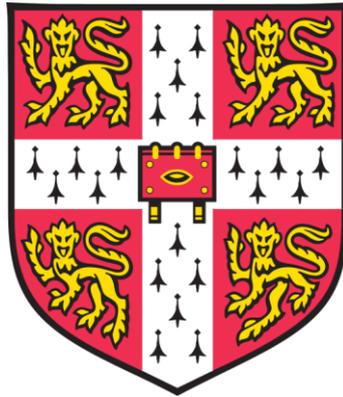


# Acquisition of metaphorical expressions by Chinese learners of English



Mengying Xia

Department of Theoretical and Applied Linguistics

Newnham College

University of Cambridge

This dissertation is submitted for the degree of  
*Doctor of Philosophy*

October 2018

人の上下に人作らぬなら

俺がその天の頂き頂く

(If one is not created above or below any other,  
I will, then, seize top of the sky for myself.)

——櫻井翔 (Sho Sakurai), 2008, ‘Hip Pop Boogie’

Translation by the author of the dissertation.

## Declaration

I hereby declare that this dissertation is the result of my own work, and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text.

Part of the material presented in Chapter 5, *The judgment of metaphorical expressions*, has previously appeared in *Psychotypology of Chinese learners of English and its influence on the acquisition of metaphorical expressions: an offline study*, *Cambridge Occasional Papers in Linguistics (COPiL)*, Volume 10, 2017.

This dissertation does not exceed the regulation length of 80,000 words including footnotes, references and appendices but excluding the bibliography.

## Acknowledgement

Suppose that, my PhD was an adventure, then, I owe so much to my compass, my dear supervisor Dr Henriëtte Hendriks. When I got lost in the middle of the ocean, she gave me courage to step on a fresh new boat, set up sail and finally reach the land. I could not imagine my journey without her; I might have sunken somewhere into nowhere.

I would also like to thank all the lighthouses in the Department of Theoretical and Applied Linguistics: Dr Dora Alexopoulou, Prof Margaret Deuchar, Dr Napoleon Katsos, Dr Teresa Parodi, Prof Ianthi Tsimpli, Dr John Williams and Dr David Willis. They marked the correct route for me to go, provided me with the weather forecast and alerts; they watched me silently when everything was fine, but never hesitated to lend me a hand when I was in trouble. I would like to thank our precious suppliers Louise Radok, Ian Smeeton, Gabby Uncles and Jane Walsh for their generous help with everyday administrative issues. I am grateful that the department kindly provided financial support for conferences, experiments and summer schools.

Throughout the years, I enjoyed sailing together with others in the same fleet, even though we went our different directions. I would like to thank all the boys and girls in the department, particularly Giulia Bovolenta, Cherry Lam, Carla Pastorino Campos, Alim Tusun and Elspeth Wilson, for we shared an office, coffee, dinner, analyses, problems and joy. Outside the office, I received continuous support from Dimitrios Alikaniotis, Rodanthi Christofaki, Dr Amy Hsieh, Yang Li, Connor Quinn, Dr Jessica Soltys, Dr Yan Tao, as well as Ruyi Dai, Dr Yanyu Guo and Dr Lulu Zhang from FAMES. I would also like to thank Wenjia Cai, Maki Kubota and their supervisor Prof Antonella Sorace at University of Edinburgh, as well as Dr Sichang Gao at Beijing Language and Culture University. My cooperation with them showed me a clearer map of the whole world.

Newnham, my lovely boathouse, took care of every part of the small boat as a wonderful mechanic team. Gratitude goes to our graduate tutor Prof Liba Taub for her warm-hearted messages when I was lost; to our tutorial officers Vicky Argent and Sarah Loveday for their encyclopaedic knowledge for dealing with everything; to our catering manager Peter Robertson for adding fuel to the engine; and to all the porters as our literal ‘guardian angels’. At Newnham I met my priceless friends, especially Roanna Kong, Aileen Lam and Janina Schupp – they made the days in Rosalind Franklin Building shining as diamonds. I would like to also thank the College for their generous financial support on travel and book purchases.

During my journey, I constantly visited new places to observe the people there and document their actions in my adventure notebook. I am extremely grateful to the people and places that welcomed my visit, including Beijing Bayi Middle School (with the assistance of Ms Qing Ma and Ms Yu Song, and all my high school teachers), Peking University, Beijing Foreign Studies University and Beijing Language and Culture University. Without them, even if I completed the trip, there would not have been anything useful on my notes.

When I was not sailing, my wireless device helped me to connect with my friends on the sea or on the land. I would like to thank Hanyang Chen, Siyi Chen, Zhi Wang, Dr Siwei Zhang and Yue Zhang for the discussions on statistics, programming and ‘Conversation with supervisors 101’; Xiaoyan Fan, Ran Li and Zhuonan Li for keeping me socially active; Zongxiao Sun for gathering interesting linguistic information; Renee Du, Xiaoying Luo and Sai Gu for the investigation on feline, canine and hamster behaviours; Cambridge University Railway Club for many thoughts on a train; Kaija, Peter and Kitty Hampson for a warm room and a nice cup of tea.

The CDs and books were my best friends when I was alone on the sea. I would like to express my deep gratitude to Mike Lee, 12dora, David Chaofan Yu for their video game commentary; to Arashi (especially Sho Sakurai) for their music and TV programme; and to others who have inspired me, or at least kept me fresh when I was disappointed.

Finally, and most importantly, I would like to thank all the people backing me up when I embarked on my adventure: my parents, Yilin Xia and Yun Huang, for their selfless support, both financially and intellectually; all my family members, especially my grandma and grandpa for their understanding; Zhiyi Yang, for his love, tolerance and in-jokes; and, Coffee and Milk, for their furry paws and warm meows as my last resort.

73,

Chris

## Abstract

This study investigates the acquisition of conventional metaphorical expressions by Chinese learners of English. A conventional metaphorical expression, following the definition of cognitive semantics, refers to the use of a conventionalised non-literal meaning of a lexical item in a multi-word phrase. For example, the word ‘attack’ in the phrase ‘attack one’s idea’, which should be interpreted as ‘to criticise somebody or something severely’, clearly departs from the literal meaning ‘to use violence to try to hurt or kill somebody’, and thus should be seen as a metaphorically used word. Consequently, the phrase ‘attack one’s idea’ is a conventional metaphorical expression.

This study explores learners’ behaviour towards and acquisition of metaphorical expressions from two major perspectives: (1) possible cross-linguistic influence in the process of acquisition and factors that could affect cross-linguistic influence; and (2) the organisation of learners’ bilingual lexicon and the status of metaphorical expressions in a bilingual lexicon. These two perspectives are considered to be the main factors that can influence learners’ acquisition of metaphorical expressions: in order to acquire a metaphorical expression, learners should be able to integrate it into the bilingual lexicon, while the process of integration can be impacted by cross-linguistic influence. Previous research has mainly been conducted on the acquisition of certain figurative expressions in a second language, predominantly idioms; however, a combination of the two perspectives and a joint analysis on the acquisition of figurative language has yet to be accomplished. This study presents a first attempt of such analysis on the acquisition of a specific type of figurative language.

The results of the experiments reported in this dissertation show that learners react differently to metaphorical expressions with different cross-linguistic availabilities (shared between Chinese and English or exclusively available in Chinese or English) but in general they encounter difficulty to achieve native-like performance when reading metaphorical expressions available in their second language. Persistent cross-linguistic influence is observed in two aspects, even among highly proficient learners: (1) learners encounter obstacles when acquiring the metaphorical expressions that are only available in their second language; and (2) learners seem to still activate the metaphorical meanings that are only available in their first language even when they read in their second language. These results altogether reflect that metaphorical expressions, regardless of cross-linguistic availability, are more difficult to acquire than literal expressions in a second language.



# Table of Contents

Declaration.....	iii
Acknowledgement.....	v
Abstract.....	vii
Table of Contents .....	ix
List of Figures.....	xiii
List of Tables.....	xv
1 Introduction .....	1
1.1 General research questions and objectives .....	1
1.2 Exploring metaphors within linguistic research .....	3
1.2.1 Metaphor in (post-)Gricean philosophy of language.....	3
1.2.2 Metaphor in cognitive and lexical semantics .....	7
1.2.3 Metaphor in psycholinguistics and experimental pragmatics.....	10
1.3 Terminological remarks .....	11
1.4 Outline of the dissertation.....	16
2 Cross-linguistic influence in second language lexical acquisition.....	18
2.1 Introduction: a short overview of studies on cross-linguistic influence .....	18
2.2 Lemmatic transfer: a general review.....	21
2.2.1 Defining and delimiting lemmatic transfer.....	21
2.2.2 Types of realisation of lemmatic transfer .....	23
2.2.3 Lemmatic transfer in the acquisition of figurative language.....	27
2.3 Factors influencing cross-linguistic influence .....	33
2.3.1 Psychotypology .....	34
2.3.2 Markedness .....	39
2.3.3 Learners' knowledge .....	43
2.4 Summary.....	45
3 Bilingual lexicon and figurative language processing in a second language.....	47

3.1	Introduction .....	47
3.2	Current proposals for the bilingual lexicon.....	48
3.2.1	Distributed (Conceptual) Feature Model/The Sense Model .....	48
3.2.2	Revised Hierarchical Model/Modified Hierarchical Model .....	56
3.2.3	Bilingual lexicon in a classroom setting .....	63
3.3	Constructing/accessing figurative language in the bilingual lexicon: the processing of figurative language .....	67
3.4	Possible routes for processing metaphorical meanings in a second language .....	72
3.5	Summary .....	78
4	Research Methodology .....	79
4.1	Research questions and hypotheses.....	79
4.2	Selection and categorisation of test materials for metaphorical expressions .....	83
4.3	Main experimental instruments .....	88
4.3.1	Acceptability judgment task (AJT).....	88
4.3.2	Self-paced reading task (SPRT).....	89
4.4	Supplementary experimental instruments .....	91
4.4.1	Lexical evaluation survey .....	91
4.4.2	Linguistic background survey and linguistic experience questionnaire .....	92
4.4.3	Proficiency test.....	93
4.4.4	Psychotypology survey .....	94
4.4.5	Backward digit span test .....	96
4.5	Participants .....	96
4.5.1	Chinese learners of English.....	96
4.5.2	Native speakers of English.....	100
4.6	Procedures .....	101
4.6.1	Acceptability judgment experiment .....	101
4.6.2	Reading experiment .....	101
4.7	Data adjustment and processing .....	102

4.7.1	Acceptability judgment score .....	102
4.7.2	The categorisation and coding of feedback sentences.....	106
4.7.3	Coding and classification of level of confidence .....	108
4.7.4	Reading time for sentences.....	109
4.7.5	Psychotypological score .....	110
4.7.6	Lexical property scores of test items .....	111
4.8	Ethical considerations .....	112
5	The judgment of metaphorical expressions.....	114
5.1	Introduction.....	114
5.2	Degree of acceptability of literal and metaphorical expressions .....	114
5.3	Judgment confidence levels and their interaction with acceptability scores .....	119
5.4	Cross-linguistic influences on the judgment of metaphorical expressions and possible influencing factors .....	122
5.4.1	Identification of traces of transfer in the acceptability judgment task .....	122
5.4.2	Influence of psychotypology on subjective transferability .....	125
5.4.3	Influence of markedness of meaning on subjective transferability .....	127
5.4.4	Influence of metaphoricalness on subjective transferability .....	129
5.4.5	Influence of general proficiency on subjective transferability .....	130
5.5	Learners' production of metaphorical expressions and alternatives in the sentence correction section.....	132
5.5.1	Four types of reaction by participants to given metaphorical expressions.....	132
5.5.2	Strategies adopted by the participants in producing feedback sentences .....	134
5.6	Summary.....	138
6	The processing of metaphorical expressions.....	140
6.1	Introduction.....	140
6.2	Reading time for literal and metaphorical expressions.....	140
6.3	Reaction time to the comprehension questions and scores for responses.....	147
6.4	Factors influencing subjective transferability and their influence on the processing of metaphorical expressions .....	153

6.4.1	Introduction .....	153
6.4.2	Influence of psychotypical distance .....	154
6.4.3	The influence of markedness of meanings .....	155
6.4.4	Influence of general proficiency .....	157
6.5	Summary .....	158
7	General discussions and conclusions .....	160
7.1	Cross-linguistic influence and the acquisition of metaphorical expressions .....	160
7.2	Factors influencing the transferability and transfer of metaphorical expressions .....	166
7.3	From asymmetry to symmetry: the difference between literal and metaphorical meanings of lexical items in second language acquisition and processing .....	172
7.4	Implications for a framework of literal and metaphorical meanings in a bilingual mental lexicon .....	176
7.5	Main contributions and directions of further research .....	183
	Bibliography .....	185
	Appendix 1: Instruction and sentences used in the acceptability judgment task .....	196
	Appendix 2: Instruction, sentences and questions used in the self-paced reading task .....	200
	Appendix 3: Lexical property scores for the critical lexical items .....	207

## List of Figures

Figure 3.1 The Distributed Conceptual Feature Model (adapted from De Groot 1992).....	49
Figure 3.2 The Sense Model (adapted from Finkbeiner and others 2004).....	52
Figure 3.3 The Revised Hierarchical Model (adapted from Kroll and Stewart 1994).....	58
Figure 3.4 The structure of <i>gongji</i> , attack and relevant concepts in an RHM fashion presentation .....	60
Figure 3.5 The Modified Hierarchical Model (adapted from Pavlenko 2009).....	62
Figure 3.6 A simplified model of bilingual lexicon with literal and metaphorical distinctions .....	73
Figure 4.1 An excerpt from the acceptability judgment task .....	89
Figure 4.2 A simplified illustration of the self-paced reading task in the experiment .....	90
Figure 4.3 An excerpt from the online lexical evaluation survey .....	92
Figure 4.4 Raw acceptability scores for different categories for five groups of participants.	104
Figure 4.5 Acceptability scores of different categories of test sentence among five groups of participants .....	106
Figure 5.1 Acceptability scores for literal and metaphorical expressions by group.....	116
Figure 5.2 Confidence scores for the six target types for each group .....	120
Figure 5.3 Distribution of coreness scores for the MB and MS expressions in the AJT, as perceived by Chinese native speakers .....	128
Figure 5.4 Distribution of the coreness scores of the LB and MB expressions in the AJT, as perceived by Chinese native speakers .....	129
Figure 6.1 Segment-by-segment reading time pattern for the native group.....	141
Figure 6.2 Segment-by-segment reading time pattern for the intermediate group.....	142
Figure 6.3 Segment-by-segment reading time pattern for the low-advanced group .....	143
Figure 6.4 Segment-by-segment reading time pattern for the high-advanced group .....	145
Figure 6.5 Segment-by-segment reading time pattern for the overseas group.....	146
Figure 6.6 Reaction times (left) and mean scores for answers for the native group .....	148
Figure 6.7 Reaction times (left) and mean scores for the answers for the intermediate group .....	149
Figure 6.8 Reaction time (left) and scores for questions for the low-advanced group.....	150
Figure 6.9 Reaction times (left) and scores for questions for the high-advanced group.....	151
Figure 6.10 Reaction times (left) and scores for questions for the overseas group.....	151

Figure 6.11 Distribution of the coreness scores for the MB and MS expressions in the SPRT, as perceived by Chinese native speakers.....	155
Figure 6.12 Distribution of the coreness scores for the MB and MT expressions in the SPRT, as perceived by English native speakers .....	156
Figure 7.1 A monolingual literal and metaphorical lexicon .....	180
Figure 7.2 A bilingual literal and metaphorical lexicon with varied strength of links .....	181

## List of Tables

Table 3.1 Two possible routes for the retrieval of the literal meaning of an L2 word.....	74
Table 3.2 Four possible routes for the retrieval of the metaphorical meaning shared between the L1 and the L2.....	74
Table 3.3 Three possible routes for the retrieval of the L1-specific metaphorical meaning from the L2 word form.....	75
Table 3.4 Three possible routes for the retrieval of the L2-specific metaphorical meaning of an L2 word .....	76
Table 3.5 Different bundles of possible routes of metaphorical meaning retrieval in L2.....	77
Table 4.1 Availability of different types of expression in the current thesis in Chinese and English.....	85
Table 4.2 A pair of sample test sentences and their structure .....	88
Table 4.3 Linguistic backgrounds of the 86 learner participants in the acceptability judgment experiment.....	98
Table 4.4 Linguistic backgrounds of the 81 learner participants in the self-paced reading experiment.....	99
Table 4.5 Comparison between the participants in the acceptability judgment task and those in the self-paced reading experiment.....	100
Table 4.6 Number of feedback sentences provided by each group of participants.....	107
Table 4.7 Distribution of relevant feedback sentences provided by each group of participants across different metaphorical conditions.....	107
Table 4.8 Distribution of psychotypical distance between Chinese and English among different groups of participants.....	111
Table 5.1 Identification of traces of transfer in the judgments of MB expressions .....	123
Table 5.2 Identification of traces of transfer in the judgments of MS expressions.....	124
Table 5.3 Identification of traces of transfer in the judgments of MT expressions.....	124
Table 5.4 Estimated percentage of traces of transfer of different metaphorical expressions in the learner groups .....	124
Table 5.5 Distribution of traces of transfer on LB and MB expressions across learners' proficiency.....	130
Table 5.6 Distribution of the types of reaction across different proficiency groups.....	133
Table 5.7 Examples of types of strategies adopted by participants in the sentence correction section.....	135

Table 5.8 Average number of relevant feedback sentences with different strategies provided by individual participants at different proficiency levels.....	136
Table 5.9 All 16 relevant feedback sentences for ‘My mother said that these books would be <i>food</i> for thought’ given by different groups of participants.....	137
Table 7.1 A pair of sample test sentences and their structure (repeated from Table 4.3).....	177
Table 7.2 Differences in degree of acceptability between literal and different metaphorical expressions.....	178
Table 7.3 Differences in processing cost between literal and different metaphorical expressions.....	179

# 1 Introduction

## 1.1 General research questions and objectives

This thesis aims to research the acquisition of conventional metaphorical expressions in a second language. More specifically, it focuses on Chinese native speakers who are acquiring English as a second language, mostly in a classroom setting. The concept of metaphor, as well as the theoretical accounts associated with the concept, have been extensively debated in contemporary linguistics, and agreement has been reached in at least one area in recent decades: metaphors, or to be more precise, metaphorical expressions, are more than a decoration of daily language use. Metaphorical expressions are widely used in all types of discourse, written or spoken, consciously or unconsciously, and have formed a great number of conventional collocations in almost every language. Even in the title of this dissertation metaphorical expressions can be observed: we can *acquire* a book or a house by paying money to a bookstore or the owner, and we can touch and feel what we have acquired; when we ‘acquire’ a language, or a particular skill, however, we probably do so without any financial cost, and certainly without knowing the shape of the language or the colour of the skill. The expansion of the meaning of ‘acquire’ is therefore metaphorical, although this may not be realised by those who use this word, or even those who have researched second language acquisition for years. It is difficult, then, to separate metaphorical expressions from other parts of language use.

This chapter serves as an overview of the dissertation. Specifically, it aims to delimit the target of the current study, namely ‘metaphorical expressions’. The current section presents the main research questions and objectives of this thesis. Section 1.2 provides a general discussion of metaphor in linguistics from several major perspectives, including philosophy of language, cognitive semantics and language processing. Section 1.3 clarifies the definitions of the relevant terminology in the dissertation. The last section of this chapter includes an overview of the following chapters.

Every language has its own metaphorical expressions; they can be conventionalised and lexicalised as an established part of the language, or be novel and creative for certain or any language users. For example, Chinese, according to Liu (2002) and Link (2013), is known to use ‘eat’ to depict all the sufferings one can imagine: you ‘eat the loss’ when you lose your wealth, ‘eat the bitterness’ when you encounter some hardship, ‘eat a lawsuit’ when you are sued by your enemy, and ‘eat a bullet’ when you, unfortunately, are sentenced to death. Such metaphorical expressions based on ‘eat’ may be beyond the understanding of English native

speakers who do not have any prior knowledge of Chinese. As a Chinese learner of English, I have also experienced such ‘metaphorical shock’ during my time in an English-speaking environment. I never thought that a person could ‘buy’ an idea, or a country could ‘buy’ its peace. Even though I have a first language that uses ‘eat’ for all sorts of life experiences, I still had no idea what ‘food for thought’ meant before I was told. Despite these dramatic differences, however, Chinese and English also share a large number of metaphorical expressions: both ‘lay the foundation’ and ‘build the structure’ of a dissertation, as if the dissertation were a skyscraper; both try to make the argument ‘clear’ and the evidence ‘strong’, as if we could visually perceive an argument or physically touch a piece of evidence. The diversity of metaphorical expressions across languages, as well as those expressions that are common to different languages, provide material for linguists to analyse how languages can reflect cultural variation and cognitive universals, while at the same time troubling those who would like to master a language other than their first.

In order to understand how metaphorical expressions are acquired in second language acquisition, and thus to help language learners understand their strengths and weaknesses in terms of metaphorical expressions, three major comparisons should be made, which will lead to the three main research questions in this thesis. The first comparison is between literal expressions and metaphorical expressions. Is there any difference in acquisition outcome between the two types of expression? Are learners in general better at one or equally good at both? If one assumes, on the basis of the discussion of literary metaphor before 1960s, that metaphorical expressions should always be considered ‘decorative’ and ‘secondary’, then such ‘secondariness’ may be the cause of learners’ receptive and productive performance with such expressions. The second comparison is between the metaphorical expressions of learners’ L1 and L2. As exemplified above, two different languages may share metaphorical expressions but will also have their own language-specific expressions. Do such differences in availability have a significant impact in terms of acquisition outcome? This comparison can be further connected to the general impact of cross-linguistic influence on second language acquisition: as cross-linguistic influence has been shown pervasively in almost all aspects of language acquisition (see Odlin 1989 for a classic summary), it is natural to assume that metaphorical expressions cannot escape the impact of cross-linguistic influence either. The third comparison is between native speakers’ and learners’ reactions to metaphorical expressions. As we will see later, both theoretical and experimental semantics have provided a good understanding of native speakers’ ability to use metaphorical expressions in daily language, which provides a baseline for us to investigate learners’ use of metaphorical expressions. Could it be that learners can achieve

native-like results without much effort? Or will it be the case that metaphorical expressions are something that non-native speakers can never fully master? The aim of this thesis is to answer the research questions that come from the three comparisons.

Furthermore, this dissertation is intended to establish the importance of metaphorical expressions as a proper topic in second language acquisition, and to associate the topic with other widely-surveyed aspects of second language acquisition, including cross-linguistic influence, bilingual vocabulary and non-literal language processing in a second language. In the domain of semantics and pragmatics, metaphorical expressions have received no less attention than other examples of non-literal language. Theoretical proposals have been raised concerning the creation, semantic construction, derivation, conventionalisation and usage of non-literal language. However, as a proportion of second language acquisition research, research into conventional metaphorical expressions and other types of non-literal language is under-represented, with a well-established research history for the acquisition of other non-literal language phenomena (particularly idioms and phrasal verbs), but a lack of in-depth, systematic investigation into the acquisition of conventional metaphorical expressions. It seems that, while idioms and phrasal verbs are generally recognised as linguistic elements that require additional instruction, there has been little research into the question of whether metaphorical expressions can be acquired autonomously by L2 learners.

It is expected that this project on the acquisition of metaphorical expressions will provide some insights of use to different groups of people: to researchers of second language acquisition, as it may improve understanding of metaphorical expressions and their status in the bilingual mental lexicon; to researchers of semantics and pragmatics, in improving understanding of the universality and cross-linguistic variability of metaphorical expressions and the influence of such cross-linguistic variation on individual language users; to instructors of a second language, allowing them to understand better the importance of metaphorical expressions and design appropriate content based on the results; and to learners of a second language, allowing them to understand their strengths and weaknesses in this area.

## 1.2 Exploring metaphors within linguistic research

### 1.2.1 Metaphor in (post-)Gricean philosophy of language

Linguistic theories regarding the linguistic realisation of metaphor began to flourish in the late 1960s when Grice (1967) first proposed the concept of *implicature*, together with the four well-known conversational maxims: quantity, quality, relevance and manner. Grice's theory

emerges against a background of compositional semantics, which suggests that the meaning of a sentence is essentially derived through the composition of the meaning of each constituent of that sentence. Compositional semantics, however, fails to explain how interlocutors 'read between the lines' and acquire the intended meanings that are not expressed by the compositional meaning. Grice aimed to provide an explanation for this by positing the idea of conversational implicature.

Grice's intention was to establish a dichotomy between 'what is said' and 'what is implicated'; the former is 'the meaning of a sentence', which should be directly derived from the composition of the meaning of each constituent of a sentence, while the context only provides referential assignment. The latter is 'the meaning of an utterance', which is the meaning intended when the speaker's intention and the contextual information are taken into consideration. In this framework, an utterance like 'the room is hot' can be a simple statement that a particular room is hot ('what is said'), or, in a particular context, a request from the speaker about opening the window or turning on the air conditioner ('what is implicated'). Grice suggests that the implicature of an utterance can be deduced from what is said using one or more of the four conversational maxims: quantity, quality, relevance and manner.

Metaphorical utterances, according to the original proposal of Grice, should be interpreted using the maxim of quality: when producing a metaphorical utterance, the speaker says something that is blatantly incorrect, and thus what is said cannot be the meaning intended. Grice (1989: 34) provides an example of how 'what is implicated' by a metaphorical utterance can be derived based on his maxim of quality: when a speaker says 'you are the cream in my coffee', the hearer will immediately know that the speaker is not intending this information literally, because a person can never be the cream in one's coffee. Therefore, what is said is blatantly false, and the hearer needs to derive the implicature of that utterance by thinking about the speaker's attitude towards 'cream in her coffee'; then the hearer will know that he is either the speaker's pride or her bane.

Grice's theory of implicature leads to several suggestions or hidden assumptions. To name a few important ones: (1) the meaning of metaphor is derived through implicature; and (2) there is expected to be a literal-first processing pattern, so that a hearer can be led to the derivation of conversational implicature when the literal interpretation is found not to be possible. Subsequent discussions of the nature of metaphorical meanings in philosophy of language, no matter what suggestions they raise, are constantly under the influence of the Gricean framework. As is usual, here they are categorised as post-Gricean approaches (including formal

semantics, truth-conditional semantics and theoretical pragmatics, but excluding cognitive semantics). This subsection will briefly introduce two major accounts that have had a significant impact on the current thesis: the indexicalist view (Stern 2000; Leezenberg 2001; Xia 2015b), and the ‘threshold of literal meaning’ view (Recanati 2004).

