of idioms. It also has to be acknowledged that the adopted design limits direct comparisons between the four sets of idioms, because of differences in the individual traits of the participants. In order to achieve the research aims and ensure the power of analysis, the following measures were taken:

- 1) Pseudo-random allocation of idioms to the four sets, prior to running the experiment. Efforts were made to balance their characteristics across the four sets (see detailed discussion in 3.3), so as to achieve approximate replication for different sets of idioms.
- 2) Random allocation of participants to the four sets of idioms, as they arrived for the experiment (see details in 3.2).
- 3) To ensure the statistical analysis reflects the structure of the data, mixed-effect regression models¹⁰ were used to investigate factors that might contribute to learning success. Such models can account for the variation both between participants and between idioms, because the random effects fit the nesting structure of the design involving 4 sets x 21 items x 6/7 participants (see further explanation in 3.6).

In addition, in order to allow extrapolation beyond the statistical analysis:

- 4) Results from statistical analysis were discussed in combination with qualitative analysis (see Chapters 5 & 6) before final conclusions were reached (in Chapter 7).

Overall, the data collection consists of three main stages: preparation, implementation, and data processing, as shown in Table 3.2. In the following sections, preparation and implementation of data collection will be introduced from sections 3.2 to 3.4. The procedure of data preparation

¹⁰ A mixed-effects model is a statistical tool which includes both fixed effects and random effects. The so-called fixed effects could correspond to treatment factors, but more generally, are parameters associated with the system under study (e.g. idiom types, number of responses). Mixed-effects models include random effects in order to account for the randomness inherent in the data, which often comes from differences in behaviour between different participants and between different items.

(or processing) prior to analysis will be explained in section 3.5. Data analysis will be briefly introduced in 3.6 and then further discussed with the interpretation and discussions of results in subsequent chapters.

Table 3.2 Data collection

I. Preparation	Selecting idioms and compiling materials	3.3
	Recruiting participants	3.2
II. Implementation	Interview 1—Learning phase	3.4
	Interview 2—Recall test & follow-up questions	
III. Data processing	Transcription and translation	3.5
	Rating responses and test results	
	Coding factors and variables	

3.2 Participants

The participants (N = 25) were Chinese EFL learners majoring in English at a prestigious high-ranking university in mainland China. They were all in their third year of study at the university and comprised 22 females and 3 males aged between 19 and 22 (median = 20). Mandarin Chinese was their only native language (L1). The participants had similar histories of EFL learning. The researcher shared these students' L1 and cultural background, and this facilitated both the interview (which was conducted in Mandarin Chinese and English) and the researcher's appreciation of sources of idiom-learning difficulties such as L1 transfer and cultural differences.

The participants all had intermediate-to-high proficiency in English. They had all passed the Test for English Majors¹¹ Band 4 (TEM-4) before the end of their second-year study. This criterion-referenced test is widely used in China to gauge the English proficiency of university undergraduate English majors in accordance with the National College English Teaching Syllabus for English Majors. The mean score of the participants on this test was 75.72 (SD =

¹¹ The Test for English Majors (TEM) is administered nationwide by National Advisory Commission on Foreign Language Teaching in Higher Education. It has two levels—Band 4 and Band 8. English majors take the TEM-4 at the end of second year and the TEM-8 near the end of their fourth (senior) year in the university.

6.52), and the scores ranged from 61 to 89 and showed a normal distribution (see Table 3.3). According to the TEM descriptors, a score of 60-69 qualifies is a pass grade, one of 70-79 is considered a good grade, and one of 80 or above is an excellent grade. Therefore, although the sample of participants was homogenous in terms of L1 background and EFL learning history, the TEM-4 grades demonstrate different levels of EFL achievement (see Table 3.4 for the TEM-4 grade distributions). Whether this made a difference to their performance in the actual learning will be explored in this study.

Table 3.3 Descriptive statistics of TEM-4 scores

	N	Minimum	Maximum	Mode	Median	Mean	SD	Skewness		Kurtosis	
TEM-4 Score	25	61	89	76	76	75.72	6.542	-.052	.464 (Std. Error)	.195	.902 (Std. Error)

Table 3.4 Distribution of TEM-4 grades

TEM-4 Grade		Number
Excellent	(80-100)	5
Good	(70-79)	17
Pass	(60-69)	3
Total		25

The participants volunteered from five parallel classes, with five students participating from each class. During recruitment, the students were informed that they would be involved in a study on vocabulary learning and that they would learn some English phrases and answer a few questions. English phrases were specified as idioms in the *Information Sheet* given to the participants (see Appendix 1) at the beginning of the interview, but the precise purpose of the study was only explained to them after data collection was completed. They all gave written consent for their data to be used for this research project (see *Participant Consent Form* in Appendix 2) in accordance with the rules by the Human Ethics Committee of Victoria University of Wellington (see *Ethics Approval No. 23363* in Appendix 3).

As mentioned above, in order to collect learners' responses to a relatively large set of idioms while at the same time keeping the time invested in the interviews manageable for the interviewees, the participants were randomly assigned to four groups in order to learn one of

the four sets of idioms. Ideally, there would be even numbers of participants for each group. However, during the recruitment stage, one student withdrew and two additional students volunteered, resulting in 7 participants in Group A and 6 in the other three groups. The participants were given a code in the form of the assigned idiom set plus their order in the participation of the interviews, e.g. A1, B2, which complied with the requirement for anonymity in the Human Ethics Approval. Their TEM-4 scores were collected on the day of the interviews with their permission (the test had been taken a few months previously).

The TEM-4 scores of each participant in the four groups are shown in Table 3.5. According to the boxplot of the distribution of the scores in Figure 3.1, Group A (median=71) was much lower than Group B (median=82), and Group C and Group D lie between these two groups (both with a median of 77). The Kruskal Wallis test showed an overall difference among the four groups ($p < .05$). In a pairwise Nemenyi test, Group A was significantly different from Group B ($p = .027$), but other pairwise comparisons did not yield significant differences. In order to assess whether this difference in TEM-4 scores might have had an impact on responses to the idioms (which of course differed across Sets A to D), the preliminary analysis below included TEM-4 scores as *random slopes* for idioms in a *mixed effects model* (MEM, introduced in 3.6). Comparison of the Transparency*Proficiency MEMs (see Chapter 6) with and without random slopes showed that these random slopes accounted for very little variance in the analysis, suggesting that the effect of proficiency is approximately the same for each idiom regardless of the set.

Table 3.5 TEM-4 scores

A1-A7	Score	B1-B6	Score	C1-C6	Score	D1-D6	Score
A1	71	B1	84	C1	79	D1	79
A2	61	B2	76	C2	84	D2	73
A3	78	B3	70	C3	70	D3	75
A4	65	B4	89	C4	78	D4	78
A5	73	B5	87	C5	69	D5	79
A6	70	B6	80	C6	76	D6	76
A7	73						
Median	71		82		77		77

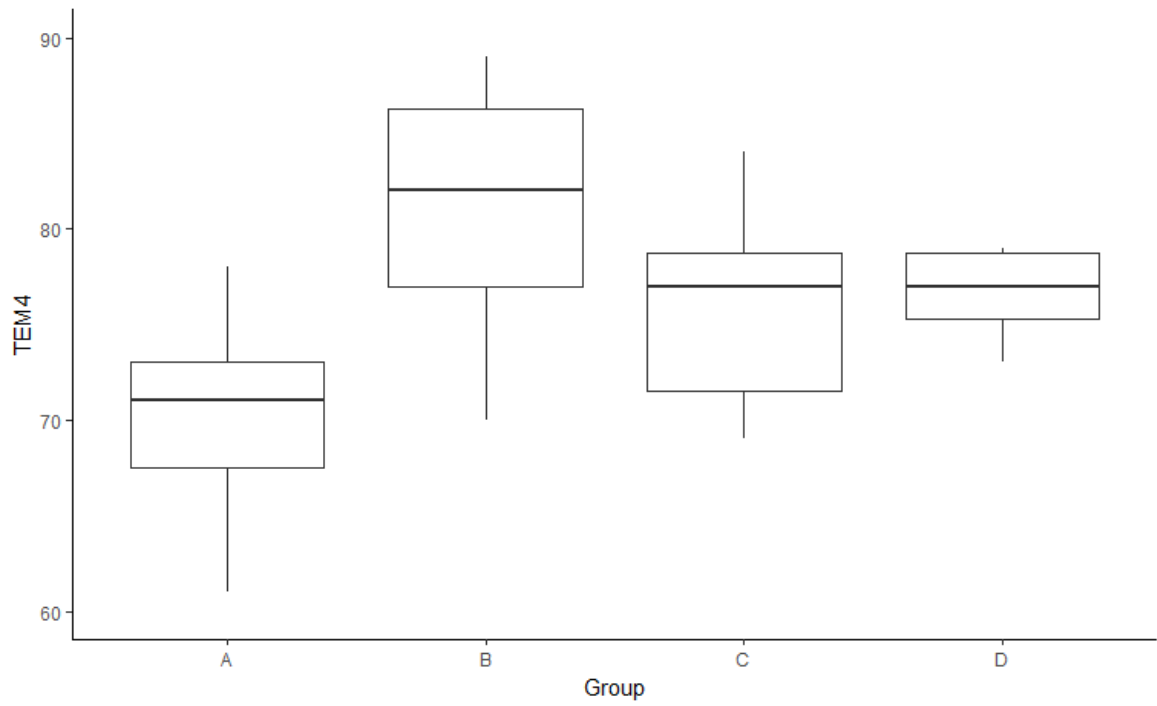


Figure 3.1 TEM-4 scores of 4 groups

3.3 Materials

3.3.1 Idiom selection

84 English idioms were sampled from observed daily communications, idiom dictionaries, and a corpus of items used by previous L2 idiom learning and teaching studies (e.g., Cooper, 1999; Boers et al., 2007; and Ramonda, 2016). All these idioms were conventionalised figurative expressions that are currently used in British English and American English. For those whose forms are slightly different in these two varieties, both forms were provided in the teaching materials, e.g. *have a green thumb/have green fingers*. The selection of these idioms followed two general criteria, **frequency** and **diversity**. Moreover, a few relatively easy idioms were deliberately included, which either appeared to be associated with a very similar L1 expression (e.g., *a hot potato*) or were relatively transparent (e.g., *in the driver's seat*; see further discussion below). These were included in order to keep the participants motivated during what was otherwise a challenging process of idiom interpretation.

To balance the sets of idioms (as explained earlier), the idioms were semi-randomly allocated into four sets (see the four sets in Table 3.6), with effort made to ensure a balanced distribution of idioms with different features across the four sets—these features were source domain, semantic transparency, and phrasal type, which will be further explained below. Given the constraints of this research, it was not also possible to match the sets for other factors, such as idiom length. Moreover, idioms containing the same words were allocated to different sets to avoid confusion that might arise for learners if they had to learn, e.g., *start from scratch* and *up to scratch*, *green thumb* and *a rule of thumb*, and *pass the baton on* and *pass the buck*.

In addition, before interviewing the 25 third-year English-major undergraduates, I tested five English-major postgraduate students individually on all 84 selected idioms. They were asked—as a benchmark group—if they recognised the idioms and to give their interpretation of their meanings, in order to ensure that the selected idioms were likely to be unfamiliar to the third-year undergraduate participants. The result showed that all the idioms were unknown to these

more advanced L2 learners (except *a dark horse*¹²), so all were included in the study.

Table 3.6 Target idioms in four sets

Set A	Set B	Set C	Set D
give the green light	take a back seat	jump the gun	come out of your shell
on the back burner	a shot in the arms	in the driving/driver's seat	flex your muscles
follow suit	beat around the bush	show someone the ropes	wet blanket
take it on the chin	a feeding frenzy	sit on the fence	cut corners
get into gear	win hands down	in the wake of something	bark up the wrong tree
(be) a sitting duck	jump in (at) the deep end	bury the hatchet	hand over fist
be par for the course	give someone the cold shoulder	teething problems	get a second wind
on the same wave length	in the doldrums	take the plunge	bite the bullet
bread and butter	ruffle someone's feathers	rub someone the wrong way	go with the flow
pull one's weight	on the same page	a rule of thumb	a can of worms
jump ship	pass (on) the baton	(hit) below the belt	give someone a let up
a dark horse	throw your hand in	a red herring	have a green thumb/ fingers
go belly up	(be) waiting in the wings	hold your horses	a hot potato
play into someone's hands	(not) up to scratch	stick your neck out	play it by ear
red tape	a loose cannon	down and out	throw your hat/cap in the ring
a drop in the bucket	leave someone high and dry	have a lot on your plate	get something off your chest
throw in the towel	be on the ropes	pass the buck	on automatic pilot
spill the beans	hit the roof/ceiling	play your cards close to your chest	tighten your belt
have/got cold feet	weigh someone down	ring a bell	on a shoestring
take the bull by the horns	let the cat out of the bag	turn over a new leaf	(start) from scratch
burn the candle at both ends	The ball is in your court.	make ends meet	pull a rabbit out of the hat

Frequency was used as the first reference criterion in selecting idioms, as it allowed me to select items that were likely to be useful to the students. Like Boers et al. (2004a, 2004b), I used the frequency bands of *Collins COBUILD Dictionary of Idioms* (1999) as an indication of relative frequency. The frequency bands in *Collins COBUILD* were based on the occurrence of idioms in *The Bank of English* (see *Collins COBUILD*, 1999). There were three bands indicated by black triangles (▲): “The commonest idioms marked with three black triangles, the next commonest with two, and the next with one” (p. xvii). Idioms with three markers are those (including their variations) occur at least once in 2 million words of the corpus; idioms with two markers are “relatively common, but not as common as the idioms with three markers”

¹² *A dark horse* was a special idiom that was found to have been brought into and become a common expression in Chinese. Although the Chinese EFL learners might know the expression, they might not be aware that it was (originally) an English idiom. Therefore, the item was still included in the experiment.

(ibid.); and idioms with one marker occur between 1 and 3 times in every 10 million words of the corpus. “[I]dioms do not receive any frequency marker at all... are expressions which are current in English, but occur less often than once in each 10 million words of corpus text” (ibid.). According to *Collins COBUILD* (1999), the majority of the selected idioms (78 out of 84) were commonly used in English, with at least one marker (see Table 3.7). Three of the remaining idioms were items which according to the dictionary were in current use but which had no frequency marker—these were *on the same page*, *wet blanket*, and *hold your horses*. The final three idioms did not appear in *Collins COBUILD* (1999) but in other dictionaries (e.g., *Oxford Idioms Dictionary*, 2013)—these were, *weigh someone down*, *hand over fist*, and *throw your hand in* (see detailed frequency band information in Appendix 4). In addition, the frequencies of the 84 idioms were also obtained from COCA, i.e., the Corpus of Contemporary American English (Davies, 2008-) in August 2016. According to the COCA frequencies, 56 out of the 84 idioms (66.7%) occur at least once in 10 million words, based on the 2016 corpus size of 520 million words (see detailed counts of occurrences in Appendix 4). This provides some reassurance that the majority of selected items are commonly used in present-day English.

Table 3.7 Frequency distribution of the 84 idioms

Frequency Band	Number
◀◀◀	34
◀◀	27
◀	17
No frequency marker	3
Not included in Collins COBUILD	3
Total	84

Diversity was the other general criterion for idiom selection. Instead of ‘cherry picking’, this study attempted to sample a variety of idioms. The diversity was reflected in three main features: source domain, semantic transparency, and phrasal type.

1) Source domain. The source domains were the identified domains of origin of the idioms in cognitive linguistic terms (see previous explanation in section 2.3.2). The figurative idioms are

derived from associations in rather specific experiential domains, and can be motivated by reference to their literal usage in those original contexts. As in Boers et al. (2007), the source domains of the selected idioms were identified and classified. This resulted in nine major domains, i.e., sports (including subdomains of boxing, horse racing, and running), travel, seafaring, war, hunting, reading, card games, animals, food and cooking. For example, the meaning of the idiom *throw in the towel* can be motivated with reference to its original, literal sense in the sports subdomain of boxing, when a boxer is down and his disappointed coach/manager may throw a towel into the boxing ring. Two idioms (i.e., *red tape* and *on a shoestring*) could not be categorised into any of the major domains and were classified as ‘Others’. More examples and the distribution of idioms according to the source domains can be found in Table 3.8.

Table 3.8 Distribution of the source domains

	Example	Set A	Set B	Set C	Set D	Total
Sports	<i>The ball is in your court.</i>	5	6	5	5	21
Boxing	<i>throw in the towel</i>	2	2	2	2	8
Horse racing	<i>win hands down</i>	1	1	1	1	4
Running	<i>jump the gun</i>	0	1	1	1	3
Travel (car/air)	<i>in the backseat</i>	2	1	2	2	7
Seafaring	<i>show someone the ropes</i>	1	2	1	2	6
Hunting	<i>bark up the wrong tree</i>	1	1	1	1	4
Animals	<i>come out of your shell</i>	4	3	4	3	14
Reading	<i>on the same page</i>	0	1	1	0	2
Card games	<i>follow suit</i>	2	1	2	1	6
Food & Cooking	<i>a hot potato</i>	2	2	1	1	6
War	<i>bury the hatchet</i>	0	1	1	1	3
Others	<i>red tape</i>	1	0	0	1	2

(Note. Some domains overlap with others, but all the domains and subdomains were distributed across sets as evenly as possible.)

2) Semantic transparency is the degree to which the meaning of an idiom can be inferred from its constituent words (see previous detailed discussion in section 2.3.3.2). It was used at this stage to gauge the level of difficulty that learners might have in inferring the meaning of an idiom based on its constituent words, as previous studies have concluded that transparent idioms are generally easier to decipher than opaque idioms. Importantly for the design of the

experiment, this measure was used to ensure a balance of transparent/easy and opaque/ difficult items in each set. The semantic transparency of the selected idioms was assessed by three native speakers of English, who were experienced ESOL teachers from UK, Canada, and New Zealand. They were asked to rate semantic transparency on an ordinal scale of *High*, *Medium*, and *Low*¹³, based on their own perception of how easily the meaning of an idiom can be inferred from its components. Note that this measure of transparency was ONLY a rough estimation of the difficulty L2 learners would experience in inferring meaning, as (a) there are marked divergences between native English-speaking teachers' and L2 learners' transparency ratings (Boers & Webb, 2015), and (b) inter-rater agreement of such a rating is often problematic¹⁴ (Boers & Webb, 2015; Ramonda, 2016; see also section 2.3.3.2). The result was that 19 out of the 84 idioms (about 22%) received absolute agreement among the three raters; 57 (about 68%) had partial agreements, as one rater had a different rating from the other two (which included 6 extreme cases of H/H/L and H/L/L), and 8 (about 10 %) had absolute disagreements. All the rating disagreements were resolved by choosing the median (see the detailed ratings by the native speakers in Appendix 5). Once the native speakers' estimate of transparency had been calculated in this way, it was subject to potential further modification based on the bilingual researcher's own knowledge of the existence of approximate equivalents in Mandarin (e.g., *tighten one's belt*, *a dark horse*, and *a hot potato*¹⁵). This was because I assumed that these equivalents would help the participants to comprehend the idioms in the actual intervention (and this assumption was verified later), and so the native speakers' estimate would be an

¹³ Note that this measure is different from the etymological transparency ratings by the L2 learners to be reported later.

¹⁴ For native speakers, even though they know how to use idioms, they do not necessarily understand all the literal underpinning of the expressions—their ratings vary due to different understandings of idioms, and to different personal interests and experiences in life. For example, based on observations in this study, Rater S, who liked sports, knew that (*be*) *on the ropes* comes from boxing, and *get your second wind* comes from running, and was apt to rate them as *High*, compared with two other raters who did not know the origins of such idioms. In addition, the raters' judgements vary according to their perspectives.

¹⁵ In Chinese, *tighten your belt* has an exact equivalent expression 勒紧裤腰带(过日子) in both literal and metaphorical sense; *a hot potato* has a similar equivalent 烫手的山芋, only replacing *potato* with *sweet potato* (山芋); *a dark horse* was the same as 一匹黑马 in Chinese, in both literal and metaphorical sense, which was found later that the Chinese expression was likely to have been borrowed from English. Likewise, *sit on the fence* and *turn over a new leaf* also have (semi)equivalent expressions in Chinese, i.e., 骑墙 and 改过自新.

under-estimation of transparency. Thus, the estimated transparency of five such idioms were modified to ‘*High*’, despite a ‘*Medium*’ or ‘*Low*’ given by the native speaker raters. Overall, one third of the selected items were of relatively high transparency, 25% were at the lower end of the scale, and the rest were in the middle. After balancing the four sets with different features, each set had 7 *high* ones, 7 to 10 *medium* ones, and 4 to 7 low ones (see Table 3.9).

Table 3.9 Distribution of the degrees of semantic transparency

	Set A	Set B	Set C	Set D	Total
High	7	7	7	7	28
Medium	10	7	10	8	35
Low	4	7	4	6	21
Total	21	21	21	21	84

In addition to the semantic transparency rating used for material preparation, an etymological semantic transparency rating—a subtler measure of learners’ perception of semantic transparency from a slightly different perspective—was collected from the learners during the interviews (see a further explanation in section 3.4.1).

3) Type of phrase. According to their syntactic function as a unit and the nature of the headword, the idioms were identified as noun phrases, verb phrases, prepositional phrases, and adjective/adverbial phrases. Out of the 84 idioms, 60 were verb phrases, 13 were noun phrases, 8 were prepositional phrases, and 3 were adjective/ adverbial phrases. Considering that idioms of certain structure(s) might be easier to interpret than others, e.g., the availability of a verb in a phrase, and that evidence also showing that prepositional phrases seem to more difficult for L2 learners to decode (Park & Chon, 2019), phrase type was also used to balance, as far as possible, the four sets of idioms (see the distribution in Table 3.10).

Table 3.10 Distribution of the phrase types

	Set A	Set B	Set C	Set D	Total
NP	4	3	3	3	13
PP	2	2	2	2	8
AP	0	1	1	1	3
VP	15	15	15	15	60
Total	21	21	21	21	84

(Note. NP stands for noun phrases, VP for verb phrases, and PP for prepositional phrases. AP includes phrases that are being used adjectivally or adverbially, i.e., *down and out*, *up to scratch*, and *hand over fist*.)

3.3.2 Etymology of idioms

In the teaching experiment, the etymology of the idioms was used as a guessing prompt for learners in the learning phase, before the idiomatic meaning and an example sentence were provided as a full explanation of the item. Since not all the idioms were included in one dictionary, and some dictionaries provide the etymological information for idioms while others do not, a number of dictionaries were used to collect the idiomatic meanings and the etymology of the selected idioms. This included *Collins COBUILD Dictionary of Idioms* (1999), *Collins COBUILD Idioms Dictionary* (2012), *Oxford Dictionary of Idioms* (1999), *Oxford Idioms Dictionary* (2013), *MacMillan English Dictionary for Advanced Learners* (2012), and the *American Heritage Dictionary of Idioms* (2003). The idiomatic meanings were collected mainly from the Collins COBUILD dictionaries (see the full list of dictionary meanings in Appendix 6), while the etymology was mainly from the Oxford dictionaries, the American Heritage Dictionary, as well as online sources, e.g., *the Free Dictionary* (retrieved September, 2016 from <https://idioms.thefreedictionary.com/>) and *the Phrase Finder* (retrieved September, 2016 from <https://www.phrases.org.uk/>). The majority of the selected idioms have a ‘straightforward’ or ‘well-established’ source from cultural-specific domains of life, allegories, or literature; a few have only an alleged etymology, e.g., *on a shoestring* from the idea that ‘debtors in British prisons would lower a shoe by its laces from a window so as to collect funds from visitors or passers-by’ (On a shoestring, n.d.), and several items have more than one possible origin (e.g., *stick your neck out* may relate to a turtle or a chick sticking its neck out). For the last case, the seemingly most relevant origin was chosen to be used (e.g., the ‘turtle’ origin for ‘stick your neck out’).

In compiling the etymological information as a prompt, a crucial principle is that the information does NOT explicitly reveal the current idiomatic meaning of the expression, so that by making their own connection between the etymology and the idiomatic meaning, learners can engage in a ‘deep’ level of processing (Boers et al., 2007; see the full list of etymology prompts in Appendix 6).

3.3.3 Example sentences and context control

Two types of contexts were created for learners to use at different stages of learning.

- 1) **A simple neutral context**—one sentence only, which provides a limited linguistic context for the idiom, e.g. *That rings a bell.* —as an initial clue for guessing;
- 2) **A richer and more supportive context** but without any revealing synonyms—an idiom was incorporated into a longer sentence or sometimes two sentences, e.g. *I've never met John Franklin, but his name rings a bell*—as an example sentence to aid fuller comprehension.

In the first stage of the interviews, it was deemed necessary to keep contextual information minimal, so that the learners would make full use of the etymological information in comprehending the idioms. On the one hand, one aim of the study is to discover the extent to which information about the literal underpinning of an idiom or an elaboration of its etymology can help learners comprehend its idiomatic meaning. Keeping the sentential context minimal was a way of avoiding variability in the interpretability of the idioms owing to the availability of contextual clues. On the other hand, providing the idiom in a simple sentence served to illustrate at least the type of syntactic phrase and whether it was about animate or inanimate object(s), and thus made the interpretation task more realistic than guessing the meaning of a completely decontextualised expression.

After the learners had made their attempts at guessing the meaning and a full idiomatic meaning was explained to them, the example sentence with a richer context was provided, to help the learners fully understand the idiom. This also concluded the complete learning round for each idiom.

Both sentential contexts were extracted from idiom dictionaries or the Corpus of Contemporary American English (COCA) or the British National Corpus (BNC), with modifications made to keep the sentences consistent in length and quality, after which two native speakers of English reviewed the sentences to ensure their authenticity. For words in the contexts (though not for the words in the idiom itself) that might be unfamiliar to participants, an L1 translation was provided, e.g., *His political career is on the ropes after the bribery (贿赂) scandal* (see the full list of example sentences in Appendix 7).

3.3.4 Format of material presentation

All the materials discussed above were presented to the participants on cue cards during the first interviews (see interview procedure in 3.4). There were two cards for each idiom. Card One had an isolated idiom on the front and the idiom in a simple neutral sentential context on the back. Card Two had its full idiomatic meaning on the front and the example sentence with a richer context on the back. The etymological prompt was provided aurally before Card Two was shown to the participants. To take *pull one's weight* as an example, the information was shown in the order of 1 to 5 (see Figure 3.2 below).

<p>1. Idiom</p> <div style="border: 1px solid black; padding: 10px; text-align: center; margin: 10px 0;"> <p>pull one's weight</p> </div> <p>[Card One—front]</p>	<p>2. The simple context</p> <div style="border: 1px solid black; padding: 10px; text-align: center; margin: 10px 0;"> <p>He needs to pull his weight.</p> </div> <p>[Card One—back]</p>
<p>3. Prompt of etymology (aural)</p>	
<p>4. The idiomatic meaning</p> <div style="border: 1px solid black; padding: 10px; margin: 10px 0;"> <p>pull one's weight</p> <p>To do one's share in a common task/ To work as hard as other people in a group</p> </div> <p>[Card Two—back]</p>	<p>5. The rich context</p> <div style="border: 1px solid black; padding: 10px; margin: 10px 0;"> <p>If he doesn't start pulling his weight, he'll lose his job .</p> </div> <p>[Card Two—back]</p>

Figure 3.2 Cue cards and the prompt

(Note. Numbers indicate the order in which the different types of information were provided to the participants)

3.4 Data collection

The experiment consisted of two phases, a learning phase and a recall phase, for which the 25 participants were interviewed individually in a quiet office space with a one-week interval between phases (see Table 3.11 for an overview). The first interview lasted from an hour to an hour and a half, including a learning session of about one hour on average where different materials and information about a set of 21 idioms were presented for the participants to learn. There were generally four steps involved for each idiom to be fully grasped, and during the interaction with the interviewer, participants' responses were elicited and collected. One week later, the second interview took about 40 minutes, including a 15-minute meaning recall test (a delayed post-test) immediately followed by a retrospective think-aloud task and then a semi-structured interview on idiom learning. Details of the procedure are explained in the following sections.

Table 3.11 Procedure of data collection

Learning phase (Interview 1)		
	Information presented	Data collected
Step 1	Idiom [Card 1 front]	Idiom and word familiarity
Step 2	Idiom in a simple sentence [Card 1 back]	Guess(es) before the prompt
Step 3	Prompt of the literal underpinning (Given verbally)	Guess(es) after the prompt
Step 4	Explanation [Card 2 front] & Example sentence [Card 2 back]	
<i>Repeat the above procedure with all the target items</i>		
	Etymological semantic transparency scale	Transparency ratings on Likert scales
Recall phase (Interview 2)		
	Tasks	Data collected
Part I	Meaning recall test	Recall answers
	Retrospective think-aloud task	Recall of the prompt & others
Part II	General questions on idiom learning	Strategies of & attitudes to idiom learning

All the interviews were audio recorded after obtaining the participants' written consent (see *Participant Consent Form* in Appendix 2). All the teaching materials were given in English and explained in Chinese whenever needed, but the interaction between the researcher and the participants in both interviews was mainly in the participants' L1 Chinese. In the think-aloud interviews, participants were allowed to use whichever language to express their thoughts, Chinese or English, that reflected their spontaneous cognitive processes and that facilitated their self-expression. Although most participants used Chinese when indicating what they thought the target idioms meant, some higher proficiency participants used English. Regardless of the language used, participants' responses were all transcribed, and the excerpts in Chinese were translated into English by the bilingual researcher (a native speaker of Chinese and near-native English speaker) to facilitate the rating of the correctness of the responses which was done together with two native English speakers (see full explanation of data processing in section 3.5). To ensure the reliability of the translation, a native speaker of Chinese with advanced-level English also reviewed the transcripts and double-checked the translations. Any unclear or unsatisfactory translation was resolved by consulting the dictionaries and through discussion between the reviewer and the researcher.

3.4.1 Learning phase

In the first interview, each participant was first briefed on the research with the *Information Sheet* (see Appendix 1) and the *Participant Consent Form* (see Appendix 2) in both English and Chinese. After they gave written consent for their data to be used for this research project by signing the forms, the interview procedure was explained to them in Chinese. The English translation of the directions was as below (also in Appendix 8 *Directions for the First Interview*):

In this experiment we are interested in what you think about when you figure out the meanings of 21 English idioms. An idiom is an expression or phrase that doesn't mean what it says—it often has a figurative meaning, which is derived from its literal meaning. I'm going to ask you to THINK ALOUD as you guess the figurative meaning of the idioms.

What I mean by thinking aloud is that I want you to tell me EVERYTHING you are thinking from the

time you first see the idiom until you tell me what it means. Some questions going through your mind after you see the idiom might be: Does a certain word give away the meaning of the idiom? How does the context help with guessing? (I will show you an example of the idiom.) Is there a similar expression in my native language? During the process, I will also provide you with some hints about the literal meaning or the etymological source of the idiom to help you work out the meaning. Please tell me anything you can think of based on the prompts I give you during the interview process. You will get to know the figurative meanings in the end.

If you are familiar with the idiom, let me know from the start. If you don't understand a word in the expression, tell me and I will explain it to you at the beginning.

After the participant indicated that they had understood the procedure, the recording started. The 21 idioms were presented to the participant on cards one after another (as shown in Figure 3.2), and the interview proceeded in the following steps:

Step 1) The idiom was presented in isolation, e.g., ‘throw in the towel’. First, the participant was asked if they were familiar with the idiom. If they knew it, they were asked to give its meaning. If it was clear from this that they already knew the idiom, they were asked if they knew its origin, and the explanation of the literal underpinning was given if necessary. Then the interview moved on to the next item. If they could not give the meaning, we moved on to Step 2. However, if the participant indicated that they were not familiar with one of the words in the idiom, the meaning of word was explained to the participant before we moved on to Step 2. Crucially, though, if the unknown word was a key word of the idiom and explaining the key word would involve explaining the literal underpinning of the whole idiom, then we moved on to Step 3, and thus the guess before the prompt was recorded as nil. The purpose of Step 1 is to check learners’ familiarity of the idioms and clarify any unknown words that blocked the interpretation.

Step 2) The idiom was presented in a simple and neutral context, i.e. one sentence without revelatory semantic clues (e.g., ‘I’m ready to throw in the towel’), and the participant was asked to guess the meaning of the expression. Providing the idiom in a sentence served to illustrate at least the type of syntactic phrase and whether it was about animate or

inanimate object(s), which made the interpretation task more realistic. It also ensured that participants would not linger on the superficial literal sense of an item (see also previous discussion in 3.3.3).

Step 3) The researcher explained the literal underpinning of the expression. This served either as a prompt for the participant to make another guess (if not the first attempt) at the figurative meaning of the idiom, or as new information confirming or clarifying their earlier correct interpretation. For example, “*Throw in the towel* comes from boxing. When a boxer is knocked to the ground, his coach or manager may throw a towel into the boxing ring.”

Step 4) The researcher explained the meaning of the idiom given in the dictionary (e.g., to throw in the towel means ‘to admit defeat or to give up’)—either as a confirmation or clarification of the participant’s interpretation in the previous step if that was correct or approximately correct, or as a correction if their previous interpretation was wrong. Then the idiom was presented in a longer sentence with a richer context so that learners could get a better understanding of its meaning and use (e.g., ‘It’s hard to find a job at the moment, but you shouldn’t throw in the towel.’) After both the literal underpinning and the actual idiomatic meaning of the idiom were clarified, the interview moved on to the next idiom.

After their set of 21 idioms had been introduced to them following the above steps, the participants were asked to do an **etymological semantic transparency rating task**. The participants were encouraged to express their thoughts and explain their reasoning as they carried out the rating in the same order as they had previously seen them. If they forgot the literal underpinning or the figurative meaning of an idiom (as sometimes happened especially after a long learning session), the meanings would be re-explained to them. Measures were taken to ensure that their ratings were based on correct and full understanding¹⁶. For example, if a participant appeared to do the rating hastily, the researcher would ask them to slow down

¹⁶ Note that while they were explicitly asked to base their rating on what they had just learned about the idioms, their ratings would still be naturally affected by their first impression of the idioms and prior knowledge of their constituent words, which could be an inherent feature of this kind of rating.

Below are the 21 idioms that we have discussed.

You may feel that the figurative (idiomatic) meaning of some idioms follows in a straightforward manner from the literal (source) meaning of the expressions. For other idioms, you may find it much less obvious how the figurative (idiomatic) meaning is derived from the literal (source) meaning. You may even feel that there is no clear link at all.

Now that you know the literal meaning as well as the figurative meaning of the idioms, please circle the number on the scale below that best represents how you feel about the semantic relationship between the literal meaning and the figurative meaning of the idiom. You can also explain your choice on the lines behind the scales.

0 = The link between the literal and the figurative meaning of the expression is not at all clear to me. Although I (now) know both the literal (source) meaning of the expression and its idiomatic meaning, I cannot see how the literal use could have given rise to the figurative use of the expression.

non-transparent¹⁷ transparent

←-----→

0	1	2	3	4	5
The link between the literal and figurative meanings is not transparent at all.	The link between the literal and figurative meanings is very vague.	The link between the literal and figurative meanings is vague.	The link between the literal and figurative meanings is somewhat clear.	The link between the literal and figurative meanings is clear.	The link between the literal and figurative meanings is very clear.

on the back burner 0 1 2 3 4 5

115

At the end of the interview, the participants were asked not to discuss individual items encountered in the interview with other students. They were informed that the second interview was also going to be about idiom learning but were not told that they would be tested.

3.4.2 Recall phase

One week later, the same participants were invited for a second interview. This interview consisted of two parts. The first part consisted of a meaning recall test followed by a retrospective think-aloud task. In the recall test, they were encouraged to write down the meanings in English, i.e. using the same language as was used when the meanings were given in the first week (see the recall test sheet of set A in Appendix 10). In the subsequent think-aloud task, they were allowed to give further explanations of their recall answers in Chinese, because they might have found it difficult to find appropriate English words to express (their understanding of) the idiomatic meanings, and in addition because grammatical errors in the learners' English might prevent the raters from evaluating the correctness of the recalled answers (see the rating procedure in 3.5.2).

During the verbal explanation of their answers, participants were also asked to say what helped them recall the meaning and whether they could recall anything else that they had learned about the idioms in the first interview. This task was intended to gauge (1) whether successful recall of idiomatic meaning coincided with recall of the literal underpinning and (2) whether other information, e.g., the example sentence or the L1 equivalent, played a part in helping learners remember the idiomatic meaning.

The second part of the session was a semi-structured interview about more general aspects of idiom learning: (a) learners' strategies to cope with unfamiliar (idiomatic) expressions, (b) methods they use to memorise such expressions, (c) their personal experience with and awareness of English idiomatic expressions, and (d) their general attitudes towards idiom learning and usage. Accordingly, a guideline of four main questions was followed, from which related questions were developed and discussed in the interview:

Q1. What do you usually do when encountering unfamiliar English idioms (e.g., ignore them, make a guess at their meaning, look them up in a dictionary)?

Q2. When and where do you remember encountering such figurative idiomatic expressions in your own learning experience?

Q3. What do you often do to remember such expressions (e.g. associate them with where you encountered them, associate them with a mental picture, compare them to a similar Chinese expression, write them down, try to use them)?

Q4. Generally, have you ever tried to use idioms once you learned them?

These questions were designed for two purposes: one is to investigate EFL learners' experiences with and awareness of idioms, their strategies and attitudes towards idiom learning, and their attitudes towards using idiom, so as to explain their learning performances and outcomes in the teaching experiment; the other is to provide a better picture of L2 idiom learning in a Chinese EFL context—the background of the teaching experiment—from the learners' perspective.

In addition to the four main questions, a set of questions were asked that concerned participants' post-experiment experience with idioms as well as their awareness of and attitudes towards them:

Q5. Since the first interview

- 1) Have you invested any effort in trying to recall the idioms which we talked about? If so, please tell me about it.
- 2) Have you noticed any of these expressions being used since our first interview?
- 3) Have you noticed any other idioms? If so, do you think you would have paid the same attention to them if we hadn't had that interview? If not, will you pay the same attention as before?

These questions also served several purposes. One is to gauge any potential extraneous factors that might arise between the learning and recall phases. For example, in answering the second question in Q5, three participants reported that one idiom they learned in the previous interview had happened to be taught in their regular English class, which means the effectiveness of the teaching experiment was confounded by the classroom teaching. The responses relating to such idioms that were re-encountered in the one-week interval between the two interviews were therefore excluded from analysis. The first question of Q5 would serve a similar purpose if a

participant deliberately put extra effort into remembering the idioms. However, if learners spontaneously recalled the idioms after the interview (e.g., a participant thought of *give a green light* at the sight of the traffic light), their responses were valid for analysis. The third question was asked to evaluate the potential effects of the teaching experiment on learners' attitudes towards idiom learning.

3.5 Data processing

The analysis of the interview data was based on transcription and translation of the interviews themselves, as well as on additional ratings of the learners' responses and on coding both of these responses and of other relevant factors. The following sections elaborate on each of these stages, in which the researcher worked as both the transcriber and translator of the interview data, and as one of the raters of the learners' responses.

3.5.1 Transcription and translation

Depending on the type of information they contained, some portions of the interviews were transcribed word-for-word while others were transcribed more selectively. Selective transcription was used for the learning phase and the retrospective think-aloud task of the recall phase. As the interactive mode yielded a large number of responses at each step, transcription included isolating participants' responses from the conversation flow and transcribing their responses when they were guessing the meaning of the idioms. The responses were organised according to the stages of the learning procedure (i.e., guesses before or after the prompt), as well as according to how many guesses had been made at each step (see the counts of guesses in 3.5.3). Word-for-word transcription was carried out for the answers to the general questions on idiom learning strategies and attitudes, in preparation for more qualitative analysis. Completed transcripts were checked carefully against the recordings, which was especially crucial for the selective transcriptions.

The interviews were translated while they were being transcribed, and the translations were subsequently carefully checked and revised if necessary. Accuracy of the translations was important, because two native English-speaking raters based their ratings on these translations

and not on the original recordings. One challenge of translating learners' responses for target idioms was rendering a large number of idiomatic Chinese expressions into English, as L2 learners often attempted to interpret L2 idioms via seemingly similar L1 Chinese idiomatic expressions. For example, many participants interpreted *go with the flow* as 随大流 (*pinyin: sui da liu*), a Chinese idiomatic expression meaning 'to follow the crowd'. This is not surprising, since L2 learners frequently resort to using their prior linguistic knowledge (including of their L1) in order to infer the meaning of a new expression. However, translating idiomatic language was often a demanding task (not all are like 随大流 which happens to have an equivalent in English). To achieve appropriate translations for this kind of response, a full explanation of both the literal and figurative meanings of the Chinese idiomatic expressions was always provided as part of the translation after consulting Chinese-to-English dictionaries (e.g., *A New Century Chinese-English Dictionary*, 2004) and English-to-Chinese dictionaries (e.g., *Oxford Advanced Learner's English-Chinese Dictionary*, 2014 and *Oxford Idioms Dictionary 牛津英语习语词典*, 2013). As noted earlier, to ensure the reliability of the translation, and importantly to ensure that there was no information lost in translation, the Chinese translations were reviewed by another Chinese-English bilingual.

For the recall phase, participants' written English answers were copied from their recall test sheets and their additional oral explanations in Chinese were transcribed and translated. It is worth noting that, because they were better able to express themselves in their native Chinese than in English, participants' written English answers sometimes differed slightly from the translation of their Chinese explanations. Note also that the explanations in Chinese were deemed a necessary and valid part of their answers. Therefore, subsequent rating of the recall answers was based on the written English answers and the translations of their Chinese explanations as a whole—this was important for raters to follow, as the purpose of the rating was to evaluate the correctness of the participants' recall of the meanings of the target idioms rather than the participants' English proficiency as reflected in their written English answers.

3.5.2 Rating the guesses and recall answers

As they were being transcribed and translated, the interview data were organised into three sections, reflecting the stages of the learning procedure: *Guess before the prompt*, *Guess after*

the prompt, and *Recall answers*. The first two of these evaluate how well learners can understand the idioms, with and without the help of information about the literal underpinnings, i.e., the ‘prompt’. The third assesses how well they can remember the meanings of the idioms a week later. To obtain a measure of learners’ understanding of and memory for the idioms, the guesses before and after the prompt and the recall answers were all marked by three raters, namely the bilingual researcher (i.e., the author) and two native English-speaking researchers. The responses of the four sets of data were rated separately, yielding 441 datapoints for Set A and 378 datapoints for each of the other three sets. In total, 1575 datapoints were rated for further analysis (see Table 3.12 for the distribution of datapoints).

Table 3.12 Distribution of the datapoints

	Guess BP	Guess AP	Recall answers	Total (N)
Set A	147 (= 21 x 7)	147	147	441
Set B	126 (= 21 x 6)	126	126	378
Set C	126 (= 21 x 6)	126	126	378
Set D	126 (= 21 x 6)	126	126	378
Total (n)	525	525	525	1575

The rating of the responses was a very demanding task. To rate meanings of idiomatic expressions is especially challenging as they contain figurative senses that can be hard to capture. In addition, the guesses were generally more diverse and difficult to rate than the recall answers, as the learners’ reactions to unfamiliar idioms varied more widely before more explanatory information was given. Overall, rating the guesses and recall answers of idiomatic meanings posed three fundamental challenges to the raters.

Challenge 1: The most common challenge was evaluating meaning based on open-ended definitions given by the learners. Compared to dictionary explanations of the idiomatic meaning of an idiom, learners’ responses were sometimes narrower (more specific), and sometimes broader (more general), especially when not much information or context was given at the early stages of the interview. In addition, there were sometimes subtle differences in the affective senses between learners’ guesses and the target meaning.

Challenge 2: A related problem was the difficulty in judging whether key aspects of the meaning of an idiom were captured in the learners' responses. The meaning of an idiom sometimes entails key information relating to different aspects of a scenario. This information was often only partially captured in learners' responses. For example, *Take the bull by the horns* means 'to act decisively and with determination and courage in order to deal with a difficult situation or problem' (Take the bull by the horns, 1999), which entails determination and bravery in the manner of one's action, as well as the difficulty of the situation. The number of the units of these key aspects (later called 'key points') and their weight in the whole meaning vary from idiom to idiom. The diversity of the idioms meant no single sweeping rule could apply to scoring responses for all the idioms (see more discussion in 3.5.2.1 and 3.5.2.2).

Challenge 3: Rating became even trickier when the responses (more often in the guesses than in the recall answers) involved Chinese idiomatic expressions. Participants sometimes attempted to interpret an unfamiliar English idiom via Chinese idiomatic expressions. This posed a challenge not only for the translator, but also for the raters. For example, in participants' responses to *go with the flow*, three common Chinese idiomatic expressions were used to interpret the idiom: (1) many participants used 随大流 (*pinyin: sui da liu*), as mentioned earlier, a Chinese idiomatic expression, which literally means to follow the main (water) flow, but its idiomatic meaning equals another English idiom *follow the crowd*, meaning 'go along with the majority, do what most others are doing' (Follow the crowd, 1997); (2) some used 随波逐流 (*pinyin: sui bo zhu liu*), a Chinese idiom (very similar to 随大流 and showing disapproval) which is equivalent to (*follow*) *like sheep* in English, and is often used to describe people who just follow other people because they have no ideas or judgement of their own; (3) one participant used the Chinese idiom 随遇而安 (*pinyin: sui yu er an*), which means to take things as they are and accommodate oneself to circumstances. Rating, in this sense, involved comparing the meanings of idiomatic expressions in the two languages, and the Chinese idiomatic expressions in learners' responses could overlap with but were not completely equal

to the target idiom, which rendered the judgement rather difficult.¹⁸

All these challenges required the raters to have a thorough appreciation of the meanings of the target idioms, correct understanding of learners' responses, and sensitive discrimination of nuances in different words and expressions, and to then provide a careful evaluation of the responses. The rating task was accomplished via joint efforts of the three raters: the bilingual researcher, who has experience in linguistic research, language teaching, and translation involving the two languages concerned, and two native-English speakers, who were experienced ESOL teachers and had an advanced understanding of the linguistics of multiword expressions as researchers in this field.

To operationalise and facilitate the rating, a rating scheme and a standard rating procedure were established through two rounds of pilot trials (see 3.5.2.1 and 3.5.2.2). A final scoring system was adopted to obtain a final score for each datapoint, based on agreement among the three raters through discussion and applying a majority rule.