The indexicalist view of metaphor originates from philosophical discussions of indexical expressions such as ‘I’, ‘here’ and ‘now’ (see Kaplan 1989). Kaplan argues that those indexical expressions do not have ‘contents’ similar to dictionary meanings, but only have certain ‘characters’ to help the interlocutors to resolve the reference of the expression. The fundamental idea is that the reference of an indexical expression is selected in a ‘bottom-up’ fashion: the character of an indexical expression leads the interlocutor to survey the possible references in the local context and assign the indexical expression to the most appropriate reference. The extension of the indexicalist view to the interpretation of metaphors treats metaphors as a special type of indexical, and suggests that the meaning of a metaphorical expression may be determined by a local context. Therefore, an alteration of local context, such as an adjacent constituent of a lexical item, can change the meaning of that lexical item and lead to the formation of a metaphorical meaning of a word. Stern (2000) proposes that a metaphorical expression always has a metaphorical character. A metaphorical expression can have several stereotypical metaphorical properties; for example, the action of ‘attack’ can involve properties like fierceness, suddenness and intention to defeat someone. The function of that metaphorical character, according to Stern, is to remind the interlocutors to select one or more of the metaphorical properties, and then construct a new meaning as the ‘appropriate metaphorical reference’. Leezenberg (2001), instead, argues that a metaphor has neither content nor character. He suggests that a metaphorical expression will evoke a set of ‘salient properties’ based on the literal concept of that metaphor, and an interlocutor will combine all the salient properties to construct the meaning of that metaphor. Xia’s (2015b) approach, which is based on Stanley’s (2002; see also Stanley and Szabó 2000) Nominative Restriction Theory, claims that every metaphorically used lexical item comes with a ‘complex structure’. The complex structure marks out one or more properties of the literal concept represented by that lexical item in the local context, so that the interlocutor can choose a suitable property based on the context.

Despite this discrepancy regarding the selection procedure for metaphorical properties, all indexicalist views of metaphor suggest that local contextual information, especially the collocation of the metaphorically used lexical items, manipulates the meaning of the metaphor, while higher level contexts do not play a further role in deciding the semantic meaning of the metaphor. The indexicalist approach has the special advantage of explaining why a

metaphorical expression can be identified and comprehended without difficulty: the hearer need only select the most appropriate meaning for a lexical item from a collection of meanings, and there is no need to construct the meaning from the literal meaning of that lexical item plus the broad context. This means it is *possible* for a hearer to comprehend a metaphorical expression as fast as comprehending a literal expression. Important for language acquisition is the idea, presented by the indexicalist approach, that the (conventional) metaphorical meaning and literal meaning of a lexical item can be distinguished by different collocations; this idea is further utilised in the current dissertation.

The second post-Gricean view to be reviewed here stems from Recanati's (2004) book *Literal Meaning*, in which he discusses different types and levels of literalness. He suggests that although all metaphors involve departure from the literal meaning of a word, some metaphors are not 'far' enough from the literal meaning to be called 'non-literal', while others may be a distant departure from the literal meaning and should definitely be seen as 'non-literal'. Recanati proposes three layers of literalness, mainly depending on the distance between the intended meaning and the compositional meaning of a sentence which can be of two types: (1) 'm-literal' for minimal-literal, that is, the intended meaning is exactly the same as the compositional meaning of the sentence; and (2) 'p-literal' for primary-literal, that is, the intended meaning of the sentence only involves minimal departure from the compositional meaning of the sentence. The two types of literalness lead to a three-way categorisation of all sentences. The first type includes sentences that are both m-literal and p-literal, such as literal sentences like 'Chris swallowed the candy'. The second type involves sentences that are neither m-literal nor p-literal, such as novel metaphorical sentences like 'Chris is a cactus'. In this sentence, the meaning of 'cactus' in the sentence clearly departs from the literal meaning as a species of plant, and the intended meaning of the whole sentence is also a clear departure from the compositional meaning of 'Chris is a cactus'. The third type contains sentences that are p-literal but not m-literal, which includes the target of the current thesis – conventional metaphorical sentences. A good example is 'the ATM swallowed my debit card', which is a conventional metaphorical expression: the meaning of 'swallowed' in the sentence clearly departs from the literal meaning of that word, because the action performed by the ATM is not the prototypical 'swallowing' by an animate object, but the meaning of the entire sentence can still be derived from the composition of each constituent. Recanati argues that, although both conventional metaphors and novel metaphors are metaphors, they can be distinguished in this paradigm, because conventional metaphors are less 'non-literal' than novel metaphors. He suggests that such a distinction implies that there is a threshold of non-literalness, where novel

metaphor occurs above the threshold and conventional metaphor below it. The idea of a non-literal threshold is borrowed and used in the current dissertation to refine the analysis of the metaphorical expressions under investigation.

### 1.2.2 Metaphor in cognitive and lexical semantics

Most, if not all, of the theories of metaphor before the Gricean account, including the seminal work by Beardsley (1962), Black (1962) and Grice himself, take one premise for their discussion of metaphors: metaphors should have a strong sense of non-literalness. In other words, in the previous theories, metaphors were always expected to be above Recanati's p-literalness threshold in order to qualify as metaphors. Since metaphors are believed always to be above the threshold, they are also believed to be marked, creative, deliberate and uncommon in daily language use; therefore, most theories of metaphor before the Gricean era tended to have a clear focus on the literary use of metaphor. The theories of metaphor based on cognitive semantics, including the Conceptual Metaphor Theory (henceforth CMT) of Lakoff and Johnson (1980; see also Lakoff 1987; Lakoff and Turner 1989) and the Blending Theory (Fauconnier 1994), have drastically changed the view that metaphors are marked in natural language. In the following I will give a brief review of CMT, some recent developments and criticism, as well as the extension of CMT into lexical semantics and the understanding of how conventional metaphorical expressions are formed, since the current study refers to the research on CMT when selecting experimental materials.

In the foundational work *Metaphors We Live By*, Lakoff and Johnson redefine the concept of metaphor in two ways. First, they suggest that metaphor is not only a linguistic phenomenon, but also a cognitive phenomenon: humans constantly use metaphor in their perception of the world, and it is a general cognitive pattern to map a concrete, familiar, touchable entity to the concept of an abstract, unfamiliar, less perceivable entity. Two levels of metaphor are clearly differentiated in CMT: conceptual metaphors, which exist in human thought and are an essential mechanism for individuals to acquire new knowledge and concepts; linguistic metaphors, which are the linguistic realisation of the conceptual metaphorical mappings. Second, since conceptual metaphors are widely available in human thought, and language is regarded as the essential reflection of human thought, the linguistic realisation of conceptual metaphors is also inevitable in human language. Linguistic metaphors are not necessarily 'deliberate' or 'creative', as is the case with the examples found in literary works; rather, any departure from the core meaning of a lexical item can be seen as 'used metaphorically'. Therefore, everyday

expressions like ‘I see what you mean’ are metaphorical, because ‘what you mean’ is an abstract entity and cannot be visually perceived, and the use of ‘see’ in that sentence is metaphorical in the sense of CMT. CMT has become one of the most influential theories of metaphors in linguistics, making a contribution to the understanding and construction of metaphor in individual languages (see Yu 1998 for Chinese examples), and the universality and cross-linguistic variation of conceptual and linguistic metaphors (e.g. Kövecses 2005). CMT is also frequently adopted in other domains of linguistic research, including pragmatics, language acquisition and psycholinguistics. The philosophical accounts of metaphor mentioned in the last section also allow for CMT to co-exist. For instance, Recanati (2004), although not explicitly stating it, treats the ‘everyday linguistic metaphors’ in CMT as examples of metaphors. In the case of ‘the ATM’ mentioned above, he regards it as a conventional metaphorical expression; meanwhile, from a CMT perspective we can see it as a linguistic realisation of personification, which is a specific type of conceptual metaphor.

Current research on metaphor in the field of cognitive semantics, although still under the influence of the original proposal of CMT in 1980s, has departed substantially from Lakoff and Johnson’s account. The original CMT receives criticism such as circular reasoning, lack of support for context-sensitivity, heavily relying on constructed examples rather than natural language evidence, etc. (see Kövecses 2017 for a list of criticism), and is no longer favoured in recent studies on metaphor. Meanwhile, scholars have been establishing new theoretical views of metaphor based on CMT, making modifications and providing empirical supports to it. For instance, a clear contrast has been observed between the original CMT proposal in 1980s and a more contemporary, dynamical view of metaphor. While the original CMT proposal has focused more on the formulaic construction of linguistic metaphors and the conceptual mappings behind them, the dynamical view can be seen as a convergence between a cognitive semantic convention and a post-Gricean view on the function of contexts. Instead of assuming that linguistic metaphors are always formulaic and stable, the dynamical view recognises both the regularities of metaphorical expressions based on conceptual metaphors and the influence of context-sensitivity when constructing metaphorical expressions in a piece of discourse (Gibbs 2017). The conceptual root of linguistic metaphor being recognised, recent theories begin to focus more on how metaphorical expressions are created, used and evolved.

The cognitive semantic view of metaphor, particularly CMT, has also been used in lexical semantics to interpret the emergence of some types of polysemy. Sweetser (1990), who uses ‘see’ as the material of her case study, proposes the idea that the different polysemous senses of ‘see’ originate from conventional linguistic metaphors. With the conventionalisation of these

linguistic metaphors, the metaphorically constructed meanings have been accepted by language users and have become part of the dictionary meanings of a lexical item. This process is also suggested by Traugott (2004) from the perspective of historical pragmatics: the metaphorical meanings of a lexical item could originally be understood as conversational implicature, which, as in the Gricean account, is derived through the literal meaning and the context, but gradually these become conventional metaphorical expressions, stabilised and lexicalised after frequent use, and finally the properties of conversational implicature disappear and the metaphorical interpretations become encoded meanings rather than *ad hoc* implicatures. It has been widely recognised (c.f. Gibbs 1995; Lakoff 1987; Sweetser 1990) that the metaphorical meaning(s) and the literal meaning of a lexical item are closely related, and lexical items that can be interpreted both literally and metaphorically should be treated as polysemy.

Overall, in this thesis, the conventional terminologies of cognitive semantics are adopted, including those from CMT and its subsequent theories, and discussion of the acquisition of metaphorical expressions upholds a clear boundary between linguistic metaphor and conceptual metaphor. All the experimental materials used in this thesis are selected from discussions of metaphorical expressions using the cognitive approach and any lexical semantic theories based on that. Note however, that cognitive views of metaphor including CMT are essentially semantic discussions of metaphor, and they do not aim to provide explanations for the acquisition of metaphorical expressions. Although implications for the use of conceptual metaphors in second language acquisition have been drawn from CMT and other cognitive views of metaphor (e.g. Littlemore and Low 2006), I question whether learners are able to utilise their knowledge of conceptual metaphors in second language acquisition. For instance, when the conceptual metaphor SEEING IS UNDERSTANDING is assumed universally, both Chinese and English allow the expression ‘a clear idea’ for an idea that is easy to understand (‘clear’ is originally used to describe visual perception), but only English allows the conventional use of the verb ‘see’ as ‘understand’, as in ‘I see what you mean’. It is unknown whether Chinese learners of English can actively make use of the conceptual metaphor SEEING IS UNDERSTANDING when they acquire the meaning of ‘see’. As this dissertation does not claim that there might be direct implications for second language acquisition from cognitive semantic view of metaphor, I will not discuss further the relationship between cognitive semantics and acquisition of metaphorical expressions, and will only use the terminologies taken from cognitive semantic views of metaphor.

### 1.2.3 Metaphor in psycholinguistics and experimental pragmatics

Investigation of the processing of linguistic metaphors has been developing since belief in the ‘literal-first’ hypothesis became widespread, well after Grice’s proposal for conversational maxims mentioned in Section 1.2.1. The ‘literal-first’ hypothesis proposes that a hearer can only proceed to the non-literal meaning of a sentence if the literal meaning of the sentence does not make any sense. Therefore, a direct inference of the hypothesis is that it will take more time for a hearer to process a metaphorical meaning than to process a literal meaning. One target for investigation in psycholinguistics and experimental pragmatics has been to examine whether the processing of metaphor is more time-consuming than the processing of literal expressions; if not, the ‘literal-first’ hypothesis may have to be rejected.

Research on the processing of linguistic metaphors was initiated by Glucksberg, Keysar and their colleagues (Glucksberg and Keysar 1993; see Glucksberg 2001 for a historical review), who compare the reading time for conventional sentential metaphors like ‘my lawyer is a shark’ and literal sentences like ‘my lawyer is a man’. The results show that, while the participants interpret both the literal and the metaphorical sentences accurately in their responses to comprehension questions, they do not spend more time on the processing of the metaphorical sentences. Later investigations, such as Chiappe and Kennedy (2001) and Utsumi (2007), arrive at the same conclusion: conventional metaphorical expressions are processed by native speakers without any additional effort. Furthermore, a series of experiments by Chiappe and Kennedy (2001) find that the processing times for metaphorical expressions vary for different individual expressions. The processing effort for linguistic metaphors correlates with the degree of conventionality of those metaphors: less conventional metaphors do require more time for a reader to read and comprehend, and more conventional metaphors evoke the same processing time as literal expressions. Recent studies utilising EEG show that less conventional metaphors trigger an N400 response in the reading process, which means that the readers perceive a semantic anomaly at the beginning of the processing, while more conventional metaphors and literal expressions do not evoke any N400 response in the processing, indicating that the readers do not perceive any semantic anomaly (Lai and others 2009).

Based on the recorded processing pattern for different types of metaphorical sentences, particularly the difference between conventional and unconventional metaphors, Bowdle and Gentner (2005) propose a theory for the development and conventionalisation of metaphor, namely ‘The Career of Metaphor’. They suggest that metaphors like ‘the girl is a rose’, depending on their conventionality, can be interpreted by language users by either (1) comparing and searching for similarities between two different concepts, resulting in an

interpretation like ‘the girl is like a rose’ or (2) categorising one concept into another superordinate concept, resulting in an interpretation like ‘the girl is a [rose in a metaphorical sense]’. When a metaphor is unconventional, a language user tends to interpret it as a comparison between two concepts. It will take more time for the user to search for the similarities between the two concepts, and thus more time will be spent processing the metaphor. When a metaphor becomes conventionalised, the language user will simply see it as a statement of categorical inclusion, and it is no longer necessary to construe the meanings in an *ad hoc* way. Therefore, the conventional linguistic metaphor can be processed in the same way as literal expressions. Here the conventionality of a linguistic metaphor does not solely depend on its linguistic form, but the context and the intended meaning of the metaphor should also be taken into consideration. A linguistic metaphor, as in ‘the girl is a rose’, can be either conventional or unconventional, depending on whether the sentence is intended to refer to her beauty (which is conventional) or her bad temper (as in the spikes on the stem of a rose, which is unconventional); therefore, it may take the language user less time to derive the meaning that the girl is beautiful, and more time to derive the meaning that the girl is easy to upset.

The Career of Metaphor, as a theory based on psycholinguistic observation, exists in parallel to the theoretical and philosophical accounts of metaphor: the mechanism behind a longer reaction time to an unconventional metaphor can be either the ‘literal-first’ hypothesis as per Grice, or other possible accounts. The main contribution of the Career of Metaphor is that for native speakers of a language, no additional time cost or cognitive effort is entailed in processing a conventional linguistic metaphor, no matter what form it is in. This means that a native speaker will spend the same time reading a literal sentence and a conventional metaphorical sentence of the same length and the same grammatical structure. This fact sets the baseline for any research regarding the processing of linguistic metaphor in a second language.

### 1.3 Terminological remarks

The topic of this thesis is ‘metaphorical expression’, that is, the linguistic realisation of a conceptual metaphorical mapping. A metaphorical expression is defined here as a grammatical multi-word phrasal structure (such as a grammatically complete VP, NP or AdjP, etc.), in which one and only one of the lexical items should be interpreted as a conventionalized metaphorical meaning in order to make the phrase meaningful. While a metaphorical expression is defined in terms of syntactic structure, such a view is essentially from formal semantics as Principle of Compositionality: a common formal semantic theory defines that the meaning of a phrasal

structure is determined by both (1) the meanings of individual lexical items in the structure; and (2) the way of structural organization (Frege 1991). In this sense, a metaphorical expression is a multi-word phrase in which one and only one word contributes to the compositional meaning using its metaphorical sense.

A metaphorical expression can be seen as a semi-fixed structure. For example, expressions like “attack the proposal”, “attack the argument” and “attack the idea” can be combined together as one semi-fixed metaphorical expressions [attack [NP]], while a possible NP should fall into the semantic domain of “ideas and thoughts” in order to trigger the intended metaphorical meaning of “attack”. Since the metaphorical meaning of a lexical item is essentially triggered by the collocation of that lexical item, words that (1) fall into the semantic domain of required collocation and (2) can form valid conventional expressions can all be filled into the blank part of the semi-fixed structure. However, when considering metaphorical expressions from an experimental view, it is difficult to examine all the possibilities of a semi-fixed metaphorical expressions at the same time. Therefore, experiments in the dissertation only examined some instances of the semi-fixed metaphorical expressions, and the author hopes to use the instances to represent all the instances in which other lexical items or phrases are filled into the blank part of a semi-fixed metaphorical expression. When describing a “metaphorical expression” in an experimental condition, it will only refer to a specific instance of that semi-fixed structure used in the experiment.

The metaphorical expressions under investigation in this thesis are all conventionalised. A pair of simple examples indicating the contrast between a literal expression and a conventionalised metaphorical expression is given in (1) below:

- (1) a. John attacked Mary with a stick.
- b. John attacked Mary’s proposal in the meeting.

The words used once literally and once metaphorically are underlined. The core and fundamental meaning of the word ‘attack’, as reflected in (1a), is ‘to use violence to try to hurt or kill somebody’ (Wehmeier 2000), and the meaning of ‘attack’ in (1a) clearly involves physical contact and conflict. Example (1b) makes use of the metaphorical meaning of ‘attack’, which is defined as ‘to criticise somebody or something severely’ (Wehmeier 2000), and the action of ‘attacking a proposal’ obviously does not involve any physical conflict, for there is not even a physical entity for ‘a proposal’. However, the metaphorical meaning of ‘attack’ in (1b) follows the definition presented by Sweetser (1990) and Traugott (2004): it is historically a derivation of the fundamental meaning of ‘attack’; it is conventionalised towards the

metaphorical use of ‘attack’; and, it is relevant to the fundamental meaning of ‘attack’. When metaphorical expressions are conventionalised, they may be considered ‘below the threshold’ according to the definition of Recanati (2004): although being metaphorical in nature, they are usually not perceived as ‘metaphorical’ as typical novel metaphorical expressions would be.

Due to the complicated use of the term ‘metaphor’ in the literature, it is largely avoided in the dissertation; on the few occasions when the exact term ‘metaphor’ appears, it only refers to a conceptual metaphor that is not linguistically expressed but conceptually embedded. When it comes to the subject of the dissertation, i.e. the conventionalised linguistic realisation of metaphor, a series of terminologies is used: ‘metaphorical expressions’, ‘metaphorical meanings (of a lexical item)’, ‘metaphorically used lexical items’, and ‘metaphoricalness’. The central terminology among the four is ‘metaphorical meanings of a lexical item’: following the Metaphor Identification Paradigm (MIP) developed by the Pragglejaz Group (2007), all the meanings that are (1) a departure of the core, fundamental meaning(s) of a lexical item but (2) associated with the core meaning(s) are defined as ‘metaphorical meanings’ of that lexical item. The MIP paradigm identifies the metaphorical meaning by excluding the concrete, precise, historically older basic meanings from the contextual meaning of a lexical item. If we strictly follow this procedure to compare the contextual meaning and basic meaning for every lexical item, the concept of ‘metaphorical meanings’ under investigation is stable and objective. This has allowed us to distinguish the conventionalised metaphorical meanings from the literal meaning(s) in a dictionary entry for a particular lexical item. In fact, all the metaphorical meanings surveyed in this study have been identified in several mainstream dictionaries as non-core meanings of polysemous words.

The concept of ‘metaphorical expressions’ is developed based on the concept of ‘metaphorical meanings’. The basic idea is that a single lexical item without any context can be neither literal nor metaphorical, as the literal or metaphorical meanings of that lexical item can only be activated by an intrasentential context (Xia 2015a). A ‘metaphorical expression’ is a multi-word expression in which the metaphorical meaning of (at least) one lexical item is activated. In example (1b), the word ‘attack’ should be interpreted as ‘to criticise somebody or something severely’, which is a conventionalised metaphorical meaning of ‘attack’. Based on this, ‘attack’ is a metaphorically used word, and the whole expression ‘attack one’s proposal’, as the activator of the metaphorical meaning of ‘attack’, is a metaphorical expression. The fact that ‘attack one’s proposal’ is a metaphorical expression confirms the metaphoricalness of ‘attack’ in that expression. While the ultimate aim of this study is to explore the acquisition of the metaphorical meanings of lexical items, it is more practical to look at learners’ reaction to

metaphorical expressions, since a metaphorical meaning cannot be activated outside the context of a metaphorical expression. It should be noted that different metaphorical expression can activate the same metaphorical meaning: ‘attack an idea’, ‘attack a proposal’ or ‘attack an argument’ can all activate the metaphorical meaning of ‘attack’, ‘to criticise somebody or something severely’, so they should all be categorised as metaphorical expressions. Such flexibility of wording shows that conventional metaphorical expressions are less idiomatic and more semantically compositional than idioms; metaphorical expressions take the form of loose collocations, since the words collocating with the metaphorically used lexical item can be changed, while a strict collocation, like an idiom, does not allow any change of words.

The term ‘conventional’ used in this dissertation in the context of ‘a conventional metaphorical expression’ expresses two ideas: first, the metaphorical meaning activated by the metaphorical expression is conventional, i.e. ‘below the threshold’ (Recanati 2004); second and more importantly, the expression itself, or the collocation of words, is also conventional and frequently seen in daily language use. Therefore, two possible types of metaphorical expressions are ruled out from the research: expressions like ‘Chris is a cactus’, in which ‘an unapproachable person’ as the intended meaning of ‘cactus’, is not conventional and thus ‘above the threshold’; and, expressions like ‘attack the notebook’, in which ‘attack’ is intended to mean ‘to criticise somebody or something severely’, since the expression is less conventional than other uses of ‘attack’ listed in the last paragraph. This dissertation will focus exclusively on the acquisition of conventional expressions that contain a lexical item that expresses its conventional metaphorical meaning.

As mentioned in Section 1.1, one of the focuses of this dissertation is the comparison between metaphorical expressions and literal expressions. It has been a tradition in research on metaphor that metaphorical expressions are compared with literal expressions, and the same approach is adopted here: literal expressions, though not the focus of the current thesis, are treated as an essential counterpart of the metaphorical expressions and a baseline of learners’ acquisition. Terminologies used for the literal expressions are constructed and defined in a way similar to the terms for metaphorical expressions above. The ‘literal meaning’ of a lexical item is the core and fundamental meaning of that lexical item, usually the meaning with the longest history (Nacey 2014); most of the time, a lexical item only has one literal meaning, but in cases like homonyms there can be more than one literal meaning for a word. In a literal expression, the literal meaning of the critical lexical item is activated, such as ‘attacked Mary (with a stick)’ in (1a), since the word ‘attack’ in the expression displays its literal meaning, or at least a meaning that is not a clear departure from the literal meaning ‘to try to hurt or defeat using

violence'. The word 'attack' in (1a), then, is a literally used lexical item, and 'attack Mary' shows its literalness by using the literal meaning of 'attack'.

Throughout this dissertation, 'being metaphorical' is defined as a binary parameter, meaning that 'literalness' and 'metaphoricalness' are used as a pair or antonym: all the expressions in the experiments or discussions should be either literal or metaphorical and not both. Note that that definition is rather an idealisation of language use. From a contextualist semantic view, daily language use allows flexibility of interpretation and the literalness or metaphoricalness of a particular expression or sentence can be achieved by the manipulation of the greater context. It is possible and sometimes feasible for an expression to be interpreted both literally and metaphorically, as in example (2), in which both (2a) and (2b) can be valid interpretations of (2):

- (2) John attacked Mary in the meeting.
  - a. John beat Mary fiercely in the meeting.
  - b. John criticised Mary harshly in the meeting.

Such ambiguity as is expressed in (2) has been extensively discussed in the indexicalist account of metaphor (e.g. Leezenberg 2001; Stern 2000; Xia 2015b) and is recognised as inevitable in language use. However, ambiguous expressions like 'attack' in (2) are excluded from the current study, since the knowledge involved in comprehending (2) and generating the metaphorical interpretation in (2b) is related to the acquisition of the metaphorical meaning of a lexical item, and even more to the ability to make contextual inferences. None of the test items in the two experiments reported in the dissertation belong to the ambiguous case; all test sentences include a clear and solid intrasentential context that leads to *either* the literal meaning of a critical lexical item *or* to one of the metaphorical meanings of that lexical item.

This thesis reports two experiments, both of which have been designed with a 2×3 format: each experiment contains two major conditions (literal/metaphorical), and under each major condition there are three minor conditions (i.e. three types of cross-linguistic transferability). Each of the sentences in the metaphorical condition contains a metaphorical expression, and each sentence in the literal condition contains a literal expression using the same critical lexical item. If a literal expression and a metaphorical expression make use of the same critical lexical item, as in 'attack Mary' and 'attack Mary's proposal', they form a 'literal/metaphorical pair', and they become each other's 'corresponding literal/metaphorical expression'. The idea of 'corresponding expressions' is at the core of the design of the experiment and will be further elaborated in Chapter 4. When a literal expression and a metaphorical expression express the

same or a similar meaning, as in ‘criticise a proposal’ and ‘attack a proposal’, they become each other’s ‘literal/metaphorical counterparts’. The appearance of ‘literal/metaphorical counterparts’ with the same meaning will be discussed as the learners’ production of metaphorical expressions is reported.

#### 1.4 Outline of the dissertation

As discussed in Section 1.1, two major aspects of language acquisition can be associated with the acquisition of metaphorical expressions: first, cross-linguistic influence, and second, the structure of the bilingual lexicon. Chapters 2 and 3 contain the literature review for the two aspects respectively: Chapter 2 discusses previous research on cross-linguistic influence (transfer), and Chapter 3 on the different theoretical frameworks that try to explain the bilingual lexicon. Chapter 2 includes several core topics related to cross-linguistic influence, such as cross-linguistic influence on the acquisition of plain lexical items and different types of figurative language, and three factors that influence the transferability of a linguistic element, including psychotypical distance, the markedness of a linguistic element and the knowledge of a learner. Chapter 3 reviews several possible frameworks of bilingual lexicon, showing their strengths and weaknesses in explaining the acquisition of the metaphorical meanings of a lexical item; it also includes a section on the processing of figurative language, which includes metaphorical expressions, focusing on the retrieval of figurative meanings when a learner reads a figurative expression. The discussion on figurative language in these two chapters will strengthen the link between general discussion of second language acquisition and studies on figurative languages. At the end of Chapter 3 a framework for the processing of metaphorical expressions is constructed as part of the research hypotheses of the dissertation.