Rater reliability is of vital importance. It represents the accuracy and consistency of ratings within and between raters (see discussion of inter-rater reliability in 3.5.3.3). A number of procedures were set up to maximise intra- and inter-rater reliability in the face of 1575 datapoints and the challenges mentioned above. Throughout the whole process, a considerable amount of research into Chinese and English words and expressions was undertaken, along with multiple discussions about the rating process.

3.5.2.1 Rating criteria and rating schemes

Before the actual ratings were obtained, two rounds of pilot trials were run to establish agreement among the three raters on the rating criteria, the rating schemes, and the rating procedure. The first round involved the rating of data from two participants in set A and two in

¹⁸ For the responses with three Chinese idioms, 隨遇而安 was later agreed by the raters as the closest interpretation of *go with the flow*, as the definition of the English idiom is 'to let things happen to you or let other people tell you what to do, rather than control what happens yourself' (Collins COBUILD, 1999)—in other words, 'be relaxed and not worry about what you should do' (Oxford Idioms Dictionary, 2013). 隨大流 and 隨波逐流 overlap with the target meaning in the sense that 'let other people tell you what to do' but they do not fit in other situations where no other people are involved, so the related responses were finally rated as half correct.

set B, in order to establish the rating schemes. These data were first rated by the three raters independently and then discussed in a group to identify potential issues and discrepancies. (This procedure was later found essential for rating a large set of datapoints and was absorbed into the actual rating procedure.)

Initially, we needed to agree on a standard explanation of each idiom that could be used as a reference explanation for that idiom. The dictionary meanings of the idioms—the same as those used in the learning phase—were adopted as the standard reference examples. Essentially, the correctness of a guess or recall response was evaluated according to how closely it resembled the dictionary explanation.

Three types of rating schemes were initially proposed: a binary scale, a 3-point scale, and a 5-point scale. The binary correct/incorrect (1/0) scheme was soon discarded, because of some of the challenges mentioned above and because L2 learners' own interpretations can be vague and obscure, and easily fall in a grey area that was incompletely correct. Therefore, the binary scheme was ruled out, and a 3-point rating scheme and a more elaborate 5-point scheme were piloted.

The 3-point scale was made up of three scores 0, 0.5, and 1 representing three categories 'correct', 'partially correct', and 'incorrect or no response' (see Table 3.13 for a full illustration). As mentioned earlier, it was found that an idiom sometimes depicts a scenario that contains more than one key aspect of information. To a large extent, the meaning can be identified as several discrete reference points, although one point might relate to another in some way. For example, *pull one's weight* means 'to work as hard as other people in a group' (Pull one's weight, 1998). It entails (a) doing one's fair share of work and (b) contributing to a group activity. If the key point(s) of the idiomatic meaning were fully captured in a learner's response, then it would be marked as 1; if that was partially captured, such a response as 'to do one's best', it would be marked as 0.5 (see further discussion on scoring in 3.5.2.2). From another perspective, if a response was related to the core meaning but was too broad or too narrow in some sense, it would also be marked 0.5. For example, *give someone a leg up* means 'to give someone an advantage' (Give someone a leg up, 1999), in other words, to help them achieve something that

they couldn't have done alone. If a response was 'to help somebody', it would be marked as 0.5, as 'help' was agreed by the raters as too general, compared to 'an advantage'.

Table 3.13 Rating scheme

1	Correct	The core meaning of the idiom was captured. <i>(The core meaning contains all the key information identified in an idiom, as per the dictionary definition.)</i>
0.5	Partially correct	Related to the core meaning of the idiom, but too narrow or too broad, or some of the key information was missing.
0	Incorrect	Irrelevant or wrong answer
	No response	No guess in the learning phase; No answer in the recall test. <i>(No guess in the learning phase was also coded as 'N/A' and treated separately in different analyses, among which wrong guess(es) and no guess in the learning phase were assumed to have different impact on memory.)</i>

The 5-point scale had two more scores 0.25 and 0.75, indicating a response was partially correct but leaning towards the wrong or right side. The advantages and disadvantages of a 5-point scale were both evident in comparison with the 3-point scale. On the one hand, it allowed a finer-grained evaluation of the responses, especially those guesses in the learning phase. When a response from a participant was very obscure or hard to judge, the 5-point scheme provided more options for the raters to better weigh the subtle nuances in meaning. On the other hand, it decreased the probability of immediate agreement among the raters, which rendered agreement on the final scores more time-consuming and harder to obtain (see the rationale of final scoring in 3.5.2.2).

Taking both aspects into account, the 3-point scale was adopted as the principal rating scheme. However, the two native English speaking raters used the 5-point scale alongside the 3-point one, so that they could make supplementary notes to facilitate the group discussions in which the final scores were arrived at. That is, these additional notes served as a reminder of the more nuanced considerations they had taken when doing their individual ratings. For example, if the three raters rated a response 0.5, 0.5, and 0 on the 3-point scale, and the one who rated 0 had a 0.25 on his/her 5-point scale rating, it might help resolve the disagreement and settle on a final score of 0.5 more easily. In this sense, the 5-point scale increased efficiency in resolving

(particularly minor) inter-rater disagreements (see more explanation in 3.5.2.3). Therefore, the raters were allowed to base their judgements on two complementary systems in their independent ratings—the 3-point scale as the principal one and the 5-point scale as a facilitative one.

In the final scoring, the scores were all determined on the basis of the 3-point scheme, and this 3-point scheme provided the data for further statistical analysis. When the ‘no responses’ are separated from the incorrect responses in 0 scores, the current scheme matches the 4-category scheme used in a related study by Ramonda (2016), in which participants’ L1 paraphrases of the target L2 idioms were rated as: *correct*, *partially correct*, *incorrect* and *blank*.

3.5.2.2 Rating procedure

Once the rating schemes were established, a second round of pilot trials was run to set up a standard procedure to train raters for the actual rating. In this second round, a subset was selected consisting of 6 of the 21 idioms in each set. These items had particularly complex meanings and/or would be highly likely to invoke the use of Chinese idiomatic expressions in the learners’ explanations. Such items are particularly likely to prove problematic for rating, and so a discussion leading to agreement as to how such items should be rated should prove an effective training procedure for the rating process as a whole. This pilot round resulted in a three-step rating procedure that was subsequently applied to all idioms:

Step 1 Identifying key points of meaning for rating

Step 2 Independent ratings by individual raters

Step 3 Final scoring based on the raters’ agreement

First of all, for each idiom, key points of meaning were identified, in order to provide an initial guideline for the subsequent rating (see *Challenge 2* in 3.5.2). To achieve this, the three raters first went through the dictionary meanings of the 84 idioms together. Through discussion, key points of the dictionary meanings were identified and weighed for each idiom, and as a result, a score was roughly divided and/or distributed among the identified key points in each idiom. There were generally three circumstances:

When the meaning of an idiom contained only one key point (e.g., *show someone the ropes* meaning ‘teach someone how to do something, especially a job’ [Show someone the ropes, 2012]), then if a learner’s response captured this key point correctly, it would be marked as 1, and if it partially captured this key information (e.g., ‘show someone how to use something’), then it would be scored as 0.5.

When the meaning of an idiom contained two key points, the scores were evaluated against the two points. This included two situations. If the raters evaluated the two points as equally important, a score was split into 0.5 for each point. For example, *pass the buck* means ‘to blame someone or make them responsible for a problem that you should deal with yourself’ (Pass the buck, n.d.), which entails two points (a) one shirks or evades their own responsibility and (b) they let other people deal with it instead. For this case, if a learner’s response only captured the first point, it would be rated as 0.5. If the raters identified one key point as the main point and the other as a subordinate one, a full score was preferably given to the main point if that was captured in a learner’s response. For example, *let the cat out of the bag* means that ‘to reveal something secret or private, often without meaning to’ (Let the cat out of the bag, 1999) The main point was to reveal a secret, and the fact that it was revealed carelessly or by mistake was a subordinate point, because the dictionary meaning contains “often”, suggesting the idiom is not always used to refer to the accidental nature of revealing the secrets. Thus, such a response as ‘to reveal a secret’ would be rated as 1.

When there were more than two key points (which rarely occurred), they were likely to be linked to some extent, and the raters had to discuss and weigh the importance of different points to decide a score for each. A special example was *pull a rabbit out of the hat*. It means ‘to surprise everyone by suddenly doing something that shows a lot of skill, often in order to solve a problem’ (Pull a rabbit out of a hat, 1998), which entailed three interrelated key points: (a) to solve a problem, (b) by using a skill/ability and (c) achieving a surprising/unexpected effect. If two out of the three points were captured in a learner’s response, e.g. ‘Someone owns a special power or ability to have something done’ or ‘to do amazing things that are beyond the expectation of other people’, the response was given a full score; if only one point were

captured, e.g. ‘to have an ability to do something’ or ‘to surprise someone’, it was given a half score.

With the key points identified as initial guidelines, the three raters’ ratings, however, still showed quite a few discrepancies in the pilot rounds, due to two main reasons related to *Challenge 1* and *Challenge 3* (in section 3.5.2). One is that the raters based their ratings on different understandings of learners’ guesses. For example, one rater may have given more attention to one aspect of the response while another focused on some other aspect(s). The other is that the English-speaking raters occasionally misunderstood responses which contained a Chinese idiomatic expression. Because these discrepancies could potentially harm the research results, **a final scoring method** was adopted, with the final score based on agreement between the three raters. In this way, the final score took into account the ratings from all three raters, with a majority rule applying when disagreements remained unresolved despite discussion (this happened at a rate of less than 2%; see full explanation in 3.5.2.3). It was hoped that this would ensure a more reliable score for each response.

In comparison to a more commonly used method, i.e., averaging scores among raters, the adopted method was a lot more time-consuming and effortful. However, in view of the challenges in translation and the difficulties in judging the learners’ responses about the idioms’ meanings, it ensured that the native English-speaking raters had as good an understanding as possible of the L2 learners’ thoughts and responses, and that a score was given with due evaluation and judgement. More importantly, it allowed discussion and resolution of potential problematic ratings when the assessment of a response was either incomplete or involved a misunderstanding, which could be hidden by averaging. For example, averaging three ratings of 0, 0, and 1; or 0, 0.5, and 1; or 0, 1, and 1, could result in a meaningless score as such patterns of ratings indicate that there might be major and unresolved disagreements among the raters (see more discussion in 3.5.2.3).

In summary, the raters went through three steps for each set of data one by one: **Step 1** identifying the key points in dictionary meanings, **Step 2** providing their independent ratings, and **Step 3** entering group discussion to agree on final scores. For independent rating, the raters were encouraged to finish rating each set within an uninterrupted period of time, so they could

remain consistent in their own ratings. For final scoring, the independent ratings were compiled by item, so that all the ratings of the same idiom could be reviewed and discussed across learners and across learning stages.

3.5.2.3 Inter-rater reliability and accounting for disagreements

Based on the established rating criteria, the 3-point scheme, and the 3-step procedure, the actual rating was undertaken, and the independent ratings from the three raters were collected for the final scoring. Krippendorff's coefficients¹⁹ of the three raters' independent ratings were calculated to gauge the inter-rater reliability before the final scoring. As shown in Table 3.14, all the coefficients of ratings for different stages indicate that the three raters had acceptable rates of agreement prior to discussion. *Guess Before Prompt* had the highest reliability coefficient among the three stages, as in most cases the learners couldn't get the correct meaning at this stage, and the incorrect answers yielded more consistent ratings of 0. In contrast, the *Guess After Prompt* varied a lot more after the information of the literal underpinning was provided, which, as a result, yielded more heterogenous ratings among the raters, and the reliability coefficient was smaller than that of *Guess Before Prompt*. The coefficient α of *Recall Answers* was higher than that of *Guess After Prompt*, which could be explained by that the learners had learned the idioms and the recall answers had a better chance of being closer to the correct idiomatic meanings, and thus were more likely to generate more consistent ratings among the raters. Overall, there was substantial agreement among the three raters in their independent ratings prior to the discussions.

Table 3.14 Inter-rater reliability

	Guess Before Prompt	Guess After Prompt	Recall Answers
Krippendorff's α of the 3 raters	0.797	0.739	0.780

¹⁹ Krippendorff's coefficient (also called Krippendorff's alpha) is commonly used for measuring inter-rater reliability. It is applicable to multiple raters with different scale metrics including ordinal scales, and can handle the missing values, e.g., no response or recall answer in this study. The value of α ranges from 0 to 1, where 0 is perfect disagreement and 1 is perfect agreement. "[I]t is customary to require $\alpha \geq .800$. Where tentative conclusions are still acceptable, $\alpha \geq .667$ is the lowest conceivable limit" (Krippendorff, 2004, p. 241).

Furthermore, among all the independent ratings (including both the guesses and the recall answers), about 60% of the cases received absolute agreement among the three raters prior to the final scoring, and the scores for these cases were therefore settled without the need for discussion. For the rest of the cases, most disagreements involved one rater rating differently from the other two. The majority of these two-versus-one cases had a discrepancy of 0.5 between two raters and the other one. These cases were relatively easy to resolve after discussion with the aid of their additional 5-point ratings (as mentioned earlier); for a few cases where one rater could not agree with the other two after discussion, the majority rule applied. Greater efforts were required in resolving cases where one rater rated a response as 1 ‘fully correct’ while another rated it as 0 ‘incorrect’. For these cases, the three raters analysed the phrases together by consulting the referenced dictionaries where the dictionary meanings were collected (see section 3.3.2), along with other dictionaries, e.g., *Oxford Idioms Dictionary* 牛津英语习语词典 (2013, a bilingual version). Subtle meanings of words or expressions in the idioms as well as the learners’ responses were carefully considered and analysed until a consensus was reached. The final scoring generally took two to three sessions—about 5 hours for each of the four sets of data.

3.5.2.4 Strengths and limitations

In summary, through transcription, translation and rating, learners’ responses to the target idioms were quantified for further analysis. The final rating results were obtained through the joint efforts of the bilingual researcher as the transcriber, translator and one of the raters, and the two native English-speaking speakers as the other two raters. In terms of the whole process, there are both strengths and limitations of such an arrangement of roles and expertise.

First, the multiple roles of the bilingual researcher, who shared the same L1 with the participants, had three main advantages for this L2 idiom teaching/learning research (the role of the researcher is indicated in square brackets in each case):

- 1) Any obstacle to comprehension encountered by L2 learners in the learning phase could be identified and tackled during the interviews, and noted for analysis, e.g., any word in the idiom that was unknown to or misunderstood by a learner [Interviewer/Teacher and Transcriber];

- 2) L2 learners' interpretations of an idiom and their reasoning in L1 could be understood and translated for evaluation [Translator];
- 3) Any misunderstanding of learners' responses by the native English-speaking raters, resulting, from example, from a peculiar Chinese idiomatic expression or its translation, could be identified, explained and resolved during final group discussions [Rater/Facilitator].

This arrangement, however, unavoidably introduced bias into the ratings. As a sympathetic Chinese-speaking learner of English who can relate to the problems that the participants faced in English, I found my original ratings typically higher than those given by the other raters. Fortunately, this bias was largely offset in the group discussions with the native English-speaking raters for the final scoring.

Second, the two monolingual English-speaking raters, who were both experienced ESOL teachers and researchers of multiword expressions, provided expert judgements from the point of view of native speakers. In a practical sense, the bilingual researcher and two native-English speakers worked together to understand and evaluate the L2 learners' understanding and learning of idioms, by bridging a bilingual gap. It has to be admitted that for data such as these that involve two completely different languages, the ideal raters would be strong bilinguals with native-level proficiency and an accurate understanding of both languages as well as metalinguistic awareness in both languages. However, in reality, there were no such ideal bilingual raters available where the data were processed.

3.6 Data analysis

After data processing, a mixture of numeric and non-numeric data was obtained from the interviews. Numeric data include the scores of participants' responses in the learning phase and the recall test, the counts of the responses, the etymological semantic transparency ratings, learners' proficiency scores, etc. Non-numeric data include transcribed and translated textual data, e.g., participants' specific responses to idioms, their recall answers on the test papers, their answers to structured interview questions; and also observational data, e.g., comments that learners made about which words they did not know or their reactions to unknown words, as well as L1 cultural or linguistic interferences. All the numeric data and much of the non-numeric data were coded for analysis. Details of coding are explained in 3.6.1.

Both quantitative and qualitative methods were employed in data analysis to address the research questions from different perspectives. Statistical analysis was conducted (a) to evaluate the overall learning process and outcome of the teaching experiment, (b) to detect factors that affect learners' interpretation and memorization of idioms under the examined teaching approach, and (c) to pinpoint cases for further in-depth investigation. An introduction of statistical tools and terms is given in section 3.6.2 below. Qualitative analysis proceeded in concert with statistical analysis to tap into the underlying factors of learners' behaviours and cognition, e.g., learning strategies, cognitive styles, and L1 culture and transfer in L2 idiom learning.

3.6.1 Coding variables and factors

Coding of variables and factors is an important part of data analysis in this study. The conduct of coding was performed at different stages of analysis. Some factors were coded during the transcription and translation of the interviews. These include familiarity of the item (i.e., whether a participant had prior knowledge of an idiom), prior lexical knowledge of the item (i.e., whether a participant indicated where there was an unknown word or a polysemous/homonymous word in the idiom—known idioms marked as 1 and unknown ones 0), L1 transfer (i.e., whether the participants came up with a Chinese idiomatic expression in their responses; see further classification later in Chapter 5 and coding in Chapter 6), and the recall of the

prompt (i.e., whether the participants could remember the etymological information in the second interviews). Some factors were coded during the analysis of participants' responses in the interviews and answers from the recall test. These include the correctness of guesses before and after the prompt and the recall answers, as well as the number of guesses made after the prompt.

All these factors can be categorised in terms of the aspects or levels they pertain to, as the codes emerged at the level of the items or cases (e.g., etymological semantic transparency), at the level of individual participants (e.g., L2 proficiency) and at the level of the process involved (e.g., correctness of recall, and the recall of the prompt). Table 3.15 provides such a summary of codes of all the variables and relevant data types.

Table 3.15 Codes and data types of factors and variables

Level	Factor	Code (label/value)	Data type
Process level	Familiarity of the item	Unknown (0/1)	Categorical
	Correctness of guesses before the prompt	GuessBP (0/.5/1)	Ordinal
	Correctness of guesses after the prompt	GuessAP (0/.5/1)	Ordinal
	Correctness of recall answers	Recall (0/.5/1)	Ordinal
	Recall of the prompt	RCOP (0/1)	Categorical
Item/case level	Prior lexical knowledge: Unknown words; Polysemy/homonymy	UnKnownW (0/1) PolyHomo (0/1)	Categorical
	L1 transfer	FFL1 (0/1)	Categorical
	Etymological semantic transparency of the items	Transparency (0-5)	Numeric/Ordinal
Learner level	L2 proficiency	TEM-4	Numeric

(Note. The label 0/1 stands for No/Yes for all relevant categorical factors.)

3.6.2 Statistical tools

Statistical analyses were conducted using a number of statistical tools run on the R platform (R Core Team, 2017). Major statistical tools and key terms are explained below. For further details of mixed effects modelling see Winter (2020).

A **Mixed Effects Model (MEM)** is a type of statistical model containing both fixed effects and random effects. **Fixed effects** are often factors that are manipulated in the experiment, but more generally, are parameters associated with the system under study, e.g., idiom types, learners'

proficiency scores, and number of responses in learning trials. MEMs include **random effects** to account for the randomness inherent in the data, which typically comes from participants and/or items. There are two types of structure within random effects, **random intercepts** and **random slopes**. Random intercepts allow for overall differences in the dependent variable between, e.g. different participants or different items. For instance, in a reaction time experiment some participants may be faster overall than others. Random slopes further allow for the fact that participants (or items) may show different behaviours across other parameters such as the sequential position of a trial in an experiment. So, in the example of a reaction time experiment, some participants might get faster with practice whereas others might get slower because of fatigue.

With the advantage of random effects structures, MEMs can handle both between-subject and within-subject data in one single analysis. This is particularly useful in this study, as two random effects—participants and idioms—are naturally nested within four groups/sets following the research design. Including both participants and idioms as random effects not only accounts for between-participant and between-item variation, but also allows for variation between the responses from the same participant across different idioms, as well as those between the responses to the same idiom from different participants.

However, the appropriate models for each analysis have to be carefully selected. **Model selection** involves statistical comparisons of models with different random and fixed effects structures (the latter will be discussed in Chapter 6). As mentioned in 3.2, one major concern in the analyses for this thesis was whether between-group differences in proficiency (TEM-4 scores) would be confounded with learning performance on the different idioms, since each group had different sets of idioms to learn. Therefore, a preliminary analysis included random slopes over proficiency for each idiom to determine if the effect of proficiency is approximately the same for each idiom. In the Transparency*Proficiency MEM (see Chapter 6), by-idiom slopes for proficiency contributed little to explaining the variance in the model, suggesting that the effect of proficiency is approximately the same for each idiom regardless of the set. Therefore, a simpler intercepts-only random effects structure was adopted for all the MEMs in this study.

Apart from the above advantages, mixed models can also handle **unbalanced designs**, i.e., different amounts of data in different conditions. For instance, in this study, the number of responses to idioms included in the analysis varied. Two factors contributed to this imbalance: (a) set A had one more participant than other sets and consequently had more responses; and (b) a small number of responses indicated that some idioms were already known by a participant prior to the teaching procedure, and these were discarded from the analysis (see detailed explanation in Chapter 4). Moreover, mixed models can perform robustly even when there are missing values in data. All these advantages make mixed-effects modelling a very powerful statistical tool.

There are two main types of mixed effects models, i.e., **linear mixed model (LMM)** and **generalised linear mixed model (GLMM)**. GLMM is an extension of LMM. The dependent variable Y of a linear model needs to be continuous, and a generalised model is more flexible, allowing also categorical and ordinal data, etc. In GLMMs, the relationship between Y and model parameters need not be linear, and there is a model function linking Y and model parameters, which is called the *link* function.

The **Cumulative Link Mixed Model (CLMM)** is a form of generalised linear mixed model for ordinal data. The dependent variables under investigation—the correctness of participants' responses either during the learning phase or in the recall test—were ordinal in nature, as they were rated with ordered categories, *incorrect/no response*, *partially correct*, and *fully correct*, although presented in the form of scores of 0, 0.5 and 1²⁰. Therefore, for statistical analyses, CLMMs were applied in mixed-effects regression analysis, using the *clmm* (cumulative link mixed modelling) function of the *Ordinal* package in R (Christensen, 2018), which allows for two random effects.

²⁰ Preliminary analyses also showed that the dependent variables and their residuals were not normally distributed, as required by more commonly used linear models, and hence LMM was ruled out.

CHAPTER FOUR: OVERVIEW OF THE LEARNING PROCESS AND OUTCOME

Introduction

This chapter examines the extent to which learners' interpretation and retention of the meanings of idioms can benefit from etymological elaboration. Using descriptive statistics (inferential statistics will follow in Chapters 5 and 6), it provides an overview of the overall learning process and outcome (section 4.2), as well as considering individual differences broken down by learner and by idiom (section 4.3). The research questions addressed here include:

RQ 1. To what extent do learners' interpretations of the meanings of idioms benefit from etymological elaboration?

RQ 2. To what extent does learners' retention of the meanings of idioms benefit from etymological elaboration?

RQ 3. Does the procedure of inferencing based on etymological elaboration have an impact on error rates in retention?

4.1 Responses counted in the analysis

Before showing the results, I will first clarify the number of responses included in the analysis and the terms to be used in the following description and subsequent discussion.

Since each of the 25 participants was presented with 21 idioms, the total possible number of *cases* to be analysed was 525, where one **case** stands for one participant's series of responses concerning a given idiom. The total number of idioms for which there is data is of course 84 (21 different idioms in each of 4 sets). For each case, there were three responses: guess(es) before the prompt, guess(es) after the prompt, and recall answers, the scores of which will be referred to as **GuessBP**, **GuessAP** and **Recall** for convenience (same as in Chapter 3). GuessAP and Recall were the major **dependent variables**, representing the learning measures relating to idiom interpretation and retention.

It is important to clarify that not all of the 525 cases were eligible for analysis. Before data analysis, **55** were excluded, resulting in **470** valid cases covering 79 idioms. Among the excluded cases, five idioms were excluded altogether; the other excluded cases concerned individual participants' reactions to one or the other idiom. There were three main reasons for exclusion, as explained below (see Table 4.1 for a summary and Appendix 12 for detailed counts).

Table 4.1 Summary of excluded cases (N)

Set	Known	Error	Re-encountered	Total
Set A	9 (7 from 1 idiom)	7 (7 from 1 idiom)	0	16
Set B	3 (3 from B5)	9 (6 from 1 idiom)	3	15
Set C	3	0	0	3
Set D	9 (3 from D3)	12 (12 from 2 idioms)	0	21
Total	24	28	3	55

The first reason is that some items happened to be **known** to one or more of the participants before the experiment. This means that they did not need to be taught to those participants and so these combinations of idioms and participants were therefore excluded from analysis. For example, 'throw in the towel' was known to participant A3 beforehand, and 'beat around the bush', 'weigh someone down' and 'let the cat out of the bag' were known to B5. All cases of 'dark horse' were excluded as this idiom was known to all seven participants. Even so, the prompts were still given to the participants as additional knowledge to keep the interview procedure consistent.

The second reason for exclusion is that some idioms happened to be **re-encountered** in class in between the two interviews and this may have influenced participants' retention of these items. Only three cases of this type occurred, i.e., 'hit the roof/ceiling' was re-encountered by B3 and B6 and 'a shot in arm' by B4. These were reported by the participants in the later part

of the second interview when the researcher asked the participants “Have you noticed any of these expressions being used since our first interview?”

Finally, some of the items involved an **error** either in the teaching procedure or resulting from a confusing etymology/dictionary meaning. This emerged during the experiment or later in the data processing. These include: (1) Three cases of errors in procedure (all in set B) were that the researcher flipped the cue cards in the wrong order and thus mistakenly presented the additional example sentences prior to the etymology prompt. (2) The etymology of *burn the candle at both ends* and the idiomatic meaning of *jump in (at) the deep end* were misrepresented by the researcher, and so all related responses to the two idioms were discarded. (3) All the cases with *bark up the wrong tree* were excluded because different idiom dictionaries seem to focus on different aspects from the original scenario where a hunting dog was mistaken about which tree a squirrel was hiding in. It is ‘doing something that will not get the result you want’ in *MacMillan English Dictionary*, ‘following the wrong course of action because their beliefs or ideas about something are incorrect’ in *Collins Cobuild Dictionary of Idioms*, and ‘being mistaken about something’ in *Oxford Idioms Dictionary*. Although all the dictionary meanings are related to each other, they do not all entail the notion of ‘blaming someone wrongly’ which was often given by the participants. The raters could not reach a consensus which was based on dictionary meanings and thus the researcher discarded the idiom from analysis. (4) All the responses to ‘on a shoestring’ were excluded because the alleged origin as ‘a shoestring let out from a prisoner cell’ was later found historically unreliable and disputed and was not a common etymology proposed for this idiom.

4.2 Overall learning gains

Based on the 470 valid cases, the learning gain in idiom interpretation and recall results can be analysed by comparing the counts of GuessBP, GuessAP and Recall. When encountering an unfamiliar idiom, learners might arrive at the correct idiomatic meaning at three different time points—before or after the etymology prompt was given, or not until the full explanation was explained to them.

In terms of initial guesses²¹ at the meaning before the etymology prompts were given, there were 56 cases (11.9% of the 470) where the participants interpreted the meanings correctly (GuessBP = 1), 54 cases (11.5%) where the interpretations were partially correct (GuessBP = 0.5); and the majority of the cases (76.6%), 360 in total, where the participants came up with either no guess or wrong guess(es) (GuessBP = 0). The following sections break down the responses according to the starting point of the correctness of participants' guess before the prompt, as a measure of their prior knowledge about the idioms. That is, the scores of GuessBP inform us whether extra information was needed in helping the idiom interpretation (GuessBP = 0 or 0.5) and where it was not (GuessBP=1). The improvement in the meaning interpretation will be analysed by comparing GuessBP and GuessAP. The Recall results in each scenario will be analysed along similar lines.

4.2.1 When initial idiom interpretation was successful (GuessBP = 1)

We first look at the cases where meanings were correctly interpreted by learners before the etymological prompts were given, which meant that an explicit etymological elaboration was not needed to assist interpretation. This group (see Table 4.2) was made up of 56 out of the total of 470 cases, and involved 25 out of the 79 idioms that remained after exclusions described above. The fact that these cases were successfully guessed before the prompt could mean that in comparison with others, they were more decomposable and/or more semantically transparent to the learners. This was supported by consistent high ratings in the etymological semantic

²¹ Note that 'guess' here does not always mean a guess, since in some cases among the 56 cases where participants interpreted the meanings correctly before the prompt, the participants were confident in their knowledge of the meaning of an idiom and so were not guessing.

transparency, with 82% scores of 5 (the maximum on the scale) and 18% of 4.

Table 4.2 Cases of correct meaning guesses before the prompt (Total N of cases = 56)

Guess	Recall	N	Examples of idioms	N by case
1	1	51	a drop in the bucket A3, A6, A7	3
			bread and butter A3	1
			on the same wavelength A7	1
			ruffle someone's feathers B1, B2, B4, B5, B6	5
			pass on the baton B1, B2, B3, B4	4
			The ball is in your court. B2, B5, B6	3
			win hands down B5, B6	2
			on the same page B3, B6	2
			give someone the cold shoulder B5, B6	2
			leave someone high and dry B1, B2	2
			hit the roof/ceiling B1	1
			in the driving/driver's seat C1, C2, C3, C5, C6	5
			have a lot on one's plate C2, C4	2
			ring a bell C1, C4	2
			sit on the fence C2	1
			teething problems C3	1
			rub someone the wrong way C6	1
			down and out C4	1
			play your cards close to your chest C4	1
			tighten your belt D1—D6	6
			a hot potato D1, D4, D5	3
			go with the flow D6	1
			pull a rabbit out of the hat D4	1
1	0.5	2	<u>weigh</u> someone down B1	1
			go with the flow D1	1
1	0	3	get into <u>gear</u> A2	1
			<u>ruffle</u> someone's feathers B3	1
			<u>weigh</u> someone down B6	1

(Note. Idioms in bold print were correctly interpreted by at least three participants out of the six/seven participants in respective groups; words underlined were unfamiliar to the relevant learners prior to the experiment.)

Among these idioms, seven were correctly interpreted by three or more participants, even though they had said the idioms were unknown to them. These were *a drop in the bucket*, *ruffle someone's feathers*, *pass on the baton*, *the ball is in your court*, *in the driving/driver's seat*, *tighten your belt*, and *a hot potato*. Many of these seven idioms have an equivalent or near-equivalent Chinese expression: *tighten your belt* has an exact equivalent Chinese figurative

expression 扎紧裤腰带 (*pinyin: zha jin ku yao dai*, which literally means ‘tighten one’s belt’) and received correct interpretation from all six participants; *a hot potato* has an equivalent Chinese figurative expression 烫手的山芋 (*pinyin: tang shou de shan yu*, which literally means ‘a hot sweet potato’); similarly, *pass (on) the baton* is also used in Chinese for the same figurative meaning; *a drop in the bucket* is similar to Chinese four-character idiom 杯水车薪 (*pinyin: bei shui che xin*), which literally means ‘trying to put out the fire on a cart-load of firewood’ and figuratively means ‘(doing) something too little or too weak to solve a problem’ (see further discussion of the influence of L1 on L2 idiom interpretation in section 5.1.2). Others seem relatively easy to interpret, e.g. *in the driving/driver’s seat*.

In this scenario, when the full explanation was given to learners by the experimenter it served as a confirmation of their own correct inference; although the note on etymology was not needed as a prompt to aid interpretation, it was provided all the same as a regular part of the teaching experiment and in order to collect the etymological transparency ratings. Unsurprisingly, in 51 out of the 56 cases (91.1%), the idioms were correctly recalled one week later. Three cases where learners completely failed to recall the meaning of an idiom (Recall score of 0) could be traced back to the inclusion of a key word that was unfamiliar to these learners prior to the experiment (see the underlined words in Table 4.2). While the meanings of the unfamiliar words were explained to the learners when they first encountered them, this was not repeated at the recall stage. One of the two cases where learners’ recalls were only partly correct was due to lack of familiarity with the word *weigh*. The other was due to the fact that the participant recalled the L1 partial equivalent of the idiom *go with the flow*, i.e., 随大流 (meaning ‘follow the crowd’), which was rated as partially correct (see previous discussion of this rating in section 3.5.2) The effects of prior lexical knowledge and L1 transfer on idiom retention will be further discussed in Chapter 6.

4.2.2 When initial idiom interpretation was unsuccessful (GuessBP = 0 or 0.5)

We now move on to the 414 cases (covering now 77 idioms) where learners could not make a fully correct inference before the prompt was given, including the cases where the learners could not come up with any (correct) guess (GuessBP = 0, N = 360) and where the learners made a partially correct guess (GuessBP = 0.5, N = 54). This is where etymological elaboration

is needed and potentially helpful for idiom interpretation. Table 4.3 shows the distribution of the correctness of guesses and recall, with a focus on cases where the recall was totally correct.

Informing the participants about the origin of the expressions was found to considerably increase the likelihood of correct inferences in the first interview. For the 54 instances where guessing had initially already been partially successful, the information about the origin of the idiom occasionally (11.1%) helped learners to arrive at a fuller comprehension. In 49.7% of the 360 cases where participants had initially failed entirely to propose a correct interpretation, they now produced correct interpretations, and in an additional 19.5% they produced partially correct ones (see factors affecting the idiom interpretation in Chapter 5).

One week later, no fewer than 65% of the items which participants had demonstrated no comprehension of at the start of the first interview were recalled correctly, and an additional 10.5% of the recalls were partially correct (38 out of 360 cases, not shown in Table 4.3). In addition, 44.5% of the items that had already elicited partially correct guesses at the very beginning were now accurately recalled. In sum, accurate meaning retention was achieved among 62.3% of the cases where learners' prior comprehension was either nil or incomplete.

Table 4.3 Overview of learning gain by cases

GuessBP score	N of GuessBP	GuessAP score	N of GuessAP	N of Correct Recall	N of Correct Recall/ N of GuessAP	N of Correct Recall/ N of GuessBP
0.5	54	1	6 (11.1%)	4	66.67%	44.5%
		0/0.5	48 (88.9%)	20	41.67%	
0	360	1	179 (49.7%)	132	73.74%	65%
		0.5	70 (19.5%)	42	60%	
		0	111 (30.8%)	60	54.05%	
Total:	414		414	258		62.3%

Moreover, when the participants recalled the precise meaning of an idiom in the second interview, this typically (91.25%) coincided with their recollection of the idiom's suggested

origin, as displayed in Table 4.4. If they failed to recall an idiom's meaning altogether, they would often (65.59%) also fail to recall its origin. These results lend support to the claim that etymological elaboration can serve as an effective mnemonic.

Table 4.4 Recall of literal underpinnings for each meaning recall score (Total N = 414)

Meaning recall	Recall of origin			No recall of origin	
	N	%		N	%
Correct	240	91.25		23	8.75
Partially correct	45	77.59	Versus	13	22.41
Failed	32	34.41		61	65.59

Additionally, among those cases where the idiomatic meanings were correctly recalled but without the recall of the origins, some participants reported that they could remember other types of information which they found very helpful (as a memory aid), such as an L1 translation that resembles the idiomatic meaning of the idioms, and/or the example sentence provided during the explanation of the idiomatic meaning; others recalled the idiomatic meaning directly. The use of L1 translation and example sentences as a memory aid was also found in other cases during the introspective interview.

4.2.3 The relationship between the correctness of interpretation and that of recall (GuessAP & Recall)

To find out whether the outcome of inferencing based on etymological elaboration has an impact on retention, I will investigate the relationship between the responses during inferencing and the recall success in several ways. In this section, I will first explore the relationship of the correctness between the guesses after the prompt and the recall answers. An analysis of the overall performance (see below) can give us a first impression of the results.

First, according to Table 4.3 (see above), the second column from the right shows that when GuessBP = 0 or 0.5, the number of correct recalls proportionally increases in line with the accuracy of the guesses after the prompt within each GuessBP response category. This pattern suggests that the scores at GuessAP could be predictive of Correct Recall. Moreover, Table 4.5

shows that for cases where the meanings of the idioms were successfully interpreted after the prompt (GuessAP = 1), 73.5% were subsequently correctly recalled, with only about 5% recalled partially correctly and a recall failure rate of 21.6%. When the inferencing after the prompt was not so successful (GuessAP = 0 or 0.5), the recall rate was only about 53%.

Table 4.5 Comparison between GuessAP vs. Recall

		Recall			
		0	0.5	1	Total
GuessAP	0	38 (33.0%)	14 (12.2%)	63 (54.8%)	115
	0.5	15 (13.2%)	40 (35.1%)	59 (51.7%)	114
	1	40 (21.6%)	9 (4.9%)	136 (73.5%)	185
	Total	93 (22.5%)	63 (15.2%)	258 (62.3%)	414

(Note. The percentages in the table are calculated according to the cases of GuessAP.)

To look at this another way, when the meanings were recalled accurately (Recall = 1, N = 258), 136 cases (52.7%) followed successful inferencing after the prompt (GuessAP = 1); and when the meaning was recalled partially correctly (Recall = 0.5, N = 63), 40 cases (63.5%) followed partially correct inferencing after the prompt (GuessAP = 0.5).

Although all the learners had learned the full meaning of idioms by the end of the first interview, the above results suggest that successful inferencing with the aid of etymological elaboration might lead to better retention, while incorrect or partial inferencing is more likely to leave inaccurate or partial memory traces. A similar pattern is also observed in the by-participant analysis (see Figure 4.1 in section 4.3.1). The statistical significance of this trend will be further investigated by means of inferential statistics and error analyses in Chapter 6.

4.3 Learning gains by learners and by idioms

In the previous sections, overall learning gains under etymological elaboration were explored grouped by performance at the before-prompt, after-prompt, and recall stages. In this section,

learning gains will be analysed at a finer-grained level—by individual learners and then by individual idioms. It is important to explore the individual differences in learning performance because a particular teaching technique might be effective for certain individuals (or items) but not others. Descriptive statistics in this section will provide a detailed picture of the learning process and outcome, and will also help identify specific learners and items that will be subjected to further detailed analysis.

It is important to note that the scores of responses 0, 0.5, and 1 provided ordinal data. That is, they represented ordered categories of wrong/nil, partially correct, and fully correct. The numbers do not represent precise values (e.g., partially correct, which was coded as 0.5, does not always mean half correct) and there were no allocated scores that were between, e.g., 0 and 0.5. Nevertheless, in order to obtain a visual overview of the data, each point in graphs such as Figure 4.1 represents a mean score calculated by averaging relevant sets of GuessBP, GuessAP and Recall scores (in the case of Figure 4.1, this is across items for each participant). Therefore, the mean scores are presented here as numerical data to show the trends of changes between different phases. Compared to the individual scores, the average scores can be considered more continuous and comparable to each other (see detailed breakdowns in Appendices 12 and 14). Figures were made to show the average change in the correctness of scores from each participant/idiom from the learning phase to the recall test (see Figures 4.1 and 4.2). Lines were drawn to link the mean scores of GuessBP, GuessAP, and Recall purely for the purpose of demonstrating the range of changes in the overall performance among different participants/idioms. The relative differences among participants/idioms were suggested by the relative height of the points and the parallel tendency of the lines joining them.

4.3.1 Learning gains by individual learners

The results of each participant are shown in Figure 4.1 for an overview. Different groups of participants are represented in different colours, and participants from the same group who were exposed to the same set of idioms have the same colour (see per-participant breakdowns in Appendix 13).

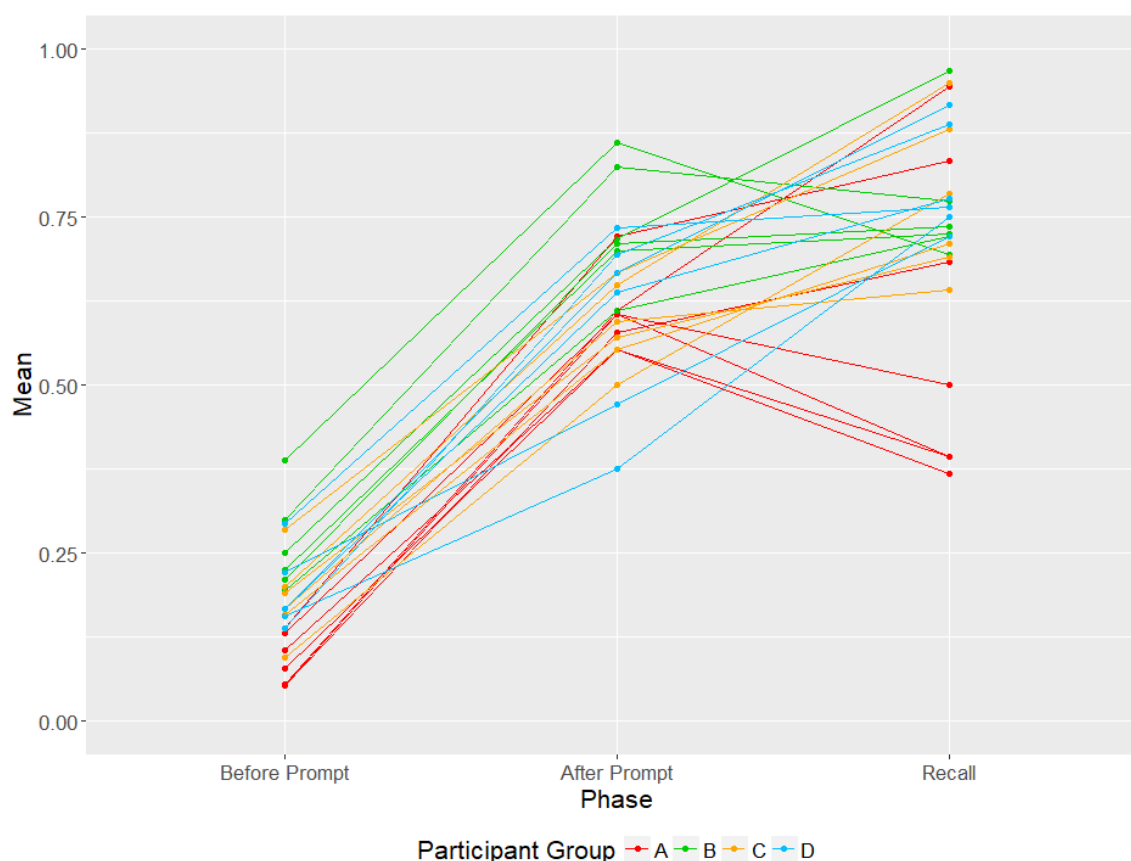


Figure 4.1 Change of mean scores between different phases for all the participants

For the interpretation/learning phase, there are many near-parallel lines between averaged GuessBP scores and GuessAP scores, which indicates that the direction and degree of improvement was similar across participants. The parallel lines display a mix of different colours, which suggests that the overall performance of participants from four different groups were rather comparable to each other. More specifically, there was an approximate 50% improvement uniformly among all the participants starting from an average GuessBP score of about 0.2 to an average GuessAP score of around 0.7 (except for D1 and D3, who showed smaller increases, see Figure 4.1.7 in Appendix 13).

If we look closer at the colour distribution, participants in group B seemed to have generally performed better than those in group A during the guessing/interpretation phase (i.e., at GuessBP and GuessAP). This could be related either to the fact that the participants in group B in general have a higher TEM-4 score (i.e., the proxy of L2 proficiency) than those in group A, or to differences between the idioms in Sets A and B. Whether the overall GuessAP scores are predicted by TEM-4 scores was explored with inferential statistics by using mixed-effects

modelling (see detailed discussion in section 5.2.2.3)—where no significant correlation was found between the two factors. Thus, the observed group difference could be related more to the differences in items rather than individuals.

At the end of the learning phase, all the participants would ideally have fully grasped the meaning of the idioms because the full idiomatic meanings were explained along with a sentence showing the usage in context. One week later, most participants obtained an average recall score above 0.5 (cf. the previous by-case analysis showing 62.3% accurate recall and 15.2% partially correct recall). The participants with average scores higher than 0.75 include A6, A3, B5, B1, C1, C2, C5, D6, D2, D5, and D4 (see detailed by-participant breakdown results in figures in Appendix 13). In other words, about half of the participants (11 out of 25) remembered more than 75% of the items that were unknown and not fully interpretable to them before the experiment (recall that the GuessBP scores shown in Figure 4.1 are for items with GuessBP = 0 or 0.5). Furthermore, as shown in Figure 4.1, 21 out of the 25 participants remembered more than half of the items.

The wide distribution of average Recall scores from about 0.375 to 1 also shows that varying degrees of learning occurred among participants from different groups. Overall, the participants who had the lowest recall scores (all below 0.5) are A1, A2, and A5 (see Appendix 13). It was shown that these three participants learned about 30% to 50% of the idioms that they had no knowledge of before the experiment, and their TEM-4 scores happened to be relatively low as well (71, 61, and 73, respectively).

Again, if we look at the colour distribution between GuessAP and Recall, group B (in green) appears to perform better than group A (in red). This appears consistent with the fact that there were more proficient EFL learners in group B than in group A. Meanwhile, many participants with higher recall scores, such as A3, B5, C2, D6 (TEM-4 scores = 78, 87, 84, 76) happened to have higher TEM-4 scores than those with lower recall scores (e.g., A1, A2, A5). However, no conclusion can be easily drawn at this point, since the variations in the recall results could also be due to relative learning ease of idioms, i.e., some idioms are more transparent than others, not to mention other interacting factors, e.g., similarity to L1. Therefore, whether L2 English proficiency is predictive of recall success will be further investigated in a quantitative

analysis in Chapter 6 by using mixed-effect modelling, which can take both between-group comparisons and variations of idioms into account. In addition, the participants who did exceptionally well in the recall test (e.g., A6, B5, C1) will be investigated further in association with their personal experiences with idioms, their learning strategies and attitudes in Chapter 6, as part of the qualitative investigation of factors affecting L2 idiom retention.