Chapter 4 describes the methodologies used in the current dissertation. Two experiments have been utilised in order to explore the three questions raised in Section 1.1: one is an offline task in which learners and native speakers are asked to rate the degree of acceptability of sentences containing metaphorical expressions, and the other is an online task in which learners and native speakers read sentences containing metaphorical expressions and their reaction patterns (reading time and general understanding) are recorded. The chapter provides an illustration of the main tasks in both experiments, as well as information on the supplementary tasks, including the proficiency test for learners and the digit span task to measure participants’ working memory. It also introduces the participants of the two experiments, and the method of data adjustment, coding and data analysis.

Chapters 5 and 6 are the two data chapters, in which the results of the judgment task and the online reading task are reported. All the analyses aim to answer the three questions in Section 1.1; therefore, the results are reported following the three comparisons: between literal and metaphorical expressions, between shared and language-specific expressions, and between native speakers and learners. The factors that influence transferability are also included in the report. In particular, traces of cross-linguistic influence on judgments of metaphorical expressions are identified in Chapter 5, and a thorough analysis of the relationship between traces of cross-linguistic influence and the three factors that influence transferability is provided. These two chapters also include a qualitative analysis of learners' production of metaphorical expressions and the literal counterparts of metaphorical expressions, and an analysis of learners' successful retrieval of the intended metaphorical meanings based on their performance in comprehension questions.

Chapter 7, as the final chapter of the dissertation, includes a general discussion, conclusions and implications. Based on the results of the experiments in the previous two chapters, answers are attempted to the fundamental research questions of this dissertation: (1) how a learner develops knowledge of metaphorical expressions when acquiring a language (here a Chinese learner acquiring English); (2) how cross-linguistic influence can affect the acquisition of metaphorical expressions; (3) how the influencing factors on transferability can affect cross-linguistic influence in the process of the acquisition of metaphorical expressions, and thus the results of the acquisition of metaphorical expressions; and finally (4) how L2 metaphorical expressions are stored in the bilingual lexicon, and whether there is any developmental change or asymmetry between literal and metaphorical expressions. At the very end of the dissertation, several possible directions for further research are given as an extension of the current study.

## 2 Cross-linguistic influence in second language lexical acquisition

### 2.1 Introduction: a short overview of studies on cross-linguistic influence

As a long-standing research topic in second language acquisition, cross-linguistic influence has received comprehensive discussion since the 1950s, even if some aspects of it still await further exploration. The phenomenon of cross-linguistic influence is referred to in SLA research literature using a number of different terms at different times. These include ‘transfer’, ‘native language influence’ and sometimes ‘interference’. Despite the variation in terminologies, the definition is clear: it is a phenomenon involving a bilingual speaker, usually a second language learner, incorporating certain elements of at least one previously known language in the production and/or comprehension of another language. In the current thesis, the two main terms, ‘cross-linguistic influence’ and ‘transfer’, are used interchangeably. When cross-linguistic influence on the acquisition of word meaning is discussed, the term ‘lemmatic transfer’ (as per Jarvis 2009) will be used.

In the field of second language acquisition, the majority of discussion is focused on the influence of the learner’s L1 on the L2, which is also the focus of the current thesis and this part of the literature review. However, other types of cross-linguistic influence, such as the influence of the L2 on the L1, and the joint influence of an L1 and L2 on an L3, or possibly an L1 and L3 on an L2, have also been extensively discussed in recent studies. The existence of cross-linguistic influence indicates that a bilingual speaker’s two languages are not stored independently of each other, but in an interconnected or (partially) overlapping system.

This thesis maintains the practice of other mainstream studies in second language acquisition in regarding cross-linguistic influence as a specific phenomenon in second language acquisition. The presence of cross-linguistic influence on a linguistic element, no matter what form it takes, indicates that a learner has not yet fully acquired the linguistic element. Cross-linguistic influence often takes the form of non-native-like production, perception or metalinguistic judgments (e.g. self-rating of confidence level by learners). When learners show a native-like production, perception or metalinguistic judgments, it may be considered that unless a linguistic element is shared between learners’ L1 and L2, cross-linguistic influence, particularly negative influence, does not appear, and learners are very likely to have acquired the linguistic element.

Early investigations of cross-linguistic influence from the L1 to the L2 are best represented by the Contrastive Analysis hypothesis (henceforth CA) proposed by Lado (1957). Lado

examined the phenomenon from the perspective of behaviourism, as developed for language by Skinner. It was widely believed at that time that knowledge of an L1 would be internalised as part of a learner's habits, and such habits would intrude into the production of the L2 and lead to errors. Lado suggested that such influences were more likely to lead to production errors if there was a clear difference between the grammar of the L1 and that of the L2; therefore, a comparison (contrastive analysis) between the L1 and L2 could reveal the linguistic elements that were most susceptible to error in the acquisition of the L2.

Such views, even though they could provide a relatively clear view of a proportion of cross-linguistic influence, have been shown to be problematic from several perspectives: The first difficulty concerns the assumption of behaviourism itself. It has been discovered that individuals do not acquire a language, either the first or the second, as a set of habits. Both children acquiring their first language and learners of a second language produce errors that are not found in the input, which means that they do not directly copy from adults or native speakers when they receive the input (see Meisel 2011 for a full historical review). Moreover, if it is assumed that a grammar is a 'habit', then cross-linguistic influence should only happen unidirectionally from the L1 to the L2, because only a grammar acquired earlier can become a 'habit'. The discovery of bidirectional influence at a relatively early stage of acquisition, namely the coexistence of L1–L2 influence and L2–L1 influence, contradicts the major prediction of the CA hypothesis. Therefore, since the mid-1970s, attitudes towards cross-linguistic influence have gradually shifted to a more developmental view, which has become the current mainstream.

With progress in relevant research and better understanding of cross-linguistic influence, it has further been discovered that grammatical differences between the L1 and the L2 do not necessarily lead to difficulties and 'errors' in L2 production by learners, and similarities do not always lead to the correct output (see Odlin 1989 for a summary). It seems that an L2 learner of language A will also replicate some errors produced by children when acquiring A as their first language, regardless of the similarities and differences between that learner's L1 and L2. Also, even if an element in the L2 is similar to the learner's L1, the learner might still make errors, and such errors cannot be predicted by the CA hypothesis. It therefore needs to be concluded that the acquisition of an L2 cannot be simply seen as being influenced by learners' behaviour. Instead, the developmental path of L2 acquisition should be taken into account and explained, by considering influence not only from the source language, but also from the target language.

Furthermore, cross-linguistic influence based on the CA tends to be seen negatively, while positive aspects of the phenomenon are largely ignored. Such bias originates from a flawed methodology: if only the differences between two grammars are emphasised and compared, the results are naturally in favour of ‘errors’ and ‘interference’ being due to the influence of L1. While these ‘errors’ are generally more prominent in observation, it is logically feasible, as has since been shown, that the elements shared between the L1 and L2 might show cross-linguistic influence as well. Influences on the same or similar elements in L1 and L2 are less ‘observable’, as often such influence leads to appropriate use of the L2 (but cf. the phenomenon of *false friends*).

In order to investigate cross-linguistic influence in a comprehensive way, one should be aware that cross-linguistic influence appears in various forms with different frequencies in the acquisition of an L2. Not only can the appearance of L1 structures in the production of the L2 be seen as transfer; but it has been discovered that avoidance of certain structures in the L2 can also reflect the influence of the L1 (Kellerman 1983). If a particular structure appears in the L2 but is absent in the L1, a learner may use other structures to express the same meaning while dismissing the L2-only structure, so as to preserve L1 grammar in the production of L2 utterances; such a learner may be less confident about the L2-only structure. A recorded case is the overuse of ‘make’ by Taiwanese learners of English: they prefer using ‘make someone upset’, rather than the direct verb form ‘upset someone’, because ‘upset’ cannot be used as a verb in Chinese and these learners follow the Chinese grammar in adding the verb ‘make’ (Wong 1983). It is difficult to record all the possible influences that an L1 has on L2 production just by investigating the interlanguage of a group of learners with similar linguistic and educational backgrounds. If possible, a comparison should be made between L2 learners with different L1 backgrounds, or between learners of a language and native speakers of that language, since both comparisons can capture the influence from a specific L1.

Several classic theories of cross-linguistic influence will be reviewed and discussed in the following, with two distinct focuses: cross-linguistic influence on the acquisition of lexical items in Section 2.2, and factors influencing cross-linguistic influence in Section 2.3. These two focuses together cover the target of this dissertation: the acquisition of metaphorical expressions should be seen as the acquisition of a particular kind of lexical item, and cross-linguistic influence in the process of acquisition may vary depending on learners’ perception and proficiency as well as the material to be acquired. The final section is a short summary of existing findings and some gaps in these studies.

## 2.2 Lemmatic transfer: a general review

### 2.2.1 Defining and delimiting lemmatic transfer

While the transfer of lexical items, or lexical transfer, can generally be defined as the presence of cross-linguistic influence during vocabulary acquisition in a second language, it is important to clarify what types of lexical transfer one might expect to observe in this thesis. In particular, with deeper understanding of cross-linguistic influence, the delicate boundary between cross-linguistic influence simply on word meanings versus influence involving conceptual change has been gradually discovered. Therefore, it is also important to clarify the scope of the term *lexical transfer*, and to establish what scope the main topic of this thesis falls within, namely transfer of metaphorical meanings of lexical items.

Jarvis (2009) argues that lexical transfer can happen at two different levels: lexeme level and lemma level. At lexeme level, the morphological and phonological features of a lexical item can be transferred to the L2, while at lemma level, the syntactic and semantic features can be transferred. Transfer at lexeme level is also referred to as ‘formal transfer’ (Ringbom 2006), since what is transferred in the process is the formal features, particularly morphological features, of a lexical item; lexemic transfer, or formal transfer, is more likely to happen between cognates. On the other hand, transfer at lemma level is referred to as ‘semantic transfer’ by Ringbom in the same article, because the most prominent transfer observed at that level is related to word meanings. Unlike lexemic transfer, lemmatic or semantic transfer can happen between cognates or non-cognates, and a similarity in form between the two words in the two languages is not always required.

Due to the general absence of cognates between Chinese and English, the focus of the current study is more on the lemmatic or semantic aspect of lexical transfer. Specifically, transfer of metaphorical meanings, which are essentially a part of the semantic properties of lexical items, is the focus of the thesis. Since the semantic aspect of a lexical item is usually seen as the link between the word form (e.g. the phonological or orthographical form) and the concepts it represents, the target of the current study is to investigate whether *such a link between the word form and a metaphorical concept* is transferrable from a learner’s first language to a second language.

It should be noted that a more fine-grained borderline should be drawn between lemmatic transfer and concept transfer, the latter drawing the attention of researchers in recent years. Language-specific concepts and the transfer caused thereby frequently appear in the language of bilingual speakers, as discussed in detail by Pavlenko (2009). To provide an example, *chashka* in Russian and *cup* in English are generally seen as translation equivalents, but *chashka*

cannot be used to refer to cups made of paper or plastic; instead, these are called *stakanchiki*, or ‘small glasses’. When an English learner of Russian misuses the word *chashka* to refer to a paper cup, not only are two lexical items ‘mixed up’, but there is also a direct transfer of the concept of CUP to Russian. While in Russian CHASHKA contains the concept of ceramic cups and STAKANCHIKI contains the concept of paper cups, a new category CHASHKA may be created by an English learner of Russian that includes both CERAMIC CUPS and PAPER CUPS, and, as a result, the form STAKANCHIKI to refer may not be used to paper cups. In this case, what is transferred is the way in which a learner categorises the individual items and formulates the conceptual categories. Hence this phenomenon is termed ‘concept transfer’ and subsequently distinguished from lemmatic transfer. To summarise briefly, lemmatic transfer affects whether a lexical item can be linked to an established concept, such as whether ‘glass’ can mean a paper cup, while concept transfer affects how different concepts in the L2 according to a learner’s knowledge of L1, such as whether glass and cup are intended to be ‘the same thing’.

It is not only literally-used words that may be linked to language-specific concepts, but also metaphorical meanings. For example, *chi* in Chinese can either describe the action ‘eat’ or be used to describe bad experiences involving ‘suffering’, because, when one experiences a loss, one can only endure it by ‘swallowing/eating the experience’. The meaning of ‘suffering’ is clearly not covered by ‘eat’ in English, and thus *chi*–‘suffering’ should be regarded as a language-specific link and any use of *eat* by a Chinese learner of English to express SUFFERING clearly reflects some type of transfer. At first glance, one might think that such transfer of metaphorical meaning might belong to concept transfer, since it is related to concepts in general. However, if we compare the examples of *chashka* and *chi*, we can see that the case of *chashka* involves recategorising paper cups in that system, which leads a learner to change the concept of *chashka*. In the case of *chi*, a learner is still fully aware that eating and suffering are two distinct concepts, and does not add or remove any subordinate concepts from the superordinate concept ‘eating’; what has happened instead is that a link has been created between *eat* and ‘suffering’. In this case, the transfer of metaphorical meaning is a case of ordinary lemmatic transfer. Therefore, it seems that transfer of metaphorical expressions is a form of lemmatic transfer rather than concept transfer.

Pavlenko herself (2009) has also analysed the type of transfer to which metaphorical expressions belong. She suggests that knowledge of metaphorical expressions belongs not to concept representation but to semantic representation, which includes implicit knowledge of ‘the mapping between words and concepts determining how many concepts and which particular concepts are expressed by a word via polysemy or metaphoric extension’ (2009: 148).

That means that she regards the metaphorical concept and the literal concept represented by a single lexical item as if they were two unrelated, independent concepts. Whether this statement is feasible from the perspective of metaphor will be discussed further in Section 3.2.2, but it is at least recognised by other researchers that transfer of metaphorical expressions is semantic, and thus can be discussed in parallel with other phenomena of lemmatic transfer and transfer of other types of figurative language.

### 2.2.2 Types of realisation of lemmatic transfer

Once we have assumed that transfer of metaphorical expressions belongs to lemmatic transfer, we can survey the possible outcomes of lemmatic transfer, whether beneficial or obstructive to the acquisition of vocabulary. Various types of realisation and the outcome of lemmatic transfer will be analysed and summarised in this section to show how lexical transfer can facilitate, complicate, or have no significant impact on, the acquisition of L2 lexical items.

Positive transfer with regards to lemmatic transfer mainly comes from transfer between cognates; however, cognates involve not only lemmatic transfer, but also lexemic transfer. The shared phonetic/orthographic form as well as identical or similar meanings in a pair of cognates can effectively facilitate the acquisition of L2 lexical items. With a proper knowledge of L1 vocabulary, a learner will easily identify the cognates in the L2 and transfer relevant knowledge to the acquisition of these lexical items. We can then draw a further assumption: when a learner observes a large number of cognates, the L1 and the L2 may be perceived as relatively close in that learner's psychotypology, which, as will be discussed in section 2.3, may boost both positive and negative transfer from the L1 to the L2. A series of investigations into the acquisition of an L3 by bilinguals (e.g. Sjöholm 1976; Ringbom 1978; Bardel and Lindqvist 2006; Leung 2005) reveals that when learners make use of their lexical knowledge of two languages to acquire L3 words, they will select the language that shares more cognates with the L3.

Although positive transfer is widely observed between cognates, it can also happen between non-cognates, even if this has been less discussed in previous studies. Positive transfer between non-cognates may happen when a learner, who attempts to acquire an L2 phrase based on lexical items that have already been acquired, assumes that a phrase that is available in the L1 is also available in the L2 if (1) the lexical items involved in the phrase have the same meaning across the L1 and the L2; (2) that phrase is also available in the L1; and (3) that phrase is not typically idiomatic, i.e. can be interpreted by compositing the semantic meanings of each lexical item..

As discussed in Section 2.1, cross-linguistic influence is generally less observable when it comes in the form of positive transfer. This is reflected in the current research on lemmatic transfer: most discussions observe a wide range of negative transfer and emphasise errors in word use due to interference from L1 vocabulary knowledge. Common problems due to negative lemmatic transfer include the phenomenon of false friends and semantic overextension. False friends involve L1 and L2 lexical items with a certain degree of morphological similarity but different meanings; a typical example is the word *sensible*, which means reasonable in English, but sensitive in French. Odlin (1989), however, proposes another type of false friend that is less typical but widely observable. This type does not involve morphological similarities but only semantic similarities, and thus, can be seen as the overextension of partial translation equivalents. In order to be classified as ‘translation equivalents’ a pair of lexical items in two languages need to share at least (part of) the lexical meaning, but it is not necessary for them to share all their meanings (for example if both words in the pair are polysemous). As a result, a learner might wrongly infer that all the meanings of the (partial) translation equivalents are shared when they are not, and then mistakenly transfer the non-shared meanings. For instance, a Finnish learner of English might use ‘spin’ to describe the ‘purring’ of a cat, because the Finnish word *kehräta* is polysemous and corresponds to the meanings of both ‘spin’ and ‘purr’ in English (Jarvis 2009).

The misuse of polysemous false friends in a second language, especially the incorrect inference that the L2 lexical item is also polysemous, might indicate an assumption that a given lexical item in a learner’s L1 and L2 forms a one-to-one correspondence. As shown in the example described above, the misuse of ‘spin’ is more like a mismatch between the word and the context in which it appears. While this can occur when the collocation is loosely formed, as in the case of ‘the purring of a cat’, it can also occur when the collocation is a fixed multi-word expression, such as in idioms. Metaphorical expressions, while residing between a loose, flexible collocation and a fixed idiom, might also encounter such false friends in the process of acquisition. It is possible for learners to assume that, if a metaphorical meaning, such as ‘suffering’ in the case of the lexical item ‘chi/eat’, is available in Chinese, it may also be possible in English. In such cases, the misuse of ‘eat’ in ‘eat some loss’ (meaning *suffer a loss*) is indeed comparable to the misuse of ‘the spinning of a cat’.

However, lemmatic transfer is not the only cause of errors. It has been proposed that the frequency of certain structures in L2 production may be different from the frequency of those structures in native speakers’ speech. Unlike qualitative differences that can be easily observed by looking for errors in L2 learners’ production, frequency differences can only be captured by

a more delicate comparison between the production of L2 learners and that of native speakers. Generally, such differences include both overproduction and underproduction (Odlin 1989). The reasons for over- and underproduction might be a lack of knowledge of appropriate L2 lexical items, or possibly the transfer of peripheral features of L1 lexical items, such as the frequency of a word or an expression in the L1. While the frequent appearance of an L1 word may lead to the overproduction of its translation equivalent in the L2, the lack of translation equivalents of an L2 word in a learner's L1 can result in underproduction. Sometimes, even if a learner has acquired a new expression in the L2, it may not be selected due to a lack of translation equivalents in the L1.

In the same article, Odlin suggests that both overproduction and underproduction should be seen as negative transfer, in much the same way as the impact of false friends, since both lead to production differences between learners and native speakers. Nevertheless, I would argue that overproduction and underproduction are not as 'negative' as the examples of false friends mentioned above, because (1) they do not directly lead to any significant semantic errors or breakdown in communication; and (2) they involve quantitative rather than qualitative differences between learners and native speakers. Although it is possible that overproduction and underproduction might make learners sound less native-like, the production influenced by them is not 'wrong'. It is not the case that a learner is not able to learn or produce a structure or expression, but it is difficult to master the more peripheral features, such as the frequency, of that structure or expression.

Other cases of lemmatic transfer that do not affect semantic meaning, as listed below, have also been observed, although sometimes it is debatable whether they are authentic cases of lemmatic transfer. While we have already seen false friends as a case of inappropriate semantic overextension, semantic overextension may also happen without the presence of a pair of (partial) translation equivalents, and sometimes without leading to any significant semantic errors. For instance, Bamgbose (1982) reports an example of semantic overextension in Nigerian L2 English: 'being away' can be phrased as 'travel', and a sentence like 'My father has travelled' can mean 'my father is away'. In such situations of semantic overextension, a lexical item can be used in L2 production to represent a meaning that is generally relevant to its original meaning, especially when a learner fails to select an appropriate L2 word and faces a lexical gap (Bamgbose 1982). This phenomenon is alternatively referred to as 'approximation' (see Blum and Levenston 1978). It should be noted that, even though a learner may use semantic extension to fill a lexical gap, this does not indicate an inability to differentiate between the two concepts. In a case where an Apache learner of English uses 'dead

food' to refer to 'rotting food', there is no evidence to show that that learner mixes the concept of 'death' and 'rot', and it is possible that that learner is fully aware that food does not have a life and can never 'die'. Therefore, semantic extension and approximation should be classified as semantic transfer rather than concept transfer.

The semantic overextension or approximation described above, as in the case of 'travel', does not necessarily lead to significant communication breakdown, as in the case of 'the spinning of a cat' does. Odlin (1989) shows reluctance to classify such semantic overextension as 'lexical transfer', because he suggests that this type of overextension is not attached to any specific language, but is universally possible between any language pairs, and even within a language. Since neither semantic overextension nor approximation is a direct reflection of the structures of either the L1 or the L2, it cannot be categorised as 'transfer from the L1 to the L2'. I would suggest, however, that such semantic overextension should be seen as lexical transfer from another perspective. The use of semantic overextension is, and must be, accompanied by a certain degree of underproduction, because the appearance of an overextended expression in a particular context always indicates that a learner has not yet fully acquired an 'appropriate' L2 expression that should be used in that context. A learner who does not know the 'appropriate' L2 expression is faced with the choice of either using the more radical strategy of transferring an L1 expression, which might lead to semantic errors, or using a language-neutral strategy and adopting an approximation, which will seem to be safer. Both of these strategies do in fact show that a learner is utilising *knowledge of L1* to fill a 'lexical gap' in L2 knowledge in different ways, and thus they should both be seen as cross-linguistic influence. The strategy of semantic overextension and approximation, as well as other language-neutral strategies (such as the use of literal expressions instead of figurative expressions), might be especially favoured when a learner finds that the intended L2 meaning is less available in the L1.

This section has briefly introduced different realisations for lemmatic transfer that have been discussed in previous research. Lemmatic transfer may lead to semantic errors in production (as in the case of false friends), as well as differences in frequency and possibly a sense of non-nativeness (as in the case of overproduction and underproduction), and sometimes to a less precise paraphrase of the intended meaning (as in the case of semantic overextension and approximation). In the current thesis, we expect to see all these possible types of realisation of lexical transfer: transfer between shared metaphorical expressions will lead to the facilitation of the acquisition of metaphorical expressions, even if what is transferred is non-cognates; transfer between different metaphorical expressions in the L1 and the L2 will lead to semantic errors in production. Furthermore, we expect to see some types of semantic extension or

paraphrase of metaphorical expressions when knowledge of L2 metaphorical expressions is absent from learners' knowledge. Although semantic extension or paraphrase is not likely to lead to any significant errors in production and comprehension, its occurrence with metaphorical expressions only may indicate that metaphorical expressions are more difficult to acquire than literal expressions, and thus need more attention from both learners and instructors.

### 2.2.3 Lemmatic transfer in the acquisition of figurative language

Several features of figurative language have further complicated transfer of these expressions: (1) figurative language often appears in multi-word structures, such as phrasal verbs and idioms; (2) figurative language (especially idioms) is generally semantically opaque, which means it is difficult to infer the meaning of a figurative expression from its components; and (3) figurative language is usually language-specific, or at least perceived as language-specific. It should be noted that in some studies the term 'formulaic expressions' may be used as an alternative to refer to some figurative expressions, especially idioms (see Cieślicka 2008 for an example). The name 'formulaic expression' indicates that such expressions consist of multiple words with a fixed order and a stable meaning across different contexts. However, the two concepts are essentially different, since a formulaic expression is not required to be figurative, and in some cases, the component that is identified as 'figurative' in an expression may be only one word, as is the case in the current dissertation. The transferability of several types of figurative language, predominantly idioms, has received intensive investigation in past decades, and this section will be devoted to a review of the progress in that area, starting with idioms.

Idioms demonstrate almost all the prominent features of figurative language mentioned above (see Grant and Bauer 2004 for a comprehensive redefinition of idioms), and that makes them seemingly less transferable from one language to another. The actual transferability of idioms has long been a focus of studies on cross-linguistic influence in the acquisition of lexical items: idioms have been the subject of well-organised surveys investigating transferability and acquisition, while other types of figurative language, including metaphorical expressions, have generally been less discussed. The transferability of idioms may, however, shed some light on the transferability of metaphorical expressions: on the one hand, with both being examples of figurative language, the two types of expression may share some similarities in terms of their figurativeness, while the distinction between these expressions and literal uses of lexical items may predict similar patterns of transfer in L2 acquisition; on the other hand, metaphorical expressions in the current dissertation do not significantly overlap with idioms, which means

that the features of the acquisition of idioms in a second language can only be seen as a reference point when we formulate hypotheses and design experiments to investigate the acquisition of metaphorical expressions.

In the view of Jordens and Kellerman (1981), idioms are mostly language-specific and semantically opaque. As a result, learners are not always aware of the corresponding expression in the L2, and consequently will avoid transferring known idioms from the L1 into the L2. While Kellerman (1983) points out that his initial proposal does not indicate that idioms are not transferrable, he suggests that idioms are indeed less transferable than the literal use of lexical items and other semantically transparent expressions. Kellerman's conclusion, however, is a rather general one that ignores differences between idiomatic expressions, and is only applicable in the comparison between the acquisition of idioms and that of other elements in second language acquisition. Within the category of idioms, it is possible that the transferability of individual idioms will still vary according to certain principles. Furthermore, the key factor measured by Jordens and Kellerman (1981) is the grammaticality of the idiomatic expressions, as perceived by L2 learners, and the acquisition of idioms does not depend only on the grammaticality issue; the more important factor is the semantic aspect of the idiom. Therefore, a grammaticality judgment test cannot fully capture whether learners transfer their knowledge of semantics in the processing and production of idioms, and an experiment should be designed to detect learners' understanding of those expressions. This need also exists for other types of figurative language, for which we expect semantic transfer to happen.