In summary, based on by-participant analysis, there was an overall moderate (about 0.5 score) improvement in idiom interpretation with the help of etymological elaboration, with similar patterns shown within groups and (relatively less so) between groups. As for meaning retention, the recall results showed a substantial learning gain—the majority of the participants remembering more than half of the idioms that they learned a week previously. In comparison with the results for the interpretations, greater variation was found in the recall results, which suggests that idiom retention is more likely to be influenced by individual differences. A descriptive analysis of individual participants helped detect L2 proficiency as a potential predictor of learning success, and identified exemplar learners for further analysis.

4.3.2 Learning gains by individual idioms

Just as the previous section considered learning gains by participants, this section explores learning gains by idioms. An overview of learning gain per idiom is shown in Figure 4.2, with each data point representing an average score across the relevant participants at different stages (see the codes of idioms in Appendix 14 and the figures of detailed per-idiom breakdowns in Appendix 15). The following text describes the different patterns observable in Figure 4.2 and figures in Appendix 15, with relevant examples.

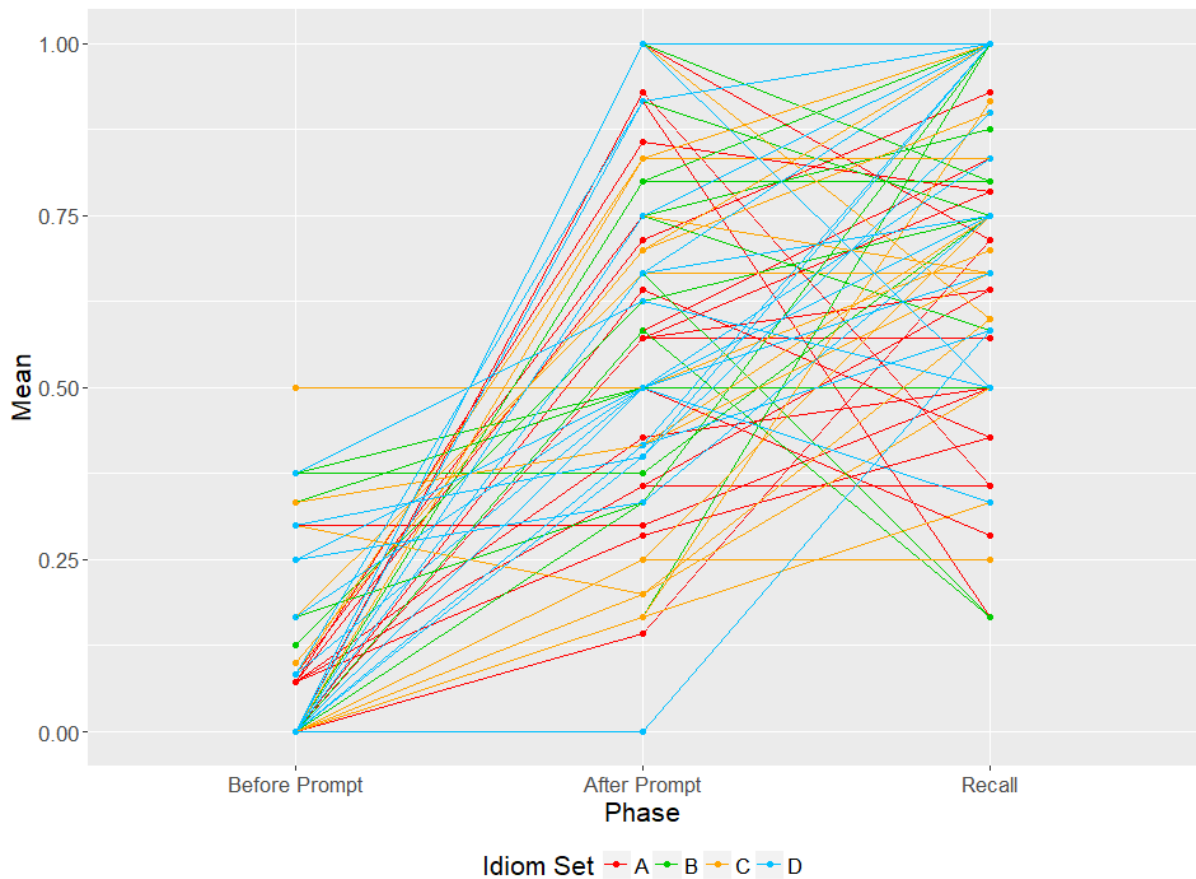


Figure 4.2 Change of mean scores between different phases for all the idioms

Comparing guesses before and after the prompts:

The clear majority of the idioms have upward lines between the mean scores for GuessBP and GuessAP, with the correctness of GuessAP widely dispersed between 0 and 1. The predominant pattern of upward lines shows that in general, informing the learners about the literal underpinning has helped them guess the meanings of the different idioms. The different slopes of these upward lines show that this has happened to different extents for the different idioms. Among the 77 idioms (after previously explained exclusions), interpretations for 20 idioms became much more accurate, with the majority of the learners proposing a correct interpretation after the etymological prompt was provided (i.e., $\text{Mean GuessAP} - \text{Mean GuessBP} > 0.75$, based on a distributional analysis of the gains). These idioms are: *follow suit*, *get into gear*, *go belly-up*, and *red tape* (in set A); *take a back seat*, *a shot in the arm*, *win hands down*, *pass on the baton*, and *let the cat out of the bag* (in set B); *in the driving/driver's seat*, *show someone the ropes*, *in the wake of*, *bury the hatchet*, *rub someone the wrong way*, *make ends meet* (in set

C); *flex your muscles*, *cut corners*, *get your second wind*, *a hot potato*, *(start) from scratch* (in set D).

Four horizontal lines show that for four idioms, the average scores of guesses after the prompt remain the same as before the prompt. This suggests that extra information such as a supporting context might be necessary for an accurate inferencing of these idioms. The idioms involved are *bread and butter* (source domain: Food), *beat around the bush* (source domain: Hunting), *play your cards close to your chest* (source domain: Card game), and *on autopilot* (source domain: Travel).

In a single case, *a rule of thumb*, the average GuessAP was worse than that the average GuessBP, suggesting that the prompt may have caused confusion.

For meaning retention:

The majority of the average Recall scores are above 0.5, showing that the majority of idioms were well remembered. 17 idioms were recalled accurately by all their respective participants (see the clearer per-idiom breakdowns in Appendix 15). These idioms are: *go belly-up* (in set A); *take a back seat*, *on the same page*, *pass on the baton* (in set B); *jump the gun*, *in the driving seat*, *show someone the ropes*, *bury the hatchet*, *down and out*, *make ends meet* (in set C); *flex your muscles*, *wet blanket*, *cut corners*, *get your second wind*, *have a green thumb*, *a hot potato*, and *throw your hat/cap into the ring* (in set D). Many of these are idioms which became easier to interpret with the aid of etymological elaboration (as shown in the previous point).

Comparing guesses after the prompt and recall results:

There were a few items whose average guesses after the prompt were higher than 0.5 but had much lower recall scores (highlighted in red rectangles in breakdown figures in Appendix 15). They are *get into gear*, *pull one's weight*, *red tape*, *in the doldrums*, *(not) up to scratch*, *ring a bell*, and *(start) from scratch*. 5 of these 7 items contain a word that was either unknown to the participants (e.g., *doldrums*, *gear*) or that had a polysemous meaning that was unfamiliar to the participants (e.g., *scratch*). In addition, four idioms received poor scores in both guessing and retention, i.e., *have/got cold feet*, *take the bull by the horns*, *ring a bell*, and *stick your neck out*.

Overall, in comparison with the per-participant results, the results by idiom showed greater diversity. This seems to suggest that the treatment benefited meaning guessing as well as recall to varying degrees largely depending on the target idioms.

4.4 Summary and discussion

To sum up, the descriptive results presented in this chapter provide an initial response to the following research questions (to be followed up by inferential statistics in Chapters 5 and 6):

RQ 1. To what extent do learners' interpretations of the meanings of idioms benefit from etymological elaboration?

RQ 2. To what extent does learners' retention of the meanings of idioms benefit from etymological elaboration?

RQ 3. Does the procedure of inferencing based on etymological elaboration have an impact on error rates in retention?

Results from all the above analyses suggest that idiom interpretation benefits substantially from etymological elaboration, or at least from paying more attention to them. Although providing the information about the literal underpinnings or the origins of the idioms does not guarantee a full understanding of the idioms, the overall results did suggest that such information can help learners reach a better understanding to various degrees. As for retention, the overall above-60% correct recall rate by case and high retention rate of the etymological information provide empirical evidence of the mnemonic effects of etymological elaboration. The results also suggest that the effectiveness of the implemented teaching technique varies across idioms (cf. Boers et al., 2007) and must be affected by many other factors.

The usefulness of etymological elaboration as an inferencing prompt seems to largely depend on the nature of the target idioms (as will be discussed in more detail in section 5.2), as greater variation was seen in the by-item analysis than in the by-participant analysis. In terms of the mnemonic effect, apart from the nature of idioms, what determines the recall success seems to relate more to the learner factors, such as their L2 proficiency—as shown in the by-participant analysis, as well as to the success rates in the inferencing procedure—as shown in the overall analysis and by-case analysis. It appeared that many incorrect or partially correct recalls were

associated with L1 interference or with insufficient prior L2 lexical knowledge. Moreover, the overall analysis and the by-idiom analysis also tell us that some idioms need extra attention and information in teaching than others; and the by-participant analysis helps to identify exemplar learners for further study.

In conclusion, a lot of learning has happened during the teaching experiment. The trends and patterns identified in this chapter lead to further investigations in the following chapters, where I examine which factors play a significant role in the learning process and outcomes and how they come into play.

CHAPTER FIVE: INVESTIGATING L2 IDIOM INTERPRETATION

Introduction

This chapter reports and discusses the findings concerning L2 idiom interpretation. It comprises two sections, which explore the factors affecting the interpretation of the meanings of L2 idioms under the etymological elaboration. The two sections take different perspectives.

The first section focuses on the problems that learners encounter at the initial stage of the learning phase (when idioms are presented in simple contexts) and during the process of interpretation. The factors involved are learners' prior L2 lexical knowledge, as well as cross-linguistic and cross-cultural influences (from L1 Chinese to L2 English). Accordingly, the research questions to be addressed include the extent to which prior lexical knowledge (RQ 4) and L1 transfer and cross-cultural differences (RQ 5) affect L2 idiom interpretation.

The second section provides inferential statistics to support and expand on the description in Chapter 4 of the facilitative effects of etymological elaboration on idiom interpretation. In particular, it explores the relationships between relevant factors, including etymological semantic transparency of idioms (henceforth *Transparency* for short), correctness of guess after the prompt (*GuessAP* for short), and L2 proficiency. Issues to be explored include: (1) how the learners evaluated the connection between the literal underpinning and the figurative meaning of the target idioms (as shown by their perceived Transparency in section 5.2.1); (2) how learners' evaluation of the connection reflects their performance in the interpretation (by a regression analysis of Transparency and GuessAP in section 5.2.2.1); (3) whether learners' perceived Transparency is associated with their L2 proficiency (by a regression analysis of Transparency and TEM-4 scores in section 5.2.2.2); and (4) whether learners' L2 proficiency is predictive of idiom interpretation with the aid of etymological elaboration (by a regression analysis of GuessAP and TEM-4 scores in section 5.2.2.3). By exploring these issues, the effectiveness of etymological elaboration is evaluated via etymological semantic transparency and correctness of guess after the prompt. The influences of L2 proficiency on these two variables are also examined.

Taken together, the two measures of L2 idiom interpretation under etymological elaboration, i.e., Transparency and GuessAP, offer two different ways to interpret analysability and semantic transparency of idioms from the L2 learner's perspective. These two measures should also relate to the factors mentioned in the first section, as L2 learners' comprehension and perception of idioms are affected by the features of the idioms as well as by the learners' own prior lexical and cultural knowledge. Moreover, Chapter 5 also serves as a prelude to Chapter 6, which investigates L2 idiom meaning retention, as many factors examined and measured for the interpretation phase are likely to have an impact on meaning retention.

5.1 Factors affecting L2 idiom interpretation

As discussed in section 2.3.3, prior L2 lexical knowledge and L1 transfer are likely to affect L2 idiom comprehension. In the following sections, I will report relevant findings from the learners' responses and from learner-teacher interactions during the learning-phase interviews.

5.1.1 Prior L2 lexical knowledge

The initial hurdle for L2 idiom comprehension observed in the interviews was participants' insufficient prior knowledge of the constituent words of the idioms. This prior knowledge includes the knowledge of both word form and meaning. As previously discussed in section 2.3.3.4, common issues at the initial stage of interpretation, i.e., when learners were asked to indicate their understanding of the idioms based on the presentation of the idioms in simple contexts, come from:

- 1) A content word that might be completely unknown to the learners, due to low frequency or obsolete usage of the word or the learners' own limited knowledge of L2 vocabulary, e.g. *doldrums* in *in the doldrums* and *hatchet* in *bury the hatchet*;
- 2) A content word that could be 'deceptively transparent' (Laufer, 1989) to the learners — where a learner seemed familiar with the word form but not with the meaning that is relevant in the target idiom. This included two similar conditions:

- (a) The content word is **homonymous** and its meaning in the idiom is less frequently used than the meaning of its homonyms, and is not known to learners. For example, 6 out of 6

participants assumed *suit* in *follow suit* refers to a formal set of clothing rather than to one of the four sets that form a pack of cards; 6 out of 6 participants did not know that *wake* in *in the wake of* refers to the track that a boat or ship leaves behind on the surface of the water, but associated it with the verb *wake* as in *wake up*.

(b) The content word is **polysemous** and has a less frequently used meaning in the idiom, which was unfamiliar to learners. For example, 6 out of 6 participants were not familiar with the meaning of *wing* in *(be) waiting in the wings* (which is the area at either side of the stage of a theatre) and mistook it as bird wings (see more examples in Tables 5.1 and 5.3).

Table 5.1 Insufficient L2 lexical knowledge by case and by idiom

	Number of cases (Total N = 470)	Number of idioms (Total N = 79)	Examples
Not guessable due to Unknown word	140 (29.8%)	40 (50.6%) (11 in set A, 8 in set B, 11 in set C, 10 in set D)	be <u>par</u> for the course, get into <u>gear</u> , in the <u>doldrums</u> , bury the <u>hatchet</u>
Misled due to Homonymy/Polysemy	62 (13.2%)	12 (15.2%) (1 in set A, 5 in set B, 4 in set C, 2 in set D)	(be) waiting in the <u>wings</u> , in the <u>wake</u> of, throw your hat into the <u>ring</u>

The frequency of occurrence of these initial issues is displayed in Table 5.1. In about 30% of the 470 valid cases²², participants failed to interpret the idiom at the beginning of the learning procedure because they did not understand a vital content word. These affect approximately 50% of the target idioms, although the number of unknown words varies for different participants and for different sets of idioms, as shown in Table 5.2. Note that 91.02% of the constituent words of the 79 idioms were in the most frequent 2000 words in BNC-COCA. Less frequent words that were unfamiliar to many participants are *frenzy*, *herring*, *plunge*, *doldrum*, *hatchet*, *chin*, *par*, and *gear*.

²² As in Chapter 4, ‘case’ here represents a case of one participant learning one item, and in specific contexts it is also used to refer to the response(s) that a participant made in learning or recalling an idiom.

Table 5.2 Unknown words by participant

Participant ID	TEM-4 score	Unknown Idioms (N)	Unknown Words (N)
A1	71	19	5
A2	61	19	10
A3	78	18	5
A4	65	19	8
A5	73	19	7
A6	70	18	8
A7	73	19	10
B1	84	20	6
B2	76	20	5
B3	70	19	5
B4	89	18	6
B5	87	16	4
B6	80	18	7
C1	79	20	6
C2	84	21	5
C3	70	19	5
C4	78	21	8
C5	69	21	5
C6	76	21	6
D1	79	18	4
D2	73	18	4
D3	75	16	2
D4	78	17	3
D5	79	18	4
D6	76	18	2
Ave. of Set A	70.14	18.71 (min 18, max 19)	7.57 (min 5, max 10)
Ave. of Set B	81	18.5 (min 16, max 20)	5.5 (min 4, max 7)
Ave. of Set C	76	20.5 (min 19, max 21)	5.83 (min 5, max 8)
Ave. of Set D	76.67	17.5 (min 16, max 18)	3.17 (min 2, max 4)
Ave. of total	75.72	18.8 (min 16, max 21)	5.6 (min 2, max 10)
Mode		18	5
Median	76	19	5

While there was a tendency for less proficient learners to report more unknown words (measured as the ratio of unknown words to unknown idioms), this was not significant ($r = -0.21$, $p = 0.3151$). Table 5.2 shows that participants in group A, which has relatively more low-proficiency achievers compared to other groups, had more unknown words on average. Participants A2, A4, A6, A7, had more unknown words than their peers in the same groups (i.e., more than 7 words), and some of their unknown words are of relatively high frequency.

As participants started to make initial guesses at the idioms, it was possible to identify content words that were deceptively transparent (in the sense of Laufer, 1989) due to homonymy or polysemy. As shown in Table 5.1, in about 13% of the cases, the relevant meaning of a homonymous/polysemous word (form) was unfamiliar to the participants, and this prevented a correct interpretation of the idiom. This issue concerns 12 (about 15%) of the 79 idioms (after previous exclusion of invalid cases). Table 5.3 (see further below) shows that these cases were often common across the participants who were exposed to the same set of idioms.

Table 5.3 Idioms containing words with an unfamiliar homonymous/polysemous meaning

Set	Idiom	N of participants (N = 7 for Set A, 6 for Set B/C/D)
A	follow <u>suit</u>	6 out of 6 (Note. One participant in Set A knew the idiom and was excluded.)
B	a <u>shot</u> in the arm	5 out of 5 (Note. One participant in Set B knew the idiom and was excluded.)
	(be) waiting in the <u>wings</u>	6 out of 6
	(not) up to <u>scratch</u>	6 out of 6
	<u>weigh</u> someone down	4 out of 5 (Note. ditto)
	The ball is in your <u>court</u>	4 out of 6
C	in the <u>wake</u> of	6 out of 6
	<u>rub</u> someone the wrong way	5 out of 6
	<u>stick</u> your neck out	5 out of 6
	turn over a new <u>leaf</u>	3 out of 6
D	<u>play</u> it by ear	6 out of 6
	throw your hat/cap into the <u>ring</u>	6 out of 6

When the above cases were noticed during the interview, the meaning of the unknown word or polysemous/homonymous word was explained to the participants immediately. It should be noted that for some idioms (e.g., *in the doldrums*, *a shot in the arm*) the etymological prompt, or information about the literal underpinning, was given to the participants in such an

explanation; and for a few participants, once they understood the unknown words or homonymous/polysemous words, they were able to arrive at a correct (or roughly correct) idiomatic meaning. For example, B6 at first took *court* in *the ball was in your court* as a place where legal matters are decided, but once the polysemous meaning of *court* (i.e., a place where games such as tennis are played) was explained to her, she was able to interpret the idiom correctly.

5.1.2 L1 transfer and cross-cultural differences

L1 is commonly found to be an influential factor in L2 vocabulary learning (Odlin, 1989; Wolter, 2006). The influence of L1 can be explained by the facts that “[adult L2 learners] already possess a well-established conceptual and lexical system, and most L2 words have a correspondent concept and translation in the adult learner’s first language (L1)” and that L2 learners activate their existing L1 linguistic and conceptual systems during L2 learning (Jiang, 2004, p. 417). In this study, L1 transfer and interference were detected throughout the learning process from participants’ guesses of the idiomatic meanings (both before and after the etymological elaboration) as well as from their reflections and autonomous comparisons between their L1 and L2 that they made after the idiomatic meanings were explained to them. As mentioned in Chapter 3, the participants spontaneously came up with a considerable number of Chinese idioms (成语 *chengyu* in Chinese, also known as four-character idioms) and other idiomatic expressions with cultural connotations during their attempts at interpreting L2 idioms. Some of the expressions are close to the target English idioms, some are comparable to some extent, and others are different in their idiomatic meanings (see further below).

As discussed in section 2.3.3.3, previously reported effects of L1 on L2 idiom learning have often been shown to vary according to how L2 idioms differ from or resemble L1 idioms. In Hu and Fong’s (2010) study with Chinese learners of English, L2 idiom interpretations were found very susceptible to the ‘workings of L1’ (Chinese): the availability of L1 equivalents affects idiom interpretation at the levels both of the whole expressions and of their constituent content words; moreover, L1 transfer is beneficial when true L1 ‘cognates’ are available, and it can be detrimental if the L2 idioms or the constituent words that they contain have ‘false

cognates'²³ or 'false friends', i.e., L1 counterparts that stimulate associations which are not shared between the two languages. Previous transfer studies also suggest that partial similarity generates more interference errors. So far, evidence found in studies of L2 idiom production suggests that more errors are generated due to partial formal similarity compared to complete difference (Irujo, 1986a; Laufer, 2000). Findings in the current study tally with all these claims from previous studies. I provide more empirical evidence of transfer errors in L2 idiom comprehension below.

Because Chinese and English do not share formal properties, L1 transfer in such a context occurs mainly at semantic and conceptual levels, and comparisons between the two languages are mainly based on the semantics²⁴ of the whole expressions and their lexical semantic components. In order to differentiate types of L1 transfer, a five-degree L1-L2 similarity scheme was used to analyse learners' Chinese responses to the meanings of the L2 idioms, and this was then used to generate four categories of equivalents based on learners' responses: *total equivalent*, *partial equivalent*, *false equivalent*, and *no equivalent*²⁵. These classifications of L1-L2 degrees of similarities and categorisation of translation equivalent are based on a synthesis of a series of studies including Charteris-Black (2002), Deignan, Gabrys, and Solska (1997), Laufer (2000), and Liontas (1999, 2002). Table 5.4 provides a summary of this scheme. The categorisation of total/partial/false/no equivalents and the estimation of whether a positive or negative influence on learning would ensue were based on the bilingual researcher's

²³ When L1 transfer is investigated, different researchers use different terms to define similarity and difference between two languages in contrastive analysis. For two languages like German and English, which belong to the same language family and thus share formal similarities, *cognate* and *false cognate* (or 'false friend') are often used to refer to a true/false resemblance or equivalent. However, for two very different languages like Chinese and English, which do not share etymological roots, and are thus much less comparable in terms of formal properties, cognate is not an appropriate term to use. Although Hu and Fong (2010) use 'false cognates' to refer to L1 counterparts that stimulate associations which are not shared between the two languages, 'false friend' might be a better term to avoid confusion.

²⁴ While L1 transfer involves cross-linguistic differences at each of lexical, semantic, and conceptual levels (Paradis 1997; Pavlenko, 1999), no further differentiation of these levels will be made here—since within the framework of Cognitive Linguistics (CL), all meanings are conceptual.

²⁵ *Equivalent* here is used in the sense of translation equivalent. In translation studies, where idioms are notoriously problematic and challenging for translators, *idiomatic equivalent* is sometimes used to refer to an equivalent in the target language to an idiom in the source language.

judgement, which was backed up by referring to the native speakers' scoring of learners' responses. Items categorised as 'total equivalents' had learner responses that were rated fully correct by the native-English raters. Those categorised as 'false equivalents' had learner responses that were rated incorrect by the raters. Those categorised as 'partial equivalents' were harder to evaluate, and this was reflected in the native-speaker ratings. The type of influence this category would have on learning was also more difficult to estimate, as discussed below.

Table 5.4 L1-L2 degrees of similarity and cross-linguistic influences

Type of influence	L1-L2 similarity	Closeness of translation
+	Same/very similar semantic or lexical components with exactly the same metaphorical association	Total equivalent
+	Different lexical components but similar metaphorical association [Type 1 partial equivalent]	Partial equivalent
+/-	Shared word or similar lexical components and somewhat similar association but with subtle differences [Type 2 partial equivalent]	
-	Shared word or similar lexical components but different associations	False equivalent
None	None	No equivalent

1) Total equivalents were identified where the target idiom had a counterpart idiom or a common fixed expression in L1 that contains the same or very similar semantic or lexical components with the same metaphorical association and conjures up the same or very similar mental image, e.g., *a hot potato*/烫手的山芋 (*pinyin: tang shou de shan yu*, literal: a hot sweet potato), *tighten one's belt*/勒紧裤腰带(过日子) (*lei jin ku yao dai guo ri zi*, literal: tighten one's belt). In this study, many of these cases (but not all) came from those idioms which were deliberately included for motivational reasons, and therefore such cases figured largely among those with GuessBP = 1 (see previous discussion in section 4.2.1) or known items (e.g., *a dark horse*) which were excluded for further analysis.

2) Partial equivalents were identified where the target idioms had a counterpart in L1 that shared similar associations in meaning. Under this category, two subtypes were differentiated.

Type 1 partial equivalents consist of different lexical components from L2, originate from different source domains and may conjure up different mental images from the target idioms. For example, *a drop in the bucket* has a similar four-word idiom in Chinese 杯水车薪 (pinyin: *bei shui che xin*), which literally means ‘to attempt to quench the fire of a cartload of firewood with one cup of water’. The two idioms share a similar figurative meaning, in terms of ‘a very small amount compared with what is needed or expected’. Another example is *start from scratch* comparable to 白手起家 (*bai shou qi jia*), which means build up one’s fortune from nothing. The two idioms are almost equivalent, although compared to *start from scratch*, the Chinese idiom has a relatively narrow sense, focusing on one’s business or fortune.

Type 2 partial equivalents share the same or similar lexical semantic components with the L2 idiom, and these share somewhat similar associations but often with subtle differences. This type differs from total equivalents in that the L1 equivalent may have a connotation or a particular cultural context that does not exist in the target L2 idiom. For example, *ruffle someone’s feathers* called up 抚逆鳞 (pinyin: *fu ni lin*) among two participants. The more correct form of this expression should be 批其逆鳞 (*pi qi ni lin*)—a Chinese idiom which literally means ‘to ruffle the inversed scales under a (Chinese) dragon’s throat (which could hurt it badly)’ and figuratively means ‘to speak one’s mind in front of the superior/emperor at the risk of upsetting or annoying him (and getting killed).’ The L1 and L2 idioms share similar lexical components and associations, except that the L1 idiom has a specific object as someone superior (because a Chinese dragon or 龙 [pinyin: long], a sacred animal in Chinese mythology, is often associated with the emperor). Another example of Type 2 partial equivalent is *go with the flow* vs. 随大流/随波逐流 (*sui da liu/ sui bo zhui liu*, literally meaning ‘to follow the crowd’) (see previous discussion in section 3.5.2). Responses with partial equivalents, especially Type 2 partial equivalents, are those that posed challenges for the translator and the raters during data processing. Their effects on the learning process are relatively more complex and there is no clear boundary between partial equivalents and total equivalents or false equivalents. Here the categorisation is based both on the researcher’s analysis and on native speakers’ ratings, where total equivalent was rated as 1, partial equivalent often as 0.5 and false equivalent as 0.

3) False equivalents were identified where a target idiom elicited an L1 idiomatic expression with a shared word or similar lexical semantic component, but the associated metaphorical meanings were different. For example, *(be) on the ropes* called up a Chinese four-character idiom 千钧一发 (*pinyin: qian jun yi fa*), which literally means thousands of units of weight hanging by a hair, and figuratively means ‘be in an extremely critical situation’. It is often used to describe moments of life and death. The participant associated *rope* with *hair* because of the common usage of ropes with hanging things, but the L1 idiom has a different meaning from *(be) on the ropes*—indicating a boxer pinned by their opponent against the ropes enclosing the boxing ring, which is used to describe being near defeat or giving up, helpless. Therefore, the L1 idiom is classified as a false equivalent. *Turn over a new leaf* was interpreted by many participants as 翻开新篇章 (*pinyin: fan kai xin pian zhang*), which has the same lexical components as the target idiom but figuratively means ‘start a new chapter in one’s life’—which is different from or actually more general than ‘behave in a better or more acceptable way than previously’. Yet it is arguable that such cases can be considered a Type 2 partial equivalent, as *turn over a new leaf* can be associated with starting a new chapter in one’s life; and a near-defeat state could be as critical as 千钧一发. Generally, compared to partial equivalents, false equivalents have more identifiable differences in lexical components and/or metaphorical associations.

Moreover, polysemy/homonymy can be a factor that contributes to L1 influence when the participants associated the wrong meanings of the polysemous word with an L1 expression. Misinterpretation of such words can call up a false equivalent. For example, the participants tended to interpret *hand* in *play into someone’s hands* as a human body part (rather than a set of cards held by one player) and relate the expression to 落入某人的手掌心 (*pinyin: luo ru mou ren de shou zhang xin*), meaning ‘fall under control of someone’ (equivalent to another English idiom *falling into someone’s hands*). The concept of 落入某人的手掌心, in the participants’ explanation, came from the famous story of “the Monkey King falling into the Buddha’s hand” in one of the four classical novels of Chinese literature *Journey to the West*.

Sometimes, the same target idiom can evoke different types of L1 counterpart in learners’ minds. *Sit on the fence* was such a case that elicited different Chinese idiomatic expressions,

including 隔岸观火 (pinyin: *ge an guan huo*), 作壁上观 (*zuo bi shang guan*), 骑墙 (*qi qiang*) and 墙头草 (*qiang tou cao*). While 骑墙 (literal: sitting on the wall; figurative: sit on the fence) is a total equivalent, 作壁上观 (literal: stand on the rampart watching the battles; figurative: be an onlooker) is a Type 2 partial equivalent. 墙头草 (literal: grass on the top of a wall which sways with every wind; figurative: someone who is likely to choose or change to either side at any time) is a Type 2 partial equivalent with a subtle difference from *sit on the fence* (which is used to describe a person's lack of decisiveness, neutrality or hesitation in choosing between two sides in an argument or a competition, or lack of courage to decide).

4) No equivalent—no counterpart of the target idiom was found in L1 responses where learners employed no or little L1 knowledge. These were often (but not always) the most difficult cases for learners to interpret without the aid of etymological elaboration. Typical examples are those that involve a western household object, e.g., *wet blanket*, *on the back burner*, and culture-specific phenomena like sports and games, e.g., *throw in the towel*, *hit below the belt*, *pass the buck*. In other words, these are the idioms “derived from a source domain that is less salient in the learner’s own culture” (Boers et al., 2004, p. 380).

In terms of matching the associative meanings (or ‘target domains’ in CL terms) between L1 and L2, total equivalents and Type 1 partial equivalents (and even Type 2 partial equivalents depending on context of use) belong to what can be called *idiomatic equivalents* in translation studies. Idiomatic equivalents also came up in learners’ self-reflection and autonomous comparisons after they learned the full idiomatic meanings, and this emergence of the idiomatic equivalents could have facilitated their understanding and in turn their memorization.

Overall, about 35% of the learners’ responses (i.e., 166 out of 470 cases) were found to contain an L1 equivalent whether a total, partial, or false equivalent; about 65% of the cases (304/470) contained no equivalent. Discrimination of the four types, partial equivalents in particular, was neither easy nor straightforward. Apart from the identified cases of L1 transfer and interference, it was also found that not all participants opted to use L1 knowledge in L2 idiom interpretation. One participant reported in retrospect that she cautiously avoided using Chinese idiomatic expressions or concepts in interpreting the English idioms, even though some turned out to be total equivalents, because she felt that “the [target] idioms should be peculiar to the English

language and culture, and the meanings of the idioms should not be related to Chinese.” This suggests that L2 learners who have cross-cultural awareness might be cautious with cross-linguistic differences, and consider idioms to be non-transferable (cf. Kellerman, 1977).

In general, based on the ratings of learners’ responses, total equivalents have a facilitating influence on the interpretation of L2 idioms, and false equivalents impede it (at least in the beginning stages of the learning process). But for the cases of partial equivalents, it is often hard to predict whether there will be a facilitative effect or not. On the one hand, whether these identified L1 influences have an enduring effect on the learning process was unclear, as such influences could be reinforced or offset by the teacher’s (i.e., the researcher’s) confirmation or corrective feedback during the interviews. On the other hand, whether the continued influence was positive or negative not only depends on the L1-L2 similarities/differences but also, and more importantly, on how learners processed the explanations that followed their initial attempt at interpretation. When the full idiomatic meaning was explained, any confusion or misunderstanding caused by cross-linguistic or cross-cultural interferences was expected to have been clarified and corrected. Whether L1 transfer and interference left an impact on the final learning outcome, i.e., the meaning recalls one week later, will be investigated in Chapter 6.

5.2 Evaluating etymological elaboration via etymological semantic transparency

After examining the factors affecting the initial stage of interpretation, i.e., when learners were asked to indicate their understanding of the idioms in simple contexts, we now turn to an exploration of three factors, i.e., two measures of analysability of idioms—etymological transparency ratings and correctness of guesses after the prompt—and learners’ L2 proficiency. The focus of this section is to investigate learners’ perceived transparency of the target L2 idioms (see previous explanation of etymological transparency in section 2.3.3.2), as well as the relationships between the three factors, through which the effectiveness of etymological elaboration in facilitating idiom interpretation is evaluated. It turns out there is a significant relationship between Transparency and GuessAP, but Proficiency is not a significant predictor of Transparency nor of GuessAP.

5.2.1 Results of etymological semantic transparency ratings

First, let's look at the results of the etymological semantic transparency ratings, through which we can get an impression of how the learners evaluated the connection between the literal underpinnings and the idiomatic meanings. Table 5.5 shows the distribution of ratings from 0 (the link is not clear at) to 5 (the link is very clear), which is based on the 414 cases where participants' initial interpretation was unsuccessful (GuessBP = 0 or 0.5)²⁶.

The participants' ratings were very unevenly spread across the six-point scale (see Table 5.5), with the lowest points, 0 to 2, selected very seldom—together only about 8%. Idioms that attracted such occasional low ratings included *on the back burner*, *take it on the chin*, *a loose cannon*, *let the cat out of the bag*, *teething problems*, *sit on the fence*, *take the bull by the horns*, *a wet blanket*, *have cold feet*, *in the wake of*, *hold your horses*, *hand over fist*, etc. The two highest points on the scale, 4 and 5, were selected the most often—together about 73.5%, which suggests that, by and large, the learners thought the origin of the idioms that they had been presented with offered a relatively clear motivation for the idioms' meanings.

Table 5.5 Counts of etymological transparency ratings (Total N = 414)

Transparency ratings	Count (N)	Percentage (%)
0	2	0.48
1	6	1.45
2	26	6.28
3	75	18.12
4	153	36.96
5	152	36.71

It is important to note, however, that there was considerable disparity among participants' judgements. It was not uncommon for idioms to receive a rating of 1 or 2 from one learner but ratings of 4 or 5 from others. Such idioms include *take a back seat*, *a loose cannon*, *let the cat*

²⁶ Recall that the total valid cases including the 414 cases (GuessBP = 0 or 0.5) and 56 cases (GuessBP = 1) are 470. This is after the exclusion of 55 cases of known, re-countered, and error items (see section 4.1). For the 56 cases where learners' initial interpretation was successful (GuessBP = 1), 10 cases received a transparency score of 4 and 46 cases received a score 5.

out of bag, etc. Making predictions about which “etymological notes” will strike individual learners as clear vs. far-fetched underpinnings of an idiom’s meaning thus appears problematic, even within a relatively homogenous group of learners. Moreover, we need to be cautious about comparisons—e.g. of *take a back seat* as eliciting an average rating of only 3.50, with *red herring* eliciting an average rating of 4.33—because they cross over from one idiom set to another and do thus not reflect the same (group of) learners’ judgements. This is also one of the reasons why it was important to analyse the data by individual responses, or by taking group membership into account in mixed-effects modelling (see below), instead of using averaged ratings (which are still provided in Appendix 16 for reference). Moreover, the non-normal distribution of the ratings provided a further reason for the use of Cumulative Link Mixed Models (CLMMs) to analyse the transparency rating scales.

In sum, the data suggest that it may be very hard for teachers, lexicographers and researchers to make reliable predictions about which motivations for idioms’ meanings will be experienced as “making good sense” by an individual learner or group of learners (cf. Boers & Webb, 2015).

5.2.2 Regression analysis

The relationships between the etymological transparency ratings (Transparency, i.e., learners’ perceptions) and the correctness of interpretation (GuessAP, i.e., learners’ performance) and whether learners’ perceptions and performances are associated with their L2 proficiency were investigated by descriptive analysis in combination with regression analysis. The regression analysis was run by using ordinal mixed-effects regression modelling, also known as *Cumulative Link Mixed Models* (CLMMs) (see more detailed explanation of mixed effects modelling and CLMM in section 3.6). This was performed using the *clmm* function in R. Transparency and GuessAP were tested as dependent variables (DVs) in different models, both treated as ordinal rather than continuous data. That is, for GuessAP, a response scored 1 is more accurate than 0.5, but it does not mean the former is twice as accurate as the latter. For the

transparency ratings²⁷ (on a 6-point scale), 4 means a higher response than 2, but we cannot be sure that 4 means a transparency rating twice as high as 2. The output and relevant terms of mixed effects regression models are explained below.

The output of a CLMM includes a model *estimate* (i.e., the regression coefficient), the *standard error* of the estimate, a *z-value* and *p-value* (the latter is shown as $\Pr(>|z|)$ in the output) for each independent variable (IV)²⁸, which is also referred to as a predicting variable or predictor. In this study, an IV is considered a significant predictor if the p-value is below the conventional threshold of 0.05. The odds ratio of the DV changing from one response category to the next can be calculated by taking the exponential of the estimate. The obtained odd ratios are added to the output of each model as *Odds ratio*. In addition, for each predictor, a likelihood ratio test was run to compare a model with the predictor to a model without it (see more detailed explanation in section 6.1.1). AIC (i.e., Akaike Information Criterion) is reported, which can be used to compare models for better fit of the data, i.e., the smaller the AIC value, the better. If the difference in AIC (Burnham & Anderson, 2002) is less than 2, the models are said to be comparable. This offers a re-confirmation of the results of the LRT.

To test the possible relationships between the three factors, i.e., Transparency, GuessAP, and Proficiency, I first ran a series of CLMMs to test the relationships between each pair of factors. A significant positive relationship was found between Transparency and GuessAP, but no significant relationship was found between Transparency and Proficiency, nor between GuessAP and Proficiency. Then I tested a more complex model, by taking Transparency as the DV²⁹ and taking GuessAP and Proficiency, and their possible interaction as the predictors. The

²⁷ Note that because CLMM considers IVs of either categorical or numeric nature, Transparency will be treated as numeric later in Chapter 6, where it is a predicting variable of the correctness of recalls. The rationale is that there is a larger number of points on the scale and the intervals between the scale values can be assumed to be equal, making it more like a continuous scale. However, GuessAP has to be converted to categorical data, as the ratings 0, 0.5, and 1 are more of categorical than numeric nature.

²⁸ Note that for categorical IVs, there will not be a single estimate, a standard error, etc., but a series of the relevant parameters for all levels of the IV except the reference (intercept) level, showing the difference of predicted values of the DV for each level compared to its value at that reference level.

²⁹ Since the transparency ratings were obtained after the guessing phase and after all the meanings had been explained to the learners, participants' ratings were most likely to be affected by the inferencing success.

results of the tested model and a series of further model comparisons with and without interaction and simple effect(s) show firstly that there was no significant interaction effect between GuessAP and Proficiency, and secondly that the model including GuessAP and Transparency as simple effects is no better than the model including GuessAP as the sole fixed effect (see detailed output in Appendix 17). Therefore, in the following report, I will only discuss models with each factor as a sole predictor. Further regression analysis in Chapter 6 will be based on the findings of this Chapter—that is the regression analysis here serves as an exploration of variables to be included in further analysis.

5.2.2.1 Etymological semantic transparency and the correctness of aided interpretation

The two measures of L2 idiom interpretation under etymological elaboration, i.e., Transparency and GuessAP, offer two different ways to interpret analysability and semantic transparency of idioms from the L2 learner's perspective.

Based on the scores of GuessBP and GuessAP in Chapter 4, it was found that some idioms were relatively transparent and easy to understand even without aid of the etymological prompt (i.e., GuessBP = 1), some were somewhat harder to get (GuessBP = 0 or 0.5 and GuessAP = 1), and others were rather opaque and impossible to comprehend (GuessBP = 0 or 0.5 and GuessAP = 0 or 0.5). Given the minimal sentential context provided to the participants in this experiment, there were three common situations where it was particularly hard for them to get a fully correct interpretation: (1) Knowing the literal underpinnings of idiom was not sufficient for a fully correct inference, especially when there was more than one key aspect/point of information entailed in the original scenario (see previous discussion about *Challenge 2* in section 3.5.2), e.g., *take the bull by the horns*, *take the plunge*. (2) The etymology involved L2 cultural knowledge or even world knowledge that was not familiar to the learners. For example, some participants who had never taken an airplane and knew nothing about how it functions found *on autopilot* hard to interpret, even with the prompt that “In aircraft, automatic pilot is a device which automatically keeps the plane on course without the need for the pilot to do much.” (3) The etymological information was too obscure for the learners to infer any figurative meaning, e.g. *a red herring*. For the above cases, especially the first scenario, an informative context of use would potentially be more conducive for identifying the target domains of the metaphorical

expression and for inferring the meaning of the whole chunk.

After explicit teaching of the idiomatic meaning which aimed at fostering a full understanding of the literal and figurative meanings, learners might find the literal underpinning understandable and relatable, and this might render idioms relatively transparent, as shown in the large proportion of high Transparency ratings in the previous section (see Table 5.5). Those idioms from culture-specific domains, such as sports, travel, historical objects/events, as well as other items which the learners originally found difficult to interpret (as discussed above), may no longer seem as opaque as they did before learning. This raises the question of how learners' perceptions of the etymological transparency of the idioms is reflected in their performance during the interpretation. This is investigated by comparing the Transparency ratings and the GuessAP scores.

Figure 5.1 shows that when the inferencing carried out with the knowledge of the literal underpinnings was correct (right-hand panel), there were proportionally more higher transparency ratings. In addition, there were much smaller numbers of lower ratings (between 0 to 2) within each of the GuessAP response categories.

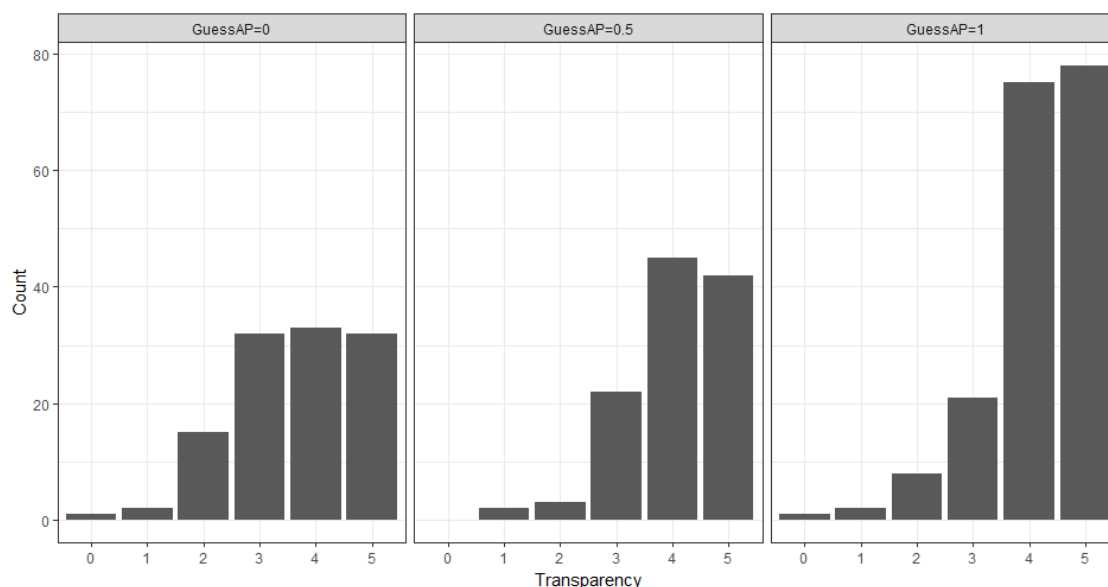


Figure 5.1 Distribution of transparency ratings with different GuessAP scores

Based on the mixed effects modelling (see output below), a significant positive relationship was found between the GuessAP scores and the Transparency ratings (both p-values less than

0.001). The likelihood ratio test shows a better fit with GuessAP in the model than without and AIC is much smaller with GuessAP included (see Appendix 17 for the full clmm output). The two estimated coefficients³⁰ in the model below show that Transparency is significantly higher when GuessAP is 0.5 than when it is 0, and the same is true when GuessAP is 1 compared to when it is 0. A further model in which the baseline value of GuessAP is changed from 0 to 0.5 (see further below), shows that Transparency is not significantly different for GuessAP values of 0.5 and 1. Overall, GuessAP contributes significantly to explaining the variance in Transparency ratings. The better the learners interpreted the meaning with etymological elaboration, the more transparent they tended to perceive the connection between the idioms' figurative meaning and their (proposed) literal underpinnings.

Model: Transparency ~ GuessAP + (1|Idiom) + (1|Participant)

Random effects:

Groups Name	Variance	Std.Dev.
Idiom (Intercept)	0.2660	0.5158
Participant (Intercept)	0.9368	0.9679

Number of groups: Idiom 77, Participant 25

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	Odds ratio
GuessAP0.5	1.0211	0.2772	3.683	p<0.001	2.78
GuessAP1	1.2912	0.2594	4.977	p<0.001	3.64

After setting the baseline of GuessAP as 0.5:

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	Odds ratio
GuessAP0	-1.0211	0.2772	-3.683	p<0.001	
GuessAP1	0.2701	0.2460	1.098	0.27231	1.31

5.2.2.2 Etymological semantic transparency and L2 proficiency

Likewise, the relationship between the etymological transparency ratings and the learners' L2 proficiency were tested with CLMM. The result below shows that Proficiency was not a significant predictor of Transparency ($p = .126$), and a negligible difference of AIC was found in model comparison with/without TEM-4 as an IV (see Appendix 17). These show that

³⁰ As an IV, the GuessAP scores were treated as categorical data, and thus the output reports two p-values and two odds ratios for comparisons between 0.5 and 0 and between 1 and 0.

learners' L2 proficiency was not a statistically significant predictor of their Transparency ratings. It is nevertheless of interest to note that the coefficient for this effect shows a negative relationship between Transparency and Proficiency, i.e. more proficient learners tend to give lower Transparency ratings. Upon examining the detailed ratings of some higher achievers and lower achievers, I find that in comparison with lower achievers, some high TEM-4 achievers gave lower transparency ratings to idioms. An avenue for future research would be to investigate whether this reflects a stronger appreciation by more proficient learners of the distance between literal and idiomatic meanings.

Model: Transparency ~ TEM-4 + (1 | Idiom) + (1|Participant)

Random effects:

Groups Name	Variance	Std.Dev.
Idiom (Intercept)	0.2772	0.5265
Participant (Intercept)	0.7004	0.8369

Number of groups: Idiom 77, Participant 25

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	Odds ratio
TEM-4	-0.04727	0.03093	-1.528	0.126	0.95

5.2.2.3 The correctness of aided interpretation and L2 proficiency

Similarly, a CLMM was performed with TEM-4 scores as a predictor and the correctness of GuessAP as the dependent variable. The result shows that the relationship was not significant ($p = 0.317$; and the AIC is slightly smaller without the predictor). This means that learners' L2 proficiency was not a significant predictor of idiom interpretation under the condition of etymological elaboration with limited contextual clues (refer back to section 4.3.1). This also seems to be in line with Johnson's (1996) finding that "complexity level in L2 metaphor interpretation appears to be quite independent from proficiency level in L2" (p. 227).

Model: GuessAP ~ TEM-4 + (1 | Idiom) + (1|Participant)

Random effects:

Groups Name	Variance	Std.Dev.
Idiom (Intercept)	0.9748	0.9873
Participant (Intercept)	0.0000	0.0000

Number of groups: Idiom 77, Participant 25

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	Odds ratio
TEM-4	0.01743	0.01741	1.001	0.317	1.02

5.3 Summary and discussion

To sum up, Chapter 5 examined the factors which are likely to affect initial L2 idiom interpretation, i.e., prior L2 lexical knowledge and L1 transfer; it investigated the effectiveness of etymological elaboration as an aid to interpretation via learners' perceived etymological semantic transparency; and it explored the relationships between etymological semantic transparency, interpretation success with the aid of etymological elaboration, and learners' L2 proficiency.

Regarding prior L2 lexical knowledge (in reply to RQ 4), it was found that in about 30% of the cases (Total N = 470), and relating to half of the target idioms, participants failed to interpret the idiom at the beginning of the learning procedure due to an unknown key content word. Meanwhile, the number of unknown words varies for different participants and for different sets of idioms, and some idioms contain very low-frequency words (e.g., *doldrums*). Moreover, lack of familiarity with the relevant meaning of homonymous or polysemous words was found to prevent successful interpretation in about 13% of the cases, involving 12 out of the 79 idioms (about 15%). Taken together, these findings point to insufficient L2 lexical knowledge as an initial obstacle to L2 idiom learning. However, after being informed of the meanings of certain unknown, polysemous, or homonymous words, as well as after receiving relevant etymological hints, some learners were able to make inferences of the idiomatic meaning that were more correct.

Regarding L1 transfer and cross-cultural differences (in reply to RQ 5), about 35% of the learners' responses (i.e., 166 out of 470 cases) referred to an L1 idiomatic expression which was either a total, a partial, or a false equivalent to the target idiom. The existence of total equivalents was indeed helpful for learners to interpret the L2 idioms correctly (as rated *fully correct* by the native-English raters), and false equivalents impeded correct interpretations (i.e. resulted in more interpretations that were rated *incorrect* by the raters). While the partial Chinese equivalents interfered with the comprehension of L2 idioms, the influence was hard

to evaluate (as was also reflected in the rating process). Moreover, in a different learning condition, a negative L1 influence could possibly be overcome by the availability of a richer situational context, as suggested by Charteris-Black (2002) and Türker (2016). Whether the L1 influence has a durable effect one week later on idiom retention will be investigated in Chapter 6 by means of an analysis of the errors in the recall results.

Regarding the etymological semantic transparency and the effectiveness of etymological elaboration, in the majority of cases (about 73.5%), the learners found the connection between the literal underpinnings and the target idiomatic meanings rather clear. This suggests that the etymological elaboration was overall well received during the interpretation phase. The uneven distribution of the Transparency ratings and the disparities between participants also show that some idioms were perceived as more transparent than others by specific learners. In other words, some idioms lend themselves better to etymological elaboration than others during the interpretation. The significant positive relationship between the Transparency ratings and the GuessAP scores, as expected, suggests that the better the learners interpreted the meaning, the more transparent they perceived the connection between the literal and figurative meanings to be. Furthermore, no significant relationship was found between learners' perceived Transparency and their L2 proficiency, nor was L2 proficiency a significant predictor of interpretation success under the examined learning condition.

While Transparency and GuessAP offer two different ways to interpret analysability and semantic transparency of idioms from the L2 learner's perspective, it is important to remember that both of these measures are potentially influenced by the L1. Since GuessAP scores were derived from the learners' responses in L1 as well as L2, L1 influence was already an integral part of the correctness of GuessAP. On the other hand, learners' etymological transparency ratings, as discussed previously, were bound to be affected not only by the features of the idioms but also by the learners' own prior lexical and cultural knowledge, as also suggested by the consistently high ratings in cases where the initial interpretation was correct due to L1 resemblances (see section 4.2.1). Finally, whether the factors examined for the idiom interpretation affect meaning retention one week later will be investigated in Chapter 6.

CHAPTER SIX: INVESTIGATING L2 IDIOM MEANING RETENTION

Introduction

This chapter investigates the factors affecting the retention of the meanings of idioms under the examined teaching technique. The first section consists of both quantitative and qualitative analysis of how well L2 idiom meaning retention is predicted by the factors involved in the study, i.e., learners' perceived transparency of idioms, learners' L2 proficiency, and the correctness of learners' guesses when they had support from etymological elaboration. This is connected to the previous exploration in section 5.2. Questions to be answered include:

RQ 6. Does learners' perceived transparency of idioms affect L2 idiom retention and how?

RQ 7. Does L2 proficiency affect L2 idiom retention and how?

Moreover, following the descriptive analysis in Chapter 4, this section further investigates RQ 3 about whether the procedure of inferencing based on etymological elaboration has an impact on retention, via first an inferential analysis and then an analysis of the recall errors. This analysis of errors concerns the factors of L1 transfer and prior L2 lexical knowledge (as identified from the participants' responses during the interpretation phase), which is linked with the exploration in section 5.1.

The second section of Chapter 6 consists of qualitative analysis of more general learning factors involved in L2 idiom learning, including learners' experiences and awareness of idioms, their learning strategies and attitudes. Relevant questions are:

RQ 8. What do the EFL learners' experiences with and awareness of idioms tell us about L2 idiom learning?

RQ 9. What does their use of strategies in comprehending and memorising unfamiliar idioms tell us about L2 idiom learning?

RQ 10. What are the learners' attitudes towards using and learning L2 idioms?

This chapter together with the previous chapter aims to provide a full examination of the Chinese EFL learners' experiences and thoughts, performances and perceptions about L2 idiom

interpretation and meaning retention with the help of etymological elaboration.

6.1 Factors affecting L2 idiom meaning retention

The dependent variable in all the analyses in this section is the correctness of meaning recall results, represented by the *Recall* scores. The independent variables are learners' perceived etymological semantic transparency of the idioms, L2 proficiency, and the correctness of guesses after the etymological elaboration (i.e., guesses after the prompt, or GuessAP, for short). As in the exploration reported in Chapter 5, Cumulative Link Mixed Models (CLMM) were used in inferential analysis to explore the effects of the above factors on the meaning recall results.

Prior lexical knowledge (regarding unknown words and polysemy/homonymy) and L1 transfer were not included in the mixed effects modelling, but they were identified from learners' responses, to gauge their possible effects on the recall errors. Table 6.1 provides a summary of all these factors and variables.

Table 6.1 Factors & analysis in investigating L2 idiom meaning retention

Analysis	Factor	Code (label/value)	Data type
DV in CLMM	Correctness of recall answers	Recall (0/.5/1)	Ordinal
IV in CLMM	Learners' perceived etymological semantic transparency of the items	Transparency (0-5) ^b	Numeric
	L2 proficiency (TEM-4 scores)	Proficiency/TEM-4	Numeric
	Correctness of guesses after the prompt	GuessAP (0/.5/1) ^b	Categorical
Analysis of errors	Prior lexical knowledge: (1) Unknown words (2) Polysemy/homonymy	UnKnownW (0/1) ^a PolyHomo (0/1) ^a	Categorical
	L1 transfer	FFL1 (0/1)	Categorical

(Note. a. The label 0/1 stands for No/Yes for all the relevant factors. b. Since independent variables in CLMM have to be either categorical or numeric, GuessAP was treated as categorical, and Transparency as numeric in the following analysis; see also explanation in section 5.2.)

6.1.1 Model selection

All the CLMMs included participants and idioms as random effects to account for differences attributed to the individual learners and items, separate to the fixed effects that were identified from the participants and idioms, such as learners' L2 proficiency and their perceived

transparency of idioms. As introduced in section 3.6, an intercepts-only random effects structure was adopted for all the mixed effects models (MEMs), after a preliminary analysis including random slopes over proficiency for each idiom showed that the effect of proficiency for each idiom was approximately the same. This was tested in the Transparency*Proficiency MEM (to be discussed later), where by-idiom slopes for proficiency contributed little to explaining the variance in the model (see previous discussion in section 3.6.2).

The statistical significance of including each predictor in CLMMs was tested by comparison of models with and without that predictor (IV) ³¹. As previously introduced in section 5.2.2, the **Likelihood Ratio Test (LRT)** is used for model comparisons. The comparison returns a likelihood ratio statistic with a chi-square distribution. The result is reported in the form (χ^2 (df) = A, $p < B$), where df = degrees of freedom, A = likelihood ratio statistic, and B relates to the p-value. If the p-value is less than 0.05, the result is said to be significant, indicating that the models are different, and the model with smaller AIC (Akaike Information Criterion) is to be preferred as a better fit³². In the description of results below, the significance of factors and/or interactions is reported in terms of such model comparisons. Further details of the model outputs can be found in the relevant Appendix.

To address RQ 6, RQ 7, and RQ 3, three variables—learners' perceived etymological semantic transparency of idioms, L2 proficiency, and correctness of guesses after etymological elaboration—were tested in a series of models to search for the best fit models for the recall results. Three hypotheses about the simple effects were as follows:

Hypothesis 1—The more transparent a learner perceives the link between the literal and figurative meanings of an idiom to be, the more likely it is that they remember it correctly.

³¹ Note that it is important to run model comparison, even if the p-value(s) of relevant predictors in a mixed effects model (MEM) are less than .05, because the summary of an individual MEM does not tell us whether a predictor (or an interaction between predictors) makes a significant contribution in accounting for variance. Although the output of a MEM has significance levels for its predictors, these are relative to a baseline setting of all the variables in the same model and do not necessarily show whether including the predictors provides significant explanation of the variance.

³² If the models are not significantly different (see previous explanation in section 5.2.2), the simpler model—with fewer predictors and/or no interactions—is preferred.

Hypothesis 2—Higher-proficiency learners produce better recall results than lower proficiency learners.

Hypothesis 3—Correct guesses are more likely to lead to correct recall.

As a significant relationship had previously been identified between Transparency and GuessAP, these two factors were treated in separate models in the following analysis, to avoid collinearity in regression models of the recall results. However, the earlier analysis of L2 proficiency and transparency (section 5.2.2.2) showed these two factors were not strongly related to one another. There is nevertheless potential for these two factors to interact in their prediction of Recall scores. For instance, those participants who have relatively higher L2 proficiency may be more inclined to engage with and remember all of the etymological elaborations proposed by a teacher, regardless of the perceived level of transparency of the etymology (refer to discussion in 2.3.3.5). Therefore, the potential interaction between Transparency and Proficiency was included as one of the predictors in the exploration, and was found to be significant (see below). Note, however, that Proficiency did not interact with GuessAP ($\chi^2(1) = 0.1005$, $p = .75$), and had no significant effect in a simple additive model ($\chi^2(1) = 2.6224$, $p = .10$; see ‘Output for 6.1.3’ in Appendix 18), and therefore was not included in a model in which GuessAP serves as a predictor of Recall. Consequently, after a series of model comparisons in searching for the best fit models, two CLMM models were selected (see below). One includes Transparency, Proficiency, and their interaction as fixed effects, and the other include GuessAP as the sole fixed effect. I will report and discuss the analysis of two models in sections 6.1.2 and 6.1.3 respectively.

6.1.2 Etymological semantic transparency and L2 proficiency

We now turn to the principal RQ 6, that is, whether learners’ perceived semantic transparency affects meaning retention of L2 idioms and how. In other words, whether learners’ appreciation of the connection between the literal underpinnings and figurative meanings of idioms matters for establishing accurate memories. From a pedagogical perspective, this will offer insights into whether the effectiveness of the etymological elaboration in facilitating retention depends on learners’ perceived transparency of idioms.

Table 6.2 gives a first impression of the relationship between the Transparency ratings and the Recall results. Because the participants' ratings were so unevenly distributed with a large number of ratings clustered at 4 and 5 (as shown in section 5.2.1), I divide them here into two categories just for preliminary, descriptive purposes: ratings 0 to 3 ("low") versus ratings 4 and 5 ("high").

Table 6.2 Meaning recall of idioms rated, by low and high transparency (total N = 414)

Transparency	Meaning recall					
	Correct		Partially correct		Wrong/nil	
	N	%	N	%	N	%
0-3	66	60.55	14	12.84	29	26.60
4-5	192	62.95	49	16.07	64	20.98

The tendency observed earlier towards a positive relationship between Transparency and interpretation (GuessAP) is not mirrored in such a tendency for Transparency and Recall, since recall scores (correct, partially correct, wrong/nil) are distributed similarly for low and high transparency sets. It is worth noting, however, that 60.55% of the cases where the items were rated relatively low (0-3) were successfully recalled by the learners. Among these cases, the literal underpinning that elicited the lowest average rating (the only item with an average rating that was below 3.00) concerned *a wet blanket*. One participant explained his low rating (0) as follows: "If it [the idiom] comes from putting out fire, I see it as something useful in a difficult or dangerous situation. This is in contrast with the negative meaning of the expression as stopping other people's enthusiasm." It seems this learner found the proposed mapping between the domains of the source and of the target improbable, because of his association of fire with danger. However, his low transparency rating did not prevent him from accurately recalling both the idiom's meaning and its proposed origin one week later. Similar examples concerned idioms such as *take it on the chin*, *take the plunge*, *throw your hat into the ring*, *get your second wind*, *play it by ear*, *a loose cannon*, etc. These suggest that etymological elaboration might not only be applied to those items that were perceived as relatively transparent, but is also useful for those that were perceived as opaque.

What Table 6.2 fails to capture, of course, is variation in recall scores due to many other potentially influential characteristics of individual idioms as well as characteristics of the individual participants. Regarding the latter (and as discussed earlier), the students' comparative success as EFL learners may be particularly relevant. Therefore, I turn to mixed-effects regression models for further exploration.

The dependent variable was the Recall scores, with the scores of 0, 0.5 and 1, representing ordered categories of wrong/nil, partially correct and fully correct. The fixed effect predictors were Transparency (i.e., etymological semantic transparency ratings) and Proficiency (i.e., the student's TEM-4 grade). Since it was assumed that the effect of perceived transparency may vary with the level of L2 proficiency (see above), I included the interaction between Transparency and Proficiency as another fixed effect. The random effect structure included by-participant and by-item intercepts. The coefficients (see below) showed that three fixed effects, Transparency, Proficiency, and their interaction were all significant in the model (all p-values less than 0.05).

Model: Recall ~ Transparency * TEM4s³³ + (1 | participant) + (1 | idiom)

Random effects:

Groups Name	Variance	Std.Dev.
idiom (Intercept)	1.1167	1.0567
participant (Intercept)	0.6505	0.8065

Number of groups: idiom 77, participant 25

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	Odds ratio
Transparency	0.26423	0.13312	1.985	0.0471	1.30
TEM-4s	0.27607	0.10951	2.521	0.0117	1.32
Transparency:TEM-4s	-0.05033	0.02412	-2.086	0.0369	0.95

Comparison of models with and without the interaction between Transparency and Proficiency (see detailed R output in Appendix 18) showed that the interaction made a significant contribution to explaining the variance in recall scores ($\chi^2(1) = 4.72$, $p < .05$). The full model with the two simple fixed effects and their interaction was therefore retained. A Type-III

³³ TEM-4s is obtained from all TEM-4 scores subtracted by the mean score 75.72, in order to get rid of the convergence error in R.

ANOVA test showed that in addition to the interaction, the simple fixed effects of Transparency ($\chi^2(1) = 4.01, p < .05$) and Proficiency ($\chi^2(1) = 7.13, p < .01$) also made significant contributions to the full model. The interaction effect is illustrated in Figure 6.1, which is generated via the *predict* function in R³⁴. Overall, these effects reflect the following patterns.

First, across proficiency levels there was an overall effect of fully correct recall being more likely for idioms perceived to have transparent literal-figurative connections (i.e. higher Transparency scores, shown by paler symbols in Figure 6.1), while unsuccessful recall was more likely for those whose proposed underpinnings were perceived to be rather obscure. This supports an affirmative answer to Hypothesis 1, i.e., Transparency matters. Second, fully correct recall was also more likely for the learners with comparatively high TEM-4 scores (towards the right of the figure). This is in line with the prediction of Hypothesis 2, i.e., higher-proficiency learners produce better recall results than lower proficiency learners.

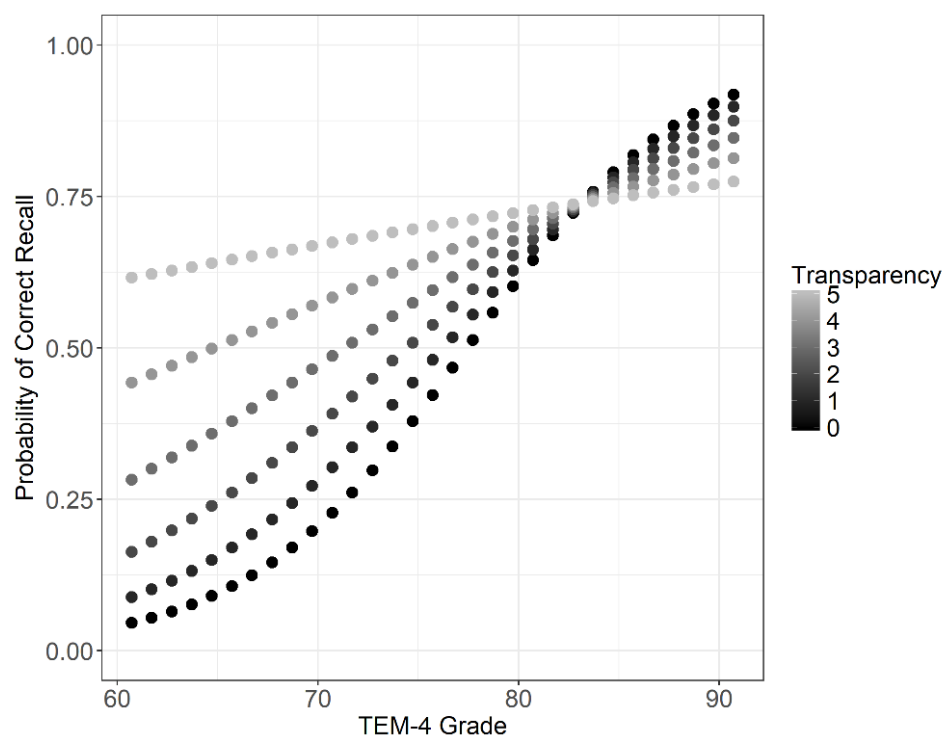


Figure 6.1 Probability of correct recall based on TEM-4 grade and Transparency

³⁴ Because the *predict* function has not been developed for *clmm*, this plot shows values derived using the *predict* function in *clmm2*-- an alternative function from the same *Ordinal* package. However, because *clmm2* only allows one random effect, I chose to keep items as the random effect since items explain more variance than participants (as shown in the output above).

The interaction effect revealed by the regression model is as follows. While the students with low TEM-4 scores performed generally more poorly than their high-achieving peers on the recall test, this was especially acute for idioms whose transparency they had rated as low.

Idioms whose literal-figurative connection they deemed highly transparent were recalled the best at this lower end of the TEM-4 grade, and there is a greater separation between the curves for the different transparency levels on the left-hand side of Figure 6.1. However, as TEM-4 scores increased, the impact of transparency gradually diminished and disappeared by grade 80 (see Figure 6.1). So, the effect of transparency as a significant predictor of recall is attributable, to a large extent, to the low achievers' data. Returning to the example of the student who had given *a wet blanket* the lowest transparency rating but nonetheless successfully recalled it, it is perhaps no coincidence that this was a student with a relatively high TEM-4 grade (78). According to the interview data, this student was clearly willing to reflect on the literal underpinning of the idiom proposed by the researcher, perhaps precisely because its motivation for the idiomatic meaning was puzzling to him. It is possibly this willingness to put a certain effort into evaluating a proposed literal-figurative connection that helped him to entrench this association in memory (recall that the student remembered both the idiom's meaning and its proposed underpinning). This assumption was later supported by follow-up interview data.

To allow comparison with previous studies that looked separately at measures of transparency (e.g., Boers et al., 2004b) and proficiency, I also explored the impact of Transparency and Proficiency as sole predictors in the Recall models. The results below show that Transparency is not a significant predictor when treated as a sole main effect ($p = .104$, the AICs are comparable with/without the predictor). At first sight, this result seems align with what Boers et al. (2004b) found in their 2003 data, when the analysis is confined to one variable, i.e., transparency. However, the result below shows that Proficiency remained a significant sole predictor ($p < .05$; the AIC is slightly smaller with the predictor than without). This further suggests that the effect of transparency was associated with learners' L2 proficiency and that it is necessary to consider multiple variables and possible interaction effects in such an analysis. The interaction is indispensable in predicting the recall results. In other words, the learning

outcomes under the examined teaching technique were affected by learners' perceived transparency of the idioms but the impact is mediated by learners' L2 proficiency level. Following this, later in section 6.2, I will explore the characteristics of exemplary learners (identified in Chapter 4) in relation to their exposure to L2 idioms, their learning strategies and attitudes.

Model 1: $RC \sim Transparency + (1 | participant) + (1 | idiom)$

Random effects:

Groups	Name	Variance	Std.Dev.
idiom	(Intercept)	1.1565	1.0754
participant	(Intercept)	0.6919	0.8318

Number of groups: idiom 77, participant 25

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	Odds ratio
Transparency	0.2088	0.1283	1.627	0.104	1.23

Model 2: $Recall \sim TEM-4s + (1 | participant) + (1 | idiom)$

Random effects:

Groups	Name	Variance	Std.Dev.
idiom	(Intercept)	1.0901	1.0441
participant	(Intercept)	0.5881	0.7669

Number of groups: idiom 77, participant 25

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	Odds ratio
TEM-4s	0.05515	0.01632	3.38	0.000724	1.06

6.1.3 The correctness of guesses with the aid of etymological elaboration

In response to RQ3, this section explores the effects of the correctness of inferencing on the recall success. Section 6.1.3.1 reports an inferential analysis of the relationship between the correctness of the recall answers (i.e., Recall) and that of the guesses with the aid of etymological elaboration (i.e., GuessAP). Section 6.1.3.2 reports an analysis of the recall errors. This analysis of errors compares the unsuccessful recall answers (including those of no recall, wrong recalls, and partially correct recalls) with their corresponding unsuccessful guesses (involving wrong and partially correct responses) made before and after the etymological elaboration, and identifies the possible memory traces and potential causes. The data involved

in both the inferential and the error analyses are the 414 cases where no correct guesses were made before the prompt (GuessBP = 0 or 0.5).

6.1.3.1 Inferential analysis of Recall and GuessAP

In Chapter 4 it was reported that the number of correct recalls increases in proportion to the accuracy of the guesses after the prompt, which suggested that the scores at GuessAP could be predictive of Correct Recall. For cases where the meanings of the idioms were successfully interpreted after the prompt (GuessAP = 1), 73.5% were subsequently correctly recalled. When the inferencing after the prompt was not so successful (GuessAP = 0 or 0.5), the recall rate was only about 53%. In this section, this observed pattern is further investigated in mixed effects modelling, where the correctness of guesses after the prompt (i.e., GuessAP) is tested as a predictor of the recall results. As an IV, GuessAP was treated as categorical data in CLMM. The CLMM results below showed that it required a change of GuessAP from 0 to 1 to make a significant change in the Recall ($p < .01$), while a change of GuessAP from 0 to 0.5 did not result in a significant increase in Recall ($p = 0.21$). When the GuessAP intercept value was set to 0.5 and a further model run, it transpired that a change of GuessAP from 0.5 to 1 also failed to result in a significant increase in Recall scores. Model comparison also shows that the original model including GuessAP is a better fit than a model without GuessAP (see the output in Appendix 18).

Model 3: Recall ~ GuessAP + (1 | participant) + (1 | idiom)

Random effects:

Groups Name	Variance	Std.Dev.
idiom (Intercept)	1.0671	1.0330
participant (Intercept)	0.6825	0.8261

Number of groups: Idiom 77, IDV 25

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	Odds Ratio
GuessAP0.5	0.4058	0.3231	1.256	0.20913	1.50
GuessAP1	0.9613	0.3149	3.053	0.00226	2.62

After setting the baseline of GuessAP as 0.5:

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	Odds Ratio
GuessAP0	-0.4058	0.3231	-1.256	0.209	

GuessAP1	0.5556	0.3067	1.812	0.070	1.74
----------	--------	--------	-------	-------	------

In addition, participants' reflections on the learning process and outcomes provide further personal evidence concerning the results presented above. Participant B3 recalled that "If I was able to guess the meaning correctly, I was often left with a stronger memory for it." Participant A2 recalled that "I often recalled the first guess that I made, which was likely to be wrong." And Participant C4 made an interesting comment, "My first impression of the idiom often mattered most. That is if I guessed it wrong, I might recall it wrong. However, if the actual meanings and my guesses contrasted strongly, then it was possible for me to remember the correct meaning." This last comment also provides a possible explanation for the special case of *wet blanket* mentioned above.

Taken together, the results presented in this section support my prediction in Hypothesis 3 that correct guesses were more likely to lead to correct recall. As a follow up, an analysis of errors was conducted to further investigate where the incorrect and partially correct recalls came from—whether they were associated with previous guesses during the learning phase or with other influential factors.

6.1.3.2 Analysis of the recall errors

The purpose of this analysis was to trace the possible reasons behind the recall errors. The identification of these error traces was conducted by comparing the unsuccessful recall responses with their respective guesses before and after the prompts. For example, one specific wrong recall for *at the wake of* was "at the early beginning of something" (i.e., the participant's response in English). It resembled and thus could be traced back to the participant's earlier guess in Chinese, "在开头" meaning 'at the beginning of.' In the same vein, one partially correct recall for *bread and butter*—"something important" resembled the participant's previous guess—"非常重要的一件事(因为吃的总是很重要)", meaning "something very important (as food is always important in life)", and this could therefore be marked as the error trace.

To further investigate the error rates and what caused the errors, I will first differentiate two scenarios: (1) no guess made before the prompt (i.e., before the etymological notes was given), and (2) one or more guess(es) made before the prompt. This is because those cases with

incorrect inferences at this stage, i.e. where $\text{GuessBP} = 0$, include both cases where no inference attempt was made and cases where wrong guesses were made. Such a differentiation will help identify whether the wrong or partially correct response could be traced back to the guesses before or after the etymological elaboration (see the two separate *None/One or more* sections of Table 6.3). The identified error traces were counted for each scenario (see the shaded columns), and represented as a number plus where it was traced back to. This is shown under the ‘error trace’ in the last line of the table. For example, *6 GuessAPs* means that there were 6 errors traceable to a guess after the prompt; *19 GuessBPs* means that there were 19 errors traceable to a guess before the prompt. In total, there were 50 error traces identified.

Table 6.3 shows that in both scenarios—whether there were guess(es) made before the prompt or not, the correct recall rates were above 60% (i.e., 62.9% and 60.7%).

Table 6.3 Tracing recall errors from GuessBPs and GuessAPs (Total case N = 414)

	Guess(es) made before the prompt									
	None (N = 112)				One or more (N = 302)					
Recall scores	1	0.5	0		1	0.5		0		
			No recall attempt	Wrong recall				No recall attempt	Wrong recall	
No. of cases (N=110)	68 (60.7%)	13 (11.6%)	26 (23.2%)	5 (4.5%)	190 (62.9%)	50 (16.6%)		20 (6.6%)	42 (13.9%)	
Error trace (N=50)	NA	6 GuessAPs	NA	1 GuessAP	NA	22 Guesses		NA	21 Guesses	
						19 GuessBPs	3 GuessAPs		14 GuessBPs	7 GuessAPs

When no guesses were made before the prompt³⁵, the participants only started to articulate guesses after the literal underpinning was given to them (see the left section of Table 6.3). In such cases, only a small number of partially correct and wrong recall answers could be traced back to the guesses after the prompt (6 out of 13 GuessAPs when Recall = 0.5; 1 out of 5 GuessAPs when Recall = 0).

³⁵ In cases where there are no guesses before the prompt, GuessBP necessarily is zero.

When the participants started to articulate their guesses before the prompt was given³⁶, then the recall error might link back to the guess(es) before and/or after the prompt (see the right section of Table 6.3). When the recall answer was partially correct, 22 out of the 50 errors (44%) could be linked to partially correct guesses, and the partially correct recall was either the recall of a guess made during trial-and-error learning or a repetition of an inaccurate inference made at that earlier stage. When the recall answer was completely wrong, 21 out of 42 errors (50%) could be linked to a wrong guess. In total, about 45.5% of the recall errors (i.e., 50 out of 110) could be traced back to unsuccessful guesses.

A further analysis considered whether the errors could be related to the presence of an L1 equivalent. A close look at the total set of 50 traceable recall errors shows that 12 of them could be traced back to items that contain a partial equivalent or false equivalent. Relevant items include previously discussed idioms, such as *go with the flow*, *turn over a new leaf*, and *play into someone's hand*. The presence of unknown words may also be important. For instance, there were 49 cases where participants failed to provide any response in the recall test, and in 33 of these cases, the idiom contained a content word which was unknown to the participants prior to the learning session (recall that the total number of cases of unguessable items due to Unknown word was 140; i.e., 23.6%). These idioms include *get into gear*, *be par for the course*, *in the doldrums*, *(not) up to scratch*, *pass the buck*, etc. Almost all the prior unknown words (see underlined) are very low-frequency words. Additionally, 9 of the 49 cases—only 5 after excluding the co-occurrences with an unknown word—contained a polysemous or homonymous word. Although L1 interference and insufficient lexical knowledge, especially the presence of a prior unknown key word, might not be the only obstacles to retaining the idiomatic meaning, this result suggests that these factors might impede L2 idiom meaning retention to some extent. In sum, the statistical and error analyses in this section reveal the potential downside of meaning guessing through trial and error.

³⁶ This is where there is one or more guess before the prompt, and where such guesses are either inaccurate, i.e., GuessBP = 0, or partially accurate, i.e., GuessBP = 0.5. Recall that the data here exclude cases where the correct interpretation was given before any prompting, i.e., GuessBP = 1.

6.2 Qualitative analysis of other factors involved in L2 idiom learning

This section reports the findings from the semi-structured interviews in the second week, which were about general factors involved in idiom learning: (a) the learners' personal experience with and awareness of English idioms (see 6.2.1), (b) their strategies in coping with unfamiliar (idiomatic) expressions when they encounter them, (c) methods they use to memorise such expressions (see 6.2.2), and (d) their general attitudes towards using and learning idioms (see 6.2.3). Finally, in-depth analysis is conducted with three learners who were identified as high achievers according to the recall test results (see 6.2.4). The question (Q) numbers given below correspond to those in the semi-structured interviews (see section 3.4.2).

6.2.1 The EFL Learners' personal experience with and awareness of L2 idioms

Q2. When and where do you remember encountering such figurative idiomatic expressions in your own learning experience?

In response to Q2, 25 participants reported their personal experience with and awareness of L2 idioms. Table 6.4 is a summary of sources of their exposure to idioms.

Table 6.4 EFL learners' sources of exposure to L2 idioms

Source of exposure	Count	Frequency within each type of source
American/British TV series & films (with subtitles)	13	Relatively often
Literary works (e.g., novels, proses) written by English native speakers in unsimplified/original versions	12	Sometimes
Translation and reading classes	9	Sometimes
Textbooks (on reading and literature)	8	Occasionally
News (e.g., on VOA/BBC), newspapers & magazines on politics/economics (e.g., in <i>The Economist</i>)	5	Sometimes
Reading comprehension tests	3	Occasionally
English speaking class given by native English speakers	1	Sometimes
English learning content on social media (e.g., apps)	1	Occasionally
Writing sample	1	Rarely

First, it is found that participants generally did not often encounter (or may not have been aware of) English idioms (or in their own words—the kind of expressions they were taught in the interviews), and that learners’ experiences with idioms varied from person to person. Several participants reported that they did not have awareness of such expressions before the interview. One participant even asked the researcher whether native speakers really use such expressions—suggesting that EFL learners might not be fully aware of the existence of idioms and thus underestimate the actual use of idioms in native-speakers’ discourse (see previous discussion in section 2.3.1). By contrast, participant B3 said she *often* encountered idioms as she read *The Economist*, a popular periodical on world politics and economics.

Second, it was reported that idioms were usually encountered in informal discourse produced by native speakers, and rarely found in written discourse. In this regard, learners’ reflections were consistent with previous corpus studies on the frequency of idioms in different genres (e.g., Moon, 1998; Simpson & Mendis, 2003). The interview data from the current study show that American and British TV series and films were the most common source of exposure for the participants, which was followed by literary works such as novels (especially the unsimplified versions—commented by two participants). Classroom teaching and textbooks were also an important source. Translation and reading classes were mentioned by 8 out of 25 participants, where they recalled that the teachers taught English idioms under the themes of animals (e.g., *hold your horses*, *dog’s day*), colours (e.g., *green-eyed monster*—similar to Chinese expression 红眼病, *pinyin: hong yan bing*, literally meaning red eyes, and figuratively meaning jealous of others’ success), and emotions (e.g., *blow off one’s lid*, *hit the roof*), and asked students to compare the expressions with Chinese idiomatic language. Newspapers and magazines were another frequent source of exposure. Moreover, it was also reported that idioms were encountered occasionally in reading comprehension tests and in assignments carried out as part of translation courses.

Overall, the interview data suggest that the EFL participants have rather limited exposure to L2 idioms inside and outside the classroom, with insufficient awareness of idioms. In further analysis below, it will be found that learners’ strategies and attitudes towards idiom learning vary depending on their personal experiences with idioms and their learning habits.

6.2.2 The EFL learners' strategies for L2 idiom comprehension and memorisation

Q1. What do you usually do when encountering unfamiliar English idioms (e.g. ignore them, make a guess at their meaning, look them up in a dictionary)?

In response to Q1, the participants predominantly chose to guess the meaning based on the context where the expression was presented. Looking the idiom up in a dictionary remained an option when guessing the meaning was not successful (i.e., when their guess did not fit the situational context). For example, A5 commented: "I prefer to first guess the meaning based on the context. I will look it up in a dictionary only when I cannot get the meaning; and I normally look up certain [single] words in the phrase in a common English dictionary, instead of looking up the whole phrase in an idiom dictionary. If I still cannot get it, I will just ignore it." Similarly, B3 reported: "I prefer inferencing the meaning [of the idiom] based on the contexts; if there is an unknown word, I may look it up in a dictionary. If the expression appears in a TV drama or a film, it will be easier to guess the meaning with the movements and the responses of the characters and the situation where it happened."

Some participants are more willing (than others) to use a dictionary when they cannot get the meaning of an idiom from the context. B2 said "My own inferences of the meaning of such expressions [i.e., idioms] are most likely to be wrong, and the wrong guesses may remain in my memory if I don't correct them by looking up the idioms in a dictionary." C1 reported that she would not let go of unfamiliar idioms when she encountered them, and would always look them up in a dictionary— "the more you know such expressions, the more you speak like a native speaker" she said. Moreover, some participants decided whether to further explore the meaning of an idiom based on whether it occurs in the textbook or in extra-curricular activities. For the former, they would pay more attention and look up an unfamiliar expression in a dictionary; otherwise, they might choose to ignore them during reading/watching.

Regarding specific dictionary usage, five participants offered three different preferences concerning the language in which they sought explanations: C1 preferred to read both Chinese and English explanations to get a full understanding of the idioms; A2 and C6 usually only looked for the Chinese explanation via an electronic dictionary on his/her phone, and

memorised the expression with its Chinese translation; B5 and D4 preferred reading the English explanation of the expression in a dictionary, rather than a Chinese explanation. D4 further commented that he did not often find that the origins of idioms were provided in dictionaries, particularly in the electronic ones. We will come back to participants B5 and C1 in the discussion of case studies in section 6.2.4.

In sum, using contextual information to infer the meaning of an unfamiliar idiom appeared to be the most frequently (and conveniently) used strategy. This result is in line with Cooper's (1998) finding in his think-aloud study of L2 idiom processing. Further explorations of the idiomatic meaning via a dictionary and particular ways of using dictionaries seem to vary according to learners' personal attitudes towards L2 idioms and perhaps towards vocabulary learning in general.

Q3. What do you often do to remember such expressions (e.g. associate them with where you encountered them, associate them with a mental picture, compare them to a similar Chinese expression, write them down, try to use them)?

In response to Q3, the participants reported strategies that they used when memorising idioms. It should be noted that responses to this question were likely to be affected by the implementation of the teaching experiment prior to this interview, but the findings may nevertheless have important implications for L2 idiom learning. Four participants reported that they never intended to remember idioms. Comments from the other students revealed the following strategies.

Table 6.5 shows that the most popular strategies were (1) associating the idiom with an L1 equivalent, (2) associating an idiom with a mental image derived from its origin, or linking it with the learner's personal life experience and/or world knowledge, and (3) associating the idiom with a specific and meaningful context. These strategies are not mutually exclusive—that is, most often participants reported more than one strategy, e.g., one may relate the idiom to a personal experience AND write it down. Still, many participants had strong preferences for some strategies over others, and very few chose all the strategies.

Table 6.5 Frequency of strategies used by the participants to memorise idioms

Strategies used to memorise idioms	Count
Associate the item with a similar Chinese expression provided there is one	13
Associate the item with a mental image (relying on the literal meaning) or with one's own life experiences and/or world knowledge	12
Associate the item with a specific meaningful context where it was encountered	9
Take notes (write the item down)	6
Others (re-encounter/repeat)	2

Some participants preferred to associate an idiom with their personal life experiences and make use of a vivid mental image to memorise the expression. For example, D2 commented that “As long as there is a supportive context, a clear vivid image of the literal sense, plus correction provided if my own inference were wrong, those [cultural or habitual] differences between my own life experiences and those peculiar to L2 countries won’t stop me from remembering the correct meaning of an idiom.” Some participants preferred to associate the idiom with the context where they encountered the expression, so that they can fully grasp the knowledge of how and where to use the idioms. Last but not least, about half of the participants reported that they would consciously or unconsciously compare the target idioms with a Chinese idiomatic expression; if there is an equivalent, then it is be much easier for them to memorise the expression. Additionally, a few participants also mentioned that seeing the idiom once might not be enough to commit it to memory, and that seeing it repeatedly or copying it down in a notebook would facilitate a more entrenched memory. These reflections from the participants also offer anecdotal evidence supporting claims about the positive impact of L1 equivalents (e.g., Charteris-Black, 2002; Liontas, 2002; Hu & Fong, 2010), as well as supporting claims about the potential benefits of copying and repetition (e.g., Stengers, Deconinck, Boers, & Eyckmans, 2016, though they found copying had no significant effect).

6.2.3 The EFL learners' attitudes towards using and learning L2 idioms

Q4. Generally, have you ever tried to use idioms once you learned them?

In response to Q4, participants reported their experiences and attitudes, reasons and concerns about *using* L2 idioms. The participants had not experienced much intentional learning of idioms (except in translation or reading classes) or discussion of the use of idioms prior to the interviews, and they demonstrated in general a very cautious attitude towards using idioms in their own speech or writing. For example, participant B5 commented that “I’m very cautious in using them...like...I only use them when I’m very familiar with the expression.” Similarly, C3 expressed her concern that “I don’t have the opportunity to use them and am afraid I won’t use them appropriately.” Such a cautious attitude is further reflected in the following patterns.

Table 6.6 Attitudes towards using L2 idioms by the EFL learners

Interest in use	ID	Interview excerpts
No interest in use	D2	<i>I have no interest in using them.</i>
	A2	<i>I don't use them in speaking, but I might understand them in reading.</i>
	A7	<i>It would be idiomatic for native speakers to use, and those expressions often require certain context to be used. In English writing, I would avoid using them and I'm afraid that I would not use them properly. When talking to other English learners, I would not use them either (because I don't think they could understand me).</i>
In speaking but not writing	B6	<i>I would not use them in writing, but I would like to use them in daily communication with others and share them with my friends, e.g., 'You are <u>the apple of my eye</u>.'</i>
	A6, B3, D1, D3	<i>I would try to use them in daily practice (by oneself) or in informal talk with friends.</i>
In writing	A4	<i>In essay writing, I tried to use the expressions, e.g., <u>cry over spilled milk</u>. I like trying on what has been newly learned and put it into use.</i>
	C6	<i>I want to use them in writing.</i>
In speaking & writing	D5	<i>I want to use idioms selectively in writing and oral presentation in class. But I would be cautious. They seem rather colloquial.</i>

First, most participants reported that they had never or seldom used idioms themselves. This was understandable, since they have limited exposure to idioms (as discussed above) and had previously learned very few idioms.

Second, there were different degrees of willingness to use idioms in different genres. Overall,

the participants were more conservative towards using idioms in writing than speaking, which suggests that learners in general have a good sense of registers where idioms might occur. Even in writing attempts, the students were cautious and chose to use idioms selectively. Typical examples are given in Table 6.6.

Third, some participants explicitly expressed a very pragmatic attitude towards using idioms, as also reflected in attitudes towards learning idioms (see below). In addition to C3 mentioned above, participant B2 commented that “I would try, but the occasions where I can use them are limited, and therefore, the chances for me to use them are low.” Some participants differentiated their preferences according to different types of idioms, i.e., the imageable vs. the non-imageable ones, and the common vs. the uncommon ones. For example, C4 reported that “I’d like to use the imageable and vivid ones, which also seem effective when used in writing and public speaking. [However] If an expression is very uncommon, then I wouldn’t use it, in case other people cannot understand it.” The last part of this claim seems to resonate with the need to compile the most common idioms for learners to learn, an issue which corpus linguists (e.g., Grant, 2007, Liu, 2003) have endeavoured to address.

In sum, despite limited exposure to idioms (possibly due to insufficient awareness of idioms) and few opportunities to use idioms (in the current EFL learning context), many participants showed a substantial amount of interest in using idioms—provided they have the opportunity to practice and use them. This would provide a rationale for enhancing idiom teaching and learning in an EFL context, especially in usage-based learning (Ellis, 2002).

Q5. Since the first interview, have you noticed any other idioms? If so, do you think you would have paid the same attention to them if we hadn’t had that interview? Will you pay the same attention as before?

The series of questions under Q5 were asked to gauge the potential impact of the teaching experiment on learners’ attitudes towards future L2 idiom learning. Regarding the third question, i.e. whether the learner would pay the same amount of attention to idioms as before, the interview data yielded 14 positive responses, such as “I will pay more attention to idioms”, and 11 neutral responses, such as “Same as before. I will not pay more attention to them”.

Among the positive ones, many participants indicated that they want to learn more idioms, and some expressed that they would like to explore the origin of idioms when they encounter them. For example:

B1—I will spend more time figuring out its possible meaning, and search for the origin that can help create an image in my mind.

B6—I will explore idioms more deeply when I see them. That is, I will find out what they mean and what they used to mean. I will note them down if I see them more often. Such expressions look useful and very interesting.

C4—I will check the origin to help me memorise an expression.

C5—The session has influenced me. I would search for the original meaning of the idioms and use simpler ways to make connections of the meanings.

D4—I will try to find out the origin of idioms and try different mnemonics to learn this kind of expressions, for example, associating with a context and their origin.

While taking these comments positively, we need to be aware of the possible halo effects from the teaching experiment in the first interviews and from the presence of the researcher, who was also the teacher/interviewer for those sessions. Nonetheless, the fact that the interviews were irrelevant to the evaluation of any of their course work and the existence of the considerable number of neutral responses suggested that the participants were fairly honest with their responses. For instance, many of the participants who held a neutral attitude and stated that they would not pay more attention also gave their reasons for this view. For example, B2 commented “I don’t think it [my attitude or attention to idioms] would change much, because the chances of meeting idioms again are low.” Another participant, A4 said she would treat idioms in the same way as how she learned set phrases or collocations.

Overall, the participants gained a substantial amount of knowledge about idioms during the teaching experiment, and the majority of them (14 out of 25) were inspired to learn more about L2 idioms. In sum, the findings about the EFL learners’ attitudes towards and experiences with idioms suggest that, while learners have demonstrated knowledge and awareness of idioms to some degree, there is still much scope for L2 teachers and practitioners to foster idiomatic competence in the EFL context (see also Lontas, 2017 and section 2.3.1 of this thesis).

6.2.4 Case analysis of outstanding individual learners

Based on the recall results, three exemplar participants with the highest scores, i.e., B5, A6, and C1, were selected for case analysis. The purpose of this case analysis is to identify what kinds of learner features—derived from personal learning experiences, strategies and beliefs—contribute to a better learning outcome under the examined teaching technique. Further, case studies of vocabulary learning strategies (VLS) could also “investigate whether specific instruction decisions [...] may lead to internalization and optimization of VLS use” (Cornell, Dean, & Tomas, 2016, p. 843).

Upon inspection of the three learners, some prominent characteristics emerge from the researcher’s observations of their performance during learning, as well as from their self-reflection in the follow-up interviews. Table 6.7 lists some characteristics of these three learners, as assessed by the researcher through the evidence provided in the interviews. Ideally, these characteristics would be measured using acknowledged tests for analytical ability, learning strategies, etc. In the absence of such tests, support for these somewhat subjective evaluations is provided in the discussion following Table 6.7. As shown in Table 6.7, not all three learners are high-achievers according to their TEM-4 scores, but all of them recalled almost all the items correctly. In the following, I will analyse the three learners in a descending order of the TEM-4 scores.