Regarding the exact transferability of different types of idiomatic expression, Irujo (1986a, 1986b, 1993) has investigated influence from the degree of cross-linguistic formal similarity of idioms on their transferability among advanced Spanish learners of English. She classifies the translation equivalents of idioms into three major categories: (1) idioms that share an identical form between the L1 and the L2, i.e. the L2 idioms can be seen as a word-to-word translation of the L1 idioms; (2) idioms that share a similarity in form between the L1 and the L2, i.e. the L2 idioms, although not the exact word-to-word translations of the L1 idioms, make use of most of the words in the L1 idioms, with one or two words altered; and (3) idioms that are totally different in their L1 and L2 forms. Such categorisation leads to a continuum of formal similarity when semantic similarity is controlled: the first category is strictly similar in form, and the third not similar at all. This categorisation thus makes it possible to investigate the influence of the formal similarity of idioms on cross-linguistic influence. Irujo (1986b) then compares learners' comprehension and production of different categories of idioms in multiple choice questions, a definition task, a discourse completion task and a translation task. In particular, the participants

are asked to complete English paragraphs with appropriate English idioms with the Spanish text given, and in the instruction and example sessions they are encouraged to use English idioms. When the proficiency of participants is controlled, they perform best when they encounter *identical* idioms; they demonstrate greater accuracy in both the comprehension and the production of such idioms, which indicates that they can make use of positive transfer to acquire the idioms. When the participants acquire *similar* idioms, they can comprehend them as well as they comprehend identical idioms; however, the production of these idioms is influenced by the L1, such that negative transfer can be observed. No transfer is detected when the idioms are in different forms, and the different idioms are more difficult to comprehend and produce, which means that difficulty in acquisition is greater than for idioms that share some elements. From such results Irujo proposes the so-called ‘Transfer Theory’ of idiom acquisition, which suggests that idioms are more likely to be acquired if they are both formally and semantically identical between the learners’ L1 and L2.

Based on the methodology and the results of previous studies (Irujo 1986b; Kellerman 1983), Irujo (1993) further explores whether highly advanced learners may avoid using idioms in L2 production if they are free to choose. She targets her study at fluent Spanish-English bilinguals who began the acquisition of English in adulthood and reside in an English-speaking environment, which means that they possess a fair knowledge of figurative language and have received sufficient exposure to native speakers’ English production. The participants were required to translate Spanish text containing idioms into English, but they were not instructed to use idioms in the translation, contrary to the previous study. As with the previous studies, incorrect word-to-word translation from L1 idioms is observed in the experiment, thereby showing a trace of L1 transfer in the production of L2 idioms. Moreover, the results show that, while fluent bilinguals are able to produce correct idioms (both intended and not intended by the experimenter) most of the time, they also make use of non-idiomatic paraphrase. Non-idiomatic paraphrase is the second method frequently used by the learners, and is widely used in situations in which the L1 and L2 idioms are different in form. Among these non-idioms, literal paraphrase is used more frequently than metaphorical non-idiomatic paraphrase. In other cases, the participants failed to produce the English idioms entirely by omitting them or giving up translating them. Irujo identifies failure of production as instances of avoidance, and points out that avoidance of L2 idioms can happen among advanced L2 learners and even fluent bilinguals living in an L2 environment. However, she suggests that the paraphrase of idiom in the translation task does not indicate that the fluent bilinguals avoided using idioms, because

such paraphrase is regarded as an alternative means to express the same meaning, and the participants may have chosen paraphrase for reasons other than communication requirements.

As well as the formal similarities between L1 idioms and L2 idioms, other factors that may influence the acquisition of idioms have also been taken into consideration. A prominent feature that may greatly contribute to discussion of figurative language acquisition is the degree of semantic transparency of an expression. The semantic transparency of an expression is the possibility of inferring the meaning of that expression from the meaning of each component item (Vega-Moreno 2007). For example, Vega-Moreno argues that language users may be able to derive the meaning of ‘spill the beans’ by composing the meaning of each component together, and thus ‘spill the beans’ is semantically transparent. On the other hand, they may fail to do so for ‘kick the bucket’, which means that ‘kick the bucket’ is semantically opaque. Irujo (1993) has already found that the degree of semantic transparency of an idiom can influence the outcome of acquisition. Overall, semantically transparent idioms are better acquired and produced by learners than semantically opaque ones. It can be further inferred that semantically transparent idioms may be more likely to be transferred from the L1 to the L2 than semantically opaque idioms, because learners might believe that a transparent idiom in their L1 may also be semantically transparent in the L2.

The concept of semantic transparency is crucial here because it may influence the acquisition of different types of figurative language. Regardless of their exact degree of semantic transparency, idioms are generally seen as semantically opaque, and such a view of opacity is observed among both native speakers and L2 learners. Therefore, unless individuals receive sufficient exposure to such opaque idioms, they may fail to understand the idioms: both native speakers and learners need sufficient input, be it through observation (native speakers and learners) or through explicit instruction (L2 learners). Metaphorical expressions, however, are figurative even if they are not always semantically opaque. As discussed in Section 1.2.2, the figurativeness of metaphorical expressions comes from a departure from the core, literal meaning of the lexical item. When a metaphorical meaning is seen as a meaning of a polysemous lexical item, the meaning of a larger constituent containing a metaphorically-used word can be derived in a compositional manner, and the only difference is the replacement of the literal meaning of that lexical item with the metaphorical meaning. This higher degree of semantic transparency in metaphorical expressions may actually ease the process of acquiring these expressions for both L1 and L2 users. According to Sweetser (1990), a native speaker is capable of deriving the metaphorical meaning(s) of a lexical item when given the contexts and the literal meaning of that lexical item, since the metaphorical meaning(s) are closely connected

to the literal meaning. In a similar fashion, in a situation in which (1) a learner has acquired the literal meaning of a lexical item, and (2) that same lexical item is presented metaphorically with appropriate contexts (i.e. a collocation that is biased towards the metaphorical interpretation), it may be possible to infer the correct metaphorical meaning of the lexical item, even if the metaphorical meaning has never been encountered before. Hence, compared with idioms, it may be easier for learners to acquire and accept metaphorical expressions, and it is possible that some learners can acquire metaphorical expressions even without guidance from instructors.<sup>1</sup>

The categorisation of idioms by Irujo has become a standard categorisation in studies of L2 idiom acquisition (e.g. Laufer 2000; Liontas 2002; Zhang 2008; Türker 2016a). The methodology used by Irujo has been applied to research on the acquisition and processing of idioms between different pairs of L1 and L2 (e.g. Liontas 2002; Zhang 2008; see Laufer 2000; Bortfeld 2003 for two similar paradigms). One of the weaknesses of the experimental methodology is obvious, particularly in the investigation of cross-linguistic influence. The production task is actually a translation task, since the text in the L1 is provided to the participants and the participants must refer to the L1 text in order to know how to complete the L2 paragraph, even if they are not told in the instruction that they should translate the idioms. The presence of L1 idioms may draw additional attention from participants to the form of L1 idioms, which may result in a greater effect from transfer in the production of L2 idioms. Therefore, alternative offline tasks have been developed and applied in this area, including a sentence completion task that is only administered in the L2 (Cieślicka 2006) and a discourse completion task in L2 conversations (e.g. Türker 2016a).

Further offline experiments on the acquisition of figurative language and the transfer thereof include the investigations of the use of phrasal verbs (Matlock and Heredia 2002), the interpretation of sentential metaphors (Littlemore 2010), and the interpretation of the linguistic realisations of HAPPY, SAD, and ANGER metaphors (Türker 2016b). Matlock and Heredia (2002) asked English learners from a variety of L1 backgrounds to complete sentences with a phrasal verb that is semantically opaque, and then compared the results with the production of native English speakers. Phrasal verbs are generally not available in learners' L1s, and in their L1s, a combination of a verb and a prepositional particle is more likely to be interpreted simply as a verb and a preposition; for instance, while the phrasal verb 'sleep in' is available in English as 'to sleep longer than usual', non-native English speakers may interpret it as 'to sleep in a

---

<sup>1</sup> While the described situation exists and we can take native speakers as an optimal example, the actual performance of L2 learners may vary due to language aptitude, sociolinguistic factors or other reasons, and it is probable that L2 learners in certain sociolinguistic backgrounds are less effective when making the inference of metaphorical meaning. See Section 3.2.4 for a review of relevant studies.

location'. The result of the sentence completion task shows that learners are significantly less willing to interpret a phrasal verb like 'sleep in' as an entire structure; instead, they tend to interpret it as verb + preposition, and produce 'sleep in his own bedroom'. That result indicates that learners are less likely to go for the non-literal meaning of a combination of a verb and a prepositional particle. However, Matlock and Heredia also report that proficient English learners were able to recognise and use phrasal verbs in the experiment, since this learner group interpreted the given expressions more as phrasal verbs than as combinations of a verb and a preposition. Türker's (2016b) study is among the first to investigate the acquisition of conventional metaphorical expressions. In his experiment, English learners of Korean are asked to interpret the meaning of 54 Korean conventional metaphorical expressions in three categories: (1) same conceptual mapping and same linguistic expression; (2) same conceptual mapping but different linguistic expression; and (3) different conceptual mapping and different linguistic expression. The results show that in general learners give the most accurate interpretation to metaphorical expressions that are both conceptually and linguistically shared between English and Korean, and the smallest number of accurate interpretations to metaphorical expressions that are both conceptually and linguistically different in English and Korean. These experiments reflect the fact that cross-linguistic similarities facilitate the acquisition not only of idioms but also of other types of figurative language, while cross-linguistic differences may create obstacles in the process of acquisition, which is exactly the same as for the acquisition of ordinary lexical items and literal expressions.

While Irujo's experimental paradigm is also applicable to research on the acquisition of metaphorical expressions, some modifications should be made in order to capture possible cross-linguistic influence accurately. First, as discussed earlier, a translation task, such as that of Irujo (1986b), or any task involving reading in one language and writing in another (e.g. Türker 2016b), may bias a learner and lead to a more significant effect from cross-linguistic influence. Such bias can be largely eliminated if the experimental task is conducted monolingually; that is, if learners receive the metaphorical expression and produce feedback in the same language, namely the L2. Secondly, although metaphorical expressions in the two languages can be divided into categories according to the form and meaning of those expressions, it should be noted that the definition of 'identical/similar' for metaphorical expressions must be different from Irujo's classification. According to her classification, two idioms are cross-linguistically 'identical' only when they can be translated in an exact word-to-word fashion. Therefore, the English idiom 'a needle in a haystack' and its Chinese counterpart 'a needle in the sea' are not identical, but only similar, because there is a mismatch between the

word ‘haystack’ and ‘sea’, even if other words in the multi-word constituents are all the same. As for metaphorical expressions, since only one word is interpreted metaphorically in the expressions, there is no requirement for the context also to be a word-for-word equivalent. If a pair of translation equivalents shares the same metaphorical meaning in the same collocation or context, they can be seen as identical metaphorical expressions; otherwise, they are essentially different. A status of ‘similar metaphorical expression’ does not exist in the current study.

By utilising the modified version of Irujo’s design in this thesis, we expect to see different effects from cross-linguistic influence in the acquisition of metaphorical expressions, which will be in line with Irujo (1986b) and will follow studies of figurative language more generally. Thus, when a lexical item can only be interpreted metaphorically in one language, we propose that the following cross-linguistic influence can be observed: if the metaphorical expression is L1-specific, when a learner is asked to judge a word-to-word translation of such an L1 expression, we expect to observe acceptance of the expression due to L1 interference; if the metaphorical expression is L2-specific, we expect to see underproduction or avoidance of the L2-specific metaphorical expression, such as rejection in acceptability judgment, omission of the expression, and possibly different types of paraphrase.

### 2.3 Factors influencing cross-linguistic influence

In a series of studies from the late 1970s to the mid-1980s, Kellerman and Jordens have investigated the strategies of transfer adopted by second language learners, particularly the conditions in which learners are more willing to transfer elements from their L1. It should be noted that, if a ‘strategy of transfer’ is being investigated, then ‘transfer’ is assumed to be consciously controlled rather than unconsciously used. Within such a framework, two types of transferability should be clearly differentiated. One is the binary, objective transferability of a linguistic element, which is decided by the linguistic elements available in a learner’s L1 and L2. If a linguistic element is available at the same time in both L1 and L2, then it is objectively transferable from the L1 to the L2. The other is the subjective transferability of a linguistic element as perceived by the individual learner in the form of a continuum. This subjective transferability of a linguistic element varies between individual learners; it is probable that for the same pair of L1 and L2, learner A may perceive a linguistic element as fully transferable while learner B may perceive it as not transferable at all. The focus of the studies by Kellerman and his colleagues, then, can be seen as an attempt to identify the possible factors influencing the subjective transferability of a certain linguistic element among different learners.

Since these seminal works on the influential factors on transfer (or cross-linguistic influence), many possible linguistic and extra-linguistic factors that may influence learners' strategies of transfer have been further analysed. In this subsection, the three main factors proposed by Jordens and Kellerman (1981), namely psychotypology, markedness and learners' knowledge, as well as further factors more recently identified, will be discussed.

### 2.3.1 Psychotypology

Psychotypology, as defined by Jordens and Kellerman (Jordens and Kellerman 1981; see also Kellerman 1983), refers to the distance between the L1 and the L2 as perceived by a learner; in other works, the concept is referred to by other terminologies, such as 'typological proximity', 'relatedness' or simply 'language distance' (see De Angelis 2007 for a summary of the terminological variation of psychotypology), even though 'psycho-', which is crucial for the concept, is not used in some of these terms. The concept of psychotypology is essentially different from the objective typological distance between two languages as investigated in linguistic typology, because psychotypology, as shown from the name, is based on each learner's psychological perception or 'feeling' of distance. Psychotypology is based on a learner's understanding of structural similarities and differences between languages, which is rooted in the learner's own metalinguistic awareness, and sometimes in folk linguistics. When a learner identifies that certain structures in the L2 are similar (or not) to the corresponding structures in the L1, or hears that the L2 is generally similar (or not) to the L1, the perceived distance between the two languages will be constructed/adjusted: it will become shorter if the two languages are thought to be similar, longer if the two languages are thought to be different. Because the concept of psychotypology is different from linguistic typology, the psychotypological distance between two languages as perceived by an individual may, or may not, correspond to the actual typological distance between them. It is also possible for learners to change the psychotypological distance between L1 and L2, when they successfully perceive the similarities or differences between the two languages.

Psychotypology contributes to transfer from the L1 to the L2 by providing a robust background for assessing the suitability of transfer between two languages. If the psychotypological distance between the L1 and the L2 is relatively close, then it is more likely that a learner will believe that the two languages share a number of features. That learner will then transfer L1 information in the production of the L2 more readily than when the two languages are assessed as psychotypologically remote. For instance, in Jordens and Kellerman (1981), Dutch native speakers tend to consider that Dutch and German are more closely related

than Dutch and English; therefore, they are more willing to transfer features from Dutch into the production of German, while they show greater suspicion of the transfer of such features into English, even if the features are actually available in English.

On the other hand, Kellerman (1983) suggests that transfer will be suppressed if a learner believes that the L1 and L2 are not sufficiently related. Ringbom (1978) examines the acquisition of English by Finnish and Swedish bilinguals, where English is perceived as closer to Swedish than Finnish. Regardless of Swedish proficiency, learners tend to transfer elements of Swedish to English, and most of the production errors reflect the influence of Swedish grammar. Finnish does not show any evident sources of transfer effect, nor does it lead to production errors. Such asymmetry in the transfer effect between the two L1s, as is explained by Ringbom (1978) and subsequently by Kellerman (1983), is due to the different psychotypological distances to English. Since Finnish is too remote from English in the perception of learners, the transfer from Finnish to English is suppressed, even if some learners have mastered Finnish better than Swedish. Ringbom's paradigm has since been widely used in research on L3 acquisition (e.g. Bardel and Lindqvist 2006; Leung 2005; Singleton 2016) to show that the perceived validity of transfer varies due to different psychotypological distances between languages, and learners always show a preference to transfer knowledge from the psychotypologically closest language, regardless of the actual proficiency of the languages (another proposed influence on transfer).

Although not explicitly mentioned in the works of Kellerman (1983, 1984; see also Jordens and Kellerman 1981) and subsequent literature, several inferences can be made about the flexible property of psychotology. Firstly, psychotology, as a subjective, perceptual factor, may differ between individual learners with the same or similar linguistic backgrounds. Even if a group of learners are from the same classroom, speaking the same L1 and learning the same L2, they may still perceive the distance between their L1 and the L2 in different ways. Such differences, following the argument of Jordens and Kellerman, can and should be reflected in the strategies of transfer adopted by these learners. Learners who perceive a shorter psychotypological distance between the L1 and L2 should be more willing to transfer elements from the L1 to the L2, whereas learners who perceive a longer psychotypological distance should suppress such transfer.

Secondly, since the psychotology perceived by learners is closely related to their observation of similarities and/or differences between L1 and L2, it is possible that different instruction methods or knowledge will affect their perception of psychotology. If an instructor explicitly emphasises the similarities (or differences) between the L1 and the L2, or

learners observe more similarities (or differences), it is likely that they will change their perception and shorten (or lengthen) the psychotypological distance between the two languages. In summary, it then can be argued that the psychotypological distance between a learner's L1 and L2 may not be static throughout the L2 acquisition process. Considering the nature of psychotypology, it would be erroneous to assume that the psychotypological distances between a pair of languages perceived by a group of learners with similar linguistic backgrounds are always static and homogenous.

Based on the concept of psychotypology, several theories regarding the transferability of linguistic elements in different languages and the 'distance' between different languages have been developed. De Angelis (2007) summarises the role of perceived distance between two languages, saying that it has been widely recognised in research on transfer; it can be seen, however, from trends in research and publication, that most of the recent discussions of psychotypology take L3 acquisition as a background, which follows the tradition of Ringbom (1978). Leung (2005) continues to use the notion of Kellerman and suggests that the initial state of an L3 comes from the psychotypologically closest language that a learner has already acquired. Jin (2009) makes a further inference, that transfer from the L2 may occur if a learner's L2 and L3 are (psycho)typologically close, even if the learner has not fully mastered the L2 when commencing acquisition of the L3. This proposal is usually called the Interlanguage Transfer Hypothesis (see Jin 2009; Westergaard and others 2016).

Focusing on L3 acquisition, Rothman and colleagues (2011; see also Rothman 2013) propose the Typological Primacy Model (henceforth TPM), suggesting that the initial state of syntactic properties available for L3 acquisition comes from either the L1 or L2, whichever is 'the closest (psycho)typological language' (Rothman and others 2011: 112; brackets added in the original text). The TPM and the original proposal of Jordens and Kellerman both suggest that transfer is more likely if a learner's L1 and L2 are perceived as being closely related. Slightly different from the original psychotypology proposal by Jordens and Kellerman, which tries to cover transfer in terms of both syntax and semantics, but does not raise any assumptions on an initial stage of acquisition, the TPM targets exclusively the transfer of syntactic properties and assumes that there is a blank initial state for L3. Before the exposure to an L3, transfer cannot occur because there is no linguistic material that can be used to draw a comparison between the future L3 and learners' L1 or L2. Rothman (2013) suggests that learners who have been exposed to an L3, even to a very limited extent, will look for a language they know that is (psycho)typologically closest to the new language, and then transfer all their knowledge of that language to build the initial state of the language. The fundamental assumption of the TPM is

that transfer to L3 comes from a single source, which seems to be in line with the results of Ringbom (1978), although it is difficult to ascertain what the initial state of the L3 is if a learner has only had limited exposure to the language. The TPM does not explain other aspects of language acquisition, but it can be assumed that the TPM can also contribute to transfer outside the scope of morphosyntax, as far as it is assumed that transfer happens not only to morphosyntax but also to other parts of language acquisition, such as phonology and lexical semantics.

Westergaard and colleagues (2016) also suggest that psychotypology may affect conscious transfer, but in a different fashion from the TPM. They suggest in their Linguistic Proximity Model (henceforth LPM) that, instead of transferring all the linguistic elements in a wholesale manner from the closest psychotypological language, a learner can be selective when deciding on the elements to be transferred. The LPM has been developed to explain the transfer observed from various sources in L3 acquisition, including both facilitative and non-facilitative transfer from both the L1 and L2, which cannot be accounted for by the TPM. The LPM proposes that that when bilingual learners acquire a third language, they will still take into consideration (psycho)typological distances from the L1 and L2 to the L3, but will transfer each property from different languages if either the L1 or L2 shares that property with the L3. The interpretation of psychotypology within the LPM, therefore, is significantly different from Kellerman's proposal and the TPM: psychotypological distance in the LPM is not measured between languages as entire entities, but between individual items in the two languages. In the LPM, it is highly probable that a learner's L1 will be psychotypologically closer to the L3 in one aspect but not in another, while the L2 will be psychotypologically closer or more distant in other aspects. In some ways this may reflect the dynamics of the psychotypology of a learner between two languages depending on the focuses of the perception of a learner, but a radical change in definition of *psychotypology* makes it more difficult to predict and explain the influence of psychotypology as shown in Jordens and Kellerman's proposal. It is also questionable whether the two *psychotypologies*, namely the distance between languages and the distance between individual linguistic elements, have different impacts on learners' L2 acquisition. In this thesis, the definition by Jordens and Kellerman is used, and the psychotypological distance between two languages will be measured.

All the proposals on psychotypology, including Jordens and Kellerman's original proposal, share the same core assumption that transfer is more likely to happen between more 'closely related' languages. That assumption covers different types of learner groups: learners with same L1 but different L2s, learners with different L1s but same L2, as well as learners of an L3.

However, all the proposals reviewed above also face three common problems. The first problem, as mentioned earlier, is related to the dynamics of psychotypology. While most, if not all, relevant studies presume that all the learners investigated in a single experiment must perceive a similar distance between a pair of languages, the real psychotypological distance as perceived by individual learners may vary considerably. By definition, psychotypology is a subjective supposition held by each learner, and we cannot expect that different individuals will hold the same belief in reality. Our lack of knowledge about individual differences in psychotypology may weaken the causal effect between psychotypology and transfer in second language acquisition: if a group of learners perceive different psychotypological distances between a pair of languages but demonstrate similar strategies in transfer, the strategies of transfer might in fact be influenced by factors other than psychotypological perception, e.g. instruction methods, knowledge of the target language, etc.

Furthermore, some proposals show explicitly a mixed use of objective typological distance and subjective psychotypological distance: they seem to take for granted that languages with a relatively short typological distance must evoke a short psychotypological distance among learners. Although it is possible to correlate psychotypological distance with actual typological distance, such correlation is not attested in any known study and thus is far from an accurate account of psychotypology. In addition, none of the proposals mentioned above makes use of a typologically verified system of typological distances between languages to support the assumptions of psychotypology. That is, in previous research, it is claimed that the psychotypological distance between a pair of languages may be inferred from actual typological distance, but without a solid theoretical foundation from typological distance between languages. The lack of any means of measuring psychotypology may lead to errors when we attempt to explain the influence of psychotypology on transfer.

The third problem, which logically follows the previous two problems, is the lack of quantification of psychotypology in general. As we have seen before, psychotypology is described qualitatively in current studies, usually in the form ‘the psychotypological distance between language A and C is generally larger than that between language B and C’. The lack of quantification of psychotypology greatly restricts the accuracy of any investigation into the relationship between psychotypology and cross-linguistic influence in the individual learner: without detailed quantification, it cannot be known if a learner will perceive the distances in different pairs of languages in different ways, or at which point the influence of psychotypology on transfer will be restricted or reduced. Only in some recent studies (e.g. Neuser 2016; Suhonen 2016) have the psychotypological distances between specific language pairs been

measured in a quantitative way (see Section 4.4.4 for a summary of current methodologies). In order to acquire a full understanding of the possible impact of psychotypology on second language acquisition in general, psychotypology should be examined and measured at the moment an experiment on cross-linguistic influence in second language acquisition is conducted.

### 2.3.2 Markedness

The term *markedness* is often only loosely defined in research on cross-linguistic comparison, not only in second language acquisition but also in other relevant areas. The problem continues in Jordens and Kellerman (1981), as well as in subsequent research by others. The ‘looseness’ of definition comes from two aspects: the scope of markedness, and its nature. When we speak of the *scope* of markedness we mean whether markedness is defined intralingually or interlingually. The terminology is mostly used intralingually, referring to the property of an element that is ‘marked’ within one language; at the same time, it can also be used interlingually to refer to the ‘cross-linguistic rarity or frequency’ of an element (Haspelmath 2006: 26). A more problematic issue is what ‘markedness’ usually refers to in linguistic investigation. Haspelmath (2006) concludes that there are twelve different definitions of ‘markedness’, including complexity, difficulty and abnormality (i.e. rarity) in all aspects of a linguistic system; all twelve types of ‘markedness’ are defined linguistically and within linguistic system(s). The definitions, however, are not fully overlapping: that an element may be morphologically complicated, such as an overt plural marker, does not indicate its rarity in daily language use; on the other hand, a rarely used expression may only be an individual word, which is not ‘marked’ structurally. The loose definition of markedness creates obstacles if we would like to quantify it and investigate its cross-linguistic influence in a quantitative manner: we must identify the ‘marked’ elements as per the definition, evaluate how marked they are, and then see how ‘markedness’ interacts with other factors in language acquisition.

If we closely examine the definition used by Jordens and Kellerman (1981; see also Kellerman 1983), we find that it is slightly different from the linguistic definition of markedness. They describe ‘markedness’ as ‘the degree of specialness to [sic] NL (=Dutch) elements within their various NL subsystems’ (Jordens and Kellerman 1981: 196). This is more like one of the *intralingual* definitions of ‘markedness’ mentioned in Haspelmath (2006), or at least a combination of several *intralingual* definitions. Unlike other linguistic definitions of ‘markedness’ summarised in Haspelmath (2006), Jordens and Kellerman’s markedness is a psycholinguistic rather than linguistic concept (see interpretation by Rutherford 1982): the

‘markedness’ of an element in a learner’s mind is decided by that learner rather than being based on the structure of the linguistic element itself. This distinction is similar to the distinction between psychotypology and linguistic typology discussed in the last subsection: the former is perceptual and subjective, while the latter is actual and objective.