Table 6.7 Characteristics of outstanding learners

ID	TEM-4	Characteristics	Mean GuessBP	Mean GuessAP	Correct Recall ³⁷
B5	87	Avid reader; strong analytical ability; making good use of various kinds of learning strategies, especially key-word method and imagery technique	0.25	0.72	15/16
C1	79	Strong associative ability & cross-cultural awareness	0.2	0.65	18/20
A6	70	Analytic visual learner (‘imager’)	0.06	0.61	16/18

³⁷ Note that the recall counts shown here are based on items that remained after the exclusions explained in Chapter 4. For B5, 3 idioms were excluded, which were known to her prior to the experiment. Similarly, A6 has one known item excluded. If the known items are counted (as were correctly answered in the recall test), then the correct recalls for B5 and A6 would be 18/19 and 17/19.

B5 is an obvious high achiever, with the second highest TEM-4 score (87) among all the participants. During the interpretation phase, B5 exhibited strong analytical ability compared to other participants in the same group. For example, for the idiom *win hands down*, she analysed that the verb *win* should collocate with nouns such as *game* or *competition*, and inferred that *hands down* should be an adverbial that determined the manner of *win* rather than the object of the verb. According to her self-report, she was an avid reader, who read extensively from English novels to world news in English, and kept a habit of reading one novel (unsimplified version) per month. Due to her extensive reading, expressions like idioms were not uncommon for her. She was also a strategic reader, who would read first for the gist and information and then the wording of the language. As mentioned earlier, when looking up an unfamiliar expression in a dictionary, she preferred to read the explanation in English instead of Chinese. She liked to make notes about unfamiliar items and memorise them by remembering a key word and associating the expression with an image.

In comparison to B5, C1 and A6 are not typical high achievers in the sample. However, their self-report and my observations of their learning process show distinct features that potentially align with etymological elaboration. C1 demonstrated a strong associative ability and cross-cultural awareness. For example, when interpreting the idiom *show someone the ropes*, once prompted that the expression came from the domain of seafaring, she quickly associated it with the Age of Discovery in European culture. From her report, idiom comprehension was difficult due to “lack of cultural knowledge” and once she learned the literal underpinning and figurative meaning, she found it easy to remember the idiom. This finding from C1 is also compatible with previous research on language-learning aptitude (Robinson, 2013; Skehan, 2015), in that associative memory ability is one of the predictors of learning success.

A6 had the lowest TEM-4 score among the three exemplar learners (70, which is at the bottom of the ‘good’ range). Yet she correctly recalled the meaning of 16 out of the 18 idioms. According to her self-report at the end of the first interview, she found that it was generally hard to guess the meaning and she was likely to guess it wrong without a prompt of etymology. In the second interview, she reported that the images of the idioms were so vivid that she could remember them very well. These comments suggest that A6 is a typical “imager” (Paivio &

Harshman, 1983), i.e., someone who is more likely to use specific mental pictures in processing metaphors. In other words, she “may have more facility in associating a whole experiential domain with one typical scene, which may then form an interactive image between different domains” (Boers & Littlemore, 2000, p. 180).

In conclusion, the three successful learning cases show that learners with strong analytic and associative abilities and the ability to create mental images are more liable to reap the mnemonic benefits of etymological elaboration.

6.3 Summary and discussion

Chapter 6 examined the factors that are likely to affect L2 idiom meaning retention under the examined learning condition, and explored possible interactions between these factors on the effects on recall.

In response to RQs 6 and 7, both learners’ perceived transparency of idioms and their L2 proficiency are significant predictors of the recall success. Regarding RQ6, if etymological elaboration helps learners’ retention of the meaning of idioms, its effect is not confined to idioms whose meaning learners find to be straightforwardly derived from the proposed origins. Rather, it appears that explanations about origins that are experienced as somewhat far-fetched can also serve this mnemonic purpose. However, the findings also indicate that this mnemonic purpose is generally served more easily in the case of idioms where the learner finds the proposed literal underpinning comparatively straightforward or plausible. More importantly, the interaction effect that emerged from the mixed effects regression model suggests that this influence of the perceived transparency of the literal-figurative connection was the greatest for learners with comparatively low TEM-4 scores, and who could by that proxy be considered the lower-proficiency or less successful EFL learners in the sample. For the higher-proficiency learners, with higher TEM-4 scores, there was no noticeable impact of the transparency variable on idiom recall.

In response to RQ 3, following up the results in Chapter 4, the trial-and-error inferencing based on etymological elaboration had an impact on error rates in retention. The inferential analysis

shows that correct guesses during the interpretation phase are more likely to lead to correct recall in the one-week-delayed post-test. The analysis of errors shows that about 45.5% recall errors could be traced back to the guess(es) before and/or after the prompt, among which some could be related to the presence of an L1 partial or false equivalent. At the same time, the presence of low-frequency unknown key words prior to the learning (e.g., *doldrums* of *in the doldrums*) was found to be a very likely cause to the failure of recall. As learning idioms with unknown component words can involve extra cognitive efforts (Zyzik, 2011), such efforts during the learning phase might not be enough for the learners to establish a well-retained form-and-meaning link, especially for those low-frequency words and under the condition that learners' attention was directed towards meaning rather than form. Although L1 interference and insufficient lexical knowledge, especially the existence of a prior unknown key word, might not be the only obstacles to retaining the idiomatic meaning, the results suggest that these factors might impede L2 idiom meaning retention to some extent. In addition to L1 and prior lexical knowledge, etymological semantic transparency also matters in this respect, because it is likely to influence the accuracy of inferences.

Overall, these findings suggest that when undertaking L2 idiom teaching in a trial-and-error manner, it is important to assist learners to make correct interpretations, and that caution should be taken to avoid incorrect inferencing, especially due to L1 interference. For those idioms with potential misleading partial or false equivalent(s), it is advised that "precedence should be given to implementations that promote retrieval over ones where students are invited to make guesses under the assumption they will remember the correct answers presented as feedback" (Strong & Boers, 2019a, p. 311).

In response to RQs 8 to 10, the second section of the current chapter has explored more general learning factors involved in L2 idiom learning, including learners' experiences and awareness of idioms, their use of strategies in comprehending and memorising unfamiliar L2 idioms, and their attitudes towards learning and using idioms. Qualitative analyses of interview data show that the EFL participants have rather limited exposure to L2 idioms inside and outside the classroom, with insufficient awareness of idioms. When encountering an unfamiliar idiom, the participants predominantly choose to guess the meaning based on the context in which the

expression occurred. Some participants would look the idiom up in a dictionary when their guess(es) did not fit the situational context. In order to memorise L2 idioms, the following strategies were most often used: (1) associating the idiom with an L1 equivalent, (2) associating an idiom with a mental image derived from its origin, or linking it with the learner's personal life experience and/or world knowledge, and (3) associating the idiom with a specific and meaningful context.

Regarding the attitudes towards learning idioms, after the learning experience in the interviews, more than half of the learners showed willingness to learn more about L2 idioms. Many of them said that they would pay more attention to idioms, and some noted that they would search for useful etymological information to assist their future learning of idioms. The participants in general showed a very cautious attitude towards using idioms in their own speech, and an even more conservative attitude towards that in writing. Even so, many participants showed a substantial amount of interest in using idioms, provided they have the opportunity to practice and use them.

Finally, a case analysis of three exemplary learners revealed features of successful learners and their useful learning strategies. The case analysis in combination with the above qualitative analyses of the general factors not only provided useful information about the Chinese EFL learners in their own learning contexts and showcased exemplary learning strategies which could be conducive to L2 idiom learning, but also offered some additional explanations to the previous quantitative findings of this study.

In conclusion, this chapter examined factors affecting L2 idiom meaning retention, i.e., the learning outcome under the examined teaching technique. These factors relate to both the features of idioms and the characteristics of learners, and the factors are intertwined with each other, as reflected in learners' performance and perception during the learning process. This means that the optimal procedure within this CL-informed teaching approach should vary from idiom to idiom, and teaching practices should be adapted to different (types of) idioms, with different learners of different proficiency levels, and more importantly adapted to specific learning contexts.

CHAPTER SEVEN: GENERAL DISCUSSION AND CONCLUSIONS

Chapters 5 and 6 have presented statistical analyses and discussion of the general patterns of the results that had previously been described in Chapter 4. Chapter 5 covered the learners' interpretations of idioms and investigated factors that affect their interpretations. Chapter 6 considered factors relevant to the retention of the meanings of the idioms over the one-week interval of this study. To provide a fuller picture, Chapter 6 also provided qualitative analysis of more general factors involved in L2 idiom learning. The current chapter provides a general discussion of these results in terms of the original research questions that were set out in Chapter 2, and draws several conclusions regarding the use of etymological elaboration in second language pedagogy, as well as pointing to avenues for future research.

7.1 Major findings and pedagogical implications

The overall results of the study reported in this thesis show that etymological elaboration can facilitate the interpretation and meaning retention of L2 idioms to a substantial degree. At the same time, the data also demonstrate that learning idioms in a second language is a complex process, involving the interplay of multiple factors. These factors include analysability/semantic transparency of the idioms, L1 transfer and cross-cultural differences, learners' prior L2 lexical knowledge, and their L2 proficiency. L2 idiom learning is also influenced by other issues, including specific teaching procedures, general learning conditions in the EFL context, and learners' attitudes and strategies. There are undoubtedly additional factors, which were not examined in this study, such as the situational or textual context where an idiom is met.

In the following sections, the major findings of this study are discussed and summarised under the first four themes below, each addressing three or more of the research questions (see section 2.4). Pedagogical implications are discussed accordingly. Finally, other pedagogical aspects are discussed.

- (1) Etymological elaboration as an inferencing prompt (RQs 1, 4, and 5)
- (2) Etymological elaboration as a mnemonic (RQs 2, and 4 to 7)

- (3) Optimising the teaching procedure of etymological elaboration (RQs 3 to 7)
- (4) L2 idiom learning in an EFL context (RQs 1, 2, and 8 to 10)
- (5) Other pedagogical aspects

7.1.1 Etymological elaboration as an inferencing prompt

While providing the information about the literal underpinnings or the origins of the idioms does not guarantee a full understanding of the idioms, such information can often help learners to infer their meaning to some extent (see also Boers et al, 2004b, 2007). The usefulness of etymological elaboration as an inferencing prompt seems to depend more on characteristics of the target idioms, as reflected in the extent of variation in the by-item analysis (Figure 4.2), which was greater than that in the by-participant analysis (Figure 4.1), as well as in the variance measures for items vs. participants as random effects in the mixed-effects model of GuessAP (see section 5.2.2.3). Such characteristics might include the semantic transparency of the figurative meaning, and the extent to which the interpretation of the items depends on the availability of a supportive context, noting that the contextual information provided in the teaching procedure examined in this thesis was rather minimal.

Some common obstacles were found to impede learners' initial interpretation of an idiom, including insufficient prior L2 lexical knowledge and L1 interferences. These obstacles needed to be addressed either before or during the implementation of etymological elaboration. In about 30% of the cases, relating to half of the target idioms, the interpretation was affected by an unknown key content word, although the number of unknown words varies for different participants and for different sets of idioms, and some idioms contain very low frequency words (e.g., *doldrums*). In about 13% of the cases, involving about 15% of the idioms (12/79), lack of familiarity with the relevant meaning of a homonymous or polysemous word in the item was found to prevent successful interpretation. However, after being informed of the meanings of certain unknown, polysemous, or homonymous words, as well as after receiving relevant etymological hints, many learners were able to make inferences about the meaning of the idiom that were at least partially correct. Regarding L1 transfer and cross-cultural differences, about 35% of the learners' responses (166/470) referred to an L1 idiomatic expression which was either a total, a partial, or a false equivalent to the target idiom. As measured by the rating

scores, the existence of total equivalents was indeed helpful for learners' correct interpretation of the L2 idioms, while false equivalents impeded correct interpretations, which tallies with the findings in Hu and Fong (2010). While the partial Chinese equivalents also interfered with the comprehension of L2 idioms, the influence was hard to evaluate. Yet, as Charteris-Black (2002) and Türker (2016) suggested, in a different learning condition, a negative L1 influence could possibly be offset by the availability of a richer situational or textual context.

Because context was kept to a minimum in the experimental procedure used in this study, the effectiveness of etymological elaboration as a prompt for interpretation is closely related to the analysability or semantic transparency of the idioms themselves. This analysability was gauged by learners' perceptions of the etymological semantic transparency of the idioms, and by their success rate in arriving at the correct interpretation with the aid of etymological prompts (i.e., GuessAP). Regarding the etymological semantic transparency, in about 73.5% cases the learners found the connection between the literal underpinnings and the target idiomatic meanings clear. This suggests that the etymological elaboration was overall a useful aid for idiom interpretation. The uneven distribution of the transparency ratings and the disparities between participants also show that some idioms were perceived as more transparent than others by specific learners. This suggests that some idioms lend themselves better to etymological elaboration than others during the interpretation exercise. Still, informing the participants about the origin of the expressions was found to considerably increase the likelihood of correct inferences. In about 50% of the 360 cases where participants had initially failed entirely to propose a correct interpretation, they now produced correct interpretations, and in an additional 19.5% they produced partially correct ones. For the 54 instances where guessing had initially already been partially successful, the information about the origin of the idiom also occasionally (11.1%) helped learners to arrive at a fuller comprehension. In addition, the significant positive relationship between the Transparency ratings and the GuessAP scores suggests that the better the learners interpreted the meaning, the more transparent they perceived the connection between the literal and figurative meanings to be. In relation to these two measures, learners' L2 proficiency was not a significant predictor of either the Transparency ratings or the interpretation success under the examined learning condition,

which aligns with claims from previous studies (e.g., Hu & Fong, 2010). As Johnson (1996) explains, “complexity level in L2 metaphor interpretation appears to be quite independent from proficiency level in L2” (p.227).

While Transparency and GuessAP offer two different ways to interpret the analysability and semantic transparency of idioms from the L2 learner’s perspective, both of them are potentially influenced by L1-L2 similarity. Since GuessAP scores were derived from the learners’ responses in L1 as well as in L2, L1 influence was already an integral part of the correctness of GuessAP. At the same time, the etymological transparency ratings were bound to be affected not only by the features of the idioms but also by the learners’ own prior lexical and cultural knowledge. This provides an explanation for the disparity in transparency ratings between native speakers and L2 learners found by Boers and Webb (2015). In turn, this indicates that native speakers’ perception of idioms might not be a reliable indicator of the analysability of idioms from a learner’s viewpoint and thus not a reliable estimation of the learning difficulty that a given idiom poses for L2 learners. As recommended by Boers and Webb (2015), pedagogy-oriented practitioners and researchers need to be more learner-minded in order to better appreciate the obstacles to L2 idiom comprehension and learning.

7.1.2 Etymological elaboration as a mnemonic

In terms of the mnemonic effects, the overall correct recall rate of more than 60% together with the high retention rate of the etymological information (about 91.25%) provide indirect but positive empirical evidence for etymological elaboration as a mnemonic. The fact that many learners recalled the literal underpinning of the idioms in conjunction with their idiomatic meaning suggests that dual coding (Paivio, 1986) through etymological elaboration facilitated the integration of idiomatic meaning into learners’ memory. In addition, the overall high retention rate could also be explained by Levels of Processing Theory (Craik & Tulving, 1975; Tulving, 1972), as the process of guessing with and without the literal underpinnings engaged learners with intensive semantic and cognitive analysis, i.e., ‘deep’ processing, which is likely to create strong memory traces. On the other hand, the etymological information enabled the learners to appreciate the link between the literal and figurative meanings motivated by

conceptual metaphors (Lakoff & Johnson, 1980). This might generate an elaborate memory trace which in turn facilitated the retrieval. Moreover, the overall correct recall rate (over 60%) of this study was much higher than the success rate (39%) of the experimental group in the first study by Boers et al. (2004b). This is probably because the post-test format used in the latter was different (a gap-fill test), but differences in teaching procedure are also likely to be an explanatory factor. The current study allowed more time for learners to explore the meaning through teacher-learner interaction. According to Craik and Lockhart (1972), retention is affected by factors such as the amount of attention devoted to a stimulus and the processing time available, and deeper processing that involves a greater degree of semantic or cognitive analysis creates memory traces that are more durable.

It was also found that the effectiveness of the implemented teaching technique varies across idioms, which is consistent with what Boers et al. (2007) found, and this must be due to many other factors. One important significant predictor is learners' perceived transparency of idioms. This means that the more transparent a learner perceives the link between the literal and figurative meanings of an idiom to be, the more likely it is that they remember it correctly. In addition to the nature of idioms, what determines the recall success seems to relate more to the learner factors, including L2 proficiency and other learner traits like cognitive abilities. Although L2 proficiency was not a predictor of the inferencing success, it was a significant predictor of the recall success. Higher-proficiency learners tend to obtain better recall results than lower proficiency learners.

First, the findings suggest that if etymological elaboration helps learners' retention of the meaning of idioms, its effect is not confined to idioms whose meaning learners find to be straightforwardly derived from the proposed origins. Rather, it appears that explanations about origins that are experienced as somewhat far-fetched can also serve this mnemonic purpose. This is supported by the finding that about 60% of the cases with relatively low (0-3) transparency ratings were successfully recalled by the learners. However, the findings also indicate that this mnemonic purpose is generally better served in the case of idioms where the learner finds the proposed literal underpinning comparatively straightforward or plausible.

Moreover, according to learners' recall results and their follow-up reflections, if there is an L1 equivalent, they often found it much easier to remember the idioms.

Second, the interaction effect that emerged from the mixed effects regression model suggests that this influence of the perceived transparency of the literal-figurative connection was the greatest for learners with comparatively low proficiency among the EFL learners in the sample. For the higher-proficiency learners, there was no noticeable impact of the transparency variable on idiom recall. This means that the learning outcomes under the examined teaching technique were affected by learners' perceived transparency of the idioms, but the impact is mediated by learners' L2 proficiency level.

Finally, the case analysis of three exemplary learners reveals that learners with strong analytic and associative abilities and the ability to create mental images are more likely to reap the mnemonic benefits of etymological elaboration. This aligns with the Dual Coding Theory and the Conceptual Metaphor Theory, in relation to the assumption that etymological elaboration draws on the benefits of mental images of a concrete scene derived from the literal underpinning and utilises the cognitive abilities behind metaphorical thinking.

7.1.3 Optimising the teaching procedure of etymological elaboration

Regarding fine-tuning and optimising the CL-informed approach, a major question raised by this study is whether it is best to engage learners in an inferencing procedure, where the literal underpinning is given to them as a basis for their inferencing, or whether it is safer to first explain the idiomatic meaning and then use the literal underpinning to make the idiomatic meaning more memorable. In order to compare the two procedures, it would clearly be ideal to run a future study that includes a between-group design. The data obtained in the current study allowed an early exploration of this question through the investigation of the effects of the correctness of inferencing on the recall success, and by examining the potential impact of the inferencing errors on the learners' retention of the correct idiomatic meaning.

On the one hand, it was found that correct interpretations of the idiomatic meaning are more likely to lead to correct recall one week later; on the other hand, about 45.5% recall errors could be traced back to the unsuccessful guesses before and/or after the prompt. Among the recall

errors, many could be traced back to the presence of an L1 partial or false equivalent. Moreover, the presence of low-frequency unknown key words prior to the learning phase was found to be a likely cause of the recall failures. Although L1 interference and insufficient lexical knowledge might not be the only obstacles to retaining the idiomatic meaning, the results suggest that these factors might have some negative impact on L2 idiom meaning retention. In addition, etymological semantic transparency also matters in this respect, because it is likely to influence the accuracy of inferences, as suggested by the significant positive relationship between the Transparency ratings and the GuessAP scores.

In general, these findings suggest that when undertaking L2 idiom teaching in a trial-and-error manner, it is important to assist learners to make correct interpretations, and that caution should be taken to avoid incorrect inferencing, especially due to L1 interference. For those idioms with potentially misleading partial or false equivalent(s), it is advised that instructions that promote retrieval should take precedence over those that invite learners to make guesses by assuming that they will remember the correct information provided in confirmative/corrective feedback (Strong & Boers, 2019a). Etymological elaboration can be used as part of the direct explanation of the idioms, in combination with an elaborate situational context that supports the idiomatic meaning. For those idioms that have an L1 true equivalent or those that are relatively transparent, etymological elaboration could potentially serve well as an inference prompt, as it should engage learners in deeper processing, while at the same time reducing the risk of inferencing errors.

Moreover, the by-idiom analysis of the study also tells us that under the influence of the above factors, the teaching of some idioms needs extra attention and information than would be the case for others. Providing more elaborate contextual support (for idioms whose literal sense is easy to understand but the idiomatic meaning is obscure, e.g., *pull the rabbit out of the hat*, *beat around the bush*), or a pictorial illustration or even a video depicting the original meaning (for idioms whose literal sense involves a motion or multiple key points, such as *hand over first*, *take the plunge*) could potentially facilitate learning.

Overall, the influential factors involved in L2 idiom learning and teaching relate to both the features of idioms and the characteristics of learners. Inevitably, these factors are intertwined

during the learning process. This means that the best procedure within a specific teaching approach should vary from idiom to idiom, and teaching practices should be adapted to different idioms as well as to different learners.

While most of the participants in this study gained knowledge of a considerable number of new idioms through an instructional procedure which engaged them with the literal underpinnings of the expressions, we need to be cautious not to oversell the proclaimed benefits of this etymological elaboration approach. As mentioned in the literature review, not all evidence to date has been favourable of its implementation (e.g., Szczepaniak & Lew, 2011). We also need to bear in mind that the present study did not include a comparison treatment. As such, it provides no evidence that learning idioms through etymological elaboration should be given precedence over other learning procedures. The jury is still out, so to speak. What we can say with a degree of confidence is that the learning gains attested here were quite substantial, at least according to a one-week delayed post-test.

Furthermore, etymological elaboration is just one of many teaching techniques informed by Cognitive Linguistics (Boers & Lindstromberg, 2009), along with many other approaches to L2 idiom learning (as discussed in section 2.3.1.2; see also Liu, 2008). It should also be borne in mind that the idioms targeted in this study are just one particular but major type of idiom, i.e., conventionalised figurative expressions which have a traceable etymology (recall that there are other types of idioms, e.g., ‘core idioms’ (Grant & Bauer, 2004), which do not have or have lost their original literal senses). In classroom teaching, various activities, tasks and exercises should be considered to suit the different needs for idiom learning, as learners appear to benefit most from instruction that incorporates a wide range of treatments that are suitable for learners of different levels of proficiency (see previous discussion of the general teaching principles in section 2.3.1.2) and for different types of idioms (Grant, 2003; Simpson & Mendis, 2003), and that foster various types of intelligence and cognitive ability (Cooper, 1998; Gardner, 1993).

7.1.4 L2 idiom learning in an EFL context

The exploration of general learning factors has informed us about the actual situation of L2 idiom learning in the Chinese EFL context, including learners’ experiences and awareness of

idioms, their use of strategies in comprehending and memorising unfamiliar L2 idioms, and their attitudes towards learning and using idioms.

The findings from the follow-up interviews show that the EFL participants have rather limited exposure to L2 idioms inside and outside the classroom, and have insufficient awareness of idioms. When encountering an unfamiliar idiom, the participants predominantly chose to guess the meaning based on the context in which the expression occurs. Some participants would look the idiom up in a dictionary when their inferences did not fit the situational context. In order to memorise L2 idioms, the following strategies were most often used: (1) associating the idiom with an L1 equivalent; (2) associating an idiom with a mental image derived from its origin, or linking it with the learner's personal life experience and/or world knowledge; and (3) associating the idiom with a specific and meaningful context. Regarding the attitudes towards learning idioms, during the interviews that followed the teaching experiment more than half of the learners showed willingness to learn more about L2 idioms. Many of them said they would pay more attention to idioms, and some expressed an inclination to search for useful etymological information to assist their future learning of idioms. When it comes to using idioms, the participants in general showed a very cautious attitude towards using them in their own speech, and an even more conservative attitude towards using them in writing. Even so, many participants showed a substantial amount of interest in using idioms, provided they have the opportunity to practice and use them.

From a pragmatic point of view, these findings suggest that a pedagogical approach to L2 idioms in an EFL context should focus more on receptive knowledge (i.e., meaning) than productive knowledge (i.e., form and use), because EFL learners will probably (and perhaps wisely) shy away from using L2 idioms themselves, unless they are very familiar with the idioms. For more advanced L2 learners and those who are learning the target language in an L2-speaking country, it might be more appropriate to promote the use of idioms and to include in classroom teaching tasks that relate to productive knowledge.

However, to take a long-term view, fostering receptive knowledge facilitates the familiarisation of L2 idioms in general and promotes processing of formulaic language in L2, which eventually may lead to more fluent and idiomatic language use. For long-term L2 learning, much could

be done by L2 teachers and course/textbook/test designers to facilitate the learning of multiword units such as idioms. For example, introducing idioms into the L2 classroom and raising learners' awareness of idioms; increasing idioms in L2 input, e.g., including idioms in textbooks, tasks, and language tests; and fostering learner autonomy through practising learning strategies (as informed by exemplary learners in this study). Idioms can also serve as a medium for introducing cross-cultural differences and raising cross-cultural awareness (Liontas, 2017). Teaching idioms in classroom will not be sufficient for second language learning, but explicit teaching can at least help learners notice and identify idioms as well as other multiword expressions. Drawing attention to idioms and encouraging learners to use etymological elaboration and other strategies to learn idioms can foster learner autonomy and lead them to pick up idiom themselves. This hopefully will help them build up confidence in understanding more authentic communications and eventually, using idioms in their own speaking and writing. These points are also reflected in Liontas' (2017) appeal for fostering idiomatic competence. In terms of L2 idiom processing, as suggested by previous studies, increasing the input and awareness of idioms should increase the possibility of the formation of 'idiom entries' (Abel, 2003) or idiomatic 'configuration' (Cacciari & Glucksberg, 1991) in L2 mental lexicon. If informing learners of information about the literal underpinnings or etymology of idioms (plus providing supportive situational contexts) could promote the internalisation of idioms, and this in turn increases learners' familiarity with idioms in general, then in the long run, it might promote a positive cycle of acquisition and bridge the gap between L1 and L2 processing.

7.1.5 Other pedagogical aspects

In the development of learning materials and in the design of textbooks, multiword expressions like idioms should be included. When semantic transparency is used as a criterion for item selection, it is important to understand L2 learners' perception of idioms so as to estimate L2 learners' difficulties in learning. As Boers and Webb (2015) advised, native-speaking teachers and textbook designers need to take into consideration factors learners bring to bear in tackling the meaning of MWEs like idioms, i.e., learners' L1 and cultural background as well as their L2 lexical knowledge.

For language testing and assessment, idioms (especially high-frequency ones) are good candidates for proficiency tests as well as for tests of metaphorical competence, since (as previously discussed in Chapter 2) the ability to understand and use idioms is one distinguishing indicator of a native-like command of a language and a reliable measure of the proficiency of foreign learners.

For the construction of idiom dictionaries, etymological notes should be considered as an important part of the explanations provided for entries, as attested both by the positive evidence of the teaching experiment and by feedback from the follow-up interviews. This may help learners grasp a better understanding of idioms and facilitate their learning. However, during the process of selecting items for this study, I found some marked differences between explanations offered by the different dictionaries I consulted. On the one hand, such different explanations suggest that an idiom may entail various focal aspects of meaning. On the other hand, different dictionaries have their own strengths and weaknesses (see more in 骆世平, 2006). Thus, it is recommended that learners make use of different kinds of dictionaries in order to understand idiomatic meanings more accurately. Learners' habits in dictionary use and their performance in the teaching experiment also suggest that learners who prefer to use bilingual or English-English dictionaries often have a better and more thorough understanding of L2 idioms. This suggests that good use of dictionaries is conducive to L2 idiom learning.

7.2 Significance of the study and methodological contributions

In addition to its empirical results, a major contribution of this thesis is its thorough review of several major issues regarding second language idiom learning: from the definition, classification, and characterisations of idioms, to the differences between L1 and L2 idiom processing, and to the status quo and various approaches to L2 idiom teaching. In this respect, this study clarified some complex issues in this area. It also clarified different measures of semantic transparency and looked into multiple factors involved in L2 idiom learning.

Second, research on second language pedagogy has so far focused on the overall effects of different teaching approaches. This study offered a different perspective and zoomed in on a particular approach to learning. By employing a think-aloud procedure, inspired by Cooper

(1999), this study invited participants to report what they were immediately aware of and thinking about during the teaching experiment. This helps us tap into learners' ongoing thinking process, in order not only to examine the process of L2 learning, but also to gauge the effectiveness of a particular teaching practice.

Third, the mixed design made use of both quantitative and qualitative methods in the analysis of the data, which permits triangulation in the interpretation of the same phenomena. The use of mixed effects modelling in the quantitative analysis also provides a robust analysis that considers individual variation among learners and items. The qualitative analysis of items and individual participants provides detailed examples and explanations for a phenomenon that is hard to capture by quantitative data alone. For future intervention studies on multiword units like idioms, item analysis and individual differences could be considered in the analysis of treatment effects in a similar manner. Overall, the method employed in this study can inform further studies on the design, materials development, and evaluation of learner responses regarding the meaning of multiword units, as well as data analysis.

7.3 Limitations of the study and suggestions for future research

While the research and the adopted methodology have many advantages as identified above, there are also a number of limitations. In addition to the methodological limitations discussed earlier (particularly the intricacies of rating the responses, discussed in section 3.5.2.4), the following paragraphs summarise the main limitations of the study and provide some suggestions for future research.

First, the number of participants was small. With the aim of refining a particular teaching approach, the total number of the target idioms was kept relatively large, but this necessitated a compromise in the number of participants, with only 6 or 7 per group. While the generalised mixed-effects modelling allowed for between-subject and between-item variation, a larger sample size would improve the statistical power of the inferential analysis.

Second, only short-term effects on meaning retention were explored. Due to time limits, I employed a one-week delayed post-test to investigate the mnemonic benefits of the teaching

technique. It would nonetheless be meaningful to explore via a longitudinal investigation whether the intervention I have described here stimulates learners' long-term engagement with L2 idioms. For example, by implementing it in course work in one semester or trimester, and by conducting a delayed post-test after a longer interval.

Third, the present study did not include a comparison treatment between etymological elaboration and other teaching approaches. It would be better to include a control group as a baseline for the treatment effects. There was also no comparison group for the effects of the teaching procedure regarding trial-and-error and errorless learning. Further investigation is needed to verify the effects of the sequence of the teaching procedure, that is, whether it is better to ask learners to hazard a guess about the idiomatic meaning after informing them of the etymology of idioms, or if it is safer to provide the idiomatic meaning upfront and then offer the etymological elaboration as a supplement to make this meaning more memorable.

Fourth, regarding the rating of learners' responses, it is possible that idiom knowledge was underestimated in my experiment. That is because we only gave full credit to learners' interpretations of the idioms if these interpretations included all the meaning components mentioned in the dictionary entry that was used as a benchmark. It cannot be ruled out that one or the other meaning component was implied rather than explicitly verbalised in a respondent's paraphrase. Even in an interview procedure that makes use of the respondent's L1, an interviewer's efforts to solicit nuanced knowledge may fail. In future (conceptual) replications, it might be worth adding a different type of test, such as a multiple-choice meaning recognition test where the respondent is required to select the meaning paraphrase that best captures the idiom's meaning.

Finally, it should be noted that the learning aim in the present experiment may in one respect be considered modest—that is, remembering the *meaning* of the idioms. For one thing, the post-test presented the learners with the English idioms again, and thus did not require them to recall the lexical makeup of the expressions. For another, it is well documented that idioms serve subtle pragmatic and evaluative functions in discourse (e.g., Moon, 1998; O'Keeffe, McCarthy & Carter, 2007) which are often not captured by dictionary definitions. Developing

expert productive knowledge of idioms would take extensive exposure to L2 natural discourse (MacArthur, 2010).

7.4 Concluding remarks

In conclusion, this study investigated factors that influence the effectiveness of etymological elaboration as a CL-informed teaching approach. Identifying these factors is a necessary step toward optimising this teaching technique. By delving into the learning process, the study offered a detailed account of the performance and perceptions of L2 learners regarding idiom learning in an EFL context. It is my hope that it will make a useful contribution to our understanding of second language idiom learning and to the development of instructional materials that help L2 learners come to grips with this elusive dimension of language.

REFERENCES

- A New Century Chinese-English Dictionary 新世纪汉英大词典*. (2004). Foreign Language Teaching and Research Publishing.
- Abel, B. (2003). English idioms in the first language and second language lexicon: A dual representation approach. *Second Language Research*, 19(4), 329-358.
- Achard, M., & Niemeier, S. (Eds.). (2004). *Cognitive linguistics, second language acquisition, and foreign language teaching*. Mouton de Gruyter.
- Alexander, R. J. (1987). Problems in understanding and teaching idiomaticity in English. *Anglistik und Englischunterricht*, 32(2), 105-122.
- Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, 62(1), 67-82.
- Ayto, J. (Ed). (2009). *The Oxford Dictionary of English Idioms* (3rd ed.). Oxford University Press.
- Baayen, H., & Lieber, R. (1996). Word frequency distributions and lexical semantics. *Computers and the Humanities*, 30(4), 281-291.
- Bachman, L. F. (1990). *Fundamental Considerations in Language Testing*. Oxford University Press.
- Bagheri, M. S., & Fazel, I. (2010). Effects of Etymological Elaboration on the EFL Learners' Comprehension and Retention of Idioms. *Journal of Pan-Pacific Association of Applied Linguistics*, 14(1), 45-55.
- Barcroft, J. (2002). Semantic and structural elaboration in L2 Lexical Acquisition. *Language Learning*, 52(2), 323-363.
- Beréndi, M., Csábi, S., & Kövecses, Z. (2008). Using Conceptual Metaphors and Metonymies in Vocabulary Teaching. In F. Boers & S. Lindstromberg (Eds.), *Cognitive Linguistic approaches to teaching vocabulary and phraseology* (pp. 65-99). Mouton de Gruyter.
- Bobrow, S. A., & Bell, S. M. (1973). On catching on to idiomatic expressions. *Memory & Cognition*, 1(3), 343-346.

- Boers, F. (2000). Metaphor awareness and vocabulary retention. *Applied Linguistics*, 21(4), 553-571.
- Boers, F. (2001). Remembering figurative idioms by hypothesising about their origin. *Prospect*, 16(3), 35-43.
- Boers, F. (2011). Cognitive Semantic ways of teaching figurative phrases: An assessment. *Review of Cognitive Linguistics*, 9(1), 227-261.
- Boers, F. (2013). Cognitive Linguistic approaches to teaching vocabulary: Assessment and integration. *Language Teaching*, 46(2), 208-224.
- Boers, F., & Demecheleer, M. (1998). A cognitive semantic approach to teaching prepositions. *ELT Journal*, 52(3), 197-204.
- Boers, F., & Demecheleer, M. (2001). Measuring the impact of cross-cultural differences on learners' comprehension of imageable idioms. *ELT Journal*, 55(3), 255-262.
- Boers, F., Demecheleer, M., & Eyckmans, J. (2004a). Cross-cultural variation as a variable in comprehending and remembering figurative idioms. *European Journal of English Studies*, 8(3), 375-388.
- Boers, F., Demecheleer, M., & Eyckmans, J. (2004b). Etymological elaboration as a strategy for learning idioms. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition and testing* (pp. 53-78). John Benjamins.
- Boer, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: putting a Lexical Approach to the test. *Language Teaching Research*, 10(3), 245-261.
- Boers, F., Eyckmans, J. & Stengers, H. (2006). Motivating multiword units: Rationale, mnemonic benefits, and cognitive style variables. In S. H. Foster-Cohen, M. M. Krajnovic & J. M. Djigunovic (eds.), *EUROSLA Yearbook*, Vol. 6 (pp. 169-190). John Benjamins.
- Boers, F., Eyckmans, J., & Stengers, H. (2007). Presenting figurative idioms with a touch of etymology: more than mere mnemonics? *Language Teaching Research*, 11(1), 43-62.
- Boers, F., & Littlemore, J. (2000) Cognitive style variables in participants' explanations of conceptual metaphors. *Metaphor and Symbol*, 15(3), 177-187.

- Boers, F. & Littlemore, J. (Eds). (2003). Cross-cultural differences in conceptual metaphor: Applied linguistics perspectives [Special issue]. *Metaphor and Symbol*, 18(4).
- Boers, F., & Lindstromberg, S. (2005). Finding ways to make phrase-learning feasible: the mnemonic effect of alliteration. *System*, 33, 225-238.
- Boers, F. & Lindstromberg, S. (2006). Cognitive Linguistic applications in second or foreign language instruction: Rationale, proposals, and evaluation. In Kristiansen, Gitte; Achard, Michel; Dirven, René; Ruiz de Mendoza, Francisco (Eds.), *Cognitive Linguistic: Current Applications, Future Orientations*. Mouton de Gruyter.
- Boers, F., & Lindstromberg, S. (Eds.). (2008a). *Cognitive linguistic approaches to teaching vocabulary and phraseology*. Mouton de Gruyter.
- Boers, F., & Lindstromberg, S. (2008b). From empirical findings to pedagogical practice. In F. Boers, & S. Lindstromberg (Eds). *Cognitive Linguistic Approaches to Teaching Vocabulary and Phraseology* (pp. 375-393). Mouton de Gruyter.
- Boers, F., & Lindstromberg, S. (2009) *Optimizing a Lexical Approach to Instructed Second Language Acquisition*. Palgrave Macmillan.
- Boers, F., & Lindstromberg, S. (2012). Experimental and Intervention Studies on Formulaic Sequences in a Second Language. *Annual Review of Applied Linguistics*, 32, 83-110.
- Boers, F., Lindstromberg, S., Littlemore, J., Stengers, H., & Eyckmans, J. (2008). Variables in the mnemonic effectiveness of pictorial elucidation. In F. Boers, & S. Lindstromberg (Eds.), *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 189-116). Mouton de Gruyter.
- Boers, F., Píriz, A. M. P., Stengers, H., & Eyckmans, J. (2009). Does pictorial elucidation foster recollection of idioms? *Language Teaching Research*, 13(4), 367-382.
- Boers, F., & Webb, S. (2015). Gauging the semantic transparency of idioms: Do natives and learners see eye to eye. In R. Heredia & A. Cieřlicka (Eds.), *Bilingual figurative language processing* (pp. 368-392). Cambridge University Press.
- Bortfeld, H. (2003). Comprehending idioms cross-linguistically. *Experimental Psychology*, 50(3), 217-230.

- Brown, C. M. (1996). Foreign language reading versus understanding: think-aloud protocols to identify L2 reading problems. *Dimension 96: Global access through languages* (pp. 20-32). Valdosta, Conference on Language Teaching.
- Brunfaut, T., & Révész, A. (2015). The Role of Task and Listener Characteristics in Second Language Listening. *TESOL Quarterly*, 49(1), 141-168.
- Burdelski, M., & Cook, H. M. (2012). Formulaic Language in Language Socialization. *Annual Review of Applied Linguistics*, 32, 173-188.
- Burnham, K., & Anderson, D. (2004). *Model Selection and Inference: A Practical Information-Theoretic Approach*. Springer-Verlag.
- Cacciari, C., & Glucksberg, S. (1991). Understanding idiomatic expressions: The contribution of word meanings. In G. B. Simpson (Ed.), *Understanding word and sentence* (pp. 217-240). Elsevier.
- Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language*, 27(6), 668-683.
- Caillies, S., & Butcher, K. (2007). Processing of idiomatic expressions: Evidence for a new hybrid view. *Metaphor and Symbol*, 22(1), 79-108.
- Cambridge Dictionary of American Idioms*. (2003). Cambridge University Press.
- 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.
- Carrol, G., Littlemore, J., & Dowens, M. G. (2018). Of false friends and familiar foes: Comparing native and non-native understanding of figurative phrases. *Lingua*, 204, 21-44.
- Chafe, W. (1968). Idiomaticity as an anomaly in the Chomskyan Paradigm. *Foundations of Language*, 4(2), 109-125.
- Charteris-Black, J. (2002). Second language figurative proficiency: A comparative study of Malay and English. *Applied Linguistics*, 23(1), 104-133.
- Chen, Y., & Lai, H. (2013). Teaching English idioms as metaphors through cognitive-oriented methods: A Case in an EFL writing class. *English Language Teaching*, 6(6), 13-20.

- Cho, K. (2010). Fostering the acquisition of English prepositions by Japanese learners with networks and prototypes. In S. Knop, F. Boers & A. Rycker (Eds.), *Fostering language teaching efficiency through Cognitive Linguistics* (pp. 259-276). Mouton de Gruyter.
- Christensen, R. H. B. (2018). Regression Models for Ordinal Data. R package version 2018.4-19. <http://www.cran.r-project.org/package=ordinal/>.
- Cieślicka, A. (2006). Literal salience in on-line processing of idiomatic expressions by second language learners. *Second Language Research*, 22(2), 115-144.
- Cieślicka, A. (2011). Suppression of literal meanings in L2 idiom processing: Does context help? *Studies in Second Language Learning and Teaching*, 1(1), 13-36.
- Cieślicka, A. (2015). Idiom acquisition and processing by second/foreign language learners. In R. Heredia & A. Cieślicka (Eds.), *Bilingual Figurative Language Processing* (pp. 208-244). Cambridge University Press.
- Cieślicka, A., & Heredia, R. (2011). Hemispheric asymmetries in processing L1 and L2 idioms: Effects of salience and context. *Brain and Language*, 116(3), 136-150.
- Collins COBUILD Dictionary of Idioms* (2nd ed.). (1999). London: HarperCollins Publishers.
- Collins COBUILD Idioms Dictionary* (3rd ed.). (2012). London: HarperCollins Publishers.
- Condon, N. (2008). How Cognitive Linguistic motivations influence the learning of phrasal verbs. In F. Boers & S. Lindstromberg (Eds.), *Cognitive Linguistic approaches to teaching vocabulary and phraseology* (pp. 133-158). Mouton de Gruyter.
- Cooper, T. C. (1998). Teaching idioms. *Foreign Language Annals*, 31(2), 255-266.
- Cooper, T. C. (1999). Processing of Idioms by L2 Learners of English. *TESOL Quarterly*, 33(2), 233-262.
- Cornell, R., Dean, J., & Tomas, Z. (2016). Up Close and Personal: A Case Study of Three University-Level Second Language Learners' Vocabulary Learning Experiences. *TESOL Journal*, 7(4), 823-846.
- Correli, L. (2006). Jazz up your English with fresh and lively idioms. *ESL Teachers Board*, 6(1), 30-32.
- Coryell, J. L. (2012). *The effectiveness of etymological elaboration as a method of teaching idioms to adolescent English language learners* (Unpublished doctoral dissertation). State University of New York, US.

- Cowie, A. P., & Mackin, R. (1975). *Oxford Dictionary of Current Idiomatic English: Verbs with Prepositions and Particles (Vol. 1)*. Oxford University Press.
- Cowie, A. P., Mackin, R., & McCaig, I. R. (1983). *Oxford Dictionary of Current Idiomatic English: Phrase, Clause and Sentence Idioms (Vol. 2)*. Oxford University Press.
- Cermak, L., & Craik, F. (Eds.). (1979). *Levels of processing in human memory*. Lawrence Erlbaum.
- Craik, F., & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671-684.
- Craik, F., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104(3), 268-294.
- Cronk, B. C., Lima, S., & Schweigert, W. (1993). Idioms in sentences: Effects of frequency, literalness, and familiarity. *Journal of Psycholinguistic Research*, 22(1), 59-82.
- Cronk, B. C., & Schweigert, W. A. (1992). The comprehension of idioms: The effects of familiarity, literalness, and usage. *Applied Psycholinguistics*, 13(2), 131-146.
- 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.
- Dąbrowska, A. (2018). *A Syntactic Study of Idioms: Psychological States in English and Their Constraints*. Cambridge Scholars Publishing.
- Davies, M. (2008-). The Corpus of Contemporary American English (COCA): 520 million words, 1990-2016. Available online at <https://www.english-corpora.org/coca/>.
- De Groot, A. M. B. (1992). Bilingual lexical representation: a closer look at conceptual representations. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp.389-412). North-Holland.
- De Knop, S., Boers, F., & De Rycker, T. (Eds.). (2010). *Fostering language teaching efficiency through cognitive linguistics*. Mouton de Gruyter.
- De Knop, S., & De Rycker, T. (Eds.). (2008). *Cognitive approaches to pedagogical grammar*. Mouton de Gruyter.
- Deignan, A., Gabrys, D., & Solska, A. (1997). Teaching English metaphors using cross-linguistic awareness raising activities. *ELT Journal*, 51 (4), 352-360.

- Elgort, I., & Warren, P. (2014). L2 vocabulary learning from reading: Explicit and tacit lexical knowledge and the role of learner and item variables. *Language Learning*, 64(2), 365-414.
- Ellis, N. C. (2002) Frequency effects in language processing. *Studies in Second Language Acquisition*, 24(2), 143-188.
- Ellis, N. C. (2012). Formulaic Language and Second Language Acquisition: Zipf and the Phrasal Teddy Bear. *Annual Review of Applied Linguistics*, 32, 17-44.
- Fernando, C. (1996). *Idioms and Idiomaticity*. Oxford University Press.
- Fernando, C., & Flavell, R. (1981). *On Idiom: Critical Views and Perspectives*. University of Exeter.
- Flores d'Arcais, G. B. (1993). The comprehension and semantic interpretation idioms. In C. Cacciari & P. Tabossi (Eds.), *Idioms: Processing, structure, and interpretation* (pp. 79-98). Lawrence Erlbaum.
- Follow the crowd. (1997). *The American Heritage Dictionary of Idioms*. Houghton Mifflin Harcourt.
- Fraser, B. (1970). Idioms within a transformational grammar. *Foundations of Language*, 6(1), 22-42.
- Fraser, B. (1974). *The Verb-particle Combination English*. Taishukan.
- Gao, L. Q., & Meng, G. H. (2010). A study on the effect of metaphor awareness raising on Chinese EFL learners' vocabulary acquisition and retention. *Canadian Social Science*, 6(2), 110-124.
- Gardner, H. (1993). *Frames of Mind: The Theory of Multiple Intelligences*. Basic Books.
- Gibbs, R. W. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory & Cognition*, 8(2), 149-156.
- Gibbs, R. W. (1983). Do people always process the literal meanings of indirect requests? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9(3), 524-533.
- Gibbs, R. W. (1984). Literal meaning and psychological theory. *Cognitive science*, 8(3), 275-304.
- Gibbs, R. W. (1985). On the process of understanding idioms. *Journal of Psycholinguistic Research*, 14(5), 465-472.