The subjectivity of markedness in Jordens and Kellerman (1981) does not simplify the question though. If ‘markedness’ here is defined as the specialness of an L1 element as perceived by a learner, the next question is the same as the one that pertaining to psychotypology: where does the sense of markedness come from? They do not provide any hint in their work, so we must make our own inference, on the assumption that the native speakers of one language do not show great variation in terms of their perception of the degree of markedness of an element. In the original work, they provide three examples to justify their argument: (1) an expression involving less productive syntactic rules is more marked than another expression that expresses the same semantic content but does not involve less productive rules; (2) an idiom, particularly a semantically opaque one, is more marked than a semantically decomposable expression; and (3) a peripheral meaning of a polysemous word is more marked than a core meaning. While some examples may correspond to certain individual definitions of markedness described in Haspelmath (2006), it is difficult to cover all the marked elements they mention using one single criterion, e.g. conceptual complexity or frequency, particularly in light of the overlapping concepts that ‘markedness’ stands for, as discussed before. For instance, a less productive syntactic rule may appear less frequently than a more productive syntactic rule, but is perhaps no more conceptually complex; on the other hand, an idiom may be semantically more opaque than a literal expression, but possibly as frequently used as the literal expression. This may eventually correspond to what Haspelmath defines as the final case of markedness: a multidimensional correlation without further specification. While it is difficult to strictly define the concept of markedness by Jordens and Kellerman, in this thesis, the methodology used to measure the degree of markedness of a linguistic element is similar to the one used by Jordens and Kellerman, in order to best follow their original proposal of markedness.

After discussing what is marked in a language (in a loose way), Jordens and Kellerman further suggest that the markedness of a linguistic element may have an impact on the subjective transferability of the element to a second language. They conclude that if an element is marked in a learner’s L1, that learner may be less willing to transfer it to the L2. Here an indication of shift from intralingual markedness to interlingual markedness can be observed: if an element is marked in the L1, a learner may believe either that it is difficult to find an identical counterpart

in the L2, or that an alternative element is available in the L2. Therefore, the learner will not transfer the linguistic element from the L1. In other words, an element that is marked intralingually may also be marked interlingually from the point of view of a learner. Although this statement does not always hold true in cross-linguistic comparison, for it is always possible for an element to be marked in language A but unmarked in language B, we should bear in mind that the definition of markedness we are discussing here is psychological rather than linguistic, and it is natural for a learner who does not have sufficient materials to make precise inferences to transfer a perception.

The markedness and subjective transferability of idioms has been extensively discussed in work by Kellerman (1979, 1983; see also Jordens and Kellerman 1981; Jordens 1986), who takes idioms as good exemplars of marked elements in a language. He briefly introduces possible influential factors on the transferability of idioms, particularly semantic transparency. If an idiom is semantically opaque, it is less likely to be transferred. This may initially sound contradictory if we compare Kellerman's result with the Transfer Theory by Irujo (1986b, 1993), which indicates that a learner may selectively transfer certain idioms from the L1 to the L2 depending on the availability of idioms, but the two results are internally connected. The selective transfer of L1 idioms can still be interpreted by the degree of markedness: it is simple and feasible to assume that not all the idioms are perceived as marked to the same degree in the L1, and less marked idioms may be perceived as being more transferable between the L1 and L2. How the degree of markedness varies between individual idioms, nevertheless, is still unknown. It may be the case, following Kellerman himself, that semantic transparency is decisive, or following Irujo (1986b) that the similarity of forms in two languages may be important, or, if applicable, the frequency of an idiom in L1 use may have a certain impact.

Jordens and Kellerman also try to implement the concept of markedness in the acquisition of different meanings of the polysemous word 'break'. That research is similar to the topic of the current thesis, namely the acquisition of the metaphorical meaning of words. Before they investigate the subjective transferability of each meaning of 'break', they first measure the degree of markedness of each meaning, which at least provides a way to quantify and evaluate the degree of markedness. The result of a card-sorting experiment (in the manner of Miller 1969) reveals that the properties of the meanings of 'break' can be reduced to two dimensions: *coreness* and *concreteness*, where coreness is the degree of prototypicality of the meaning, and concreteness refers to the literalness or imaginability of the meaning. They state that between the two dimensions, coreness, or prototypicality, should be used to indicate the (un)markedness of a meaning. More importantly from the perspective of the study of metaphor, they point out

that the coreness and concreteness of a meaning do not *always* correlate: an abstract meaning, such as ‘break one’s heart’, can be core, while a concrete meaning, such as ‘someone’ voice breaks’, can be peripheral. The follow-up examination of subjective transferability also shows that it is the core meanings and not the concrete meanings that are more likely to be transferred from Dutch to English by Dutch learners of English. They then conclude that the core meanings of a polysemous word are less marked and more likely to be transferred, while peripheral meanings are more marked and learners tend not to transfer them. They also argue that a learner knows intuitively that some meanings of a lexical item seem to be more ‘universal’ than other meanings, and will rely on such intuition in forming a personal perception of the transferability of a meaning.

At first glance, the result above may easily be interpreted as implying that abstractness/metaphoricalness does not play a significant role in influencing the subjective transferability of the meaning of a polysemous word, but a cautious examination is always required to confirm that point. The most prominent problem with the statement, which is essential to the current study, is the degree of correlation between metaphoricalness and coreness, if we follow Jordens and Kellerman in taking coreness as the indicator of markedness. Jordens and Kellerman divide the meanings of ‘break’ into two ranks to show that there are clear differences between the rank of coreness and that of concreteness, but it seems that coreness and concreteness are indeed weakly correlated in the survey, since both the concrete-but-peripheral meanings and the abstract-but-core meanings are relatively rare (see Figure 1, Jordens and Kellerman 1981: 210). Without further details, it is difficult to assert that coreness and concreteness are two independent factors and concreteness (or, in the current study, metaphoricalness) does not influence the subjective transferability of any meaning at all. In the current study, the methods used by Jordens and Kellerman will be adopted selectively to investigate (1) if there is any correlation between markedness (indicated by the degree of coreness, see Section 4.4.1 for methodology) and literalness and (2) if markedness can be used to predict the subjective transferability of metaphorical expressions and the outcome of the acquisition of metaphorical expressions.

In order to relate this discussion to other parts of their arguments, Jordens and Kellerman finally discuss the interaction between psychotypology and markedness in their work. They clearly demonstrate that both psychotypology and degree of markedness should be treated as continuums, and these two continuums can jointly predict the subjective transferability of a particular linguistic element between a specific pair of L1 and L2 (and for a particular learner if the individual difference of psychotypology is taken into consideration). The interaction

provides great insight into the detection of transfer among learners, while at the same time, the nature of a continuum shows that both psychotypology and degree of markedness should be quantified for further discussion, which has only been half accomplished in previous investigations.

### 2.3.3 Learners' knowledge

The final influential factor in transfer strategy proposed by Jordens and Kellerman is the so-called 'knowledge of the L2 learner'; although in a later revision Kellerman abandoned it as a major influential factor (Spolsky 1985), we feel that this factor is nevertheless an important one in second language acquisition. The logic of this factor is the most straightforward among the three major factors: if learners possess, or believe that they possess, knowledge of an element in the L2, they no longer need to adopt strategies to deal with it. The 'knowledge' defined here includes two parts: the actual knowledge that has already been acquired, and the 'assumed knowledge' that is believed by the learner to have been acquired. Acquired knowledge can be achieved from either intentional or incidental learning, while assumed knowledge can come from a learner's guess, deduction, inference or 'intuitional guess'. Although the role of assumed knowledge is not confirmed by a learner in the process of acquisition, it is likely that the assumed knowledge is derived from some acquired knowledge of the learner, for a piece of knowledge, even 'assumed', cannot come out of nowhere. The advantage of defining 'knowledge' this way, as already discussed in Jordens and Kellerman (1981), is that a learner cannot fully differentiate the two parts of the acquisition process. While current methodologies of assessment can largely differentiate implicit and explicit knowledge (see Ellis 2005), it is still difficult to differentiate actual and assumed knowledge, even if we assess a learner's metalinguistic knowledge (i.e. ask for an explanation for particular judgments); therefore, a good way is to combine the two parts of knowledge, as proposed by Jordens and Kellerman.

To continue with Jordens and Kellerman's proposal, it can be predicted that the effect of cross-linguistic influence from L1 will be generally weakened throughout the process of second language acquisition. This remains in line with the general description of cross-linguistic influence in Section 2.1: if a learner has fully mastered a linguistic element, cross-linguistic influence from the L1 will eventually disappear. However, considering the different types of cross-linguistic influence (or transfer), the 'weakened' effect is not always universally observable. On one hand, a learner is less likely to make any mistakes that originate from the grammar of the L1, and it would be methodologically easy to observe the disappearance of

negative transfer; on the other hand, it is not known whether positive transfer still plays a role when that learner reaches a higher level of proficiency, because the effect of positive transfer and the reflection of knowledge are similar in linguistic form if we *only* examine the learner's linguistic output. It can be seen from the analysis above that evidence for 'knowledge' (including both actual and assumed) and positive transfer is not clearly separated: when a learner correctly produces a linguistic element shared between L1 and L2, it may be evidence that knowledge of the linguistic element has been acquired, or simply that what is known from the L1 has been mapped to the L2, unless justification is required for the production. The boundary between 'the utilisation of transfer strategy' and 'the adoption of (assumed) knowledge', therefore, cannot be effectively reflected solely by the output of a learner. In fact, Jordens and Kellerman realise this problem, and they try to blur the boundary by pointing out that the transfer strategy always departs from some type of knowledge. It seems to be difficult for them, as well, to establish a definite boundary between knowledge and transfer in their experiments without referring to the metalinguistic descriptions provided by the learners. At a later stage, Kellerman eliminates this factor from his discussion, which might ultimately indicate that the borderline between '(assumed) knowledge' and 'transfer' is not definable for him.

However, from the perspective of a learner, transfer only happens when he or she does not have the relevant knowledge, or at least assumes that this is the case, and the difference between '(assumed) knowledge' and 'transfer' can eventually be reduced to the difference between the 'existence' and 'non-existence' of that knowledge; this means that even a naïve learner might make a distinction between the existence and non-existence of knowledge. Therefore, if we include a metalinguistic description of a learner, especially in terms of confidence level in producing a linguistic element, in an analysis of output, it may be possible to establish what knowledge is already acquired or assumed to have been acquired by a learner, and what comes from transfer. If a learner is confident in judging or producing a linguistic element, and is able to justify the output that is produced, then that judgment or production is more likely to come from (assumed) knowledge; on the other hand, if that learner is not confident in judging or producing a linguistic element, and cannot refer to metalinguistic knowledge to explain the output, this may suggest a reliance on cross-linguistic influence when producing the output.

As well as the problem of the blurred borderline between positive transfer and (assumed) knowledge, another problem can be identified if we consider the methodology used to examine a learner's knowledge. In existing studies, a learner's knowledge is often reflected by L2 proficiency, and it is generally assumed that the higher a learner's proficiency is, the more

knowledge is possessed and assumed. While that measurement is feasible, proficiency itself is occasionally measured by the length of exposure to the L2 in a classroom setting (see Experiment 1 of Jordens and Kellerman 1981), or, in some cases, the major of a group of college student participants (e.g. Zhang 2008). It seems that when developing the original proposal, Kellerman aimed to establish a correlation between three factors: a learner's knowledge (both acquired and assumed), proficiency level and the length of L2 instruction received. In a strict sense, such correlation cannot provide a full picture of a learner's knowledge, although it can be a fall-back if a more accurate test is not applicable. The use of the length of L2 instruction to reflect a learner's knowledge is rather problematic. Longer exposure to L2 classroom instruction can only reflect that a learner has been taught more, whereas a learner's knowledge from non-guided acquisition, as well as assumed knowledge, is left unevaluated. If 'knowledge' here is defined as a combination of acquired and assumed knowledge, a measurement of acquired knowledge exclusively does not provide a strong foundation for the entire argument. To take a further step and to associate it with the topic of the current thesis, different aspects of L2 knowledge may develop in an imbalanced way as proficiency rises, which may not be accounted for in the correlation. It is probable that some L2 knowledge is less accessible, or even less learnable, in the process of acquisition, and it may remain underdeveloped even though a learner has received years of L2 instruction and achieved relatively high proficiency. Therefore, it cannot be simply hypothesised that, with a rise in general L2 proficiency, a learner will gradually give up strategies of transfer in all aspects of the L2. While in the current thesis the knowledge of the learner is still considered as one of the major factors that influence strategies of transfer, whether the rise of proficiency will lead to the development of knowledge of metaphorical expressions will also be one of the focuses of the current thesis. At the same time, a possible methodological solution will be proposed to differentiate positive transfer and (assumed) knowledge in a learner's judgment.

## 2.4 Summary

This chapter has served as the first part of the literature review, mainly focusing on the phenomenon of cross-linguistic influence in the acquisition of lexical items. While there is a lack of comprehensive discussion of the acquisition of metaphorical expressions in the L2 in the current literature, we can refer to previous studies on the acquisition of ordinary lexical items and idioms, since metaphorical expressions seem to reside between the two on the continuum of semantic transparency. The general predictions of cross-linguistic influence on metaphorical expressions seem to be in line with other studies on lemmatic transfer in general.

Overall, we may expect to see both positive and negative lemmatic transfer in the acquisition of metaphorical expressions, depending on the cross-linguistic availability of the metaphorical expressions in learners' L1 and L2. While learners can acquire the shared metaphorical expressions better, they may misuse the metaphorical expressions only available in their L1 in a fashion similar to false friends.

The factors influencing cross-linguistic influence are also discussed in this chapter. The proposal of Jordens and Kellerman, together with follow-up studies, shows that cross-linguistic influence is subject to factors other than cross-linguistic similarity (or difference) in expressions, and these factors, including psychotypology, markedness and a learner's knowledge, should also be taken into account if we intend to provide a full exploration of possible cross-linguistic influence on the acquisition of metaphorical expressions. Apart from the exact influence of the three factors, which is crucial for the current thesis, this part of the review also shows that there is a lack of proper methodology to estimate these influential factors. Possible methodological solutions, including the estimation of learners' psychotypology and the measurement of markedness of a linguistic element, will be presented in Chapter 4.

### 3 Bilingual lexicon and figurative language processing in a second language

#### 3.1 Introduction

It has been widely recognised that it is in the nature of the acquisition of a lexical item in a second language that a learner will associate the word form in the L2 with the concept that the word represents, and in some cases, with the word form in the L1 as well. A learner who comes across a word that has not previously been acquired needs to activate the link between the word form (either phonological or orthographical) and the relevant concept. The organisation of lexical items in two languages and the concepts in a learner's mind is generally called the bilingual lexicon, and is a crucial topic in research on second language acquisition. A number of models of bilingual lexicon have been developed focusing on the links between the concepts and the vocabulary of a learner's two languages, but most restrict their scope of discussion to the acquisition of the literal meaning of each lexical item.

I would suggest that the acquisition and comprehension of metaphorical expressions is in line with the acquisition and comprehension of a lexical item in general. This is because, as discussed in Section 1.2.2, the conventionalised metaphorical use of a lexical item is already regarded as a particular type of word meaning in the current discussion of lexical semantics, which is essentially the same as the literal meaning(s) of a lexical item. Given that conventionalised metaphorical meanings, as investigated in the current thesis, have an equal status to the literal meanings of lexical items in theoretical discussions, such equality can and should be reflected in a model of word acquisition and retrieval as well. In order to retrieve and acquire a conventional metaphorical meaning of an L2 lexical item, a learner needs to pass through the L2 word form in order to establish the link between the metaphorical concept and the L2 expression.

Considering the status of a metaphorical meaning of a word in current theoretical semantic frameworks, I argue that the possible route of such meaning retrieval, as well as predictions on the time a reader will spend on such meaning retrieval, can and should be elaborated within a general lexical acquisition and retrieval model. In the case of an L2 learner acquiring and retrieving a metaphorical meaning of an L2 lexical item, it will be necessary to add the metaphorical meaning of that lexical item to the literal meaning(s) and integrate them into the current bilingual lexicon. The entire process of metaphorical meaning retrieval, therefore, relies on the way in which the meaning is stored in the learner's mental lexicon, in which the influence of the first language may also play a role. The learner may access the metaphorical meaning of

an L2 word via its translation equivalent in the L1, or via the literal meaning of that L2 word, or in other ways. We can thus develop a series of assumptions regarding the online comprehension of metaphorical expressions based on current works on the bilingual mental lexicon.

In this chapter, I will firstly review various models of bilingual lexicon. While all of them focus on the links between the L1, the L2 lexicon and concepts, and some of them on the differences between access to abstract and concrete words, few works have investigated how a metaphorical meaning can be stored in the mental lexicon and accessed in the reading process. I would suggest that, although the current frameworks are not satisfactory for the discussion of metaphorical expressions, a combination of the theories can provide some insights and lead to a more precise prediction regarding the comprehension of metaphorical expressions by looking at the possible routes from the word form to the correct meaning. I will also show how L2 classroom instruction may affect learners' bilingual lexicon, together with the implications it has on the acquisition of metaphorical expressions. The second half of this chapter includes a review of the processing of figurative language in a second language, including the processing of idioms and phrasal verbs; they can be used as possible references to the processing of another type of figurative language, namely metaphorical expressions. Finally, I will elaborate on how a joint analysis based on several models of bilingual lexicon can provide some insights into the storage and retrieval of metaphorical meanings of lexical items. An understanding of this storage and retrieval should also allow clearer predictions to be made about possible cross-linguistic influence.

## 3.2 Current proposals for the bilingual lexicon

### 3.2.1 Distributed (Conceptual) Feature Model/The Sense Model

The Distributed Feature Model (henceforth DFM) was first proposed by De Groot (1992; see also Van Hell and De Groot 1998) and further developed by De Groot and Kroll (1997) into the Distributed Conceptual Feature Model (henceforth DCFM). The essence of De Groot's proposal is that the concepts that match a lexical item are never a unified, indivisible entity, but a series of concepts represented as interconnected nodes. In earlier frameworks of semantic memory (c.f. Hierarchical Network Model of Semantic Memory by Quillan 1966; Collins and Loftus 1975), it is proposed that a word can be associated with several relevant concepts, and the meaning of that word, therefore, consists of the relations between that word and other relevant words. 'Concepts' in those frameworks are illustrated in a Late-Wittgenstein style (see Wittgenstein 1953): there is no clear boundary between two words or two concepts, although

some ‘core, stable’ features that are relatively independent of contexts can be shared across speakers and languages. In some situations, a single concept may vary depending on the contexts in which it appears, and the meaning of a word may largely depend on the context.

Taking this as an assumption, De Groot further elaborates on the possibilities of word activation in a bilingual network. She suggests that, given that a word can be linked to several conceptual representations (hence the framework is called ‘distributed conceptual’), a pair of translation equivalents may share the full set of conceptual representations, or only part of the set, as illustrated in Figure 3.1 below. It is also possible for a pair of translation equivalents to refer to slightly different concepts in the two languages, because some of the meaning representations are not shared between them. De Groot also states that it is very likely that some concepts may be represented by a single lexical item in a learner’s L1 but not in the L2; that does not indicate that the learners cannot express the equivalent of such a lexical item in general, but simply that such concepts are not representable in a single lexical item in the L2. The different proficiency level of learners’ languages, as well as learners’ different understandings of the conceptual representation of lexical items, can often lead to an asymmetry in reaction time and accuracy for a pair of translation equivalents, and moreover, an imbalance in the construction of the bilingual lexicon by a L2 learner.

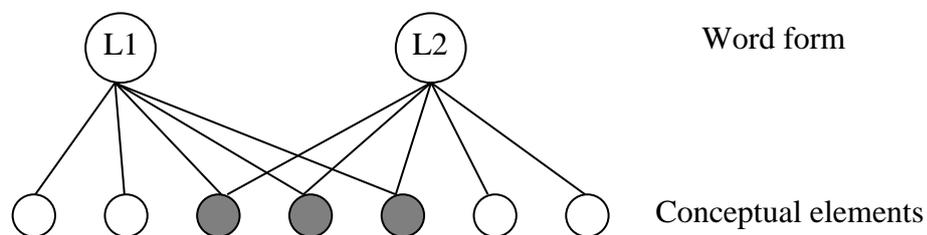


Figure 3.1 The Distributed Conceptual Feature Model (adapted from De Groot 1992)

As shown in Figure 3.1, a pair of translation equivalents in two different languages, either cognates or non-cognates, can each connect to a series of conceptual elements. Among the conceptual elements shown at the bottom, some may usually be experientially based, relatively stable across different contexts, and therefore recognised as ‘core’ elements. Other elements may be more context-dependent and can be constructed in an *ad hoc* manner. The conceptual elements are linked to either of the translation equivalents. A few of them in grey are shared between the two words, which means that those concepts are more language-neutral. The white elements, on the other hand, are generally more language specific. In the situation indicated by

Figure 3.1, the pair of translation equivalents has a partial meaning overlap; it is possible in other cases for two translation equivalents to have a full overlap of conceptual features, or only a slight overlap.

De Groot proposes that the degree of overlap of conceptual features across a learner's L1 and L2 is a decisive factor in the learner's speed and quality of performance in translation tasks. If a pair of translation equivalents share more conceptual features, the learner is more likely to perform better in the translation task by reacting faster to the pair and providing more accurate answers. De Groot does not discuss the link between the word forms of the L1 and L2; instead, she vaguely hints that the connection between L1 and L2 words is via the shared conceptual features: if the L1 and L2 words share more conceptual features, the connection between the two words is stronger. From our perspective, it should be expected that a pair of translation equivalents will be more transferable if they share more conceptual features, because there is a stronger, more stable link between the L1 and L2 words. A learner is more likely to receive cross-linguistic influence from the L1 when accessing the L2 words with a stronger conceptual overlap, because the L1 words are more likely to be co-activated.

In the DCFM, the degree of overlap of conceptual representations between a pair of translation equivalents can be attributed to three major factors, namely the cognate status, the part of speech, and the concreteness of a word; De Groot suggests that the influence of these three factors is intuitive and observable. Cognates are considered to share more conceptual features than non-cognates due to the historical relationship between cognates, and the shared morpheme inherited from a common parent language by a pair of cognates can indicate that the concept associated with the morpheme is also shared. Concrete words are considered to share more conceptual features than abstract words, because two concrete words are more likely to refer to one single concrete entity in the real world without any significant difference, but the concepts referred to by the abstract words would be less obvious in the real world, and it would be more difficult to evaluate the similarity of concepts cross-linguistically. The effect of part of speech comes from a further investigation in Van Hell and De Groot (1998), in which the participants were asked to perform a translation task on both nouns and verbs. The results reveal that nouns are generally translated faster than verbs, which is interpreted by Van Hell and De Groot as showing a difference in degree of overlap of conceptual features between the two parts of speech.

In a more recent modification of the DCFM, namely the Sense Model by Finkbeiner and colleagues (2004; see also Xia and Andrews 2015), the same framework of distributed conceptual features has been used to illustrate the storage of polysemous words in the bilingual

lexicon. Finkbeiner and colleagues point out that the DCFM does not reasonably explain the asymmetry between L1–L2 and L2–L1 priming in a translation task, and they attribute this difference to the asymmetry of polysemy status in a learner’s L1 and L2. It can easily be recognised that many lexical items are polysemous; in some cases this may result from the conventionalised use of metaphorical expressions, as is discussed in the current thesis, while other cases may be due to historical reasons. Finkbeiner and his colleagues assume that only a small number of meanings may be determinate for a pair of translation equivalents to be ‘equivalent’, while the rest of the meanings are not ‘equivalent’ and are sometimes even without any counterpart in the other language. Therefore, as shown in Figure 3.2 below, learners may well be aware of every possible meaning of a polysemous word A in their L1, but only know a couple of the most prominent meanings of the translation equivalent B in their L2. Among the meanings of the translation equivalent B, it is more likely that the basic meaning(s) that makes A and B a pair of translation equivalents will be known: for example, among all the meanings of the word ‘attack’, a Chinese learner of English is most likely to know it as ‘to try to hurt or defeat using violence’, because that meaning matches the meaning of *gongji*, the translation equivalent of ‘attack’ in Chinese. At the same time, the learner may not be aware of other meanings of the L2 word B; those meanings may include not only meanings that are exclusively available in the L2, but also some peripheral meanings shared between the L1 and the L2. The above assumption of the Sense Model also explains why the priming effect from the L1 to the L2 is significantly stronger than the priming effect from the L2 word to the L1 between a pair of translation equivalents: as indicated in Figure 3.2, the meaning(s) of the L2 word that a learner knows is mostly the meaning(s) shared between the L1 and L2 words, and the presence of the L1 word can always activate all the meanings of the L2 word known by the learner; however, the presence of the L2 word can only activate part of the L1 meanings (in the graph 20%), hence a weaker priming effect. Finkbeiner and colleagues (2004) observe that, while the priming effect from L2 words to L1 semantic categories is robust, there is a lack of priming effect from L2 words to L1 individual words in a translation task. The Sense Model suggests that this difference is due to the lack of knowledge of polysemy in L2: learners can always activate the full L1 meaning of a semantic category if they attempt to find the semantic category of an L2 word, and thus show a robust L2–L1 priming effect; however, they can only partly activate L1 meanings when translating individual words, and thus the priming effect is absent in a translation task.

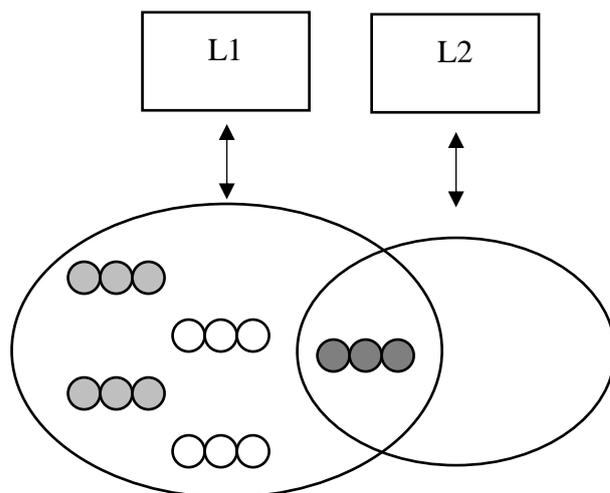


Figure 3.2 The Sense Model (adapted from Finkbeiner and others 2004)

To summarise, the concepts, features, relationships between individual words and conceptual features, and the cross-linguistic differences in conceptual representations illustrated in the DCFM (as well as the subsequent Sense Model) perfectly fit the requirements of an investigation into the acquisition of metaphorical meanings, especially if we would like to explore the difference between metaphorical meanings and literal meanings of the same lexical item in a learner's bilingual lexicon. De Groot adopts a Late-Wittgenstein view in her model, presuming that a word is not linked to a single, unified and static concept all the time, but can shift between several concepts that are largely similar but with fine-grained, minor differences. The Sense Model also directly points out that the framework is applicable to polysemous lexical items, either monolingual or bilingual. Those views further coincide with a few works on the emergence of conventionalised metaphorical meanings from different aspects. On the one hand, a cognitive semantic approach to metaphor (c.f. Sweetser 1990; see Lakoff and Johnson 1980 for a more extensive and theoretical argument) would stress that there is a natural link between the literal and metaphorical concepts represented by a single word. The logic here is simple: people will use a single word to refer to two distinct concepts only if they perceive that the two concepts share much in common. On the other hand, radical contextualist approaches, such as those of Leezenberg (2001), Recanati (2004) and Stern (2000), follow the view of Wittgenstein and suggest that the metaphorical interpretation of a lexical item is triggered by the context in which it appears. Although the two branches offer distinct proposals for the actual functions of a metaphor, they do not contradict or exclude each other in terms of the relationship between the metaphorical meaning and the literal meaning of a word. They also jointly support the possibility that the metaphorical meanings and the literal meanings can be

represented by several conceptual features in semantic memory, which shows congruence with the DCFM (or the Sense Model). Depending on the context, either the conceptual features related to the metaphorical meanings are activated, or those related to the literal meanings.