- Gibbs, R. W. (1986). Skating on thin ice: Literal meaning and understanding idioms in conversation. *Discourse Processes*, 9(1), 17-30.
- Gibbs, R. W. (1987). Linguistic factors in children's understanding of idioms. *Journal of Child Language*, 14(3), 569-586.
- Gibbs, R. W. (1990). Psycholinguistic studies on the conceptual basis of idiomaticity. *Cognitive Linguistics*, 1(4), 417-451.
- Gibbs, R. W. (1991). Semantic Analyzability in Children's Understanding of Idioms. *Journal of Speech and Hearing Research*, 34(3), 613-620.
- Gibbs, R. W. (1994). *The Poetics of Mind: Figurative Thought, Language, and Understanding*. Cambridge University Press.
- Gibbs, R. W. (1997). Metaphor in idiom comprehension. *Journal of Memory and Language*, 37, 141-154.
- Gibbs, R. W. (2002). A new look at literal meaning in understanding what is said and implicated. *Journal of Pragmatics*, 34(4), 457-486.
- Gibbs, R.W. (2017). *Metaphor wars: Conceptual metaphors in human life*. Cambridge University Press.
- Gibbs, R. W., Bogdanovich, J. M., Sykes, J. R., & Barr, D. J. (1997). Metaphor in idiom comprehension. *Journal of Memory and Language*, 37, 141-154.
- Gibbs, R.W., & Nayak, N.P. (1989). Psycholinguistic studies on the syntactic behavior of idioms. *Cognitive Psychology*, 21(1), 100-138.
- 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.
- Gibbs, R. W., & O'Brien, J. E. (1990). Idioms and mental imagery: The metaphorical motivation for idiomatic meaning. *Cognition*, 36(1), 35-68.
- Gibbs, R. W., & Steen, G. J. (Eds) (1999). *Metaphor in Cognitive Linguistics: Selected papers from the 5th International Cognitive Linguistics Conference*, Amsterdam, 1997. John Benjamins.
- Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, 8(3), 183-206.

- Giora, R. (2002). Literal vs. figurative language: Different or equal? *Journal of Pragmatics*, 34(4), 487-506.
- Giora, R. (2003). *On our mind: Salience, context, and figurative language*. Oxford University Press.
- Giora, R., & Fein, O. (1999). On understanding familiar and less-familiar figurative language. *Journal of Pragmatics*, 31(12), 1601-1618.
- Give someone a leg up. (1999). In *Collins COBUILD Dictionary of Idioms* (2nd ed.). HarperCollins Publishers.
- Glucksberg, S. (1989). Metaphors in conversation: How are they understood? Why are they used? *Metaphor and Symbol*, 4(3), 125-143.
- Glucksberg, S. (1993). Idiom meanings and allusional content. In C. Cacciari & P. Tabossi (Eds.), *Idioms: Processing, structure, and interpretation* (pp. 3-26). Lawrence Erlbaum.
- Glucksberg, S., Brown, M., & McGlone, M. S. (1993). Conceptual metaphors are not automatically accessed during idiom comprehension. *Memory & Cognition*, 21(5), 711-715.
- Glucksberg, S., Keysar, B., & McGlone, M. S. (1992). Metaphor Understanding and Accessing Conceptual Schema: Reply to Gibbs (1992). *Psychological Review*, 99(3), 578-581.
- Grant, L. (2003). *A Corpus-Based Investigation of Idiomatic Multiword Units* (Unpublished doctoral dissertation). Victoria University of Wellington, New Zealand.
<http://hdl.handle.net/100 63/327>
- Grant, L. E. (2007). In a manner of speaking: Assessing frequent spoken figurative idioms to assist ESL/EFL teachers. *System*, 35(2), 169-181.
- Grant, L., & Bauer, L. (2004). Criteria for redefining idioms: Are we barking up the wrong tree? *Applied Linguistics*, 25(1), 38-61.
- Grant, L., & Nation, P. (2006). How many idioms are there in English? *International Journal of Applied Linguistics*, 151, 1-14.
- Gu, Y. (2003). Fine brush and freehand: The vocabulary learning art of two successful Chinese EFL learners. *TESOL Quarterly*, 37(1), 73-104.

- Gu, Y. (2013). Vocabulary learning strategies. In Carole A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Wiley-Blackwell.
- Guo, S. (2007). Is idiom comprehension influenced by metaphor awareness of the learners? A case study of Chinese EFL learners. *Linguistics Journal*, 3(3), 148-166.
- Guo, S. (2008). Differential Effects of Etymological Elaboration and Rote Memorization on Idiom Acquisition in College EFL Learners. *CELEA Journal*, 31(6), 52-62.
- Gyllstad, H., & Wolter, B. (2016). Collocational processing in light of the phraseological continuum model: Does semantic transparency matter? *Language Learning*, 66(2), 296-323.
- Halliday, M. A. K. (1973). *Explorations in the Functions of Language*. Edward Arnold.
- Halliday, M. A. K. (1985). *An Introduction to Functional Grammar*. Edward Arnold.
- Herdina, P., & Jessner, U. (2002). *A Dynamic Model of Multilingualism: Perspectives of change in psycholinguistics*. Multilingual Matters.
- Hockett, C. F. (1958). *A Course in Modern Linguistics*. Macmillan.
- Howarth, P. (1996). *Phraseology in English academic writing*. Max Niemayer Verlag.
- Hu, Y.-H., & Fong, Y.-Y. (2010). Obstacles to CM-guided L2 idiom interpretation. In S. D. Knop, F. Boers, & A. D. Rycker (Eds.), *Fostering Language Teaching Efficiency through Cognitive Linguistics* (pp. 293-313). Mouton de Gruyter.
- Hunt, K. W. (1970). Syntactic maturity in schoolchildren and adults. *Monographs of the Society for Research in Child Development*, 35(1), 1-61
- Hymes, D. (1971). *On Communicative Competence*. University of Pennsylvania Press.
- Idiomaticity. (n.d.) In *MacMillan Dictionary* (British English version). Retrieved April 5, 2019, from <https://www.macmillandictionary.com/dictionary/british/idiomaticity>
- Irujo, S. (1986a). Don't put your leg in your mouth: Transfer in the acquisition of idioms in a second language. *TESOL Quarterly*, 20(2), 287-304.
- Irujo, S. (1986b). A piece of cake: Learning and teaching idioms. *English Language Teaching*, 40(3), 236-242.
- Irujo, S. (1993). Steering clear: Avoidance in the production of idioms. *International Review of Applied Linguistics in Language Teaching*, 31(3), 205-219.

- Johnson, M. (1987). *The body in the mind: the bodily basis of meaning, imagination, and reason*. The University of Chicago Press.
- Johnson, J. (1989). Factors related to cross-language transfer and metaphor interpretation in bilingual children. *Applied Psycholinguistics*, 10(2), 157-177.
- Johnson, J. (1996). Metaphor interpretations by second language learners: children and adults. *The Canadian Modern Language Review*, 53 (1), 219-241.
- Johnson, J., & Rosano, T. (1993). Relation of cognitive style to metaphor interpretation and second language proficiency. *Applied Psycholinguistics*, 14(2), 159-175.
- Karlsson, M. (2013). Quantitative and qualitative aspects of L1 (Swedish) and L2 (English) idiom comprehension. *Studies in Second Language Learning and Teaching*, 3(2), 293-319.
- Katz, J., & Postal, P. (1963). Semantic interpretation of idioms and sentences containing them. *Quarterly Progress Report of the MIT Research Laboratory of Electronics*, 70, 275-282.
- Kellerman, E. (1977). Towards a characterisation of the strategy of transfer in second language learning. *Interlanguage Studies Bulletin*, 2(1), 58-145.
- Kellerman, E. (1979). Transfer and non-transfer: Where we are now. *Studies in Second Language Acquisition*, 2(1), 37-57.
- Keysar, B., & Bly, B. (1995). Intuitions of the transparency of idioms: Can one keep a secret by spilling the beans? *Journal of Memory & Language*, 34(1), 89-109.
- Khonbi, Z. A., & Sadeghi, K. (2017). Improving language learners' idiomatic competence: Does mode of teaching play a role? *Iranian Journal of Language Teaching Research*, 5(3), 61-79.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.
- Konopka, A., & Bock, K. (2009). Lexical or syntactic control of sentence formulation? Structural generalizations from idiom production. *Cognitive Psychology*, 58 (1), 68-101.
- Kövecses & Szabó (1996). Idioms: A view from Cognitive Semantics. *Applied Linguistics*, 17(3), 326-355.
- Kremmel, B., Brunfaut, T., Alderson, J. C. (2017). Exploring the role of phraseological knowledge in foreign language reading. *Applied Linguistics*, 38(6), 848-870.

- Krippendorff, K. (2004). *Content Analysis: An Introduction to Its Methodology*. Sage.
- Kroll, J. F. (1993). Accessing conceptual representations for words in a second language. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 53-81). John Benjamins.
- Kroll, J. F. & Sholl, A. (1992). Lexical and conceptual memory in fluent and nonfluent bilinguals. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 191-204). North-Holland.
- Kuiper, K., Columbus, G., & Schmitt, N. (2009). Acquiring phrasal vocabulary. In S. Foster-Cohen (Ed.), *Advances in language acquisition* (pp. 216-240). Palgrave Macmillan.
- Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. The University of Chicago Press.
- Lakoff, G. (1993). The contemporary theory of metaphor. In A. Ortony (Ed.), *Metaphor and thought* (pp. 202-251). Cambridge University Press.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.
- Langacker, R. (1987). *Foundations of Cognitive Grammar: Theoretical prerequisite*. Stanford University Press.
- Langacker, R. (1991). *Foundations of Cognitive Grammar. Vol. II: Descriptive Application*. Stanford University Press.
- Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. In J. Coady, & T. Huckin (Eds.), *Second Language Vocabulary Acquisition: A Rationale for Pedagogy* (pp. 20-34). Cambridge University Press.
- Laufer, B. (1989). A factor of difficulty in vocabulary learning: Deceptive transparency. *AILA Review*, 6 (Vocabulary Acquisition), 10-20.
- Laufer, B. (2000). Avoidance of idioms in a second language: The effect of L1-L2 degree of similarity. *Studia Linguistica*, 54(2), 186-196.
- Lennon, P. (1998). Approaches to the teaching of idiomatic language. *IRAL, International Review of Applied Linguistics in Language Teaching*, 36(1), 11-30.
- Let the cat out of the bag. (1999). In *Collins COBUILD Dictionary of Idioms* (2nd ed.). HarperCollins Publishers.

- Levorato, (1993). The acquisition of idioms and the development of figurative competence. In C. Cacciari & P. Tabossi (Eds.), *Idioms: Processing, structure, and interpretation* (pp. 101-128). Lawrence Erlbaum.
- Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: a longitudinal case study. *Journal of Second Language Writing*, 18(2), 85-102.
- Li, T. F. (2009). Metaphor, image, and image schemas in second language pedagogy. Lambert Academic Publishing.
- 李炜. (2008). *Words, words, words 咬文嚼字学英语*. 北京工业大学出版社.
- 李玉平. (2008). *English Idioms and their Cultural Origin 英语习语及其文化源流*. 南开大学出版社.
- Libben, M. R. & Titone, D. A. (2008). The multidetermined nature of idiom processing. *Memory & Cognition*, 36(6), 1103-1121.
- Lindstromberg, S., & Boers, F. (2005). From movement to metaphor with manner-of-movement verbs. *Applied Linguistics*, 26(2), 241-261.
- Lindstromberg, S., & Boers, F. (2008a). The mnemonic effect of noticing alliteration in lexical chunks. *Applied Linguistics*, 29(2), 200-222.
- Lindstromberg, S., & Boers, F. (2008b). Phonemic repetition and the learning of lexical chunks: The power of assonance. *System*, 36(3), 423-436.
- Liontas, J. I. (1999). *Developing a pragmatic methodology of idiomaticity: The comprehension and interpretation of SL vivid phrasal idioms during reading* (Unpublished doctoral dissertation). The University of Arizona, Tucson, AZ.
- Liontas, J. I. (2002). Context and idiom understanding in second languages. In S. H. Foster-Cohen, T. Ruthenberg, & M.-L. Poschen (Eds.), *EUROSLA Yearbook: Annual conference of the European Second Language Association* (pp. 155-185). John Benjamins.
- Liontas, J. I. (2003). Killing two birds with one stone: Understanding Spanish VP idioms in and out of context. *Hispania*, 86(2), 289-301.
- Liontas, J. I. (2015a). Developing idiomatic competence in the ESOL classroom: A pragmatic account. *TESOL Journal*, 6(4), 621-658.

- Liontas, J. I. (2015b). Straight from the Horse's Mouth: Idiomaticity Revisited. In R. Heredia & A. Cieřlicka (Eds.), *Bilingual Figurative Language Processing* (pp. 301-340). Cambridge University Press.
- Liontas, J. I. (2017). Why Teach Idioms? A Challenge to the Profession. *Iranian Journal of Language Teaching Research*, 5(3), 5-25.
- Littlemore, J. (2009). *Applying Cognitive Linguistics to Second Language Learning and Teaching*. Palgrave Macmillan.
- Littlemore, J. (2015). *Metonymy: Hidden Shortcuts in Language, Thought and Communication*. Cambridge University Press.
- Littlemore, J., & Low, G. (2006a). *Figurative thinking and foreign language learning*. Palgrave MacMillan.
- Littlemore, J., & Low, G. (2006b). Metaphoric competence, second language learning, and communicative language ability. *Applied Linguistics*, 27(2), 268-294.
- Littlemore, J., Chen, P. T., Koester, A., & Barnden, J. (2011). Difficulties in Metaphor Comprehension Faced by International Students Whose First Language is not English. *Applied Linguistics*, 32(4), 408-429.
- Liu, D. (2003). The most frequently used spoken American English idioms: A corpus analysis and its implications. *TESOL Quarterly*, 38(4), 671-700.
- Liu, D. (2008). *Idioms: Description, comprehension, acquisition, and pedagogy*. Routledge.
- Longman Dictionary of English Idioms*. (1979). Harlow & London: Longman Group Limited.
- 骆世平. (2006). 英语习语研究 A Research on English Idioms. 上海外语教育出版社.
- MacArthur, F. (2010). Metaphorical competence in EFL: Where are we and where should we be going? A view from the language classroom. *AILA Review*, 23, 155-173.
- MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101(4), 676-703.
- MacDonald, M. C., & Seidenberg, M. S. (2006). Constraint satisfaction accounts of lexical and sentence comprehension. In M. J. Traxler, & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 581-611). Elsevier.
- MacMillan English Dictionary for Advanced Learners* (2nd ed.). (2012). Macmillan Education.

- Makkai, A. (1972). *Idiom Structure in English*. The Hague by Mouton.
- Makkai, A. (1993). Idiomaticity as a reaction to *L'Arbitraire du Signe* in the universal process of semeio-genesis. In C. Cacciari & P. Tabossi (Eds.), *Idioms: Processing, structure, and interpretation* (pp. 297-324). Lawrence Erlbaum.
- Malt, B. C., & Eiter, B. (2004). Even with a green card, you can be put into pasture and still have to work: Non-native intuitions of the transparency of common English idioms. *Memory & Cognition*, 32(6), 896-904.
- Mäntylä, K. (2004). *Idioms and Language Users: The Effect of the Characteristics of Idioms on their Recognition and Interpretation by Native and Non-Native Speakers of English* (Unpublished doctoral dissertation). University of Jyväskylä, Finland.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*. 25(1), 71-102.
- Martinez, R., & Murphy, V. A. (2011). Effect of Frequency and Idiomaticity on Second Language Reading Comprehension. *TESOL Quarterly*, 45(2), 267-290.
- Martinez, R., & Schmitt., N. (2012). A Phrasal Expressions List. *Applied Linguistics*, 33(3), 299-320.
- McCarthy, M. (1992). English idioms in use. *Revista Canaria de Estudios Ingleses*, 25, 55-65.
- McCarthy, M. (1998). *Spoken language and applied linguistics*. Cambridge University Press.
- McLay, V. (1987). *Idioms at Work*. Language Teaching Publications.
- McPherron, P., & Randolph, P. T. (2014). *Cat Got Your Tongue? Recent research and classroom practices for teaching idioms to English learners around the world*. TESOL Press.
- Moon, R. (1998). *Fixed Expressions and Idioms in English: A Corpus-based Approach*. Clarendon Press.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Heinle and Heinle.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Nattinger, J., & DeCarrico. J. (1992). *Lexical Phrases and Language Teaching*. Oxford University Press.

- Noroozi, I. & Salehi, H. (2013). The Effect of the Etymological Elaboration and Rote Memorization on Learning Idioms by Iranian EFL Learners. *Journal of Language Teaching and Research*, 4(4), 845-851.
- Nunberg, G. (1978). *The pragmatics of reference*. Indiana University Linguistics Club.
- Nunberg, G., Sag, I., & Wasow, T. (1994). Idioms. *Language*, 70(3), 491-538.
- Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge University Press.
- O’Keeffe, A., McCarthy, M., & Carter, R. (2007). *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge University Press.
- Olson, G. M., Duffy, S. A., & Mack, R. L. (1984). Thinking-out-loud as a method studying real-time comprehension processes. In D. E. Kieras & M. A. Just (Eds.) *New methods in reading comprehension research* (pp. 253-286). Lawrence Erlbaum.
- Oxford Advanced Learner's English-Chinese Dictionary* (8th ed.). (2014). 商务印书馆; Oxford University Press (China).
- Oxford Dictionary of Idioms* (1999). Oxford University Press.
- Oxford Idioms Dictionary 牛津英语习语词典* (2nd ed./ English-Chinese ed.). (2013). Oxford University Press; Foreign Language Teaching and Research Publishing.
- Paivio, A. (1971). *Imagery and verbal processes*. Holt, Rinehart, and Winston. (Reprinted 1979, Lawrence Erlbaum)
- Paivio, A. (1986). *Mental representations: A Dual Coding approach*. Oxford University Press.
- Paivio, A. (2006). *Mind and its evolution: A dual coding theoretical interpretation*. Lawrence Erlbaum.
- Paivio, A., & Desrochers, A. (1980). A dual-coding approach to bilingual memory. *Canadian Journal of Psychology*, 34(4), 388-399.
- Paivio, A., & Desrochers, A. (1981). Mnemonic techniques in second-language learning. *Journal of Educational Psychology*, 73(6), 780-795.
- Paivio, A., & Harshman, R. (1983). Factor analysis of a questionnaire on imagery and verbal habits and skills. *Canadian Journal of Psychology*, 33(1), 17-28.

- Pan, X. (2019). The effectiveness of the Conceptual Metaphor Approach to English idiom acquisition by young Chinese learners. *Metaphor and the Social World*, 9(1), 59-82.
- Park, J., & Chon, Y. V. (2019). EFL learners' knowledge of high-frequency words in the comprehension of idioms: A boost or a burden? *RELC Journal*, 50(2), 219-234.
- Pass the buck. (n.d.) In *Cambridge Advanced Learner's Dictionary & Thesaurus*. Cambridge University Press. Retrieved September 12, 2016, from <https://dictionary.cambridge.org/dictionary/english/pass-the-buck>.
- Peaty, D. (1983). *Working with English Idioms*. Pearson Education.
- Pelletier, F. J. (1994). The Principle of Semantic Compositionality. *Topoi*, 13 (11), 11-24.
- Peterson, R. & Burgess, C. (1993). Syntactic and semantic processing during idiom comprehension: Neurolinguistic and psycholinguistic disassociation. In C. Cacciari & P. Tabossi (Eds.), *Idioms: Processing, structure, and interpretation* (pp. 201-225). Lawrence Erlbaum.
- Peterson, R., Burgess, C., Dell, G., & Eberhard, K. M. (2001). Dissociation between syntactic and semantic processing during idiom comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(5), 1223–1237.
- Pollio, H., Barlow, J., Fine, H., & Pollio, M. (1977). *Psychology and the poetics of growth: figurative language in psychology, psychotherapy, and education*. Lawrence Erlbaum.
- Pollio, H., & Smith, M. (1980). Metaphoric competence and complex human problem solving. In R. P. Honeck & R. R. Hoffman (Eds.), *Cognition and figurative language* (pp. 365-392). Lawrence Erlbaum.
- Pull a rabbit out of the hat. (1998) In *Cambridge International Dictionary of Idioms*. Cambridge University Press.
- Pull one's weight. (1998) In *Cambridge International Dictionary of Idioms*. Cambridge University Press.
- Qi, Y., & Ding, Y. (2011). Use of formulaic sequences in monologues of Chinese EFL learners. *System*, 39, 164-174.
- R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Ramonda, K. (2016). *The role of semantic transparency and metaphorical elaboration*

- through pictures for learning idioms in a second language* (Unpublished doctoral dissertation). University of Birmingham, UK. <http://etheses.bham.ac.uk/6689/>
- Robinson, P. (2013). Aptitude in second language acquisition. In Carole A. Chapelle (ed.), *The Encyclopedia of Applied Linguistics*. Wiley-Blackwell.
- Ruhl, C. (1989). *On Monosemy: A study in linguistic semantics*. State University of New York Press.
- Schweigert, W. A. (1986). The comprehension of familiar and less familiar idioms. *Journal of Psycholinguistic Research*, 15(1), 33-45.
- Show someone the ropes. (2012). In *MacMillan English Dictionary for Advanced Learners*. Macmillan Education.
- Simpson, R., & Mendis, D. (2003). A corpus-based study of idioms in academic speech. *TESOL Quarterly*, 37(3), 419-441.
- Sinclair, J. M. (1987) Collocation: A progress report. In R. Steele & T. Thomas (Eds.), *Language topics: Essays in honor of Michael Halliday, II* (pp. 319-331). John Benjamins.
- Siyanova-Chanturia, A. (2015a). Collocation in beginner learner writing: A longitudinal study. *System*, 53, 148-160.
- Siyanova-Chanturia, A. (2015b). On the ‘holistic’ nature of formulaic language. *Corpus Linguistics and Linguistic Theory*, 11(2), 285-301.
- 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.
- Siyanova-Chanturia, A., Conklin, K., & van Heuven, W. (2011). Seeing a phrase “time and again” matters: The role of phrasal frequency in the processing of multiword sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(3), 776-784.
- Siyanova-Chanturia, A., & Martinez, R. (2015). The idiom principle revisited. *Applied Linguistics*, 36 (5), 549-569.
- Skehan, P. (2015). Foreign language aptitude and its relationship with grammar: A critical overview. *Applied Linguistics*, 36(3), 367-384.

- Skoufaki, S. (2008). Conceptual metaphoric meaning clues in two L2 idiom presentation methods. In F. Boers, F. & S. Lindstromberg (Eds.) *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 101-132). Mouton de Gruyter.
- Smith, L. P. (1925). *Words and Idioms*. Constable & Company.
- Snider, N., & Arnon, I. (2012). A unified lexicon and grammar? Compositional and non-compositional phrases in the lexicon. In S. Gries & D. Divjak (Eds.), *Frequency effects in language* (pp. 127-163). Mouton de Gruyter.
- Sperber, D. & Wilson, D. (1986/1995). *Relevance: Communication and Cognition*. Blackwell.
- 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.
- Steinel, M., Hulstijn, J., & Steinel, W. (2007). Second language idiom learning in a paired-associate paradigm: Effects of direction of learning, direction of testing, idiom imageability, and idiom transparency. *Studies in Second Language Acquisition*, 29(3), 449-484.
- Stengers, H. (2007). Is English exceptionally idiomatic? Testing the waters for a lexical approach to Spanish. In F. Boers, J. Darquennes, & R. Temmerman (Eds.) *Multilingualism and Applied Comparative Linguistics: Pedagogical Perspectives* (pp. 107-125). Cambridge Scholars Publishing.
- Stengers, H., Deconinck, J., Boers, F., & Eyckmans, J. (2016) Does copying idioms promote their recall? *Computer Assisted Language Learning*, 29(2), 289-301.
- Strässler, J. (1982). *Idioms in English: A Pragmatic Analysis*. Verlag.
- Strong, B., & Boers, F. (2019a). Error in trial and error: Exercises on phrasal verbs. *TESOL Quarterly*, 53(2), 289-319.
- Strong, B., & Boers, F. (2019b). Weighing up exercises on phrasal verbs: Retrieval versus trial-and-error practices. *Modern Language Journal*, 103(3), 562-579.
- Suñer, F. (2018). The interplay of crosslinguistic differences and context in L2 idiom comprehension. *Research in Language*, 16(4), 495-513.

- Suñer, F. (2019). Exploring the relationship between stay-abroad experiences, frequency effects, and context use in L2 idiom comprehension. *Journal of Language Teaching and Research, 10*(1), 35-44.
- Swinney, D. A., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning & Verbal Behavior, 18*(5), 523-534.
- Szczepaniak, R., & Lew, R. (2011). The role of imagery in dictionaries of idioms. *Applied Linguistics, 32*(3), 323-347.
- Tabossi, P., Fanari, R., & Wolf, K. (2008). Processing idiomatic expressions: Effects of semantic compositionality. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(2), 313-327.
- Tabossi, P., Fanari, R., & Wolf, K. (2009). Why are idioms recognized fast? *Memory & Cognition, 37*(4), 529-540.
- Tabossi, P., & Zardon, F. (1993). The activation of idiomatic meaning in spoken language comprehension. In C. Cacciari & P. Tabossi (Eds.), *Idioms: Processing, structure, and interpretation* (pp. 145-162). Lawrence Erlbaum.
- Take the bull by the horns. (1999). In *Collins COBUILD Dictionary of Idioms* (2nd ed.). HarperCollins Publishers.
- The American Heritage Dictionary of Idioms*. (2003). Houghton Mifflin Harcourt.
- Titone, D. A., & Connine, C. M. (1994a). Comprehension of idiomatic expressions: Effects of predictability and literality. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(5), 1126-1138.
- Titone, D. A., & Connine, C. M. (1994b). Descriptive norms for 171 idiomatic expressions: Familiarity, compositionality, predictability, and literality. *Metaphor and Symbol, 9*(4), 247-270.
- Titone, D. A., & Connine, C.M. (1999). On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics, 31*(12), 1655-1674.
- Titone, D., Columbus, G., Whitford, V., Mercier, J., & Libben, M. (2015). Contrasting bilingual and monolingual idiom processing. In R. Heredia & A. Cieřlicka (Eds.), *Bilingual Figurative Language Processing* (pp. 171-207). Cambridge University Press.

- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381-402). Academic Press.
- Türker, E. (2016). The role of L1 conceptual and linguistic knowledge and frequency in the acquisition of L2 metaphorical expressions. *Second Language Research*, 32(1), 25-48.
- Vasiljevic, Z. (2015). Effects of Etymology and Pictorial Support on the Retention and Recall of L2 Idioms. *Electronic Journal of Foreign Language Teaching*, 12(1), 35-55.
- Vega-Moreno, R. (2001). Representing and processing idioms. *UCL Working Papers in Linguistics*, 13, 73-107.
- Vega-Moreno, R. (2003). Relevance Theory and the construction of idiom meaning. *UCL Working Papers in Linguistics*, 15, 303-323.
- Vega-Moreno, R. (2005). Idioms, transparency, and pragmatic inference. *UCL Working Papers in Linguistics*, 17, 389-425.
- Verspoor, M., & Boers, F. (2013). Cognitive Linguistics of Second Language Acquisition. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Wiley-Blackwell.
- Verspoor, M., & Lowie, W. (2003). Making sense of polysemous words. *Language Learning*, 53(3), 547-586.
- Vespignani, F., Canal, P., Molinaro, N., Fonda, S., & Cacciari, C. (2009). Predictive Mechanisms in Idiom Comprehension. *Journal of cognitive neuroscience*, 22(8), 1682-1700.
- Warmington, M., & Hitch, G. J. (2014). Enhancing the learning of new words using an errorless learning procedure: Evidence from typical adults. *Memory*, 22(5), 582-594.
- Weinert, R. (1995). The role of formulaic language in second language acquisition: A review. *Applied Linguistics*, 16(2), 180-205.
- Weinreich, U. (1969). Problems in the analysis of idioms. In Puhvel, J. (Ed.), *Substance and Structure of Language* (pp. 23-81). University of California Press.
- Win hands down. (2012). In *Collins COBUILD Idioms Dictionary* (3rd ed). HarperCollins Publishers.
- Winter, B. (2020) *Statistics for Linguists: An Introduction Using R*. Routledge.
- Wolter, B. (2006). Lexical network structures and L2 vocabulary acquisition: The role of L1 lexical/conceptual knowledge. *Applied Linguistics*, 27(4), 741-747.

- Wood, M. M. (1986). *A definition of idiom*. Indiana University Linguistics Club.
- Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge University Press.
- Wray, A. (2008). *Formulaic language: Pushing the boundaries*. Oxford University Press.
- Wright, J. (1999). *Idioms Organiser*. Language Teaching Publications/Heinle, Thomson.
- Wu, X., Chen, B., & Huang, L. (2006). 中国学生对英语习语的理解: 习语类型与二语水平的作用 [Processing of English idioms by Chinese EFL learners: Effects of idiom type and proficiency level]. *外语教学与研究/Foreign Language Teaching and Research*, 3, 196-201.
- Yasuda, S. (2010). Learning phrasal verbs through conceptual metaphors: A case of Japanese EFL learners. *TESOL Quarterly*, 44(2), 250-27
- Yorio, C. A. (1980). Conventionalized language forms and the development of communicative competence. *TESOL Quarterly*, 14(4), 433-442.
- Yorio, C. A. (1989). Idiomaticity as an indicator of second language proficiency. In K. Hyltenstam & L. Obler (Eds.), *Bilingualism across the Lifespan: Aspects of Acquisition, Maturity and Loss* (pp. 55-72). Cambridge University Press.
- Zhang, L. (2009). *The effect of etymological elaboration on L2 idiom acquisition and retention in an online environment (WebCT)* (Unpublished master thesis). Iowa State University, US. <https://lib.dr.iastate.edu/etd/11126>
- Zyzik, E. (2011). Second language idiom learning: The effects of lexical knowledge and pedagogical sequencing. *Language Teaching Research*, 15(4) 413-433.

Appendix 1: Information Sheet for Participants



Investigating Second Language Idiom Learning

INFORMATION SHEET FOR PARTICIPANTS

Thank you for your interest in this project. Please read this information before deciding whether or not to take part. If you decide to participate, thank you. If you decide not to take part, thank you for considering my request.

Who am I?

My name is *Xinqing Wang* and I am a Doctoral student in Applied Linguistics at Victoria University of Wellington. This research project is work towards my dissertation.

What is the aim of the project?

This project aims to investigate Chinese EFL (English as Foreign Language) learners' interpretation and learning strategies of English idioms (like *rain cats and dogs*, *couch potato*, *paint the town red*), and the effectiveness of a teaching approach. This research has been approved by the Victoria University of Wellington Human Ethics Committee [*Ethics Approval: 23363*].

How can you help?

If you agree to take part, I will interview you in a spare classroom individually. I will ask you questions about your understandings of some English idioms and responses to the teaching technique. There will be two interviews; approximately 60 minutes each time. I will record the interviews and write it up later. You can stop the interview at any time, without giving a reason. You can withdraw from the study by November 26, 2016. If you withdraw, the information you provided will be destroyed or returned to you. It is guaranteed that the research is separate from your course work and is in no way related to your assessment and will have no direct impact on your grade.

What will happen to the information you give?

This research is confidential. I will not name you in any reports, and I will not include any information that would identify you. Only my supervisors and I will read the notes or transcript of the interviews. The interview transcripts, summaries and any recordings will be kept securely and destroyed 3 years after the research ends.

What will the project produce?

The information from my research will be used in my PhD dissertation. You will not be identified in my report. I may also use the results of my research for conference presentations, and academic reports. I will take care not to identify you in any presentation or report.

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

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

- choose not to answer any question;
- ask for the recorder to be turned off at any time during the interview;
- withdraw from the study up until November 26, 2016, when all data are collected;
- ask any questions about the study at any time;
- receive a copy of your interview recording;
- read over and comment on a written summary of your interview;
- be able to read any reports of this research by emailing the researcher to request a copy.

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

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

Student:

Name: Xinqing Wang

University email address:
Aileen.Wang@vuw.ac.nz

Phone: 86-15589579991

Supervisor:

Name: Frank Boers

Role: Associate Professor

School: Linguistics and Applied Language
Studies

Email: Frank.Boers@vuw.ac.nz

Phone: 064 4 463 6014

Human Ethics Committee information

If you have any concerns about the ethical conduct of the research, you may contact the Victoria University HEC Convener: Associate Professor Susan Corbett. Email susan.corbett@vuw.ac.nz or telephone +64-4-463 5480.



Investigating Second Language Idiom Learning **第二语言习语学习调查研究**

INFORMATION SHEET FOR PARTICIPANTS **参与调研信息介绍**

感谢您对此项目感兴趣，在决定是否参与前请阅读以下信息。如果您决定参与，感谢您的合作；如果决定不参与，感谢您的关注。

研究者介绍

王新清，惠灵顿维多利亚大学应用语言学博士生。这项研究是其博士研究项目的一部分。

研究目标

这项研究目标有二：一是调查中国英语学习者对英语习语的理解程度和学习策略，二是探索一种习语教学方法的可行性和效果（习语例如 rain cats and dogs, couch potato, paint the town red 等）。该研究已通过惠灵顿维多利亚大学人类伦理道德委员会审批立项 [编号为：23363]。

研究过程与参与者的贡献

研究过程由对参与者的两次一对一访谈组成，每次时长约 60 分钟，两次访谈间隔为一周，地点是安静的教室或办公室。访谈中，您将接触到一些英语短语表达并回答相关问题。访谈将通过录音保存，以便之后整理访谈笔记。录音过程中，您可以随时暂停。2016 年 11 月 26 日前，可选择撤出您的访谈信息；若撤出，您所提供的信息数据会损毁或返还给您。此项研究作为一项课外调研不会影响到您的任何课业评估和成绩。

访谈信息保存与处理

研究收集的信息为保密文件。参与者姓名不会在任何研究报告中出现，可识别个人的信息受到保护不会公开。访谈的笔记和录音转写只有研究者本人和导师可见。访谈录音、笔记和总结将被妥善保管，并在研究结束三年后销毁。

研究结果呈现

访谈所获信息将用于研究者博士论文。研究结果会在会议展示和学术报告中呈现，任何汇报过程中，研究者会确保访谈参与者的个人信息不被识别。

如接收邀请参与研究调查，参与者的权利：

接受邀请参与调研为自愿行为。作为参与者您有以下权利：

- 选择不回答任何问题
- 访谈中任何阶段要求停止录音
- 2016年11月26日数据收集完毕前，申请撤回所有提供的信息
- 获得参与访谈的录音副本
- 阅读和评论访谈笔录总结
- 通过邮件向研究者索取和阅读研究所得报告副本

如有任何疑问，可以通过以下联系方式咨询：

学生：

王新清

学校邮箱： Aileen.Wang@vuw.ac.nz

国内 电话： 86-15589579991

导师：

Frank Boers

语言与应用语言学系 副教授

Email: Frank.Boers@vuw.ac.nz

Phone: 064 4 463 6014

Human Ethics Committee information 人类道德伦理委员会信息

如对此研究的伦理道德行为有任何疑问，可与维多利亚大学人伦委员会负责人 Susan Corbett 副教授联系。邮箱是 susan.corbett@vuw.ac.nz， 电话是 +64-4-463 5480。

Appendix 2: Participant Consent Form



Investigating Second Language Idiom Learning

CONSENT TO PARTICIPATE IN RESEARCH

This consent form will be held for 3 years.

Researcher: *Xinqing Wang*
School of Linguistics and Applied Language Studies
Victoria University of Wellington

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

I understand that:

- I may withdraw from this study at any point during data collection and before November 26, 2016, and any information that I have provided will be returned to me or destroyed.
- The information I have provided will be destroyed 3 years after the research is finished.
- Any information I provide will be kept confidential to the researcher and the supervisor. I understand that the results will be used for a PhD report and a summary of the results may be used in academic reports and/or presented at conferences.
- My name will not be used in reports. I understand that if the researcher wants to publish anything that might identify me, she will ask for my permission.
- I would like a summary of my interview: Yes ☐ No ☐
- I would like to receive a summary of the research findings and have added my email address below. Yes ☐ No ☐

Signature of participant: _____

Name of participant: _____

Date: _____

Contact details: _____



Investigating Second Language Idiom Learning **第二语言习语学习调查研究**

CONSENT TO PARTICIPATE IN RESEARCH **参与调研同意书**

(文件保存时间为三年)

研究者信息： 王新清 Xinqing Wang
惠灵顿维多利亚大学
语言学与应用语言学系

- 我已阅读《参与调研信息介绍》并了解项目内容。有关疑问已得到明确解答。我明白，如有疑问我可以随时提问。
- 我同意参与录音访谈。

我明白：

- 在调查研究过程中任何时候及 2016 年 11 月 26 日前，我可以撤回我的访谈信息，我所提供的任何信息将返回给我本人或被销毁。
- 我所提供的所有信息将在这项研究完成三年后被销毁。
- 我所提供的所有信息会被妥善保管，仅对研究者本人及其导师可见。我明白信息将用于博士论文的学术探讨中，研究结果可能会在学术报告或在学术会议中呈现。
- 我的姓名不会出现于任何研究报告中。我明白如果研究者在任何学术发表中提及识别我本人的信息，必须经由我许可。
- 我需要一份访谈总结 是 ☐ 否 ☐
- 我需要研究结果总结并在下面提供了邮箱地址 是 ☐ 否 ☐

参与者签名： _____

参与者姓名： _____

参与日期： _____

联系方式： _____

Appendix 3: Human Ethics Approval Confirmation



Phone 0-4-463 5480
Email susan.corbett@vuw.ac.nz

MEMORANDUM

TO	Aileen Wang
COPY TO	Frank Boers
FROM	AProf Susan Corbett, Convener, Human Ethics Committee
DATE	23 September 2016
PAGES	1
SUBJECT	Ethics Approval: 23363 Investigating Second Language Idiom Learning

Thank you for your application for ethical approval, which has now been considered by the Standing Committee of the Human Ethics Committee.

Your application has been approved from the above date and this approval continues until 1 November 2019. If your data collection is not completed by this date you should apply to the Human Ethics Committee for an extension to this approval.

Best wishes with the research.

Kind regards

A handwritten signature in black ink, appearing to read 'Susan Corbett', written over a horizontal line.

Susan Corbett
Convener, Victoria University Human Ethics Committee

Appendix 4: Frequencies of the 84 idioms in COCA and Collins COBUILD

Idiom	Frequency count of the figurative use in COCA	Frequency band in Collins COBUILD (1999)
In the wake of something	4818	◀◀◀
(Start) from scratch	2034	◀◀◀
Make ends meet	1035	◀◀◀
Red tape	924	◀◀◀
Follow suit	787	◀◀◀
On the same page	614	No ◀
On a shoestring	397	◀◀◀
Be waiting in the wings	287	◀◀◀
Bread and butter	272	◀◀◀
On the back burner	263	◀◀◀
Have/get cold feet	215	◀◀
Down and out	213	◀◀◀
A rule of thumb	212	◀◀◀
Take a back seat	195	◀◀◀
(Be) par for the course	191	◀◀
Jump the gun	186	◀◀
A drop in the bucket	185	◀◀
Be on the ropes	184	◀◀
A red herring	177	◀◀◀
Play into someone's hands	175	◀◀◀
Throw in the towel	161	◀◀◀
A feeding frenzy	159	◀◀
Rub someone the wrong way	157	◀
Jump ship	156	◀◀
Flex your muscles	154	◀◀◀
A can/bag of worms	153	◀◀
Bite the bullet	149	◀◀◀
Ruffle someone's feathers	140	◀◀
Ring a bell	131	◀
Give the green light	124	◀◀◀
A loose cannon	123	◀◀
Tighten our belt	122	◀◀◀
Give someone a leg up	120	◀◀
Cut corners	115	◀◀◀

In the doldrums	115	◀◀◀
Take the plunge	105	◀◀◀
Take it on the chin	100	◀◀
A dark horse	99	◀◀◀
Pull one's weight	96	◀◀
A shot in the arm	92	◀◀◀
(Hit) below the belt	89	◀
Leave someone high and dry	81	◀◀
Go belly-up	78	◀◀◀
A hot potato	74	◀◀
Have a lot on your plate	72	◀◀
Wet blanket	69	No ◀
Pass the buck	68	◀◀◀
Spill the beans	66	◀◀
Go with the flow	66	◀◀
Pass (on) the baton	64	◀
Come out of your shell	61	◀◀◀
The ball is in your court	60	◀◀
(Be) a sitting duck	59	◀
Stick your neck out	56	◀◀
Play your cards close to your chest	55	◀
Sit on the fence	54	◀◀◀
Bury the hatchet	50	◀
Beat around the bush	49	◀
Hand over fist	48	Not included
Get something off your chest	43	◀◀
Give someone the cold shoulder	42	◀◀◀
Throw your hat/cap in the ring	41	◀
Bark up the wrong tree	39	◀
Show someone the ropes	38	◀
Win hands down	35	◀◀
Turn over a new leaf	32	◀◀
Play it by ear	31	◀
Have a green thumb/ fingers	20	◀
Take the bull by the horns	17	◀
Teething problems	17	◀◀◀

Hold your horses	14	No ◀
Burn the candle at both ends	14	◀
Get into gear	14	◀◀◀
On the same wavelength	10	◀◀
Hit the roof/ceiling*	8	◀◀
Pull a rabbit out of the hat	7	◀◀
In the driving/drivers' seat	6	◀◀◀
Get your second wind	3	◀
(Not) up to scratch	2	◀◀◀
Jump in (at) the deep end	2	◀◀
Let the cat out of the bag	1	◀
On automatic pilot	1	◀◀◀
Weigh someone down	0	Not included
Throw your hand in	0	Not included

(*Note.* The list is arranged in descending order of the frequency counts in COCA in 2016.)

Appendix 5: Semantic transparency by three native English speakers

	Idioms	Rater M	Rater S	Rater D	Final
1	Come out of your shell	H	L	H	H
2	Give the green light	H	M	H	H
3	Spill the beans	L	L	L	L
4	Throw in the towel	M	L	M	M
5	Show someone the ropes	L	L	L	L
6	Jump the gun	M	M	H	M
7	Have a lot on your plate	M	M	H	M
8	Sit on the fence	M	H	M	<u>H</u>
9	Be waiting in the wings	H	L	L	L
10	Play it by ear	L	M	H	M
11	Bury the hatchet	M	M	L	M
12	Have/get cold feet	M	L	L	L
13	Bark up the wrong tree	M	M	M	M
14	Bite the bullet	L	L	L	L
15	Let the cat out of the bag	M	L	L	L
16	Pull a rabbit out of the hat	M	H	M	M
17	Red tape	L	M	H	M
18	Beat around the bush	L	L	L	L
19	Get your second wind	M	H	M	M
20	On the same wavelength	H	H	M	H
21	On the same page	M	H	H	H
22	Cut corners	H	H	M	H
23	In the wake of something	M	M	L	M
24	Take the bull by the horns	M	M	H	M
25	Pull one's weight	H	M	M	M

26	Weigh someone down	H	H	H	H
27	Hit the roof/ceiling	L	M	L	L
28	Give someone a leg up	M	M	L	M
29	Tighten our belt	M	L	M	<u>H</u>
30	Take a back seat	M	H	H	H
31	In the driving/drivers' seat	H	H	H	H
32	Have a green thumb/ fingers	M	M	L	M
33	(Not) up to scratch	L	M	L	L
34	Burn the candle at both ends	L	L	L	L
35	Pass the buck	L	L	L	L
36	A feeding frenzy	H	M	H	H
37	A drop in the bucket	H	M	H	H
38	A red herring	L	L	L	L
39	Ring a bell	M	H	M	M
40	(Be) a sitting duck	H	H	M	H
41	Teething problems	H	M	M	M
42	Jump ship	H	H	L	H
43	Flex your muscles	H	M	H	H
44	A loose cannon	M	M	M	M
45	Get something off your chest	H	M	M	M
46	Play your cards close to your chest	H	H	H	H
47	Stick your neck out	M	H	M	M
48	On the back burner	M	M	M	M
49	Go with the flow	H	H	M	H
50	Hand over fist	M	M	L	M
51	Leave someone high and dry	L	L	L	L
52	Pass the baton on	H	H	M	H
53	Wet blanket	M	L	L	L

54	Win hands down	M	H	L	M
55	A rule of thumb	M	L	L	M
56	In the doldrums	M	M	L	M
57	Rub someone the wrong way	H	H	M	H
58	Ruffle someone's feathers	H	H	L	H
59	(Hit) below the belt	M	H	H	H
60	A dark horse	L	M	L	<u>H</u>
61	Down and out	H	H	M	H
62	Bread and butter	L	M	M	M
63	A shot in the arm	M	H	L	M
64	On a shoestring	L	M	L	L
65	A can/bag of worms	M	L	L	L
66	Throw your hand in	M	H	L	M
67	Jump in (at) the deep end	H	H	M	H
68	Make ends meet	L	L	L	L
69	Go belly-up	M	M	M	M
70	Get into gear	M	H	M	M
71	Take the plunge	M	H	L	M
72	Follow suit	H	H	L	H
73	Play into someone's hands	M	H	L	M
74	Give someone the cold shoulder	M	M	L	M
75	A hot potato	M	H	M	<u>H</u>
76	On automatic pilot	H	H	H	H
77	Take it on the chin	M	M	M	M
78	Be on the ropes	L	H	L	L
79	Turn over a new leaf	M	M	L	<u>H</u>
80	The ball is in your court	M	H	L	M
81	Hold your horses	M	M	M	M

82	(Start) from scratch	L	M	L	L
83	Throw your hat/cap in the ring	L	M	L	L
84	(Be) par for the course	M	L	L	L

Note. Underlined are modified ratings for five idioms which have equivalent expressions in Chinese.

Appendix 6: Dictionary meanings and example sentences

Idiom	Simple context	Idiomatic meaning	Rich context
give the green light	He gave us the green light.	To give official approval for something to be done or allow it to happen.	The government gave the green light to our project.
on the back burner	It's been put on the back burner.	[Of something] on hold or suspended temporarily/ not getting or needing immediate attention	She put her career on the back burner after getting married.
follow suit	They had to follow suit.	To follow someone else's example.	If other stores lower their prices, we'll have to follow suit .
take it on the chin	Paul just took it on the chin.	To accept unpleasant events bravely and without complaining	Paul really took it on the chin when the boss shouted at him today.
get into gear	I gradually got into gear.	Start working, or start something working efficiently	It took me a while to get into gear when I came back to work after my holiday.
(be) a sitting duck	He was a sitting duck.	An obvious target (something/ someone) that is very easy to attack and criticize	Any government that openly adopted such a strategy would be a sitting duck for the opposition (反对派).
be par for the course	That was par for the course.	What is normal or expected in any given circumstances.	Grandma didn't sleep much last night but that is par for the course for her.
on the same wavelength	We're on the same wavelength.	Understanding each other well because they share the same attitudes, interests, and opinions.	We used to complete each other's sentences because we were on the same wavelength .
bread and butter	It was my bread and butter.	The most important or only source of your income.	I worked as a bartender for a year, and it was the tips that were my bread and butter .
pull one's weight	He needs to pull his weight.	To do one's share in a common task/ To work as hard as other people in a group	If he doesn't start pulling his weight , he'll lose his job.

jump ship	He has jumped ship.	To quit or resign (especially when there is difficulty with the job)	The original star of the TV drama jumped ship after the first season to a science fiction series.
a dark horse	Smith is a dark horse.	If you describe as a dark horse, you mean that very little is known about them, although they may have recently had success or may be about to have a success.	William had briefly been a dark horse candidate for President in 1908.
go belly-up	It went belly-up.	Go bankrupt 破产、倒闭	General Motors and U.S. Steel almost went belly-up during the economic crisis of 1983.
play into someone's hands	We played into their hands.	To make a foolish mistake or act in the way that someone wants you to act, so that they gain an advantage over you	If we allow terrorists to disrupt our lives we're just playing into their hands .
red tape	There is too much red tape.	Official rules and procedures which seem unnecessary and cause delay	Don't let red tape prevent you from getting the treatment you deserve.
a drop in the bucket	This is a drop in the bucket.	An insufficient amount in comparison with what is required	The charity was given 5000 dollars, but it was just a drop in the bucket compared to what was needed.
throw in the towel	I'm ready to throw in the towel.	To admit defeat, to give up	It's hard to find a job at the moment, but you shouldn't throw in the towel .
spill the beans	I did not spill the beans.	To reveal a secret often by accident or with intentions that are not good	He assures me he did not spill the beans about the surprise birthday party.
Have/got cold feet	She got cold feet.	To suddenly become nervous about doing something that you have planned or agreed to do.	Everything was fine until the last minute, when she suddenly got cold feet .
take the bull by the horns	You have to take the bull by the horns.	To act decisively and with determination and courage in order to deal with a	I decided to take the bull by the horns and get a real job.

		difficult situation or problem.	
burn the candle at both ends	She is burning the candle at both ends.	To exhaust oneself, especially by being up late and getting up early to work	I no longer burn the candle at both ends and I feel better.
take a back seat	Frank decided to take a back seat.	To occupy an inferior decision; let other people make the decisions or have control	Frank decided to take a back seat during the planning of the project.
a shot in the arm	It gave us a shot in the arm.	a strong positive influence which gives encouragement or energy	The Rio Olympics has been a shot in the arm to the restaurant business in Brazil.
beat around the bush	You're beating around the bush.	To speak of many things related to the most important issue but never talk about the issue	This is no time to beat around the bush ; we need to make a decision now.
a feeding frenzy	It started a feeding frenzy.	A situation where people try to get as much information as possible about an event, or to make as much profit as they can from it, especially in an unpleasant way	Her sudden tragic death started a feeding frenzy in the media.
win hands down	They will win hands down.	To win very easily	If a general election were held now, the Conservative Party would win hands down .
jump in (at) the deep end	She jumped in at the deep end.	To do something new and difficult without help or preparation	I think you will learn the job faster if you jump in at the deep end .
give someone the cold shoulder	Tom gave her the cold shoulder.	To deliberately ignore someone (often abruptly)	Nancy was sure many people would give her the cold shoulder at the party.
in the doldrums	We are in the doldrums.	(Economy/ business) nothing new is happening and it is not doing well; (someone) is very inactive/depressed	When the economy is in the doldrums , every business feels the effects.
on the same page	They are no longer on the same page.	Having the same understanding or amount of knowledge; thinking in a similar way	We made a decision quickly because we were on the same page .

pass the baton (on)	It's time to pass the baton.	To give responsibility for something important to another person	The president wants to pass the baton to someone who shares his values.
throw your hand in	He threw his hand in.	Stop doing something or taking part in something, especially because you are not successful	If I fail again this time, I'll throw my hand in .
(be waiting) in the wings	She is waiting in the wings.	Stop doing something or taking part in something, especially because you are not successful.	There is another presidential candidate waiting in the wings .
(not) up to scratch	It is not up to scratch.	At the good standard that is expected or needed	If his work doesn't come up to scratch , get rid of him.
a loose cannon	He is a loose cannon.	A person whose actions are unpredictable and uncontrollable, liable to cause damage	The presidential candidate is a loose cannon who has insulted many people.
Leave someone high and dry	He was left high and dry.	To leave you alone and without any help	All my workers quit and left me high and dry .
be on the ropes	They are on the ropes.	To be very close to defeat or failure	His political career is on the ropes after the bribery (贿赂) scandal.
hit the roof/ceiling	He hit the roof.	To become very angry (usually with shouting at someone)	The teacher hit the roof when I forgot to do my homework.
ruffle someone's feathers	His words ruffled her feathers.	To upset or annoy someone	A frank discussion of politics is likely to ruffle a few feathers .
weigh someone down	It weighed her down.	To worry or depress someone/ to make someone feel tired or weak	Surprisingly, all her troubles didn't weigh her down .
let the cat out of the bag	He let the cat out of the bag.	To reveal something secret or private, often without meaning to.	By the time I realized I had let the cat out of the bag , it was too late.
The ball is in your court.	The ball is in his court now.	You have to do something before any progress can be made in a situation because it is your turn, move or play.	I've told him he can have his job back if he apologizes. The ball is in his court now.
jump the gun	Don't jump the gun.	To act too soon before the right, proper or expected time.	The government doesn't want to jump the gun

			by announcing the policy now.
in the driving/driver's seat	They were in the driver's seat.	To be in control of a situation	The radicals (激进派) were in the driver's seat of the party for six years.
show someone the ropes	He showed her the ropes.	To teach someone how to do something, especially a job	The manager spent the whole morning showing the new secretary the ropes .
sit on the fence	They are sitting on the fence.	Avoid deciding between two sides of an argument, discussion, quarrel, etc. (used to express disapproval of someone)	We can't sit on the fence in such a situation; we must make a decision.
in the wake of something	It happened in the wake of the war.	Following / as a result of (an unpleasant event)	Lending standards have been tightened in the wake of the financial crisis.
bury the hatchet	Let's bury the hatchet.	To agree to forget arguments and disputes in order to become amicable again	The time has come for both sides to bury the hatchet .
teething problems	There have been teething problems.	Problems that you experience in the early stages of an activity	The Council admits there have been teething problems with the new voucher system (付款凭单制度).
take the plunge	She took the plunge.	To do something important or difficult that you have been thinking about doing for a long time	Those women who took the plunge and ran for election had great success.
rub someone the wrong way	He rubs many people the wrong way.	To annoy or irritate somebody.	His remarks about raising taxes rubbed a great many people the wrong way .
a rule of thumb	I do that as a rule of thumb.	A general rule about something which you can be confident will be right in most cases	The rule of thumb is that, when in doubt, ask someone what they prefer to be called.
(hit) below the belt	He hit below the belt.	To unfairly insult someone by using cruel, inappropriate or unfair comments that are too personal; (unfair or cruel)	Will he hit below the belt in his first major presidential debate of his life?

a red herring	That's a red herring.	Something that distract people's attention from the main subject, problem or situation that they should be considering	About halfway through the book the butler seemed to be the murderer, but that turned out to be a red herring .
hold your horses	Can you hold your horses, please?	To wait and be patient	I am asking you to hold your horses and hear me out.
stick your neck out	He is sticking his neck out.	To take a risk; It means to say or do something bold	He stuck his neck out for the deal (交易), hoping to make some big money.
down and out	He is down and out.	(Of a person) without money, a job, or a place to live,	He was down and out, and living on the street.
have a lot on your plate	I have a lot on my plate.	To have a lot of things to do/deal with	I have a lot on my plate at work and can't help you this time.
pass the buck	She is passing the buck.	To blame someone or to make them responsible for a problem that you should deal with yourself	When the boss makes mistakes, she always tries to pass the buck .
play your cards close to your chest	She always plays her cards close to her chest.	To not tell people what you are thinking or planning	In business, it is worth playing your cards close to your chest .
ring a bell	That rings a bell.	To cause someone to remember something or for something to seem familiar	I've never met John Franklin, but his name rings a bell .
turn over a new leaf	He plans to turn over a new leaf.	To completely change your behaviour or attitude in order to be better than you were	He turned over a new leaf and stopped drinking.
make ends meet	They can't make ends meet.	To earn and spend equal amounts of money (or to have just enough money to pay for the basic expenses).	The company is expected to raise the wages of the staff who struggle to make ends meet .
come out of your shell	Finally, she came out of her shell.	If you come out of your shell, you become less shy and more talkative and friendly.	It was the best experience in my entire life, and made me come out of my shell .
flex your muscles	They are flexing their muscles.	To act in a way that shows power	The trade union (工会) flexed its muscles by

		or strength	calling the workers to strike
wet blanket	She was often a wet blanket.	A person who discourages enjoyment or enthusiasm	Anne was being a real wet blanket at the celebration dinner last night.
cut corners	They are cutting corners.	To do something in the easiest, cheapest, or fastest way, by ignoring the rules, often harming the quality	Companies are cutting corners to remain competitive in the market
bark up the wrong tree	You're barking up the wrong tree.	To completely misunderstand a situation; to make a false assumption about a situation	You're barking up the wrong tree if you think I want to go to church with you.
hand over fist	They are making money hand over fist.	[For money and merchandise to be exchanged] very rapidly and continuously	Almost all the oil companies were making money hand over fist .
get your second wind	I've just got my second wind.	To have increased energy or strength after feeling tired or weak	I was tired early this evening, but now I've got my second wind .
bite the bullet	I decided to bite the bullet.	To accept something difficult and suffer in silence	I know it's difficult but you need to bite the bullet and carry on.
go with the flow	I usually go with the flow.	To let things happen to you or let other people tell you what to do, rather than trying to control what happens yourself.	When I'm with this group, I just relax and go with the flow .
a can of worms	That's a can of worms.	A situation which causes a lot of trouble for you when you start to deal with it	The UK have opened a can of worms by leaving the EU.
give someone a leg up	They gave him a leg up.	To give someone an advantage.	His father gave him a leg up into a management position.
have a green thumb/green fingers	My mother has a green thumb.	To be good at gardening/to have the ability to make plants grow well.	My mother has a green thumb and the plants in our house are always flourishing.
a hot potato	That is a hot potato.	A subject or problem that is very controversial and most people would rather not have to deal with it.	Recent environmental legislation has made the issue of coastal development a hot potato .

play it by ear	I'm going to play it by ear.	To deal with things as they happen, rather than following a plan or previous arrangement.	I think right now I'll play it by ear and see what happens.
throw your hat/cap in the ring	He threw his hat in the ring.	To announce that you intend to compete for something, especially a political position	She lost the election for mayor after an opponent threw his hat in the ring at the last moment.
get something off your chest	You need to get it off your chest.	To tell someone about your feelings or emotions, especially after not speaking for a long time	He got it off his chest when a friend offered to listen.
on automatic pilot	She is on automatic pilot today.	To do something without thinking about what you are doing (usually because you have done it many times before or you are very tired)	By the second week of the election campaign she was giving all her speeches on automatic pilot .
tighten your belt	We have to tighten our belts.	To spend less money than you once did; to use less of something	We have to tighten our belts because I lost my job.
on a shoestring	It was done on a shoestring.	(To do something) using very little money.	The film was made on a shoestring budget.
(start) from scratch	They started from scratch.	Start a task from the beginning	Starting from scratch , they grew their business into a global enterprise.
pull a rabbit out of the hat	Can he pull a rabbit out of the hat?	To surprise everyone by suddenly doing something that shows a lot of skill, often in order to solve a problem	Your opponent will gain control of the company, unless your lawyers can pull a rabbit out of the hat for you.





Appendix 7: Etymological prompts of the idioms

	Idiom	Etymological prompt
1	Come out of your shell	<i>Shell</i> refers to a turtle shell.
2	Give the green light	<i>Green light</i> refers to the traffic light.
3	Spill the beans	From a voting system used in ancient Greek: White beans indicated positive votes and black beans negative, and votes had to be unanimous. If the collector 'spill the beans' before the vote was complete and a black bean was seen, the vote was halted.
4	Throw in the towel	From boxing: When a boxer is knocked to the ground, his coach or manager may throw a towel into the boxing ring.
5	Show someone the ropes	Of nautical origin, where knowing the cords that were necessary to operate a ship and its sails was an essential maritime skill.
6	Jump the gun	From sports contests that are started by firing a gun. It refers to an athlete in a race who starts running before the starter has fired the gun.
7	Have a lot on your plate	Have a lot food to finish on one's plate
8	Sit on the fence	To straddle a fence between two different properties.
9	Be waiting in the wings	In the theatre, the wings are the sides of the stage that cannot be seen by the audience, where actors wait until it is their turn to step on the stage.
10	Play it by ear	From playing the music without reference to printed notation.
11	Bury the hatchet	From an American Indian tradition that hatchets were buried by the chiefs of tribes when they came to a peace agreement.
12	Have/get cold feet	Having a physical feeling of being unable to move forward or frozen in place, as if the feet are encased in ice.
13	Bark up the wrong tree	From raccoon hunting: Hunting dogs bark at the bottom of trees where they think their prey is, but the prey may have escaped to another tree.
14	Bite the bullet	From the practice of having a wounded soldier clench a bullet in his/her teeth as a way to cope with the extreme pain of a surgical procedure without anesthetic.
15	Let the cat out of the bag	From an old trick where one person pretended to sell a piglet in a bag to another, although the bag really contained a cat, which was discovered when the buyer got home and opened the bag.
16	Pull a rabbit out of the hat	This alludes to the magician's trick of pulling a rabbit out of an empty hat.
17	Red tape	Historically in Europe, legal and official documents were bound with red tape.
18	Beat around the bush	From hunting: In medieval times, hunters hired men to beat the area around bushes with sticks in order to flush out game taking cover underneath. They avoided hitting the bushes directly because this could sometimes prove dangerous.

19	Get your second wind	<i>Wind</i> here means 'breath' or 'easy or regular breathing'. This expression comes from running; after feeling out of breath at the beginning of a race, you later find it easier to breathe.
20	On the same wavelength	<i>Wavelength</i> refers to that of radio waves which carry a broadcast.
21	On the same page	As if reading from the same page (in order to understand what is being discussed).
22	Cut corners	From driving: When you come to a sharp turn in the road, instead of going all the way to the corner, you go straight across and cut off the corner to save time by shortening the distance of travel.
23	In the wake of something	<i>Wake</i> refers to the track of waves left by a ship or other object moving through the water.
24	Take the bull by the horns	From a common, but dangerous, practice of wrestling with steers in the American West. To control a bull or a steer, the cowhand would first have to catch it. Trying to grab the neck or legs of a dangerous creature like this was not an option. The only solution was to take a deep breath and directly grab the bull by the horns and then pull it to the ground.
25	Pull one's weight	From rowing, where each crew member must pull on an oar at least enough to propel himself or herself.
26	Weigh someone down	To burden something by attaching additional weight or placing it on top.
27	Hit the roof/ceiling*	Reach the top height of a room.
28	Give someone a leg up	When a rider needs help mounting a large horse, he might ask someone for a leg up. That person will then create a foothold by cupping both hands so that the rider can use this to step up and get into a position to get his leg up and over the horse's back.
29	Tighten our belt	As a method to stave off hunger.
30	Take a back seat	The back seat in contrast to the driver's seat in a vehicle
31	In the driving/drivers' seat	Being in the position to drive and control the vehicle.
32	Have a green thumb/ fingers	<i>Green fingers</i> refer to the green stains you'll get under the fingernails and on the hands from handling and cutting lots of plants.
33	(Not) up to scratch	In the days of bare-knuckle fighting, bouts took place within a large circle drawn on the bare ground. The referee would draw a line called the "scratch" in the centre of the ring. If a contestant was knocked down, he had to stand up and walk to the line within a certain time limit before the match could continue. If he could not do this, it was said that he was "not up to the scratch" and his opponent would win the fight.
34	Burn the candle at both ends	A candle lit at both ends will burn up more than twice as quickly. In 18th-century England, the phrase literally meant: You'd be wasting a valuable candle if you attempted to burn it at both ends.
35	Pass the buck	In poker, the buck was a marker or object which was passed to the person whose turn it was to deal the next hand. This person could

		either keep the marker or pass it on, in order to avoid dealing and being responsible for declaring the first stake.
36	A feeding frenzy	A feeding frenzy is originally an occasion when a group of sharks or other fish attack and eat something. If hungry animals have a feeding frenzy, they become very excited by the smell of food and fight each other to get a share of it.
37	A drop in the bucket	A drop (of water) into an empty bucket
38	A red herring	From the custom of using the scent of a smoked, dried (red) herring to rain dogs to hunt. The fish smells so strong that it can even distract bloodhounds from hunting the hares.
39	Ring a bell	There are many bells that ring to remind or instruct us to do things: doorbell (open the door), telephone bell (pick up the phone), school bell (come to class), toaster bell (take out the toast), and the clothes dryer bell (take out the clothes).
40	(Be) a sitting duck	From hunting, where a hunter can shoot a duck that remains in one spot, in contrast to one in flight.
41	Teething problems	When babies are getting their first set of teeth, they are often in pain and cry a lot.
42	Jump ship	Of nautical origin: If sailors jump ship, they leave their ship without permission and do not return. (Sailors when joining a ship are required to "sign on" for the duration of the voyage. They are "bound" to the ship and ordinarily cannot leave it until the voyage ends or the ship returns to home port.)
43	Flex your muscles	Athletes stretch and tighten their muscles before a race, a fight, a game, etc.
44	A loose cannon	Of nautical origin, referring to a cannon mounted on the deck of a sailing ship, which if dislodged during combat or a storm could cause serious damage to both vessel and crew by sliding about.
45	Get something off your chest	From card playing: You usually conceal your hand of cards from the other player by holding them against your chest.
46	Play your cards close to your chest	From card playing: as if one were playing cards and not permitting anyone to see any of the cards.
47	Stick your neck out	When a turtle sticks its neck (and head) out, it becomes more vulnerable (in the open) to predators. But a turtle must leave the safe haven of its shell to eat and such.
48	On the back burner	From cooking on a stove with several burners of varying heat: food cooking at a lower temperature on a back burner receives or requires less frequent attention than that cooking at a high temperature on a front burner.
49	Go with the flow	The <i>flow</i> refers to the ebb and flow of tides
50	Hand over fist	Of nautical origin: referring to sailor moving his hands steadily one over the other while pulling in a rope or raising a sail.

51	Leave someone high and dry	Of nautical origin: referring to ships that were beached. The word 'dry' implies that, not only were they out of the water, but had been for some time and could be expected to remain so.
52	Pass the baton on	The phrase refers to a relay race in which one runner literally passes a baton to the next runner.
53	Wet blanket	A wet blanket is a means to smother/extinguish fires.
54	Win hands down	From horse races: Jockeys need to keep a tight rein in order to encourage their horse to run. Anyone who is so far ahead that he can afford to slacken off and still win can drop his hands and loosen the reins - hence winning 'hands down'.
55	A rule of thumb	It comes from the fact that people, e.g., tailors and carpenters, often used their thumbs to estimate measurement.
56	In the doldrums	Doldrums are the part of the world's seas near the equator where there is little wind, making it difficult to sail. (The sailing ships were often becalmed/easily stranded due to lack of wind).
57	Rub someone the wrong way	From stroking an animal against the lie of its fur.
58	Ruffle someone's feathers	Displaying ruffled feathers is a defense mechanism that enables a chick to look larger and more aggressive. To actually ruffle a bird's feathers, which normally you would do only on a living bird, you disturb their nicely arranged position on the bird's body. Birds don't like their feathers ruffled by anyone but themselves.
59	(Hit) below the belt	From boxing: It refers to the rule that forbids boxers from hitting each other below the waist.
60	A dark horse	From horse racing: It was first recorded in 1831 in a book by Benjamin Disraeli. He described an incident in a horse race that a dark-colored horse "which had never been thought of...rushed past the grand stand in sweeping triumph."
61	Down and out	From boxing, where a boxer is knocked down and stays down for a given time, thereby losing the bout.
62	Bread and butter	As a basic food (in western culture), bread spread with butter.
63	A shot in the arm	The phrase refers to an injection of a drug.
64	On a shoestring	One story is that debtors in British prisons would lower a shoe by its laces from a window so as to collect funds from visitors or passers-by.
65	A can/bag of worms	Bait stores routinely sold cans of worms to fishermen, who often discovered how easy it was to open them and how difficult it was to close them. Once the worms discovered an opportunity to escape, it became nearly impossible to keep them contained.
66	Throw your hand in	When you have no chance of winning a game of cards, you throw your hand (=cards) into the middle of the table.
67	Jump in (at) the deep end	From swimming: This phrase refers to the deep end of a swimming pool, where it is too deep to stand.

68	Make ends meet	Making ends meet at the bottom lines of the income and losses columns in the bookkeeping.
69	Go belly-up	This alludes to the posture of a dead fish in the water.
70	Get into gear	From driving a car: to move the lever that controls the gears into a position that allows the vehicle to begin moving.
71	Take the plunge	From diving in a body of water, i.e., to cast or thrust oneself forcibly or suddenly into water.
72	Follow suit	A suit is one of the four groups of playing cards:     . This phrase comes from card games in which one must play a card from the same suit as the one led. If you follow suit when you are playing a card game, you put down a card with the same type of symbol on it as the card put down by the person before you.
73	Play into someone's hands	From card playing: A part of the game's strategy is to force your opponent to play certain cards. If you manage to do so, then she or he is playing into your hands, giving you an advantage.
74	Give someone the cold shoulder	A shoulder is a cut of meat which includes the upper part of the animal's front leg. This expression refers to a medieval practice where important guests were given roast meat. Less important people were only given cold meat left over from previous meals.
75	A hot potato	If pass a hot baked potato from one person to another—no one wants to be left holding it for very long, because it will burn their fingers.
76	On automatic pilot	In aircraft, automatic pilot is a device which automatically keeps the plane on course without the need for the pilot to do much.
77	Take it on the chin	From boxing, where a boxer takes a direct and full brunt of a punch to the chin.
78	Be on the ropes	From boxing, where a weakened boxer is forced back to the ropes of the boxing ring and leaning against them for support.
79	Turn over a new leaf	A leaf is a fresh clean page. The expression alludes to turning the page of a book to a new page.
80	The ball is in your court	The court here refers to an area in which you play sports such as tennis. In the game of tennis, if the ball is in your court then it is your turn to hit the ball.
81	Hold your horses	Historically related to horse riding or travelling by horse, it literally means to keep your horse(s) still.
82	(Start) from scratch	From the sporting world, where 'scratch' has been used since the 18th century to describe a starting line that was scratched on the ground.
83	Throw your hat/cap in the ring	The ring here refers to a boxing ring—a circular space in a crowd of onlookers. The idiom comes from a custom in America that one who wished to challenge a boxer would throw his hat into the ring as a way of having his challenge noticed by the crowd.
84	(Be) par for the course	In golf, par is the number of strokes a good player would expect to take in order to get it into all the holes on a particular golf course.

Appendix 8: Directions for Interview One

访谈指南

In this experiment we are interested in what you think about when you figure out the meanings of 21 English idioms. An idiom is an expression or phrase that doesn't mean what it says. It often has a figurative meaning, which is derived from its literal meaning. I'm going to ask you to THINK ALOUD as you guess the figurative meaning of the idioms.

What I mean by *thinking aloud* is that I want you to tell me EVERYTHING you are thinking from the time you first see the idiom until you tell me what it means. Some questions going through your mind after you see the idiom might be: Does a certain word give away the meaning of the idiom? How does the context help with guessing? (I will show you an example of the idiom) Is there a similar expression in my native language? During the process, I will also provide you with some hints about the literal meaning or the etymological source of the idiom to help you work out the meaning. Please tell me anything you can think of based on the prompts I give you during the interview process. You will get to know the figurative meanings in the end.

If you are familiar with the idiom, let me know from the start. If you don't understand a word in the expression, tell me and I will explain it to you at the beginning.

Appendix 9: Semantic Transparency Rating Form (Set A)

Instructions

Below are the 21 idioms that we have discussed.

You may feel that the figurative (idiomatic) meaning of some idioms follows in a straightforward manner from the literal (source) meaning of the expressions. For other idioms, you may find it much less obvious how the figurative (idiomatic) meaning is derived from the literal (source) meaning. You may even feel that there is no clear link at all.

Now that you know the literal meaning as well as the figurative meaning of the idioms, please circle the number on the scale below that best represents how you feel about the semantic relationship between the literal meaning and the figurative meaning of the idiom. You can also explain your choice on the lines behind the scales.

5 = The link between the literal and the figurative meaning of the expression is very clear to me. Given the literal use of the expression, it is easy for me to see how the figurative use is derived from it.

0 = The link between the literal and the figurative meaning of the expression is not at all clear to me. Although I (now) know both the literal (source) meaning of the expression and its idiomatic meaning, I cannot see how the literal use could have given rise to the figurative use of the expression.

non-transparent

transparent



0	1	2	3	4	5
The link between the literal and figurative meanings is not transparent at all.	The link between the literal and figurative meanings is very vague.	The link between the literal and figurative meanings is vague.	The link between the literal and figurative meanings is somewhat clear.	The link between the literal and figurative meanings is clear.	The link between the literal and figurative meanings is very clear.

give the green light	0	1	2	3	4	5	_____
on the back burner	0	1	2	3	4	5	_____
follow suit	0	1	2	3	4	5	_____
take it on the chin	0	1	2	3	4	5	_____
get into gear	0	1	2	3	4	5	_____
(be) a sitting duck	0	1	2	3	4	5	_____

be par for the course	0	1	2	3	4	5	_____
on the same wavelength	0	1	2	3	4	5	_____
bread and butter	0	1	2	3	4	5	_____
pull one's weight	0	1	2	3	4	5	_____
jump ship	0	1	2	3	4	5	_____
a dark horse	0	1	2	3	4	5	_____
go belly-up	0	1	2	3	4	5	_____
play into someone's hand	0	1	2	3	4	5	_____
red tape	0	1	2	3	4	5	_____
a drop in the bucket	0	1	2	3	4	5	_____
throw in the towel	0	1	2	3	4	5	_____
spill the beans	0	1	2	3	4	5	_____
Have/got cold feet	0	1	2	3	4	5	_____
take the bull by the horns	0	1	2	3	4	5	_____
burn the candle at both ends	0	1	2	3	4	5	_____

Name: _____

Age: _____

Gender: ☐ male ☐ female

TEM4 score: _____

Appendix 10: Meaning Recall Test Sheet (set A)

Please write down the meaning of the following idioms.

give the green light	_____
on the back burner	_____
follow suit	_____
take it on the chin	_____
get into gear	_____
(be) a sitting duck	_____
be par for the course	_____
on the same wavelength	_____
bread and butter	_____
pull one's weight	_____
jump ship	_____
a dark horse	_____
go belly-up	_____
play into someone's hands	_____
red tape	_____
a drop in the bucket	_____
throw in the towel	_____
spill the beans	_____
Have/got cold feet	_____
take the bull by the horns	_____
burn the candle at both ends	_____

Appendix 11: General questions for the second interview (record sheet)

General questions to answer:

- 1) What do you then do when encountering unfamiliar idioms (e.g. ignore them, make a guess at their meaning, look them up in a dictionary)?
- 2) When and where do you remember encountering such figurative idiomatic expressions in your own learning experience?
- 3) What do you often do to remember them (e.g. associate them with where you encountered them, associate them with a mental picture, compare them to a similar Chinese expression, write them down, try to use them)?
- 4) Since our first interview,
 - Have you invested any effort in trying to recall the idioms which we talked about? If so, please tell me about it.
 - Have you noticed any of these expressions being used since our first interview?
 - Have you noticed any other idioms?
If so, do you think you would have paid the same attention to them if we hadn't had that interview? If not, will you pay the same attention as before?
- 5) Generally, have you ever tried to use such figurative idiomatic expressions once you learn it?

Name _____ Date _____

Grammar _____ Vocabulary _____

Appendix 12: Excluded items by participants

Set	Item	Participant	Exclusion	Total cases	
Set A	bread and butter	A6	Known	9	16
	throw in the towel	A3			
	dark horse	A1—A7			
	burn the candle at both ends	A1—A7	Error	7	
Set B	beat around the bush	B5	Known	3	15
	weigh someone down	B5			
	let the cat out of the bag	B5			
	hit the roof/ceiling	B3, B6	Re- encountered	3	
	a shot in the arm	B4			
	beat around the bush	B4	Error	9	
	on the same page	B5			
	throw your hand in	B6			
	jump in (at) the deep end	B1—B6			
Set C	make ends meet	C1, C3	Known	3	3
	a rule of thumb	C3			
Set D	wet blanket	D3	Known	9	21
	have a green thumb	D2, D4, D6			
	a hot potato	D3			
	(start) from scratch	D1, D3, D4, D5			
	bark up the wrong tree	D1—D6	Error	12	
	on a shoestring	D1—D6			

Appendix 13: Breakdown figures of learning gains by participants

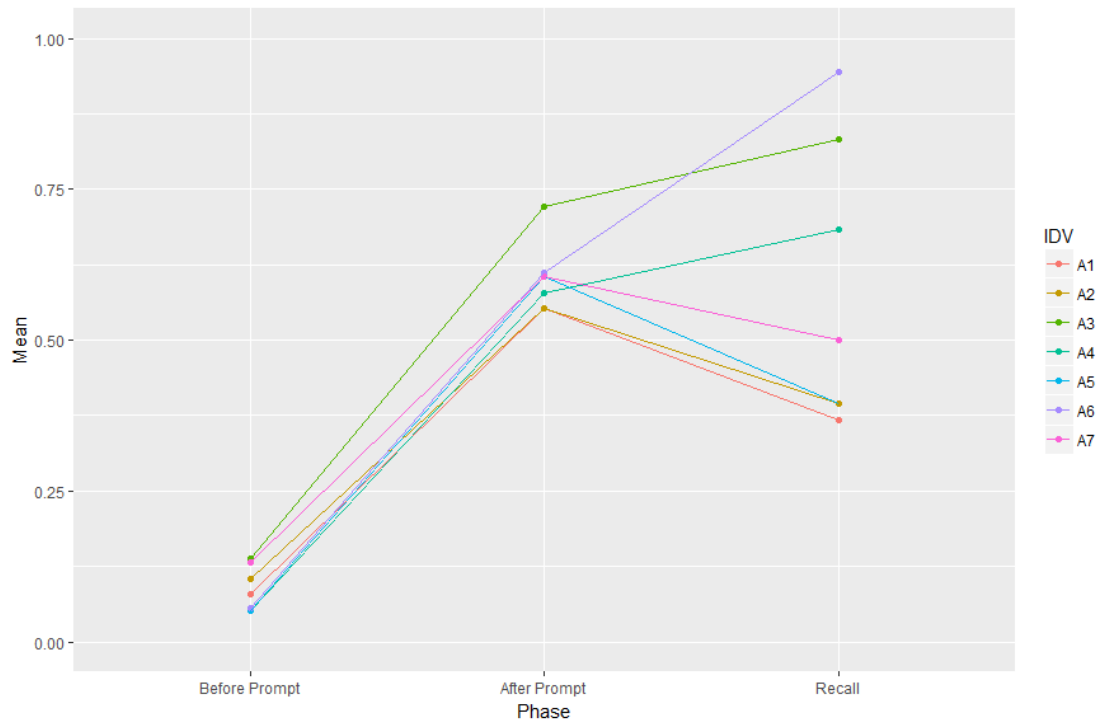


Figure 4.1.1 Change of mean scores between different phases for group A

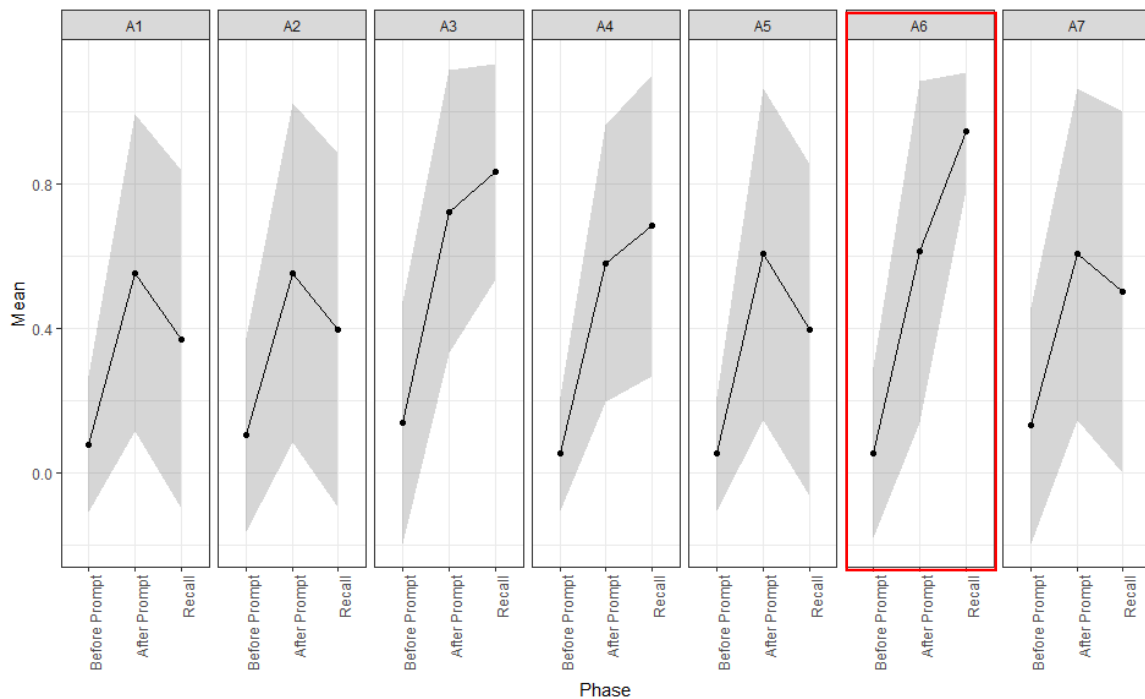


Figure 4.1.2 Individual change of mean scores³⁸ between different phases for group A

³⁸ Standard deviations are added to the means as ribbons in the figure.

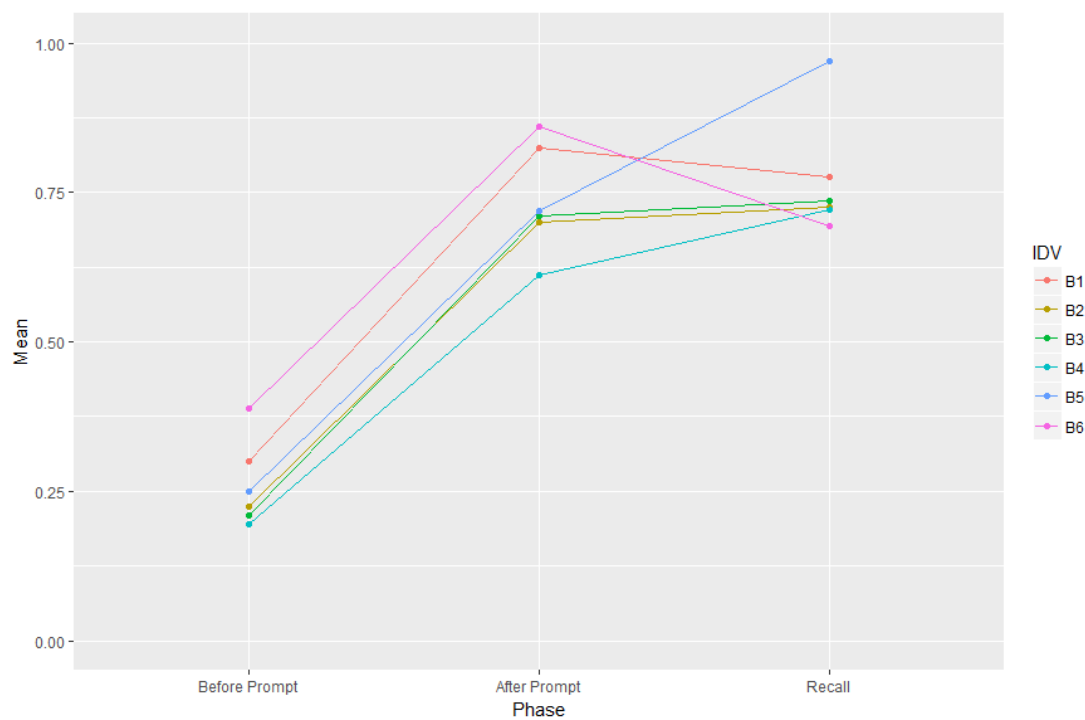


Figure 4.1.3 Change of mean scores between different phases for group B

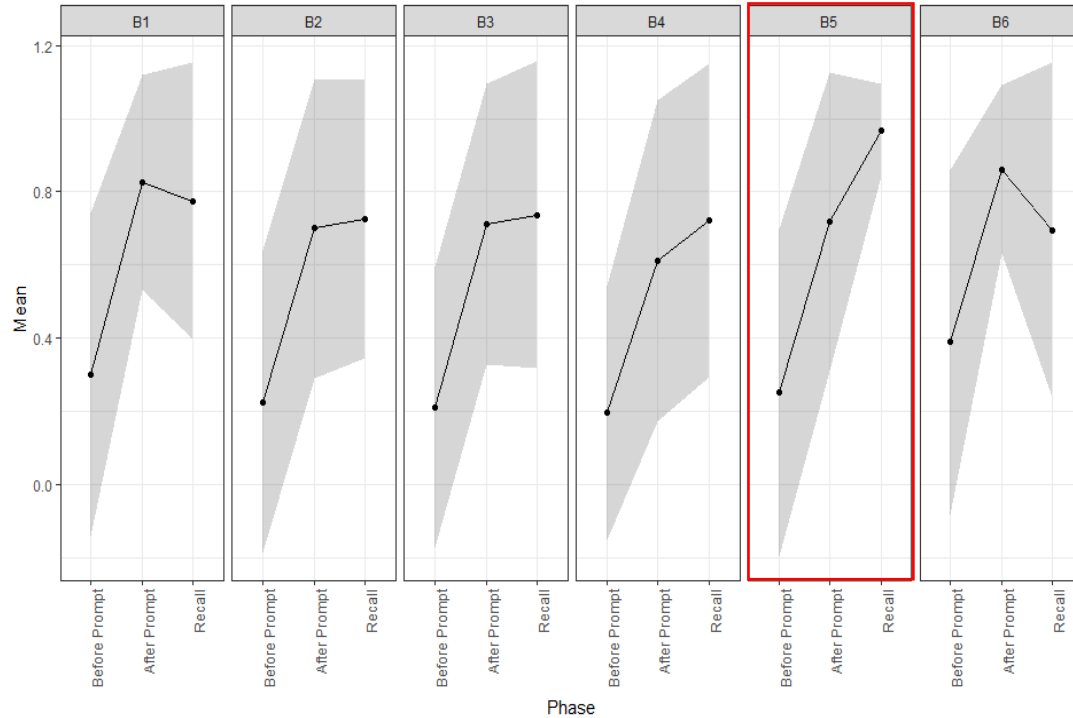


Figure 4.1.4 Individual change of mean scores between different phases for group B

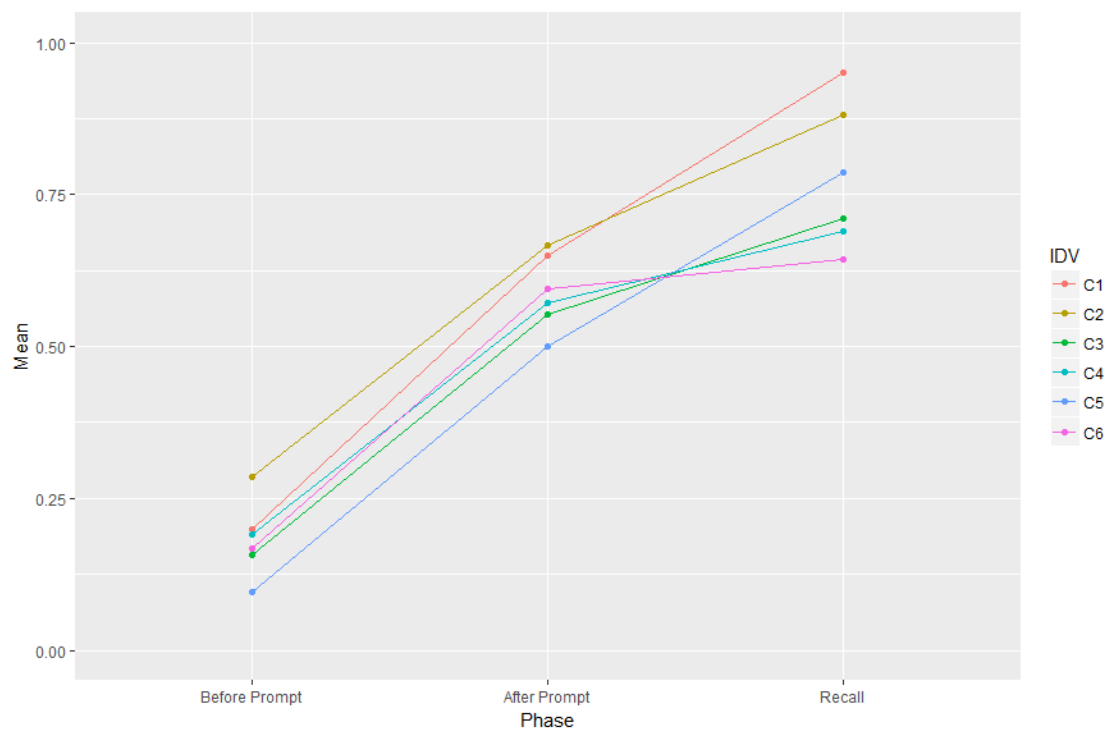


Figure 4.1.5 Change of mean scores between different phases for group C

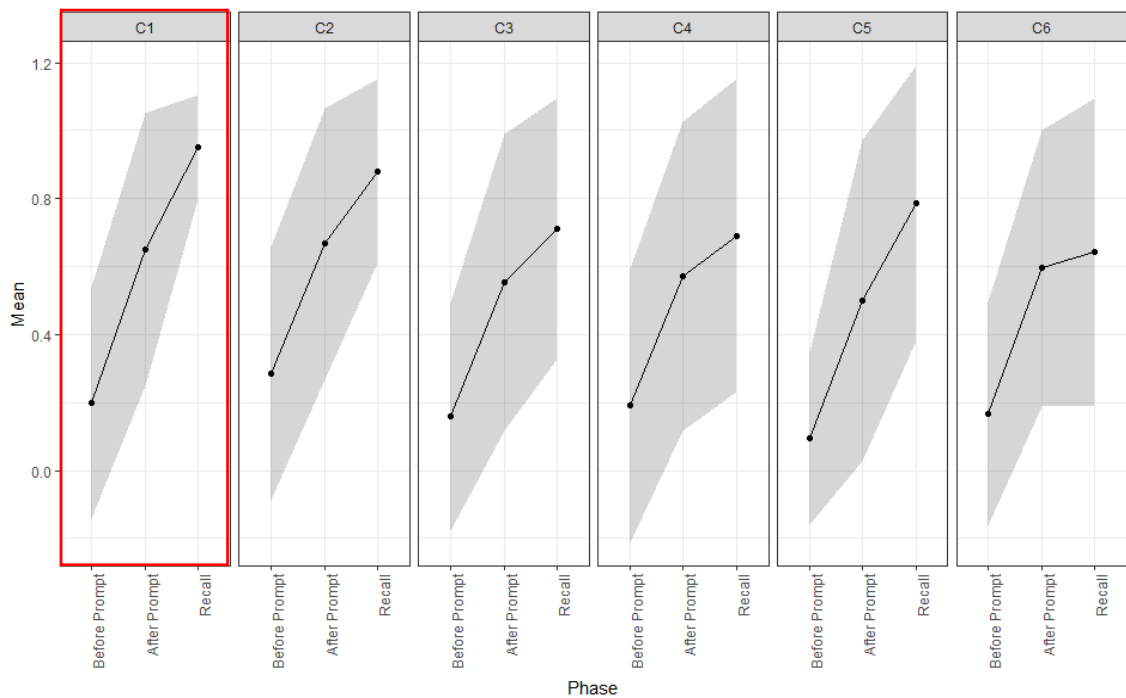


Figure 4.1.6 Individual change of mean scores between different phases for group C