The influence of concreteness on the distribution of conceptual features, although not the most important statement in the whole framework, leaves sufficient space for an explanation of the acquisition and real-time processing of metaphorical expressions. Due to the nature of their experiments, neither De Groot nor Finkbeiner and colleagues discuss in detail the fact that concreteness may vary according to the meanings (or ‘senses’ in Finkbeiner and others 2004) of a single word. In their research, words are simply divided into two categories, either concrete or abstract. This is because in a word translation task, lexical items appear individually and there is no context to indicate the concrete or abstract meaning(s) of that word. In real language use, however, a polysemous word can appear with either a more concrete or a more abstract meaning, depending on the context in which it is used. This will be particularly prominent in experiments in which we contrast the metaphorical use and the literal use of a single word. By definition, the metaphorical use of a word is the linguistic outcome of mapping a concrete conceptual domain to an abstract one (Lakoff and Johnson 1980; Lakoff and Turner 1989); or, to take an Aristotelian view (Arist. *Rhet* III.21, 1475b1-30, trans. Butcher 1961) the metaphorically used word is the ‘transference’ of a concrete word to an abstract context. Both definitions indicate that, when a word involves conventional uses in both a literal and a metaphorical way, it can be linked to both concrete and abstract conceptual representations. The concrete concepts reside in their own concrete domain, associated with the literal meanings of that word, and the abstract concepts associated with the metaphorical meanings are involved in cross-domain mapping. Therefore, it may also be expected that the processing difference between concrete and abstract words indicated in the DCFM will be replicated in the concrete and abstract meanings of a single word.

Although it is intuitively feasible and has been empirically verified that a concrete pair will share more conceptual features than an abstract pair, one could argue that the effect of cognate status and the effect of part of speech are not the only issues that influence the degree of overlap in conceptual features. Cognates usually share phonological and orthographical features, which makes them stand out from translation equivalents with different spellings and/or pronunciation. A learner may show a faster and more accurate reaction when translating a cognate pair, simply because the two words look and sound similar, while conceptual features may not be accessed instantaneously in the process, and therefore may not play the role suggested by De Groot. The effect of part of speech on De Groot’s translation task is not clear either. Although Van Hell

and De Groot (1998: 195) suggest verb pairs take people more time to translate due to their ‘less dense conceptual representation’, it is also possible that the processing mechanism of verbs differs from that of nouns. It is well recognised that more time is taken to read and react to a verb than to a noun in word matching experiments, and there is a lack of evidence for different conceptual representations (Gomes and others 1997). Therefore, we cannot ignore the possibility that more time may be needed to translate a verb pair than to translate a noun pair even if there is no difference in the density of conceptual representation. The difference in processing time could have its roots in the processing of syntactic information rather than in conceptual or semantic information. In short, we cannot fully differentiate the influence of conceptual representation and other influences of word classes on semantic processing in a general psycholinguistic task. The two problems will be avoided in the current study: first, Chinese and English, being two typologically unrelated languages, do not share many true cognates, and only non-cognate translation equivalents are used in the study; second, the possible effect of part of speech is minimised in the online self-paced reading task, because all the target words in that experiment are verbs. In the offline acceptability judgment task, the factor of part of speech will be included as a possible fixed factor.

De Groot, as well as Finkbeiner and colleagues, further outline how the concreteness of words can affect the conceptual representations between a pair of translation equivalents. Although neither of them explicitly links such differences to a possible trace of cross-linguistic influence, a series of feasible connections can be made regarding the concreteness of a pair of translation equivalents, the degree of overlap in conceptual elements of that pair, and the transferability of that pair. As has been discussed above, concrete pairs share more conceptual elements than abstract pairs; if a pair of translation equivalents shares more conceptual elements, then learners are more likely to use the L2 translation equivalent in the way they use the L1 word, i.e. to transfer the use of the L1 word to the L2. Thus, it may be expected that a pair of concrete translation equivalents will be more transferable than a pair of abstract translation equivalents. If a further step is taken and the assumption is extended to a pair of concrete and abstract meanings of a single word, it may then be expected that the concrete, literal meaning shared by a pair of translation equivalents will be more transferable than the abstract, metaphorical meaning. Due to the influence of concreteness, we can assume that even if the abstract, metaphorical meaning is also shared between the pair, it will still be less transferable than the concrete, literal meaning, which may lead to an asymmetric performance when a learner is asked to judge or process the abstract, metaphorical use. Note that this seems to contradict the proposal of Jordens and Kellerman (1981), who claim that the transferability of

meaning of a lexical item is not significantly influenced by the concreteness of that meaning. In the current thesis, this point will be further explored.

While it has been well received, problems have been identified in the DCFM in recent discussions. Although De Groot has developed a rather consistent system to describe the difference of conceptual mapping between a pair of translation equivalents, the method used in the DCFM to represent the conceptual difference, namely the feature-based approach, is a rather primitive and outdated model in cognitive psychology (Pavlenko 2009). When one perceives a concept, or connects a concept to a lexical item in a language, one does not always decompose the concept into a variety of features, or operate the conceptual features in a parametric way in order to formulate a concept. Instead, one may construct prototypes to represent a concept, and fine-tune those prototypes to context-specific concepts according to the given context. This discrepancy can actually be reflected in the area of metaphorical expressions: the literal meaning of a lexical item is usually the prototype of a series of concepts, and the metaphorical meaning of that lexical item is modified based on the literal meaning, the context in which it appears, and the conceptual domain mentioned in the context. If a framework for the mental lexicon is intended to reflect the delicate differences in matching concepts and lexical items between languages in the human mind, it should take the form of conceptual representation into consideration. In the current thesis, the model of conceptual representation favours the prototype model.

Another point that may indicate a need for further modification is that the DCFM does not directly reflect the developmental path of L2 learners and the role of cross-linguistic influence in the developmental process. The original DFM proposal has been established as a possible model of bilingual lexicon for a balanced, proficient bilingual, and from Figure 3.1 it can be seen that an L2 lexical item can and should be accessed via conceptual mediation, i.e. constituting a direct link from conceptual to lexical level. However, that is not usually the case, at least not for every lexical item; Dufour and Kroll (1995) report that conceptual mediation is common among more fluent bilinguals, while less fluent bilinguals may have limited direct access to the conceptual level from the lexical level. As we will see in the next section, less proficient learners are considered to rely on the lexical link between a pair of translation equivalents from L1 and L2 in order to access the conceptual meaning of the L2 lexical item. This means the DCFM may not represent the bilingual lexicon of a less proficient learner, and the structure of the bilingual lexicon of a less proficient learner may be critically different from the picture shown in the DCFM. The Sense Model, on the other hand, assumes that a developmental trend can be observed when a learner acquires more meanings of a polysemous

word, but it does not (aim to) explain which meaning(s) are more likely to be acquired, and, more specifically, how a learner decides where a meaning of a polysemous word should be placed in the mental lexicon. When a learner acquires a shared metaphorical meaning, will it be successfully placed in the area shared by L1 and L2? When a metaphorical meaning that is only available in the L1 is encountered, will it be transferred to the L2? When a metaphorical meaning that is only available in the L2 is encountered, will the link between that meaning and the L2 word form be successfully established? Those questions are yet to be answered in the thesis.

### 3.2.2 Revised Hierarchical Model/Modified Hierarchical Model

While De Groot notes that a lexical item does not always map to a solid, unified entity on a conceptual level, most of the proposals of the bilingual lexicon do not emphasise this feature. Instead, they pay significantly more attention to what framework can best capture the relationship between three elements: the concept (as an entire, inseparable entity), the L1 word, and the L2 word. In this way, the L1 and the L2 words are usually recognised as a pair of translation equivalents without any additional need to specify their degree of ‘equivalence’. This is due to the tradition of discussions on the bilingual lexicon from the era of Kolers (1963), who particularly focuses on the debate over the compound or coordinate organisation of bilingual lexicon. Recent discussions, including all the frameworks reviewed in this chapter, generally favour a compound organisation of a bilingual lexicon, in which the L1 word and L2 word are connected with a single concept. A sequential bilingual who acquires an L2, has already acquired an L1, and has a comprehensive structure that organises the L1 lexicon and relevant concepts. In acquiring the L2 vocabulary, there are two major possibilities for integrating the new L2 words into the mental lexicon. Either the L2 word can be linked to the L1 translation equivalent, which is called ‘word association’; or the L2 word may be directly attached to the concept, without any assistance from the L1 translation equivalent, and that is called ‘concept mediation’.

However, the results of experiments and even simple retrospective analysis both reveal that neither of the frameworks can solely represent the structure of the bilingual lexicon of a sequential bilingual. Previous studies show that more and less proficient learners may rely on concept mediation to different extents in a single lexical decision task (e.g. Dufour and Kroll 1995). This not only shows that concept mediation is not fully available to all second language learners, but also indicates that word association is not the universal method of bilingual lexicon organisation either: the presence of concept mediation among more proficient learners indicates

that they may not rely on word association to make lexical judgments. Furthermore, as has been discussed before, it is probable that a learner may acquire an L2 word for which the associated concept is actually not available in the L1, which will lead to the absence of the link between the L1 and L2 words. Therefore, a third possibility for the organisation of the bilingual lexicon is a combination of the two basic assumptions, namely word association and concept mediation, depending on learners' proficiency as well as the availability of concepts and lexical items in different languages.

The Revised Hierarchical Model (henceforth RHM) of Kroll and Stewart (1994) aims to integrate the two basic assumptions and provide a comprehensive picture of the bilingual lexicon across L2 acquisition. The RHM is modified based on an old version of the Hierarchical Model, an alternative form of the shared storage system; a brief illustration is given below in Figure 3.3. Three main components are usually considered to be involved in the RHM: a larger L1 vocabulary indicated by a larger rectangle, a smaller L2 vocabulary indicated by a smaller rectangle, and a group of individual concepts. One prominent feature of the RHM is that the direction and strength of links between any two components are explicated in the framework, which leads to a possible elaboration of the developmental trend of the bilingual lexicon and more precise predictions of experimental results. The links between the L1 and L2 vocabulary stand for word association and display clear asymmetry: the link from L2 to L1 is stronger than the link the other way around. This follows the asymmetry shown in bidirectional translation tasks in which translation from L2 to L1 is usually faster than from L1 to L2. Both the L1 and L2 lexicons are linked to a group of concepts, but the strength of the links is significantly different. While the link between the L1 lexicon and the concepts is solid and strong, the link between the L2 lexicon and concepts, which represents concept mediation, is relatively weaker. The weaker link indicates that concept mediation always exists, even for less proficient learners, but its function and scope may be more restricted compared with the link between the L1 lexicon and concepts. With a growth in proficiency of the L2, it is expected that the link from L1 to L2 word forms (backward translation) and the link between L2 word forms and concepts will gradually be strengthened.

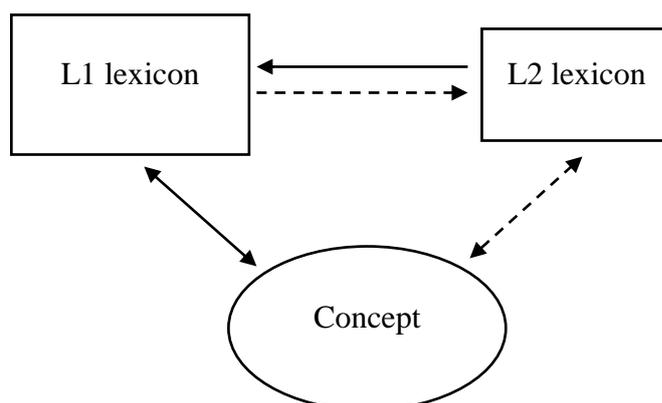


Figure 3.3 The Revised Hierarchical Model (adapted from Kroll and Stewart 1994)

The RHM provides feasible explanations for two major concerns in research on second language vocabulary acquisition that previous forms of the hierarchical model fail to cater for. The most prominent advantage of the RHM is that it illustrates the developmental order of the second language lexicon within its framework. Such an illustration can be used to explain a series of developmental changes (e.g. Dufour and Kroll 1995) observed in the actual learning process or among learners at different levels of proficiency, either in a classroom setting or in more naturalistic environments of vocabulary acquisition. The RHM recognises that word association and concept mediation co-exist in the bilingual lexicon, and assumes that the association between an L2 word and its L1 translation equivalent is inevitable, even if the L1 is not used as the language of instruction. While all three major components always exist in the bilingual lexicon, what changes with the progress of a learner is the strength of each link in the framework, rather than the presence or absence of each link. The differences in performance caused by different proficiency levels or different acquisition environments, therefore, are explained as a quantitative matter rather than a qualitative one.

The other strength of the RHM, which is more closely related to second language processing, is that it considers the retrieval of a lexical item in second language comprehension. From a general psycholinguistic point of view, the comprehension of a word can be seen as a real-time reactivation of the link between the word form and its corresponding concept. Only when the relevant concept is matched to the word form, can the word or larger constituent be fully understood by the processor. Therefore, the reaction time to a word, no matter what language it is in, can be used to indicate the length and/or strength of the link between that word and the concept it is linked to in the framework. It is generally and naturally assumed that a learner always minimises the temporal and cognitive cost of word comprehension; in the RHM, such

minimisation of cost is accomplished by taking ‘the fastest route possible’, which can be either the shortest link or the strongest link, depending on availability. When reading an L2 word, a more proficient learner can directly link the word and the corresponding concept together, but a less proficient learner may rely on a detour via the L1 word to establish that link because the link between the L2 word and the concept is not strong enough to minimise the temporal cost. Therefore, it can be hypothesised that a more proficient learner can react faster to a word that has been learnt, while a less proficient learner may take a substantially longer time to react to a word, even if taking ‘the fastest route’ possible in the bilingual lexicon (c.f. Athanasopoulos 2015).

The shortcomings and gaps in the RHM have been discussed extensively in previous literature. Criticism comes from several points of view, such as the problem of asymmetry in terms of the lexical link between the L1 and the L2 lexical items, and the over-simplified structure of the bilingual lexicon (Brysbaert and Duyck 2010). Like other models of bilingual lexicon, except the DCFM, the RHM assumes that a lexical item can only map onto one solid concept, and thus emphasises more the overall structure of the bilingual lexicon, in terms of the two sets of vocabulary and a set of concepts. Without further specification, one can only assume that in the RHM each pair of translation equivalents can, and can only, connect to an individual concept, and two translation equivalents (or near-equivalents when full equivalents are not available) will always connect to the exact same concept. Such assumptions of word-concept mapping, according to Brysbaert and Duyck (2010), largely ignore the case of homonyms, homophones and polysemy widely spread across languages, because those lexical items involve one-to-more rather than one-to-one mappings between word form and concept.

The problem of representation of metaphorical expressions in the bilingual lexicon is in line with the over-simplification discussed by Brysbaert and Duyck. Basically, the RHM cannot effectively reflect the acquisition of several meanings of a polysemous word, and one needs to duplicate the concept or create additional concepts to ensure a learner has stored all the meanings of that polysemous word. For a metaphorical meaning of a word, the problem may be even more complicated, if we consider the relationship between the literal meaning and the metaphorical meaning of that word. To clarify the current situation, we can refer to a pair of translation equivalents: *attack* in English and *gongji* in Chinese, both of which can be used literally to describe an aggressive physical act and metaphorically to describe an aggressive verbal argument against a theory. Both the literal meaning and the metaphorical meaning, thus, are shared between the two languages. Obviously, the metaphorical meaning is related to the

literal meaning, and both meanings should be connected to the word form of the lexical item in each language.

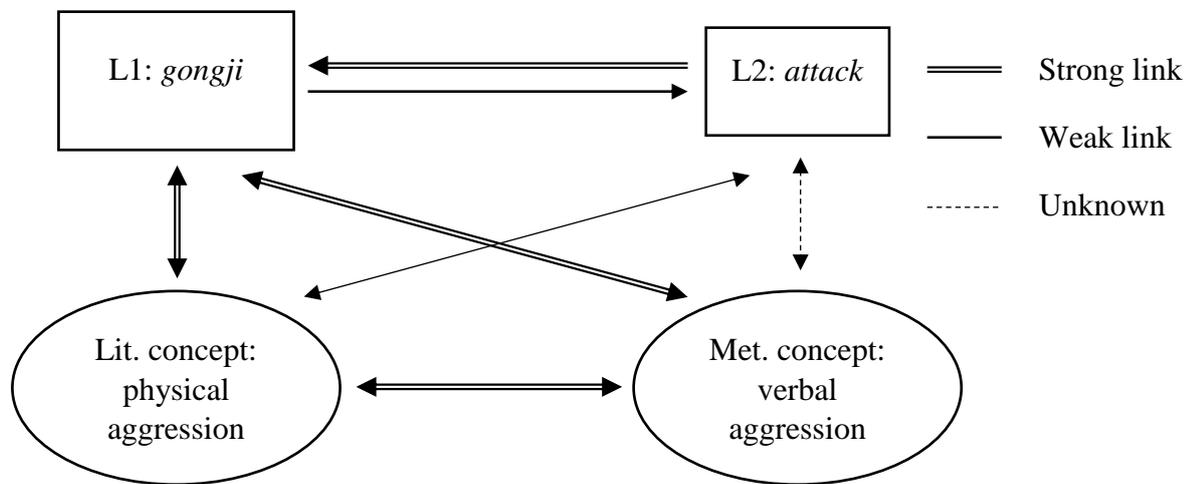


Figure 3.4 The structure of *gongji*, *attack* and relevant concepts in an RHM fashion presentation

Figure 3.4 above shows the structure of two-word forms and two concepts for a Chinese learner of English. The line weights indicate the length of each link. The double lines indicate established strong links, and the single lines indicate established weak links. The dashed lines indicate that the links are ‘unknown, to be confirmed’: this means we do not know if the learner has established those links, or how strong they are. The link between the two concepts means that they are associated with each other, while it is not clear whether a learner is aware of that. For the majority of words, the acquisition of the metaphorical meaning in the L2 usually happens no earlier than the acquisition of the literal meaning of the word in the L2<sup>2</sup>. Therefore, by the time learners experience the meaning, they will know that both the literal and metaphorical concepts are linked to the L1 word form *gongji*, and will already have established the links between the L2 word forms and the literal concept. They may also have subconscious knowledge that the two types of *aggression* on a conceptual level are linked. The RHM, however, cannot make any prediction regarding the situation above, because it does not give any explanation of how an existing concept and an existing L2 word form might be connected at the time of acquisition. In the view of the current version of the RHM, the metaphorical concept should be treated as a new concept, and the lexical items in the two languages should

<sup>2</sup> In the current thesis, all the words investigated in the experiments are known to the learners in terms of literal meanings, since the literal meanings already appear in the learners’ textbook. The status of the metaphorical meaning is unknown.

be treated as novel words that the learner has no knowledge of; this means that a polysemous word is seen as a homonymous word with two irrelevant meanings. In the RHM, the learner cannot make use of previous L2 word knowledge in the acquisition of a related word meaning of a known word. This view is problematic for metaphorical expressions with a polysemous nature; as is discussed in Section 1.2.2, a language user may be able to infer the metaphorical meaning from the literal meaning if given sufficient contextual clues and instructions. We can therefore assume that it is at least possible for a learner to make use of the link between the L2 word form and the literal concept to establish the link between the L2 word form and the metaphorical concept. The mechanism of the RHM blocks the route without verifying its probability, while one target of the current research is to examine the availability of the route in different conditions.

Furthermore, in a similar way to the problem mentioned above, we can see that the RHM can predict neither how a learner will suppress a metaphorical meaning if it is only available in the L1, nor how that learner will acquire a metaphorical meaning that is only available in the L2. That problem leads to another shortcoming of the RHM: the RHM generally assumes that all concepts are shared cross-linguistically, and thus fails to recognise the existence of language-specific concepts. As explained in Section 2.2.1, Pavlenko (2009) has extensively discussed the existence of language-specific concepts and the transfer of those concepts among bilingual speakers. She then proposes the Modified Hierarchical Model (henceforth MHM), to illustrate how conceptual differences between two languages can be represented in the bilingual lexicon. The most significant difference between the MHM and the RHM is that the former demonstrates how conceptual categories can be linked to different word forms; this clearly illustrates how conceptual transfer happens in second language acquisition. The strong point of the MHM is that it proposes ‘conceptual restructuring’ as a special type of transfer from L1 to L2, in which a learner can transfer some L1-specific conceptual categories when establishing the link between the L2 word and the shared categories. This can explain why an English learner of Russian may transfer the conceptual category represented by *cup* to *chashka*, as mentioned in Section 2.2.1. Conversely, it may also be possible to link some L2-specific categories to the L1 word form at some stages of L2 acquisition, which will lead to conceptual transfer from L2 to L1. The acquisition of L2 lexical items, can thus be seen as a type of conceptual restructuring, and with progress in L2 acquisition the learner is gradually able to categorise concepts in an L2-specific fashion. An illustration of the MHM is shown below in Figure 3.5.

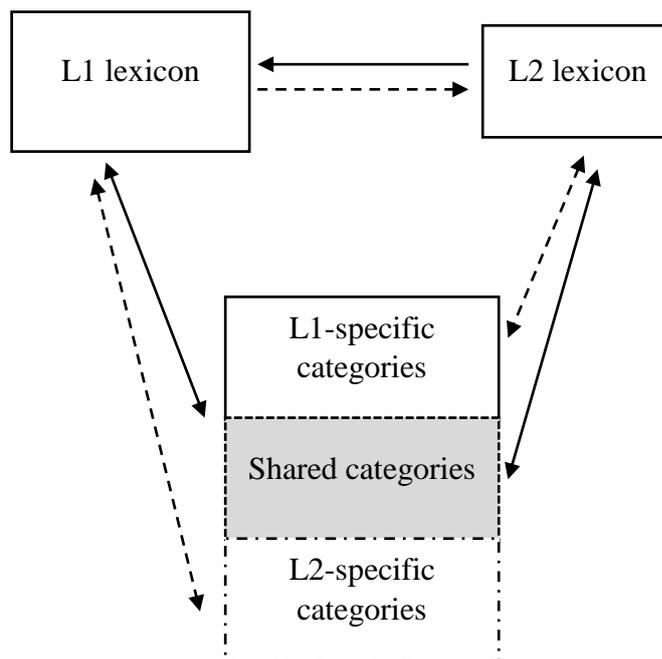


Figure 3.5 The Modified Hierarchical Model (adapted from Pavlenko 2009)

Considering the main emphases and the structure of the MHM, it can be seen as an integrated and updated version of both the DCFM and the RHM. It can reflect both the conceptual difference between a pair of (near) translation equivalents, and the developmental trend of the bilingual lexicon. However, as discussed in Section 2.2.1, transfer of metaphorical meaning is a form of lemmatic transfer rather than concept transfer, and thus it cannot be accommodated in the MHM. While Pavlenko (2009) suggests that the literal and metaphorical concepts of a single lexical item should be treated as if they were two unrelated, independent concepts, previous discussion in Section 1.2.2 has shown that these two meanings are actually relevant to each other from the perspectives of both lexical and cognitive semantics. An additional proposal is still needed to elaborate the acquisition of metaphorical expressions and the relationship between the literal and metaphorical meanings (or concepts) of a lexical item.

To summarise, both the RHM and the MHM clearly illustrate how learners with different proficiency can process words in different ways, but the acquisition of metaphorical expressions and other polysemous lexical items is left untouched. The neglect of one-to-many word-concept mapping makes it impossible to explain using the RHM, while the MHM realises the problem but avoids it since it is not the major topic or the main motive of Pavlenko's proposal. While other aspects of the bilingual lexicon have been discussed in detail and these two frameworks could serve as a basis for a developmental view of a bilingual lexicon framework, the interaction

between the bilingual lexicon and semantic transfer is not fully explored, and this is particularly related to the acquisition of metaphorical expressions in a second language. More discussion is needed to explain how a learner can connect a pair of translation equivalents to a number of different but related concepts, and how cross-linguistic influence happens in the process of acquisition.

### 3.2.3 Bilingual lexicon in a classroom setting

While previous frameworks for the bilingual lexicon reviewed in the chapter are generally applicable to language acquisition in different environments, it is worth noting that this thesis largely focuses on the acquisition of metaphorical expressions in a classroom learning setting, specifically the English classroom in China, where English is not used in other environments outside the classroom. It is possible that some features of classroom learning may affect the formation of a bilingual lexicon, which will make the acquisition of metaphorical expressions different from acquiring literal expressions. Therefore, the difference between classroom learning and other types of lexical acquisition should be taken into consideration in order to make better predictions for the outcome of acquisition. This subsection is a review of the features of lexical acquisition in the classroom setting, with a particular focus on the English learning environment in China.

Currently in Beijing, Shanghai and other provincial capitals, young learners will start their English courses in a formal setting in primary education, usually when they are between six to nine years old (Grade 1 to Grade 4). In some less developed areas, the onset of English courses may be delayed until the ages of twelve or thirteen, when learners enter secondary school (Grade 7). Some learners may receive earlier exposure to English, either in formal or informal ways; the age of pre-school exposure of children in provincial capitals is usually between three and five.

At the age of twelve, students begin to receive secondary education, with three years compulsory (Grade 7 to Grade 9) and three years optional (Grade 10 to Grade 12), but only those who finish Grade 12 can take the national joint college entrance examination (NJCEE) and be qualified for admission to universities. Although a choice of several second languages is available across the country and in the NJCEE, most secondary schools will settle on English; students will continue to learn English daily in these secondary schools. The English teachers will make use of both English and Chinese in the classroom, while other courses, which are generally not related to English, are completely in Chinese. By the time of the English exam in the NJCEE and the start of tertiary education, a learner will usually have been learning English

for 6 to 12 years and will have been exposed to English on a daily basis for no less than six years in total.

At tertiary level, all the students, irrespective of their majors, are required to attend compulsory English courses and pass the Band 4 English test in order to become qualified to receive a college degree (Ma 2009). Students with an English major will receive more professional training in English language, literature and linguistics, and it is recommended that they pass the Professional Band 8 English test. As well as the compulsory English courses, in some universities, students with other majors can choose English courses as their electives. When proceeding to graduate study (including MPhil, PhD and other doctoral degrees), candidates are expected to pass the National Graduate Examinations (NGEs), in which English is a major choice among a variety of second language exams. It should be noted, however, that English is not the official working language in the majority of tertiary education establishments, and tertiary education students in China, even the English majors, will not use English extensively in their daily communication with their peers. English has never received any institutional status in mainland China and is still seen as a 'foreign language' rather than a typical 'second language'.

Ma (2009) presents in detail some advantages, as well as a larger number of constraints, in English classroom teaching in China, particularly in terms of vocabulary acquisition. She concludes that, although highly motivated, Chinese learners do not receive a satisfactory outcome in the acquisition of English vocabulary. The culture of learning in general, as well as the historical methodology of vocabulary teaching and the lack of experienced, professional English teaching personnel, leads to an unhelpful situation in English vocabulary learning among Chinese students: the acquisition of vocabulary is consistently emphasised and students pay great attention to it, but methods are restricted simply to memorisation and grammar-translation tasks, while other skills relevant to the inference of word meanings and figurative language are largely neglected, due to outdated instruction methods and limited exposure to language materials. Ma does not extend her discussion to the acquisition of metaphorical uses of words in English specifically; therefore, it is not clear whether the Chinese learners of English are able to derive metaphorical meanings of lexical items, even if they are not explicitly taught such knowledge in a classroom setting. Nevertheless, it can be inferred from the current situation, especially from the lack of systematic contents related to the metaphorical uses of words and the execution of figurative thinking (see Kecskés 2000; Littlemore and Low 2006 for examples) in English textbooks and courses, that Chinese learners of English in a classroom

setting may show a gap in terms of metaphor comprehension and production, and this is a great obstacle to higher proficiency of English.

Jiang (2000) identifies two major constraints on the acquisition of lexical items in a classroom setting, which is also the main background of the current thesis. He characterises classroom learning as a process that lacks ‘sufficient and highly contextualised input in the target language’ (Jiang 2000: 49). The lack of natural input, in his view, will create obstacles to learners in their retrieval of semantic, syntactic and morphological information autonomously. From the perspective of lexical semantics, this constraint could indicate that, through the classroom learning process, learners may (1) only successfully acquire some of the meanings/uses of a lexical item that have been fully mastered by the native speakers, and (2) be less able to utilise the contextual information in a sentence to derive the possible meaning of a word in it.

The other constraint is the role of the L1 in the process of L2 acquisition. Although, as discussed in Chapter 2, cross-linguistic influence is inevitable in second language acquisition, the role of L1 may be even more prominent and influential in a second language classroom. English, as a second or foreign language, is learnt after a learner has acquired Chinese and has established the connection between a series of concepts and the corresponding Chinese word. Therefore, a learner who has acquired a word in English will tend to establish a lexical link between L1 and L2 translation equivalents, exactly as described in the RHM (Kroll and Stewart 1994). That may lead to a transfer of meanings from L1 to L2: the learner might connect the conceptual elements or bundles that are only available in the L1 to the L2 lexical item, while at the same time not being aware of the conceptual elements that are only available in the L2, since there is nothing to transfer, which means there might be a failure to acquire a certain part of the L2. In addition, a learner’s ability to extract information from contextual cues in the L2 is also largely suppressed when the learner relies too much on the L1. It may be more difficult to infer the meaning of a new lexical item from the given context, or to summarise the morphological or syntactic features of the lexical item. In a similar fashion, the learner may be less sensitive to the metaphorical meanings of a lexical item and may fail to derive a possible metaphorical meaning from the literal meaning of the lexical item and the contextual information, even if the literal meaning of the lexical item has already been acquired. The same problem is identified in the teaching process as well: English teaching in the classroom, especially in primary and secondary education, also frequently utilises Chinese as the medium language. The teachers generally rely on translation equivalents to teach the words in the L2 (Jiang 2004), especially at the initial stage when learners are relatively young (e.g. in primary

education). Therefore, the connection between the L2 word and its translation equivalent in L1 is relatively strong, while the link between the L2 word and its own semantic specification is weaker. The transfer from L1 words to L2 words will be more obvious in that situation.

Both of these constraints, as analysed above, are likely to lead to obstacles to the successful acquisition of metaphorical expressions. A lack of sufficient input may lead to a lack of exposure to the metaphorical uses of some lexical items; as a consequence, it is highly possible that a learner will understand the literal meanings of the lexical items but fail to accept the metaphorical uses. At the same time, a weakened ability to utilise contexts and draw inferences may result in an inability to derive a valid interpretation of metaphorical expressions, and thus a learner may be unable to understand and acquire a metaphorical expression on his/her own. The reliance on translation equivalents in classroom teaching may also trigger some unnecessary transfer from the L1 to the L2, especially for those who have only limited cumulative exposure to L2 materials.

However, there is still some space left for the acquisition of certain metaphorical expressions. According to Jiang (2007), linguistic knowledge of the L2 possessed by the learners in a classroom setting comes mainly from three different sources: learners' own knowledge of the L1, knowledge of the L2 through cross-linguistic transfer; some internalised knowledge gained through exposure to L2 language materials and interaction; and knowledge explicitly taught by the instructors. If some metaphorical expressions appear in learning materials, a learner may be able to identify, memorise and correctly use them, even without realising that those expressions are metaphorical. Or, in a similar fashion, if instructors point out that a direct translation of an L1-specific metaphorical expression is not acceptable in the L2 (e.g. the notion of 'Chinglish'), then a learner might avoid using it intentionally. Overall, whether a learner can successfully acquire a metaphorical expression mainly depends on the classroom input, i.e., the contents of textbooks and explicit teaching. That means that if a metaphorical expression has previously appeared in teaching materials, it is more likely to be acquired and accepted by a learner, compared with metaphorical expressions that are not explicitly taught in language courses. On the other hand, when a learner receives some metaphorical expressions as input in a non-guided situation, it is less possible to take these expressions in, to make native-like judgments of them, or even to use them efficiently in production.

Two major hypotheses can be drawn from the analysis above: (1) compared to native speakers of the target language, who should accept both the literal and the metaphorical meanings of a lexical item, L2 learners will be more reluctant to accept the metaphorical

meanings than the literal meanings, and this will be more obvious if the metaphorical meanings are only available in the L2; and (2) L2 learners, especially those with less knowledge of the L2, will demonstrate a certain degree of negative transfer when they are asked to comprehend and produce some expressions that are available in the L1 but impossible in the L2. Those hypotheses can be tested only if a comparison is made between a group of learners and a group of native speakers.

### 3.3 Constructing/accessing figurative language in the bilingual lexicon: the processing of figurative language

The processing of figurative language, including metaphor, idioms and phrasal verbs, by second language learners has only recently been stressed, even though the transferability of figurative language in second language acquisition and learners' general performance in the acquisition of L2 figurative expressions has been extensively studied in the past few decades (see Chapter 2.2 for a comprehensive review). The study of figurative language processing includes two major topics: (1) how figurative language is stored in a learner's mental lexicon; and (2) whether a learner can achieve a native-like processing pattern when processing figurative language. This subsection aims to provide a brief review of previous research on figurative language processing in the L2 to explore the status of figurative language in the bilingual lexicon and the possible patterns of figurative language acquisition and processing by L2 learners.

The reading time of an expression or a sentence structure is regarded as an effective indicator of processing difficulties in second language processing (Clahsen and Felser 2006): a longer reading time usually indicates that a learner may experience some difficulties, possibly including complex syntactic structures, lexical retrieval difficulties, ambiguity resolution, etc. In this subsection, processing difficulties presented by different types of figurative language are again considered to be reflected in prolonged reading or reaction time. As we will see in the rest of this subsection, the reading time for a figurative expression is associated with the 'meaning-making' of that expression, both for native speakers and for L2 learners. In a broader sense, the meaning of a multi-word figurative expression can be either constructed or accessed. 'Constructing' here refers to a more complicated process of meaning-making, involving searching, selecting and adjusting the meanings of individual constituents of a figurative expression. 'Accessing' refers to a more direct process, which means that a figurative meaning is readily available and a reader can ascertain the meaning without complicated selection, activation and possibly inhibition of incorrect meanings. Although, from the description, it

could be concluded that accessing should be faster than constructing, they form a continuum of processing time: some people can quickly construct the meaning of a figurative expression, while others may slowly access the meaning of an expression. In the current thesis, especially in this part of the review, it is suggested that a longer reading time is more likely to involve construction and a shorter reading time is more likely to involve direct access.

A long-standing question in figurative language processing, whether L1 or L2, is the validity of the 'literal-first' hypothesis (c.f. Section 1.2.1 and Section 1.2.3), which is a typical process of construction involving (1) initial activation of an invalid meaning, (2) re-construction of a valid meaning from several possible ones, and (3) inhibition of invalid meanings. However, previous research on L1 figurative language processing has shown that native speakers do not take additional steps when accessing the meaning of an idiomatic expression or a conventional metaphorical expression, as indicated by the fact that native speakers do not take longer to process figurative expressions (see Glucksberg 2001 for a summary). In particular, when the reading times for an idiomatic expression in a biased literal context and a biased idiomatic context are compared, it will be discovered that native speakers spend significantly less time reading and understanding the idiomatic use than the literal use. Thus, it is widely assumed that idioms are stored in a holistic manner in the mental lexicon of native speakers, and native speakers prefer accessing the idiomatic meaning to the literal meaning when they see a string of words that can be interpreted idiomatically. One step further, native speakers may bypass the literal meaning and directly access the figurative interpretation of an expression when processing it (i.e. the Direct Access view; see Gibbs 1984). Alternatively, it is still possible that native speakers are able to activate the literal and figurative meanings in parallel, or activate a set of primitive conceptual elements yet to form a specific meaning of that lexical item (e.g. Carston 2002). Then, depending on the salience provided by the context, native speakers will access the most salient or context-relevant meaning, be it literal or metaphorical (see Giora 1997, 2002 for Graded Salience Hypothesis and figurative language comprehension).

However, due to the difference between the L1 and the L2, and particularly the lack of metaphorical competence in L2 (Kecskés 2000; Littlemore and Low 2006), it is possible that L2 learners are not as sensitive as native speakers in terms of access of non-literal meanings. While 'literal-first' may not be the case with figurative language processing among native speakers of a language, it is still possible that L2 learners will need to spend more time reading and understanding the figurative meaning of an expression in their L2, including idioms, proverbs, phrasal verbs and metaphorical expressions. In this case, a longer reading time involves the composition of literal meaning and a selection and adjustment in order to derive

the figurative meaning. It is even possible that learners may perceive the figurativeness of a figurative expression differently from native speakers (see Boers and Webb 2015 on semantic transparency of idioms), which eventually leads to a different processing pattern. At the same time, the reading time for L2 figurative language might be influenced by factors other than figurativeness itself. If we consider the assumptions of idiom acquisition, particularly Transfer Theory (Irujo 1986, 1992; see Chapter 2.2 for a summary), we can expect to see that learners will display a varied reading time for idioms: they may take less time to read idioms that are identical or similar to those in their L1 and L2, while they may take more time to read idioms that are different in their L1 and L2. The L2 proficiency of a learner may also be expected to influence the processing of figurative language, which is the same as the influence from proficiency on the processing of other structures in an L2: more proficient learners may be expected to process figurative language faster than less proficient learners.

Previous studies on figurative language processing by L2 learners show conflicting patterns in different situations. Cieślicka (2006) investigates the processing of English idioms by Polish learners of English in a lexical priming test, and the results indicate that learners favour the literal meaning of English idioms when they first hear them in the experiment. She reports that learners show a faster reaction for a target if its primer is a literal expression, while they take more time to react to a target after hearing an idiom. Cieślicka suggests that the longer reaction time after the presentation of an idiom is due to the activation of the literal meaning of that string before learners finally arrive at the idiomatic meaning. The result of this experiment is largely in line with the 'literal-first' assumption of idiom processing (e.g. Gibbs and Nayak 1989; Glucksberg 2001). Cieślicka then argues that, unlike in the case of native speakers who can directly access the idiomatic meaning, the literal meaning of an expression is always more salient among L2 learners, and L2 learners always need to construct the idiomatic meaning. Similar results have been observed in other studies (see Siyanova-Chanturia and others 2011 for an example of an eye-tracking study): L2 learners show a pattern of idiom-processing distinct from that of native speakers: they treat idioms in the same way as novel expressions, and tend to derive the compositional meaning of an idiom before accessing the figurative interpretation.

However, a series of studies by Heredia and his colleagues on the processing of phrasal verbs and idioms by learners of English reveals that learners are able to show a reading pattern similar to that of native speakers of English and access the figurative meaning of an expression first. Matlock and Heredia (2002) examine the reading time for English idiomatic phrasal verbs and literal verb phrases (both in the form of verb + preposition) by native speakers of English,

early and late English learners. They find that learners of English, especially learners who start learning before 12, show a reading pattern similar to the native speaker group, and spend significantly less time reading the idiomatic phrasal verbs than reading the literal expressions. Heredia and colleagues (2007) examine the processing of different types of idiom in isolation, and in literal, figurative or unbiased contexts by Spanish learners of English, to check the validity of Irujo's Transfer Theory in real-time idiom processing. While Transfer Theory indicates that idioms that are identical or similar in L1 and L2 will be processed faster by L2 learners than those that differ across the languages, the results of Heredia and colleagues show a contradictory pattern. Different idioms in English and Spanish actually trigger faster reactions among learners (1) when they appear in isolation; (2) when the context is not biased; or (3) when the context is biased to the figurative interpretation of the string. Identical or similar idioms are processed more slowly than the different idioms in the three conditions above. They then argue that idioms are stored differently in learners' mental lexicon, depending on the form of the idioms. The idioms with different forms in learners' L1 and L2 are stored separately and holistically in their mental lexicon, which is similar to the storage of idioms of native speakers, and learners can directly access the idiomatic meanings when processing those phrases. In contrast, identical or similar idioms require the learners to reconstruct them as they are accessed, which may trigger inter-language competition and ultimately slow down the processing (Heredia and others 2007). The observations of Heredia and colleagues indicate that L2 learners do not always adopt the 'literal-first' strategy in the processing of figurative language; in particular, highly proficient learners can identify the figurative uses of language immediately, and show a native-like processing pattern.

The conflicting results of different studies reveal that figurative language processing is a complicated phenomenon involving multiple factors; the tasks used in the studies, the test items and even the participants' native language might affect the result. A survey by Titone and colleagues (2015) on idiom processing by English-French bilinguals shows that the figurativeness of an idiom, its general familiarity, and cross-linguistic similarity between learners' L1 and L2 can all influence the speed at which it is processed and the accuracy of results, but the decomposability of an idiom does not directly influence the processing of that idiom. It should be noted that these factors are not independent of each other: a pre-test survey shows that the cross-linguistic similarity of an idiom in two languages has a significant positive correlation with the familiarity of that idiom, and both of the factors positively correlate significantly with its decomposability. This means that if an idiom is familiar to learners, it is usually semantically more decomposable, i.e. learners will find it easier to derive the meaning

of the idiom from its semantic composition, when the idiom looks similar cross-linguistically between learners' L1 and L2 (see also Libben and Titone 2008). The correlation between the factors might be used to explain why learners show conflicting performance across different experiments: idioms with a higher degree of cross-linguistic similarity are semantically more decomposable, so a learner may be more willing to use the literal meaning of each component to derive the figurative meaning of the idiom, which will lead to the early activation of the literal meanings of the string. On the other hand, less familiar idioms are not always decomposable, and learners may need to memorise the idiomatic meanings separately and holistically, which may result in the direct retrieval of the figurative meaning of those idioms.

While idiom comprehension by L2 speakers has been widely surveyed in online experiments, few past studies have focused on the processing of conventionalised metaphorical expressions. García and colleagues (2015) point out that the mechanism of metaphor comprehension and idiom comprehension might be essentially different. According to them, comprehension of metaphor may require more contextual information, because it is possible that metaphorical expressions, either conventional or unconventional, are not stored in learners' mental lexicon in a holistic way, and learners may always need to carry out *ad hoc* construction in order to establish the meaning of the metaphorical expressions. On the other hand, metaphorical expressions are semantically more decomposable than some idioms, and learners may therefore rely on compositional meaning to construct the metaphorical meaning of those expressions by default. It remains largely unknown how learners process and understand conventional or unconventional metaphorical expressions, and this thesis will explore the processing of conventional metaphorical expressions to show whether it resembles the processing of idioms or is more akin to the processing of literal expressions, or perhaps somewhere in between the two situations.

It should be noted that different experimental tools may also affect learners' performance. One noticeable effect is the so-called online/offline difference, which may be used to explain the distinction between online experiments measuring the reaction time and offline experiments examining the judgments made by learners. Offline tasks are generally considered a method to examine the explicit knowledge of learners (Sanford and others 2004), and if an offline task involves production, it also requires learners' active vocabulary. Online tasks, on the other hand, are considered to tap into learners' implicit knowledge (Mitchell 2004); usually online tasks involve perception but not production, only requiring learners' passive vocabulary. Differences in results between online and offline tasks may not necessarily be in conflict: it is possible for learners to be able to understand figurative language intuitively, but to lack

confidence or enough explicit knowledge to make active judgments, so they may even deliberately avoid figurative language even when they are given sufficient time to produce an utterance. In the current thesis, both online and offline methods will be used to explore the acquisition and processing of metaphorical expressions in second language acquisition.

### 3.4 Possible routes for processing metaphorical meanings in a second language

Taking the available models for the bilingual lexicon and the general processing pattern of figurative language as a foundation, we can build a potential model to capture the acquisition, storage and accessing of metaphorical meanings in the bilingual lexicon. Based on that model, several possible routes of metaphorical meaning processing can be derived, depending on the availability of the metaphorical meanings in the two languages.

A simplified model for the storage of literal and metaphorical meanings in general can be seen below in Figure 3.6, which shows a modified version of the MHM, while stressing the availability of different types of metaphorical expressions rather than the conceptual differences. Both the L1 and L2 lexicon are connected to three types of concepts: the literal concept<sup>3</sup>, the shared metaphorical concept, and the language-specific metaphorical concept. The bold grey links show the connection between the literal and metaphorical concepts on a conceptual level; they are grey because these links are not systematically available to learners who have no prior knowledge of conceptual metaphors, but it is possible for learners to build up a connection in an *ad hoc* manner given sufficient contextual information. The heavier lines between the L1 lexicon and relevant concepts, as well as from the L2 lexicon to the L1 lexicon indicate that these links are relatively strong, while the dashed links between the L2 lexicon and relevant concepts show that concept mediation may exist under limited conditions. In particular, a learner may not be fully aware of the existence of L2-specific metaphorical concepts, and the links to that set of concepts, both the lexicon-concept link and the between-concept link, are therefore exceptionally weak.

---

<sup>3</sup> For the sake of simplicity, the literal concepts are treated as being fully shared between the L1 and the L2 in the current thesis. However, as pointed out by Pavlenko (2009), there may be language-specific literal concepts.

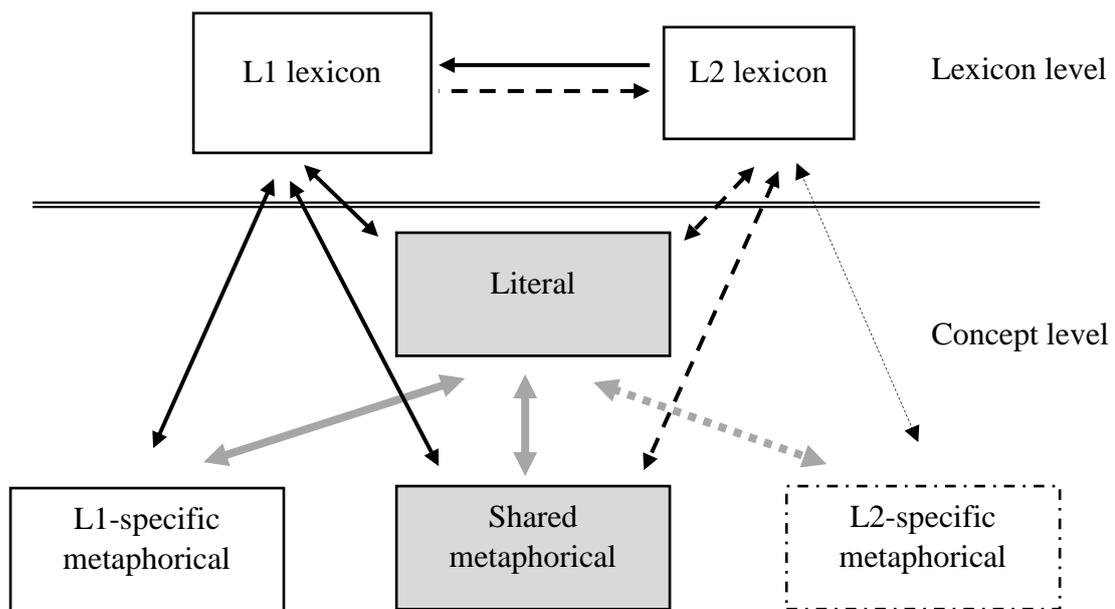


Figure 3.6 A simplified model of bilingual lexicon with literal and metaphorical distinctions

Since the processing of idiomatic expressions and other types of figurative language among native speakers of a variety of languages indicates that native speakers of a language can process figurative expressions in the same manner as literal expressions, we take the processing of metaphorical meanings by native speakers as a baseline. We assume that native speakers will spend the same amount of time processing the literal meaning and the metaphorical meaning of a lexical item: this means that native speakers will establish the link directly between the word form and the conventional metaphorical meanings of that lexical item, and they will not rely on other routes to retrieve the metaphorical meaning.

There are two ways for L2 learners to retrieve the literal meaning of an L2 word when they see the L2 word form, listed in Table 3.1. They may either directly pursue the link between the L2 word form and the literal concept, namely via concept mediation ( $LR_1$ ), or they may take an additional step and establish the link via the L1 word form, namely via word association ( $LR_2$ ).  $LR_1$  is structurally less complex than  $LR_2$ , since  $LR_1$  takes a shorter route and has fewer steps. A less complex route is generally considered faster than a more complex route, but it should be noted that speed of retrieval may vary even for the same route. Since the widespread appearance of concept mediation is largely recognised as an indicator of higher L2 proficiency (see Dufour and Kroll 1995), it can be assumed that  $LR_1$  may be used if a learner is at medium- or high-level proficiency, while  $LR_2$  is more likely to be used by a less proficient learner. These two possibilities are taken as the fundamental assumptions for the retrieval of metaphorical

meanings. Also, it is assumed that the processing of literal expressions is always consistent: if LR<sub>2</sub> is selected by a learner, we expect to observe word association in all cases of meaning retrieval, because that means that proficiency is not yet sufficient to allow concept mediation.

Table 3.1 Two possible routes for the retrieval of the literal meaning of an L2 word

Possibilities	Route	Complexity	Proficiency
LR <sub>1</sub>	L2 form–Literal	Low	Medium–high
LR <sub>2</sub>	L2 form–L1 form–Literal	Medium	Low

For the retrieval of the shared metaphorical meaning, a learner may have four main possibilities. The first possible route (MBR<sub>1</sub>) is the same as the metaphorical processing of a native speaker. This means the learner can directly establish the link between the L2 word form and the shared metaphorical meaning, and thus has already integrated the metaphorical meaning as part of the mental lexicon. The second possible route (MBR<sub>2</sub>) assumes that the learner will firstly adopt word association, and then directly pursue the link between the L1 word form and the shared metaphorical meaning. The third possible route (MBR<sub>3</sub>) assumes that the learner may first arrive at the literal concept via concept mediation from the L2 word form, and then arrive at the shared metaphorical meaning via the link between the literal and metaphorical concept. The last possibility (MBR<sub>4</sub>) takes the longest route: learners will first go through concept mediation to reach the literal concept; when they will find that the literal concept does not provide a valid explanation of the whole expression, they will not make any inference of the metaphorical meaning from the literal meaning. Instead, they will revert to the L1 word form, and then access the shared metaphorical expressions via the L1 word form.

Table 3.2 Four possible routes for the retrieval of the metaphorical meaning shared between the L1 and the L2

Possibilities	Route	Complexity	Proficiency
MBR <sub>1</sub>	L2 form–Shared metaphorical	Low	Medium–high
MBR <sub>2</sub>	L2 form–L1 form–Shared metaphorical	Medium	Low
MBR <sub>3</sub>	L2 form–Literal–Shared metaphorical	Medium	Medium–high
MBR <sub>4</sub>	L2 form–Literal–L1 form–Shared metaphorical	High	Medium–high

When a learner is forced to retrieve a metaphorical meaning that is only available in L1 from an L2 word form, for instance, to comprehend an impossible expression like ‘eat the loss’ (meaning ‘suffer the loss’), it will first be noted that there is no direct link from the L2 word form to the L1-specific metaphorical concept. It is generally unlikely for the learner to receive such kind of input in their exposure to the L2, and this condition is set up only for a contrast with the processing of possible metaphorical expressions in the L2. In this condition, the learner needs to construct the L1-specific metaphorical meanings from the L2 input. In an offline task, such as acceptability judgment, the appearance of these meanings will lead to rejection; in an online task, however, under the time pressure, a learner is given no choice but forced to construct the intended meanings. One possibility (MSR<sub>1</sub>) is that the learner will first access the L1 word form by word association, and then arrive at the L1-specific metaphorical concept via the link from the L1 form. The other main possibility is that the learner will first access the literal concept by concept mediation, and there will then be a second divergence: use can be made of the link at conceptual level, and the learner may reach the L1-specific metaphorical concept by referring to the literal concept (MSR<sub>2</sub>); or, upon finding that the literal concept does not provide a feasible explanation, may return to the L1 word form, and then use the lexical-concept link to arrive at the L1-specific metaphorical concept (MSR<sub>3</sub>). Compared with the first two possibilities, it may take more time for the learner to construct the L1-specific metaphorical meaning from the L2 word form in these conditions in MSR<sub>3</sub>, since there is a necessity to return to the L1 word form.