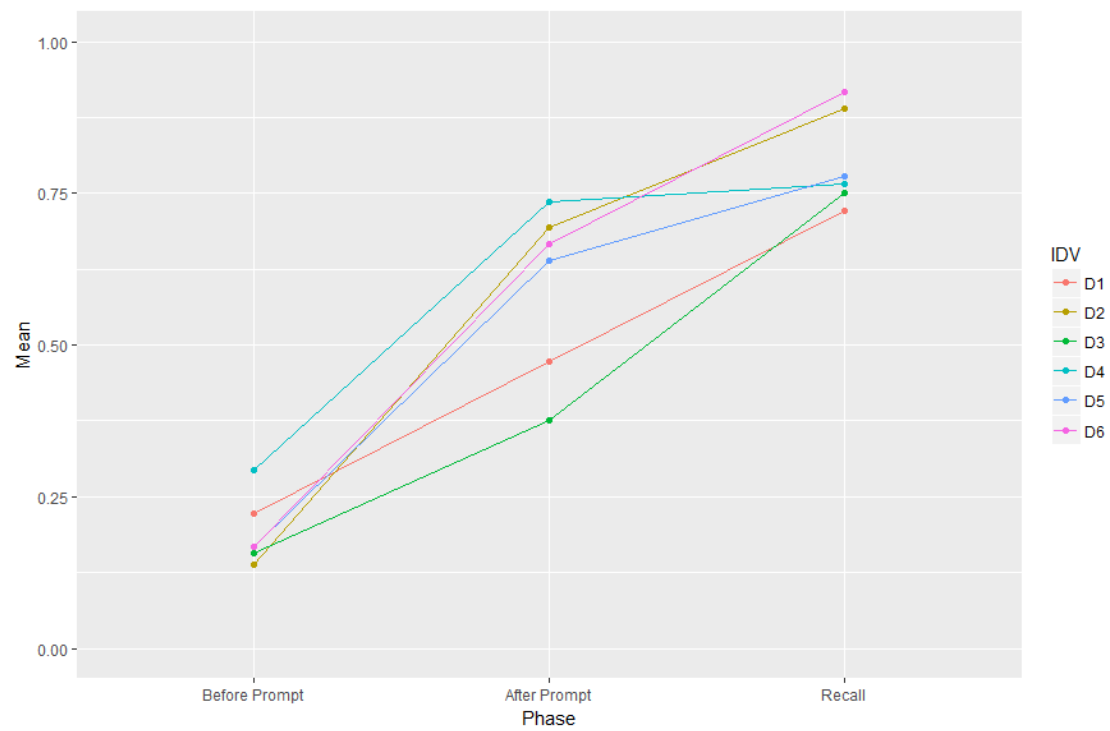


Figure 4.1.7 Change of mean scores between different phases for group D

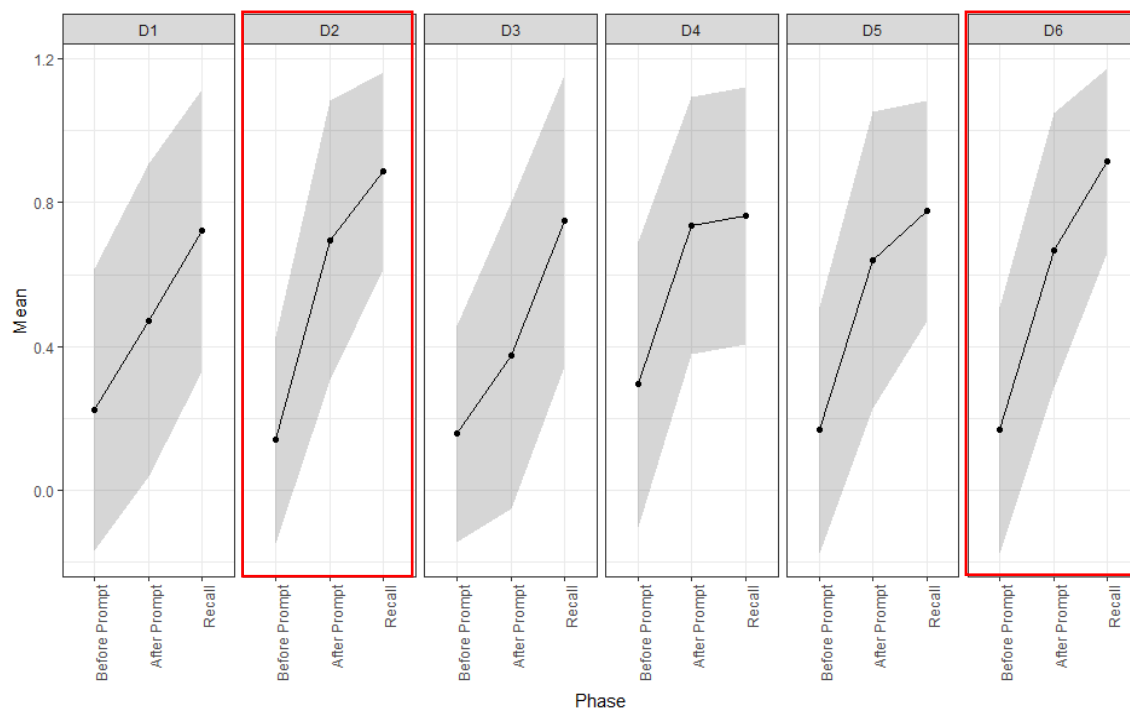


Figure 4.1.8 Individual change of mean scores between different phases for group D

Appendix 14: The codes of idioms

Code	Set A	Code	Set B
a01	give the green light	b01	take a back seat
a02	on the back burner	b02	a shot in the arm
a03	follow suit	b03	beat around the bush
a04	take it on the chin	b04	a feeding frenzy
a05	get into gear	b05	win hands down
a06	(be) a sitting duck	b07	give someone the cold shoulder
a07	be par for the course	b08	in the doldrums
a08	on the same wavelength	b09	on the same page
a09	bread and butter	b10	pass on the baton
a10	pull one's weight	b11	throw your hand in
a11	jump ship	b12	be waiting in the wings
a13	go belly-up	b13	(not) up to scratch
a14	play into someone's hands	b14	a loose cannon
a15	red tape	b15	leave someone high and dry
a16	a drop in the bucket	b16	be on the ropes
a17	throw in the towel	b17	hit the roof/ceiling
a18	spill the beans	b19	weigh someone down
a19	have/got cold feet	b20	let the cat out of the bag
a20	take the bull by the horns	b21	The ball is in your court.
	Set C		Set D
c01	jump the gun	d01	come out of your shell
c02	in the driving/driver's seat	d02	flex your muscles
c03	show someone the ropes	d03	wet blanket
c04	sit on the fence	d04	cut corners
c05	in the wake of something	d05	hand over fist
c06	bury the hatchet	d07	get your second wind
c07	teething problems	d08	bite the bullet
c08	take the plunge	d09	go with the flow
c09	rub someone the wrong way	d10	a can of worms
c10	a rule of thumb	d11	give someone a leg up
c11	(hit) below the belt	d12	have a green thumb
c12	a red herring	d13	a hot potato
c13	hold your horses	d14	play it by ear
c14	stick your neck out	d15	throw your hat/cap into the ring
c15	down and out	d16	get something off your chest
c16	have a lot on your plate	d17	on automatic pilot
c17	pass the buck	d20	(start) from scratch
c18	play your cards close to your chest	d21	pull a rabbit out of the hat
c19	ring a bell		
c20	turn over a new leaf		
c21	make ends meet		

Appendix 15: Breakdown figures of learning gains by idioms

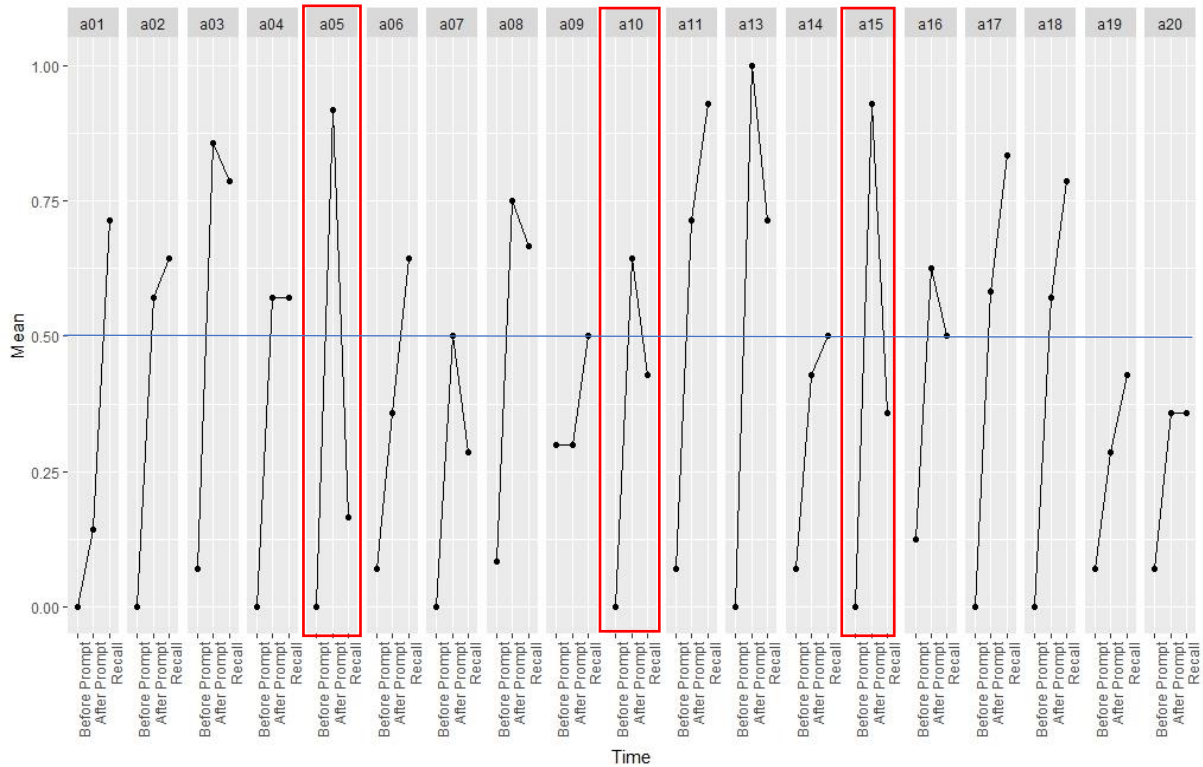


Figure 4.2.1 Change of mean scores over time for idioms in set A

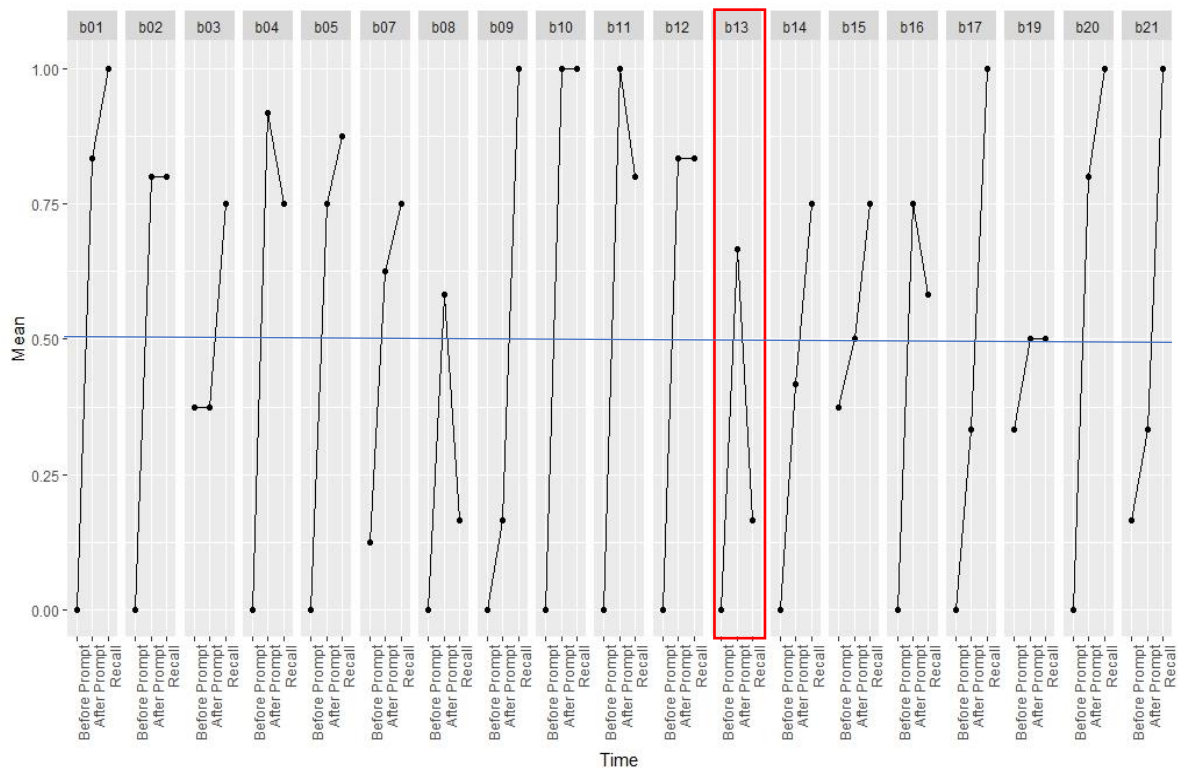


Figure 4.2.2 Change of mean scores over time for idioms in set B

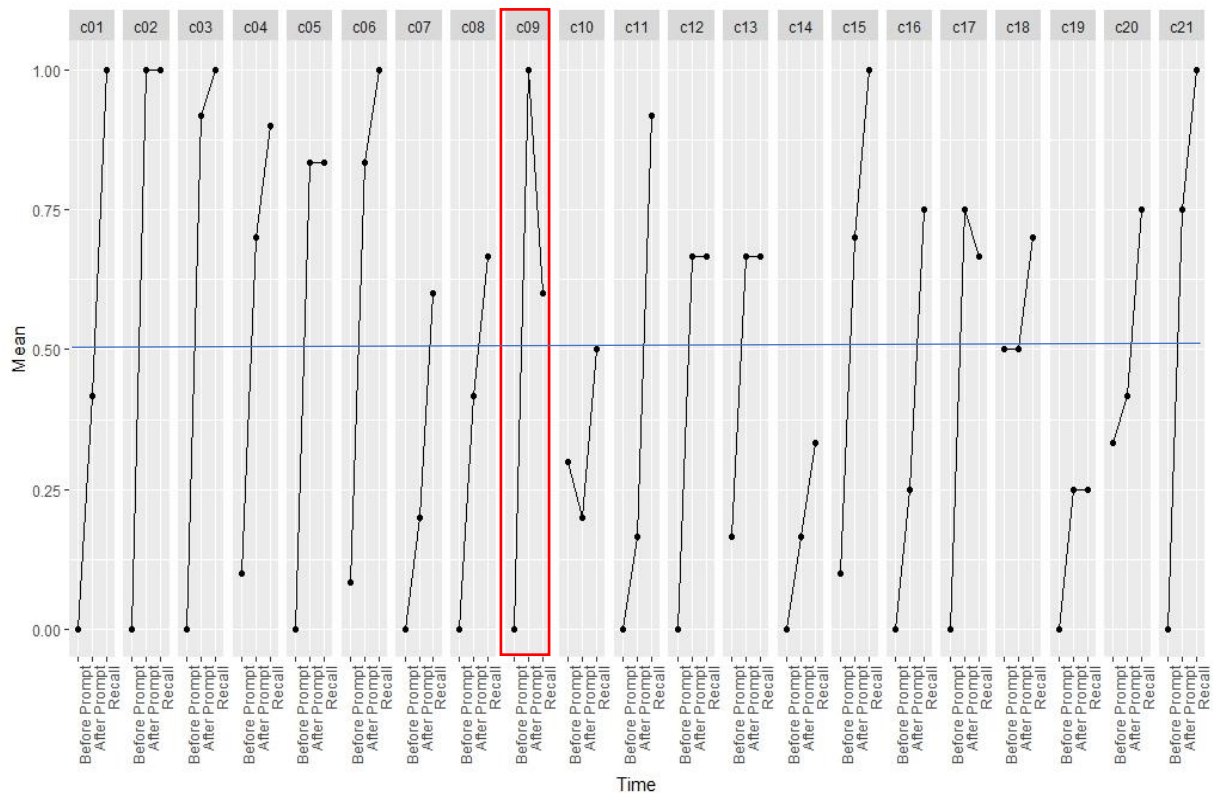


Figure 4.2.3 Change of mean scores over time for idioms in set C

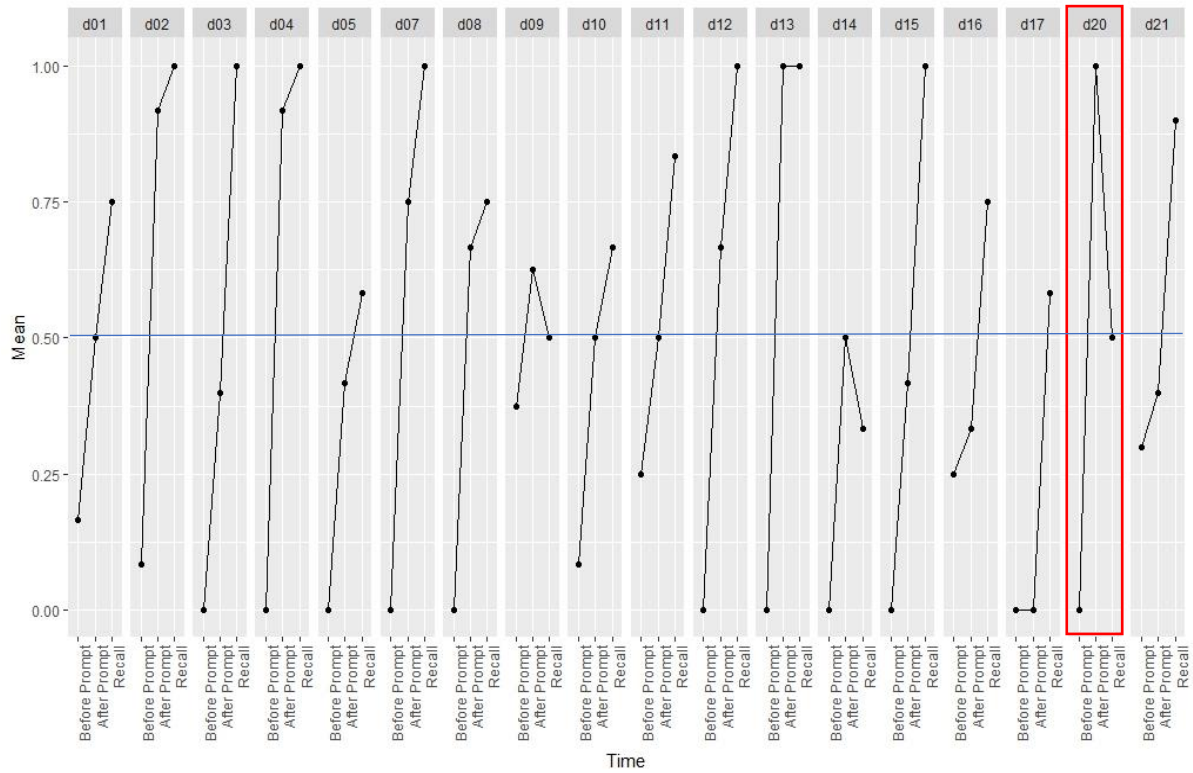


Figure 4.2.4 Change of mean scores over time for idioms in set D

Appendix 16: Descriptive data of etymological transparency ratings & correctness of responses of all the idioms (Total N = 470³⁹)

SET	IDIOM	N	BP ⁴⁰ =1	BP= 0.5	BP= 0	TRANS ⁴¹ Mean	SD	Recall ⁴² Mean	SD
A	go belly-up	7	0	0	7	4.86	0.38	0.71	0.49
A	a drop in the bucket	7	3	1	3	4.71	0.76	0.71	0.49
A	on the same wavelength	7	1	1	5	4.71	0.76	0.71	0.49
A	get into gear	7	1	0	6	4.43	0.79	0.14	0.38
A	red tape	7	0	0	7	4.43	0.53	0.36	0.48
A	follow suit	7	0	1	6	4.29	0.49	0.79	0.39
A	pull one's weight	7	0	0	7	4.29	0.76	0.43	0.53
A	bread and butter	6	1	3	2	4.17	0.75	0.58	0.49
A	throw in the towel	6	0	0	6	4.17	0.75	0.83	0.41
A	(be) a sitting duck	7	0	1	6	4.14	0.69	0.64	0.48
A	give the green light	7	0	0	7	4.14	0.69	0.71	0.49
A	spill the beans	7	0	0	7	4.14	0.90	0.79	0.39
A	have/got cold feet	7	0	1	6	4.00	1.15	0.43	0.19
A	jump ship	7	0	1	6	4.00	1.00	0.93	0.19
A	play into someone's hands	7	0	1	6	4.00	1.15	0.50	0.50
A	be par for the course	7	0	0	7	3.57	0.98	0.29	0.39
A	on the back burner	7	0	0	7	3.57	1.40	0.64	0.48
A	take the bull by the horns	7	0	1	6	3.57	0.98	0.36	0.48
A	take it on the chin	7	0	0	7	3.14	1.57	0.57	0.53
B	ruffle someone's feathers	6	6	0	0	4.67	0.52	0.83	0.41
B	The ball is in your court.	6	3	1	2	4.67	0.82	1.00	0.00
B	win hands down	6	2	0	4	4.67	0.82	0.92	0.20
B	be on the ropes	6	0	0	6	4.67	0.52	0.58	0.49
B	weigh someone down	5	2	2	1	4.60	0.55	0.40	0.42
B	pass the baton on	6	4	0	2	4.50	0.55	1.00	0.00
B	give someone the cold shoulder	6	2	1	3	4.50	0.55	0.83	0.26
B	leave someone high and dry	6	2	3	1	4.33	0.82	0.83	0.26
B	in the doldrums	6	0	0	6	4.33	0.82	0.17	0.41
B	beat around the bush	4	0	3	1	4.25	0.50	0.75	0.29
B	be waiting in the wings	6	0	0	6	4.17	0.41	0.83	0.41
B	a shot in the arm	5	0	0	5	4.00	0.71	0.80	0.45
B	throw your hand in	5	0	0	5	4.00	1.00	0.80	0.45

³⁹ This table covers all the 470 cases with 79 idioms (i.e., 55 invalid cases involving known items, re-encountered items and errors were excluded from analysis).

⁴⁰ BP = Correctness of guesses before the prompt/GuessBP

⁴¹ TRANS = Etymological semantic transparency/Transparency

⁴² Recall = Correctness of recall responses

B	a loose cannon	6	0	0	6	3.83	1.47	0.75	0.27
B	on the same page	5	2	0	3	3.80	1.30	1.00	0.00
B	let the cat out of the bag	5	0	0	5	3.60	1.67	1.00	0.00
B	(not) up to scratch	6	0	0	6	3.50	1.05	0.17	0.41
B	take a back seat	6	0	0	6	3.50	1.38	1.00	0.00
B	hit the roof/ceiling	4	1	0	3	3.25	1.50	1.00	0.00
B	a feeding frenzy	6	0	0	6	3.17	0.75	0.75	0.42
C	in the driving/driver's seat	6	5	0	1	4.83	0.41	1.00	0.00
C	play your cards close to your chest	6	1	5	0	4.83	0.41	0.75	0.27
C	down and out	6	1	1	4	4.67	0.82	1.00	0.00
C	rub someone the wrong way	6	1	0	5	4.67	0.52	0.67	0.52
C	turn over a new leaf	6	0	4	2	4.67	0.82	0.75	0.27
C	(hit) below the belt	6	0	0	6	4.50	1.22	0.92	0.20
C	make ends meet	4	0	0	4	4.50	0.58	1.00	0.00
C	ring a bell	6	2	0	4	4.33	1.21	0.50	0.55
C	a red herring	6	0	0	6	4.33	0.82	0.67	0.41
C	bury the hatchet	6	0	1	5	4.33	0.52	1.00	0.00
C	have a lot on your plate	6	2	0	4	4.00	0.89	0.83	0.41
C	sit on the fence	6	1	1	4	4.00	1.55	0.92	0.20
C	pass the buck	6	0	0	6	4.00	0.89	0.67	0.52
C	jump the gun	6	0	0	6	3.83	1.17	1.00	0.00
C	stick your neck out	6	0	0	6	3.83	0.75	0.33	0.52
C	take the plunge	6	0	0	6	3.83	0.98	0.67	0.52
C	teething problems	6	1	0	5	3.67	1.75	0.67	0.52
C	hold your horses	6	0	2	4	3.67	1.51	0.67	0.26
C	in the wake of something	6	0	0	6	3.67	1.03	0.83	0.41
C	a rule of thumb	5	0	3	2	3.40	0.89	0.50	0.50
C	show someone the ropes	6	0	0	6	3.33	0.82	1.00	0.00
D	tighten your belt	6	6	0	0	5.00	0.00	1.00	0.00
D	a hot potato	5	3	0	2	5.00	0.00	1.00	0.00
D	have a green thumb	3	0	0	3	4.67	0.58	1.00	0.00
D	go with the flow	6	2	3	1	4.50	0.84	0.58	0.20
D	(start) from scratch	2	0	0	2	4.50	0.71	0.50	0.71
D	a can of worms	6	0	1	5	4.17	0.75	0.67	0.41
D	flex your muscles	6	0	1	5	4.17	0.98	1.00	0.00
D	give someone a leg up	6	0	3	3	4.17	0.98	0.83	0.26
D	pull a rabbit out of the hat	6	1	3	2	4.00	0.89	0.92	0.20
D	bite the bullet	6	0	0	6	4.00	0.63	0.75	0.42
D	get your second wind	6	0	0	6	4.00	0.89	1.00	0.00
D	cut corners	6	0	0	6	3.83	0.75	1.00	0.00
D	on automatic pilot	6	0	0	6	3.83	0.98	0.58	0.49
D	come out of your shell	6	0	2	4	3.67	1.03	0.75	0.27
D	get something off your chest	6	0	3	3	3.67	0.52	0.75	0.42
D	hand over fist	6	0	0	6	3.67	1.21	0.58	0.38

D	throw your hat/cap into the ring	6	0	0	6	3.67	1.03	1.00	0.00
D	play it by ear	6	0	0	6	3.17	0.98	0.33	0.52
D	wet blanket	5	0	0	5	2.80	1.92	1.00	0.00
	TOTAL	470	56	54	360				

Appendix 17: CLMM output for Chapter 5

```
> Datasub<-subset(Data,GuessBP!=1)
> Model_test1<-clmm(Transparency ~ GuessAP * TEM4 + (1|Participant) + (1|Idiom),
  Datasub)
> summary(Model_test1)
```

Cumulative Link Mixed Model fitted with the Laplace approximation

formula: Transparency ~ GuessAP * TEM4 + (1 | Participant) + (1 | Idiom)

data: Datasub

link	threshold	nobs	logLik	AIC	niter	max.grad	cond.H
logit	flexible	414	-502.39	1028.77	1045(3151)	1.00e-03	1.2e+07

Random effects:

Groups	Name	Variance	Std.Dev.
Idiom	(Intercept)	0.2495	0.4995
Participant	(Intercept)	0.8069	0.8983

Number of groups: Idiom 77, Participant 25

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
GuessAP0.5	-3.62032	3.18737	-1.136	0.2560
GuessAP1	-2.23511	2.91940	-0.766	0.4439
TEM4	-0.09434	0.04131	-2.284	0.0224 *
GuessAP0.5:TEM4	0.06218	0.04234	1.469	0.1420
GuessAP1:TEM4	0.04748	0.03896	1.219	0.2229

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:

	Estimate	Std. Error	z value
0 1	-12.246	3.206	-3.819
1 2	-10.829	3.146	-3.442
2 3	-9.203	3.128	-2.943
3 4	-7.516	3.114	-2.414
4 5	-5.479	3.100	-1.768

```
> Model_test2<-clmm(Transparency ~ GuessAP + TEM4 + (1|Participant) + (1|Idiom),
  Datasub)
> summary(Model_test2)
```

formula: Transparency ~ GuessAP + TEM4 + (1 | Participant) + (1 | Idiom)

data: Datasub

	link	threshold	nobs	logLik	AIC	niter	max.grad	cond.H
logit	flexible		414	-503.59	1027.17	530(1593)	4.69e-05	4.0e+06

Random effects:

Groups	Name	Variance	Std.Dev.
Idiom	(Intercept)	0.2580	0.5080
Participant	(Intercept)	0.8161	0.9034

Number of groups: Idiom 77, Participant 25

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
GuessAP0.5	1.03564	0.27669	3.743	0.000182 ***
GuessAP1	1.30243	0.25892	5.030	4.9e-07 ***
TEM4	-0.05690	0.03288	-1.730	0.083564 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:

	Estimate	Std. Error	z value
0 1	-9.458	2.605	-3.630
1 2	-8.042	2.532	-3.177
2 3	-6.416	2.509	-2.557
3 4	-4.733	2.497	-1.896
4 5	-2.703	2.486	-1.087

> anova(Model_test1, Model_test2)

Likelihood ratio tests of cumulative link models:

	formula:	link:	threshold:
Model_test2	Transparency ~ GuessAP + TEM4 + (1 Participant) + (1 Idiom)	logit	
		flexible	
Model_test1	Transparency ~ GuessAP * TEM4 + (1 Participant) + (1 Idiom)	logit	
		flexible	

	no.par	AIC	logLik	LR.stat	df	Pr(>Chisq)
Model_test2	10	1027.2	-503.59			
Model_test1	12	1028.8	-502.39	2.4012	2	0.301

##The interaction between GuessAP and TEM4 is not significant.

> Anova.climm(Model_test1, type = 3)

Analysis of Deviance Table (Type II tests)

Response: Transparency

	LR	Chisq	Df	Pr(>Chisq)
GuessAP		1.3204	2	0.5167

```
TEM4          0.0000  1    1.0000
GuessAP:TEM4  2.4012  2    0.3010
```

```
> Anova.clm(Model_test1, type = 2)
```

Analysis of Deviance Table (Type II tests)

Response: Transparency

```
          LR Chisq Df Pr(>Chisq)
GuessAP    27.4750  2  1.081e-06 ***
TEM4        2.8706  1  0.09021 .
GuessAP:TEM4 2.4012  2  0.30101
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```
> Model_test3<-clmm(Transparency ~ GuessAP + (1|Participant) + (1|Idiom), Datasub)
```

```
> summary(Model_test3)
```

Cumulative Link Mixed Model fitted with the Laplace approximation

formula: Transparency ~ GuessAP + (1 | Participant) + (1 | Idiom)

data: Datasub

```
link threshold nobs logLik AIC      niter      max.grad cond.H
logit flexible 414 -505.02 1028.05 534(1605) 5.63e-05 8.4e+01
```

Random effects:

```
Groups Name      Variance Std.Dev.
Idiom (Intercept) 0.2660  0.5158
Participant (Intercept) 0.9368  0.9679
```

Number of groups: Idiom 77, Participant 25

Coefficients:

```
          Estimate Std. Error z value Pr(>|z|)
GuessAP0.5  1.0211    0.2772   3.683 0.00023 ***
GuessAP1    1.2912    0.2594   4.977 6.45e-07 ***
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:

```
          Estimate Std. Error z value
0|1 -5.1671    0.7589 -6.808
1|2 -3.7509    0.4478 -8.376
2|3 -2.1243    0.3191 -6.657
3|4 -0.4426    0.2873 -1.541
4|5  1.5884    0.2991  5.311
```

```
> anova(Model_test2, Model_test3)
```

Likelihood ratio tests of cumulative link models:

	formula:	link:	threshold:
Model_test3	Transparency ~ GuessAP + (1 Participant) + (1 Idiom)	logit	flexible
Model_test2	Transparency ~ GuessAP + TEM4 + (1 Participant) + (1 Idiom)	logit	flexible

	no.par	AIC	logLik	LR.stat	df	Pr(>Chisq)
Model_test3	9	1028.0	-505.02			
Model_test2	10	1027.2	-503.59	2.8706	1	0.09021

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##Difference between AICs was negligible. Simpler model (Model_test3) was kept.

Output for 5.2.2.1

```
> Model1 <- clmm(Transparency ~ GuessAP + (1|Participant) + (1|Idiom),
data=Datasub)
```

```
> summary(Model1)
```

Cumulative Link Mixed Model fitted with the Laplace approximation

formula: Transparency ~ GuessAP + (1 | Participant) + (1 | Idiom)

data: Datasub

	link	threshold	nobs	logLik	AIC	niter	max.grad	cond.H
logit flexible	414	-505.02	1028.05	534(1605)	5.63e-05	8.4e+01		

Random effects:

Groups	Name	Variance	Std.Dev.
--------	------	----------	----------

Idiom	(Intercept)	0.2660	0.5158
-------	-------------	--------	--------

Participant	(Intercept)	0.9368	0.9679
-------------	-------------	--------	--------

Number of groups: Idiom 77, Participant 25

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
GuessAP0.5	1.0211	0.2772	3.683	0.00023 ***
GuessAP1	1.2912	0.2594	4.977	6.45e-07 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:

	Estimate	Std. Error	z value
0 1	-5.1671	0.7589	-6.808
1 2	-3.7509	0.4478	-8.376
2 3	-2.1243	0.3191	-6.657
3 4	-0.4426	0.2873	-1.541
4 5	1.5884	0.2991	5.311

```
> Model1.0 <- clmm(Transparency ~ 1+(1|Particiapnt) + (1|Idiom), data=Datasub)
> anova(Model1,Model1.0)
```

Likelihood ratio tests of cumulative link models:

	formula:	link: threshold:
Model1.0	Transparency ~ 1 + (1 Participant) + (1 Idiom)	logit flexible
Model1	Transparency ~ GuessAP + (1 Participant) + (1 Idiom)	logit flexible

	no.par	AIC	logLik	LR.stat	df	Pr(>Chisq)
Model1.0	7	1050.9	-518.45			
Model1	9	1028.0	-505.02	26.859	2	1.471e-06 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Output for 5.2.2.2

```
> Model2<-clmm(Transparency ~ TEM4 + (1|Participant) + (1|Idiom), data=Datasub)
> summary(Model2)
```

Cumulative Link Mixed Model fitted with the Laplace approximation

formula: Transparency ~ TEM4 + (1 | Participant) + (1 | Idiom)

data: Datasub

link	threshold	nobs	logLik	AIC	niter	max.grad	cond.H
logit	flexible	414	-517.32	1050.65	397(1194)	2.06e-04	4.0e+06

Random effects:

Groups	Name	Variance	Std.Dev.
Idiom	(Intercept)	0.2772	0.5265
Participant	(Intercept)	0.7004	0.8369

Number of groups: Idiom 77, Participant 25

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
TEM4	-0.04727	0.03093	-1.528	0.126

Threshold coefficients:

	Estimate	Std. Error	z value
0 1	-9.424	2.470	-3.816
1 2	-8.009	2.392	-3.349
2 3	-6.413	2.367	-2.709
3 4	-4.829	2.355	-2.051
4 5	-2.902	2.344	-1.238

```
> anova(Model2, Model1.0)
```

Likelihood ratio tests of cumulative link models:

	formula:	link: threshold:
Model1.0	Transparency ~ 1 + (1 Participant) + (1 Idiom)	logit flexible
Model2	Transparency ~ TEM4 + (1 Participant) + (1 Idiom)	logit flexible

	no.par	AIC	logLik	LR.stat	df	Pr(>Chisq)
Model1.0	7	1050.9	-518.45			
Model2	8	1050.6	-517.32	2.2543	1	0.1332

Output for 5.2.2.3

```
> Model3 <- cllmm(GuessAP ~ TEM4 +(1|Participant) + (1|Idiom), data=Datasub)
> summary(Model3)
```

Cumulative Link Mixed Model fitted with the Laplace approximation

formula: GuessAP ~ TEM4 + (1 | Participant) + (1 | Idiom)

data: Datasub

link	threshold	nobs	logLik	AIC	niter	max.grad	cond.H
logit	flexible	414	-428.75	867.50	179(550)	1.26e-02	8.9e+05

Random effects:

Groups	Name	Variance	Std.Dev.
Idiom	(Intercept)	0.9748	0.9873
Participant	(Intercept)	0.0000	0.0000

Number of groups: Idiom 77, Participant 25

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
TEM4	0.01743	0.01741	1.001	0.317

Threshold coefficients:

	Estimate	Std. Error	z value
0 0.5	0.1549	1.3240	0.117
0.5 1	1.5612	1.3259	1.177

```
> Model3.0 <-clmm(GuessAP ~ 1 +(1|Participant) + (1|Idiom), data=Datasub)
> anova(Model3, Model3.0)
```

Likelihood ratio tests of cumulative link models:

	formula:	link:	threshold:
Model3.0	GuessAP ~ 1 + (1 Participant) + (1 Idiom)	logit	flexible
Model3	GuessAP ~ TEM4 + (1 Participant) + (1 Idiom)	logit	flexible

	no.par	AIC	logLik	LR.stat	df	Pr(>Chisq)
Model3.0	4	866.5	-429.25			
Model3	5	867.5	-428.75	1.0046	1	0.3162

Appendix 18: CLMM output for Chapter 6

Output for 6.1.2

```
> library(ordinal)
> Datasub<-subset(Data,GuessBP!=1)
> Datasub$TEM4s<-Datasub$TEM4-75.72 #TEM4s is obtained from TEM4 subtracted by the mean score (to get rid of the convergence error)
```

Model

Cumulative Link Mixed Model fitted with the Laplace approximation

formula: Recall ~ Transparency * TEM4s + (1 | Participant) + (1 | Idiom)
data: Datasub

link	threshold	nobs	logLik	AIC	niter	max.grad	cond.H
logit	flexible	414	-354.44	722.87	420(1662)	1.08e-04	1.3e+04

Random effects:

Groups	Name	Variance	Std.Dev.
Idiom	(Intercept)	1.1167	1.0567
Participant	(Intercept)	0.6505	0.8065

Number of groups: Idiom 77, Participant 25

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
Transparency	0.26423	0.13312	1.985	0.0471 *
TEM4s	0.27607	0.10951	2.521	0.0117 *
Transparency:TEM4s	-0.05033	0.02412	-2.086	0.0369 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:

	Estimate	Std. Error	z value
0 0.5	-0.6140	0.5834	-1.052
0.5 1	0.3776	0.5825	0.648

##ANOVA Test III for the significance of simple factors in the interaction model

```
> library(RVAideMemoire)
> Anova.clmm(Model, type = 3)
Analysis of Deviance Table (Type II tests)
```

Response: Recall

```
              LR Chisq Df Pr(>Chisq)
Transparency      4.0080  1  0.045284 *
TEM4s             7.1260  1  0.007597 **
Transparency:TEM4s 4.7228  1  0.029766 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model 1

Cumulative Link Mixed Model fitted with the Laplace approximation

formula: RC ~ Transparency + (1 | Participant) + (1 | Idiom)

data: Datasub

```
link threshold nobs logLik AIC niter max.grad cond.H
logit flexible 414 -358.54 727.09 199(802) 7.74e-06 2.2e+02
```

Random effects:

```
Groups      Name      Variance Std.Dev.
Idiom       (Intercept) 1.1565  1.0754
Participant (Intercept) 0.6919  0.8318
Number of groups: Idiom 77, Participant 25
```

Coefficients:

```
              Estimate Std. Error z value Pr(>|z|)
Transparency  0.2088      0.1283   1.627   0.104
```

Threshold coefficients:

```
              Estimate Std. Error z value
0|0.5 -0.8663      0.5641 -1.536
0.5|1  0.1144      0.5620  0.204
```

```
> Model0<-clmm(Recall ~ 1 + (1|Participant) + (1|Idiom), Datasub)
```

```
> anova(Model0, Model1)
```

Likelihood ratio tests of cumulative link models:

```
      formula:                                link: threshold:
Model0 RC ~ 1 + (1 | Participant) + (1 | Idiom)      logit flexible
Model1 RC ~ Transparency + (1 | Participant) + (1 | Idiom)logit flexible
```

```
      no.par   AIC logLik LR.stat df Pr(>Chisq)
Model0      4 727.76 -359.88
Model1      5 727.09 -358.54 2.6731 1  0.1021
```


Model 2

Cumulative Link Mixed Model fitted with the Laplace approximation

formula: Recall ~ TEM4s + (1 | Participant) + (1 | Idiom)

data: Datasub

```
link threshold nobs logLik AIC niter max.grad cond.H
logit flexible 414 -358.42 726.84 284(1059) 1.39e-01 2.0e+00
```

Random effects:

Groups	Name	Variance	Std.Dev.
Idiom	(Intercept)	1.0901	1.0441
Participant	(Intercept)	0.5881	0.7669

Number of groups: Idiom 77, Participant 25

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
TEM4s	0.05515	0.01632	3.38	0.000724 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:

	Estimate	Std. Error	z value
0 0.5	-1.68839	0.01958	-86.21
0.5 1	-0.71819	0.01958	-36.68

```
> Model0<-clmm(Recall ~ 1+(1|Participant) + (1|Idiom), Datasub)
```

```
> anova(Model0, Model2)
```

Likelihood ratio tests of cumulative link models:

	formula:	link:	threshold:
Model0	RC ~ 1 + (1 Participant) + (1 Idiom)	logit	flexible
Model2	RC ~ TEM4s + (1 Participant) + (1 Idiom)	logit	flexible

	no.par	AIC	logLik	LR.stat	df	Pr(>Chisq)
Model0	4	727.76	-359.88			
Model2	5	726.84	-358.42	2.9198	1	0.0875 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Output for 6.1.3

Model selection starting from the interaction model

```
Model_RC<-clmm(RC~GuessAP*TEM4s + (1|Participant) + (1|Idiom), Datasub)
```

```
> summary(Model_RC)
Cumulative Link Mixed Model fitted with the Laplace approximation
```

```
formula: RC ~ GuessAP * TEM4s + (1 | Participant) + (1 | Idiom)
data:    Datasub
```

```
link threshold nobs logLik AIC niter max.grad cond.H
logit flexible 414 -353.59 725.18 691(2644) 8.00e-04 3.6e+02
```

Random effects:

```
Groups      Name      Variance Std.Dev.
Idiom       (Intercept) 1.0366  1.018
Participant (Intercept) 0.6052  0.778
```

Number of groups: Idiom 77, Participant 25

```
Coefficients:      Estimate Std. Error z value Pr(>|z|)
GuessAP0.5         0.373074  0.326215  1.144 0.25277
GuessAP1           0.928049  0.318414  2.915 0.00356 **
TEM4s              0.056892  0.042953  1.325 0.18533
GuessAP0.5:TEM4s   0.004481  0.048507  0.092 0.92640
GuessAP1:TEM4s    -0.014580  0.045650 -0.319 0.74943
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:

```
      Estimate Std. Error z value
0|0.5 -1.2064    0.3175 -3.800
0.5|1 -0.2202    0.3064 -0.719
```

ANOVA Test II for the significance of interaction in the interaction model

```
> library(RVAideMemoire)
> Anova.clmm(Model_RC, type = 3)
Analysis of Deviance Table (Type II tests)
```

```
Response: RC
LR Chisq Df Pr(>Chisq)
GuessAP      8.7767  1  0.003051 **
TEM4s        2.2076  1  0.137331
GuessAP:TEM4s 0.1005  1  0.751259 ##Interaction not significant
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##Move on to a simple effect model

```
> Model_RC1<-clmm(RC~GuessAP + TEM4s + (1|Participant) + (1|Idiom), Datasub)
> summary(Model_RC1)
```

Cumulative Link Mixed Model fitted with the Laplace approximation

formula: RC ~ GuessAP + TEM4s + (1 | Participant) + (1 | Idiom)

data: Datasub

link	threshold	nobs	logLik	AIC	niter	max.grad	cond.H
logit	flexible	414	-353.68	721.37	363(1423)	9.71e-04	2.6e+02

Random effects:

Groups	Name	Variance	Std.Dev.
Idiom	(Intercept)	1.0409	1.0202
Participant	(Intercept)	0.6023	0.7761

Number of groups: Idiom 77, Participant 25

Coefficients: Estimate Std. Error z value Pr(>|z|)

GuessAP0.5	0.37814	0.32230	1.173	0.24070
GuessAP1	0.93999	0.31392	2.994	0.00275 **
TEM4s	0.05264	0.03230	1.630	0.10316

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:

	Estimate	Std. Error	z value
0 0.5	-1.1996	0.3133	-3.829
0.5 1	-0.2142	0.3025	-0.708

```
> anova(Model_RC1,Model13)
```

Likelihood ratio tests of cumulative link models:

	formula:	link:	threshold:
Model13	RC ~ GuessAP + (1 Participant) + (1 Idiom)	logit	flexible
Model_RC1	RC ~ GuessAP + TEM4s + (1 Participant) + (1 Idiom)	logit	flexible

	no.par	AIC	logLik	LR.stat	df	Pr(>Chisq)
Model13	6	721.99	-354.99			
Model_RC1	7	721.37	-353.68	2.6224	1	0.1054

negligible differences in AICs, stick to the simple Model 3

Model 3

```
> Datasub$GuessAP= as.factor(Datasub$GuessAP)
> Model13 <- clmm(Recall~ GuessAP+ (1|Participant) + (1|Idiom), data=Datasub)
> summary(Model13)
```

Cumulative Link Mixed Model fitted with the Laplace approximation

formula: Recall ~ GuessAP + (1 | Participant) + (1 | Idiom)

data: Datasub

	link	threshold	nobs	logLik	AIC	niter	max.grad	cond.H
	logit	flexible	414	-354.99	721.99	309(1240)	2.06e-04	4.2e+01

Random effects:

Groups	Name	Variance	Std.Dev.
--------	------	----------	----------

Idiom	(Intercept)	1.0671	1.0330
-------	-------------	--------	--------

Participant	(Intercept)	0.6825	0.8261
-------------	-------------	--------	--------

Number of groups: Idiom 77, Participant 25

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
GuessAP0.5	0.4058	0.3231	1.256	0.20913
GuessAP1	0.9613	0.3149	3.053	0.00226 **

```
> Model0<-clmm(Recall~ 1+ (1|Participant) + (1|Idiom), data=Datasub)
```

```
> anova(Model3, Model0)
```

Likelihood ratio tests of cumulative link models:

	formula:	link: threshold:
Model0	Recall ~ 1 + (1 Participant) + (1 Idiom)	logit flexible
Model3	Recall ~ GuessAP + (1 Participant) + (1 Idiom)	logit flexible

	no.par	AIC	logLik	LR.stat	df	Pr(>Chisq)
Model0	4	727.76	-359.88			
Model3	6	721.99	-354.99	9.7723	2	0.00755 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1