Table 3.3 Three possible routes for the retrieval of the L1-specific metaphorical meaning from the L2 word form

Possibilities	Route	Complexity	Proficiency
MSR <sub>1</sub>	L2 form–L1 form–L1-specific metaphorical	Medium	Low
MSR <sub>2</sub>	L2 form–Literal–L1-specific metaphorical	Medium	Medium–high
MSR <sub>3</sub>	L2 form–Literal–L1 form–L1-specific metaphorical	High	Medium–high

When a learner tries to retrieve a metaphorical meaning that is only available in the L2 from an L2 word form, it will be noted that a simple word association cannot solve the problem because the L1 word form is not directly connected to the L2-specific metaphorical concept. The first possibility (MTR<sub>1</sub>) is to make use of concept mediation and travel directly through the

link between the L2 word form and the L2-specific metaphorical concept; that is the least complex route, but there needs to be sufficient knowledge of metaphorical expressions in the L2 to establish and maintain the link. The second possibility (MTR<sub>2</sub>) is that the learner may first use concept mediation to access the literal concept of that word, and then reach the L2-specific metaphorical meaning via the link between the literal and metaphorical concept. Alternatively, if it is necessary to rely on word association at the beginning of processing, it will only be possible to retrieve the literal concept after word association, and then a link between the literal and metaphorical concept will still be needed (MTR<sub>3</sub>).

Table 3.4 Three possible routes for the retrieval of the L2-specific metaphorical meaning of an L2 word

Possibilities	Route	Complexity	Proficiency
MTR <sub>1</sub>	L2 form–L2-specific metaphorical	Low	High
MTR <sub>2</sub>	L2 form–Literal–L2-specific metaphorical	Medium	Medium–high
MTR <sub>3</sub>	L2 form–L1 form–Literal–L2-specific metaphorical	High	Low

These possible routes can theoretically come in bundles as shown in Table 3.5. The logic is that the retrieval of the metaphorical meaning does not conflict with the retrieval of the literal meaning, and a learner always takes the shortest possible route. Therefore, if a learner can make use of concept mediation consistently for one type of literal or metaphorical expression, it is assumed that word association will not be used in other types. Bundle 1 is a case where word association is extensively used, in which a learner must achieve access to the L1 word form before any further processing. The proficiency level of a learner may be relatively low, while the complexity of the route is generally high. Bundles 2 to 5 are cases where concept mediation is generally used, but the destinations of concept mediation vary depending on the availability of links in the model in Figure 3.6. A learner using the Bundle 2 strategy is considered not to be able to infer the shared metaphorical meaning directly from the literal meaning of the word and the given context, so it is necessary to revert to the L1 word form. The improvement in the Bundle 3 strategy is that a learner can infer the shared metaphorical meaning from the literal meaning and the given context, which is more like the ‘literal-first’ assumption of figurative language processing. In Bundle 4 and 5 strategies, a learner can bypass the literal meaning and directly link the shared metaphorical meaning to the L2 word form. Bundle 5 is the strategy

that resembles the word-processing of native speakers of the target language, because in that case a learner can establish the links directly to the metaphorical meanings that are available in the L2, just as with the links to the literal meaning of that word. No ‘literal-first’ effect will be observed in the Bundle 5 strategy in cases of shared metaphorical meanings or L2-specific metaphorical meanings.

Table 3.5 Different bundles of possible routes of metaphorical meaning retrieval in L2

	Bundle 1	Bundle 2	Bundle 3	Bundle 4	Bundle 5
Proficiency requirement	Low	Medium	Medium to high	Medium to high	High
Literal	LR <sub>2</sub>	LR <sub>1</sub>	LR <sub>1</sub>	LR <sub>1</sub>	LR <sub>1</sub>
Shared metaphorical	MBR <sub>2</sub>	MBR <sub>4</sub>	MBR <sub>3</sub>	MBR <sub>1</sub>	MBR <sub>1</sub>
L1-specific metaphorical	MSR <sub>1</sub>	MSR <sub>3</sub>	MSR <sub>2/3</sub>	MSR <sub>2/3</sub>	MSR <sub>2/3</sub>
L2-specific metaphorical	MTR <sub>3</sub>	MTR <sub>2</sub>	MTR <sub>2</sub>	MTR <sub>2</sub>	MTR <sub>1</sub>
General complexity	High	Medium to high	Medium	Low to medium	Low

In terms of the general complexity of processing of each bundle, the reading times of different metaphorical expressions, as well as literal expressions can be utilised to speculate which strategy a learner is using in the processing of different metaphorical expressions. If a learner does not show a significant difference in reading time when processing a literal expression and a metaphorical expression, it can be inferred that there is no need to refer to the literal concept in order to obtain access to the metaphorical concept. Similarly, if a learner spends significantly longer processing a metaphorical expression than processing a literal expression, it can be inferred that an additional path has been taken to reach the metaphorical meaning. A self-paced reading experiment is designed to examine the reading time of literal and metaphorical expressions by native speakers of English and Chinese learners of English, and the results will be discussed in Chapter 6.

### 3.5 Summary

This chapter has served as the second part of the literature review and has mainly focused on the organisation of the bilingual lexicon of L2 learners, further restricted to sequential L2 learners in a classroom setting. The processing pattern for other types of figurative expression by L2 learners, such as idioms and phrasal verbs, is also reviewed in the chapter. This review provides some insights into possible processing patterns for metaphorical expressions. Specifically, the feasibility of the ‘literal-first’ hypothesis in figurative language processing among L2 learners remains a core question in terms of the acquisition and processing of metaphorical expressions, especially because the teaching of our learners will have promoted literal usage before metaphorical meanings.

Although none of the currently available models of bilingual lexicon is fully applicable to the acquisition of metaphorical expressions, we can construct a hybrid model that can reflect (1) the difference between concrete, literal meanings and abstract, metaphorical meanings, and (2) the difference between learners at different levels of proficiency. Based on the hybrid model, we can then generate a series of possibilities regarding the storage and access of metaphorical meanings of a lexical item and make further predictions regarding the reading and comprehension of metaphorical expressions by L2 learners. Some of these hypotheses generated by the hybrid model contradict the predictions derived based on the literature on cross-linguistic influence in Chapter 2, an issue that will be discussed in detail in Chapter 7, after the reporting of the experimental results.

## 4 Research Methodology

### 4.1 Research questions and hypotheses

Based on the literature review in the previous chapters, it can be observed that there might be conflicting views regarding the acquisition of metaphorical expressions. To start to tackle the problem, we will take the premise of CMT by Lakoff and Johnson (1980), and assume that the metaphorical meaning of a lexical item is more abstract than the literal meaning of that lexical item. If this assumption stands, we can then compare the acquisition of a more concrete meaning and a more abstract meaning of a polysemous word. While Jordens and Kellerman (1981) suggest that the presence of metaphoricalness in the meaning of a lexical item *may not* play an important role, from the viewpoint of the construction of the bilingual lexicon, however, particularly the DCFM of De Groot (1992), whether the meaning of a lexical item is abstract, or metaphorical, may be a crucial factor in influencing the transferability of that meaning (c.f. Section 3.2.1).

The seemingly conflicting theories lead us to question the possible cross-linguistic transferability of metaphorical expressions in the process of the acquisition of English by Chinese learners. More generally, we should ask whether learners perceive metaphorical expressions as transferable from learners' L1 to the L2, which can be reflected by the degree of acceptability of metaphorical expressions across different conditions. The logic here follows Odlin's (1989) observation on the lexical acquisition of L2: a learner may directly transfer the usage of an L1 lexical item in the L2 by using the L2 translation equivalent of that lexical item, even if there is a possibility of semantic mismatch (see Section 2.2.2 for examples). It is likely that the same situation applies to the acquisition of concrete and abstract meanings of a lexical item: if a learner assumes that the literal meaning of an L1 lexical item is transferable to the L2, they might assume that the metaphorical meaning(s) of the same lexical item is transferable to the L2 as well; however, if the learner believes that the metaphorical meaning(s) and the literal meaning of a lexical item should be treated differently, the transfer of the metaphorical meaning may be rejected.

Observations on the transferability of lexical items from the L1 to the L2 (as in Chapter 2), together with the current frameworks for the acquisition/organisation of the bilingual lexicon (as in Chapter 3) and factors recognised in the literature as influencing transferability, might lead to a paradigm that can be used to examine the transferability of metaphorical expressions in L2 acquisition. When we take a pair of translation equivalents from L1 and L2, in order for

them to qualify as translation equivalents, we assume that some, if not all, literal meaning(s) of that lexical item are transferable from the L1 to the L2. At the same time, we can compare potential differences between that shared literal meaning and the metaphorical meaning(s) of the same lexical item to examine whether they are transferable to the same extent. Moreover, it can also be observed whether the learners treat the metaphorical meanings and the literal meaning of the same lexical item in the same manner in the construction of the bilingual lexicon. The transferability of the different meanings of a single lexical item can be reflected by the following two factors: the degree of acceptability of an expression as perceived by the language learners, and the reaction time to that given expression. Given this scenario, the current thesis aims to answer the following questions:

- How do Chinese learners of English acquire the conventional metaphorical meaning of a lexical item in their second language? Does any form of cross-linguistic influence appear in the process of acquisition? If so, how does cross-linguistic influence affect the acquisition of metaphorical expressions?
- How do Chinese learners of English retrieve a conventional metaphorical meaning of a lexical item in their second language in real-time processing? Are they able to establish the link directly from the word to the metaphorical meaning, namely by retrieving the metaphorical meanings in the same way as retrieving the literal meanings? Or, do they require any additional process in order to understand the metaphorical meaning?
- How do Chinese learners of English deal with metaphorical expressions with different availability and transferability across languages in offline judgment and online reading tasks? This includes (1) metaphorical expressions that are transferable; (2) metaphorical expressions that are not transferable because there is no corresponding element in the L2; and (3) metaphorical expressions that are not transferable because there is no corresponding element in the L1.
- How do the three factors mentioned in Jordens and Kellerman (1981) as potentially influencing transfer guide the acquisition of the metaphorical expressions? To remind the reader, the factors are (1) psychotypological distance between Chinese and English; (2) the markedness of an expression; and (3) the learner's knowledge of English.

The learners' performance will be compared with that of the native speakers of English. A series of hypotheses regarding the native speakers' reaction to metaphorical expressions in English is listed below as the benchmark for the learners' performance:

- Native speakers of English should accept all metaphorical expressions available in English, and should reject metaphorical expressions that are not available in English. They should accept the literal meaning and the metaphorical meaning of the same word to a similar degree, if both types of meanings are available in English.
- Native speakers of English should retrieve the metaphorical meanings of an English word as quickly as the literal meaning, and they should not hesitate after reading such metaphorical expressions. Therefore, they should spend the same amount of time reacting to the metaphorical use and the literal use of a target lexical item, and they should answer the comprehension questions with similar speed and accuracy.
- Native speakers of English should not show understanding of metaphorical meanings that are not available in English, and they should show significant hesitation after reading such metaphorical expressions. Therefore, compared with the literal use of target lexical items, they should spend significantly more time reading the Chinese-specific metaphorical uses of target words and answering the questions related to these metaphorical uses. The accuracy of their answers related to these metaphorically used expressions should also be lower than for answers to the literal use of the same lexical items.

A series of hypotheses regarding learners' reaction to metaphorical expressions in English is listed as follows:

- In an offline judgment task, Chinese learners of English should show different reactions to different types of metaphorical expressions listed as follows, depending on the availability of the expressions and their proficiency levels:
  - All learners can potentially accept metaphorical expressions available in both Chinese and English, but they should accept the metaphorical meanings less readily than the literal meanings of a same word.
  - More proficient learners should reject metaphorical expressions that are only available in Chinese when presented in a word-to-word translation in English, while less proficient learners might (partially) accept such expressions;
  - More proficient learners should accept metaphorical expressions that are only available in English and not possible in Chinese, while less proficient learners might reject such expressions.

- In a real-time processing task, Chinese learners of English should show different reading patterns for the three different types of metaphorical expression, depending on the availability of the expressions across different languages and their proficiency levels:
  - All learners should read the sentences containing metaphorical expressions available in both Chinese and English according to a similar pattern to that shown when they read the sentences containing the literal meanings of the same lexical item. They should spend the same amount of time answering the questions on the literal and metaphorical expressions, and the answers should be of similar accuracy.
  - More proficient learners should show significantly more hesitation after reading the metaphorical expressions that are only available in Chinese than after reading the literal expressions and spend significantly more time answering the questions on the metaphorical expressions than those on the literal expressions. The accuracy of answers related to the metaphorical expressions should be significantly lower than for answers related to the literal expressions among more proficient learners. Less proficient learners should show less hesitation than more proficient learners; they should also answer the questions regarding the meanings of the metaphorical expressions faster and more accurately (i.e., as if a translation equivalent exists) than more proficient learners.
  - More proficient learners should read the metaphorical expressions only available in English following a similar pattern to that in which they read the sentences containing the literal expressions. They should spend the same amount of time answering the questions on literal and metaphorical expressions, and the answers should be of similar accuracy. Less proficient learners should show more hesitation than more proficient learners; they should also answer the questions on the metaphorical expressions more slowly and less accurately than more proficient learners.
- As for transfer, we assume that the following traces should exist to show that a learner transfers knowledge of a metaphorical expression rather than completely acquiring that expression from scratch:
  - Learners will accept metaphorical expressions shared between Chinese and English with relatively low certainty, because a shared expression is transferable and they are not confident in justifying their judgment using their own knowledge;

- Learners will accept metaphorical expressions that are available exclusively in Chinese, irrespective of their level of confidence, because a Chinese-specific metaphorical expression is available in their L1 but not transferable;
  - Learners will reject metaphorical expressions that are available exclusively in English, irrespective of their level of confidence, because an English-specific metaphorical expression is not available in their L1 and there is nothing for them to transfer.
- Considering the influence of psychotypology on transfer, learners who perceive English as a language close to Chinese should allow transfer from Chinese to English more frequently compared with those perceiving English as a language distant from Chinese.
  - Considering the influence of markedness, learners should transfer knowledge of the metaphorical expressions that are rated ‘unmarked’ more than knowledge of those metaphorical expressions that are rated ‘marked’ under the criteria of Jordens and Kellerman (1981) (c.f. Section 2.3.2).
  - Considering the influence of knowledge of the target language, which can be roughly estimated by the proficiency of individual learners, learners with less knowledge and lower proficiency should show more instances of negative transfer than those with more knowledge and higher proficiency.
  - The three influencing factors, namely psychotypology, markedness and knowledge of a target language, may also affect a learner’s reading time for a metaphorical expression in different ways.

#### 4.2 Selection and categorisation of test materials for metaphorical expressions

The experiments in the dissertation made use of two main instruments: an acceptability judgement task in order to investigate learners’ explicit knowledge of and attitude towards metaphorical expressions when given sufficient time to process those expressions, and a self-paced reading task in order to investigate learners’ reading pattern and possible ways of retrieval of metaphorical meanings in a real-time reading setting. The acceptability judgement task aimed to answer the research questions related to cross-linguistic influence on the acquisition of metaphorical expressions, and the self-paced reading task aimed to answer the research questions related to the organisation of metaphorical meanings in a bilingual mental lexicon. All the test materials used in the main experimental instruments as reported in later sections

took the format of individual sentences (henceforth ‘test sentences’), and each sentence contained a critical lexical item to be tested. Each critical lexical item appeared in a pair of test sentences in an experimental task: in each case, in one sentence the intended meaning of the lexical item was metaphorical, and in the other it was literal. All the critical lexical items were selected from the established literature on metaphors in Chinese and English (particularly from Lakoff and Johnson 1980; Liu 2003; Link 2013; Yu 1996). The lexical items used in the acceptability judgment task were partly different from those in the self-paced reading task: the former involved an equal number of nouns and verbs, while the latter involved verbs only.

For the main experimental tasks, three categories of metaphorical expression were designed: (1) the Metaphorical-Both (MB) category, including metaphorical expressions that were available in both Chinese and English with the same meaning and that could be translated in a word-to-word manner without the comprehensibility in the two languages being affected; (2) the Metaphorical-Source (MS) category, which included metaphorical expressions only available in the source language Chinese, where a word-to-word translation to English is believed to be incomprehensible to English native speakers; and (3) the Metaphorical-Target (MT) category, which included metaphorical expressions that were only available in the target language English, and where a word-to-word translation of these expressions to Chinese is believed to be incomprehensible to Chinese native speakers. This classification corresponded to the different possible levels of transferability in the acquisition of metaphorical expressions: the expressions in the MB category were generally transferable; those in the MS category were not transferable because there was no word-to-word corresponding expression in the target language; the expressions in the MT category were not transferable either, since learners did not have any equivalents of the expressions in their L1. To provide a baseline for the results of acquisition and the patterns of processing, three categories of literal expression were also designed and examined in the experiments: (1) the Literal-Both (LB) category, the literal, baseline counterpart of the MB category; (2) the Literal-Source (LS) category, the literal, baseline counterpart of the MS category; and (3) the Literal-Target (LT) category, the literal, baseline counterpart of the MT category. The LB and MB, LS and MS and LT and MT categories made use of the same group of critical lexical items. The availability of the different types of expression in Chinese and English is shown in Table 4.1 below.

Table 4.1 Availability of different types of expression in the current thesis in Chinese and English

	Availability in Chinese	Availability in English
MB (Metaphorical-Both)	Yes	Yes
MS (Metaphorical-Source)	Yes	No
MT (Metaphorical-Target)	No	Yes
LB (Literal-Both)	Yes	Yes
LS (Literal-Source)	Yes	Yes
LT (Literal-Target)	Yes	Yes

In the acceptability judgment task, each category consisted of eight different critical lexical items, including four nouns and four verbs. The critical lexical items are listed in Appendix 1, together with their part of speech and general frequency of each of them (with the same part of speech but including both literal and metaphorical meanings) in the British National Corpus (henceforth BNC). The 24 critical lexical items were either one-syllable or two-syllable words ( $M=1.208$ ,  $SD=0.415$ ) and their orthographic lengths were between 3 to 7 letters ( $M=4.458$ ,  $SD=1.503$ ). The general frequency of the lexical items was not fully controlled since that would lead to difficulty finding appropriate expressions, but frequency was included as a factor in the data analysis and will be analysed for possible influence on the acquisition of metaphorical expressions.

The test sentences, including the MB or MT, expressions were extracted from the established literature on metaphor, the BNC and selected works of English literature. Small modifications were made to ensure that (1) the sentences were of similar length; and (2) the intermediate learners could understand each word without referring to a dictionary. The sentences that included MS expressions that were originally available in Chinese were selected from literature on Chinese metaphors and then translated by the experimenter. The 24 sentences that included literal uses of the critical lexical items were also selected from the BNC and selected English literature works. Additionally, 26 fillers were included in the task, half of which were grammatical, half ungrammatical. After modification and translation, all the sentences were examined and verified by three native English consultants to confirm that (1) all the sentences in the MB and MT categories, literal counterparts and grammatical fillers were acceptable to native English speakers; (2) all the ungrammatical fillers were indeed unacceptable; and (3) all the sentences in the MS category were unacceptable. Note that in the

main experiment, it was shown that the judgments varied between individual participants in the native speaker group, and some of the test items, even though verified by the consultants, were found not to be generally acceptable to all participants. Three Chinese native speakers with high proficiency English also examined and verified the sentences to confirm that (1) all the sentences containing MB expressions had exact word-to-word translation equivalents in Chinese; (2) none of the sentences containing MT expressions could be understood after a word-to-word literal translation to Chinese; and (3) all the sentences containing the literal uses of the critical lexical items could be understood after a word-to-word literal translation to Chinese.

For the self-paced reading task, only verbs were selected as critical lexical items to construct the test sentences. As explained below, this was done to ensure that the reading time could be compared across test items. The process of selection and construction of test sentences was similar to the process for the acceptability judgment task, but the length of the critical lexical items and the structure of test sentences was further controlled to fit the standard requirements of a self-paced reading task. This is also why all the metaphorical expressions and the counterpart literal phrases of the critical lexical items were constructed as verb phrases in the form of Verb-Article-Noun, beginning with the critical lexical item (a verb). A pair of examples is provided below to demonstrate the structure.

a) Literal: built a balcony

Metaphorical: built an argument

In example (a), the underlined verb was the critical lexical item, which was followed by the noun phrase that included an article. The noun served as the direct object of the verb. When participants read the noun, they were expected to know whether the verb was used literally or metaphorically. By using verb phrases only, it was possible to measure reading time across similar structures across items.

The critical lexical items are listed in Appendix 2. All of them appeared in the sentence in the past tense form. The MS and MT categories consisted of six different verbs, and the number was doubled for the MB category, making twelve different verbs. All the critical lexical items were monosyllabic, and the average orthographic length was 5 letters ( $SD=1.38$ ). Due to the restriction on part of speech and the number of syllables, only certain lexical items appeared in both the AJT and the SPR. One might argue that differences in critical lexical items in the two experiments might lead to varied results, which might make the conclusion presented in the current thesis less convincing. Nevertheless, it should be noted that the question under investigation in the experiments is not the acquisition of a definite list of lexical items, but the acquisition of a type of meaning for the lexical items (i.e. metaphorical meaning). Theoretically,

any meaning that fits the requirement of the investigation could and should be investigated, but only a sample of meanings and lexical items was selected due to the limitations of the experiments. Therefore, in the statistical processing of the experimental data, the difference between individual lexical items should and will be treated as ‘noise’ to avoid the phenomenon ‘language as a fixed factor’ (Barr and others 2013), and the difference in test items should no longer be an influential factor under such treatment.

144 sentences were constructed and presented to the participants in the self-paced reading task. 24 sentences contained a metaphorical expression based on the 24 critical lexical items listed above, and every metaphorical expression appeared only once; 24 sentences contained an expression that included the literal use of the target words, and the remaining 96 sentences were fillers. Each sentence consisted of 12 words. In both metaphorical and literal sentences, the first three words formed the subject of the sentence in the form The-Adjective-Noun; the fourth to sixth words included the metaphorical or literal use of the target lexical item; the final six words provided the contextual information and enough space for a spill-over effect<sup>4</sup>. The filler sentences were in the same format, but the only difference was that they did not contain any critical lexical item. Except for the differences in noun phrases and necessary contextual information to support the literal or metaphorical meanings, the rest of the sentences shared a highly similar structure. For example, the parts of speech of the words immediately following the metaphorical expressions were the same, and the same words were utilised when the context allowed; when different words needed to be used to fit different contexts, the lengths of words were controlled. An example of a pair of test sentences with an illustration of their structure is shown in Table 4.2.

---

<sup>4</sup> *Spill-over effect* refers to the phenomenon of the participant showing hesitation after the display of the target items.

Table 4.2 A pair of sample test sentences and their structure

The	famous	doctor	built	a	balcony	for	his	home	near	the	city
The	famous	doctor	built	an	argument	in	his	article	in	the	book
W01	W02	W03	W04 <sup>5</sup>	W05	W06 <sup>6</sup>	W07	W08	W09	W10	W11	W12
Segment 01			Segment 02			Segment 03			Segment 04		
Subject			Literal/metaphorical expression			Spill-over area			Wrapping-up area		

### 4.3 Main experimental instruments

#### 4.3.1 Acceptability judgment task (AJT)

All the test sentences in the acceptability judgment task were presented with an 11-point Likert scale of acceptability (henceforth ‘the score scale’), a blank for sentence correction and a 5-point Likert scale of confidence (henceforth ‘the confidence scale’). An example of the presentation of a test sentence is given below in Figure 4.1. For each sentence, participants were asked to give a ‘score for the sentence’ on the score scale, to indicate how confident they were when they gave the score on the confidence scale, and when appropriate, to provide a sentence showing corrections in the blank space provided. The range of the score scale varied from 0 to 10, resembling the typical scoring system in secondary schools in China, so naïve participants could easily understand the concept of acceptability judgment by converting ‘the degree of acceptability’ to ‘the score that the sentence should receive’.

<sup>5</sup> W04 was the position of the main verb in all the sentences. In a metaphorical or literal sentence, W04 was the position in which a critical lexical item appeared, either metaphorically or literally.

<sup>6</sup> W06 was the disambiguation position. Upon reading as far as that part, a participant would know whether W04 was metaphorical or literal.

40. My mother said that these books would be food for thought.

0	1	2	3	4	5	6	7	8	9	10	
Not at all confident			Not very confident			Neutral		Confident		Very confident	

41. The mother held a belly of gas because her son failed in the exam.

0	1	2	3	4	5	6	7	8	9	10	
Not at all confident			Not very confident			Neutral		Confident		Very confident	

Figure 4.1 An excerpt from the acceptability judgment task

A brief instruction with examples was provided in the introductory part of the acceptability judgment task: participants were asked to imagine that they were marking the compositions for their classmates, and they should give a score to each sentence. It was also explicitly stated in the instructions that the participants were not allowed to refer to any sort of dictionary, and they should follow their intuition for the entire task. The instruction aimed to help secondary school students who participated in the experiment to understand their task in the experiment, as well as to ease any anxiety or fear of authority when they received materials from the experimenter. Although this setting might guide some participants to judge the sentences in a prescriptive way, they were encouraged to utilise their intuition for in their judgments.

Since this experiment was conducted as a paper-and-pen test, it was difficult to use an automated randomised questionnaire. Therefore, two versions of the AJT questionnaire with two different sequences of test sentences were provided to avoid any sequencing effect. Each participant was assigned a version randomly.

#### 4.3.2 Self-paced reading task (SPRT)

All the test sentences in the self-paced reading task were divided evenly into two counterbalanced lists, in such a way that the metaphorical use and the literal use of a same critical lexical item would not appear in the same list. The experiment was programmed using E-Prime 2.0 Standard version, with the sequence of the two sentence lists randomised, and the sentences within each list also randomised. Using such an arrangement, for a single critical verb, half of the participants would see the literal use first, and the other half would see the

metaphorical use first. This was also intended to make sure any potential priming effect would be eliminated in the modelling.

In the experiment, the sentences were presented in a word-by-word fashion in a moving window paradigm. A simplified illustration of the self-paced reading paradigm is shown in Figure 4.2 below. At the beginning of each sentence, the participant would see a fixation marker ‘+’ on the left of the laptop screen, which would indicate the position of the first word. After the spacebar was pressed, the fixation marker would disappear and the first word (‘The’ in every sentence) would appear at the position of the fixation marker. As the participant pressed the spacebar, the next word would appear immediately adjacent to the previous word, and the previous word would disappear. The participant could not go back to read the previous words and could only proceed to the next word. After each sentence, a comprehension question would appear, for which the participant would press the F key (with a ‘Y’ sticker, for a ‘yes’ answer) or the J key (with an ‘N’ sticker, for a ‘no’ answer) on the laptop keyboard to answer. The use of comprehension questions was intended (1) to draw the participant’s attention to the meaning of the sentences and prevent keys being pressed randomly; and (2) to detect whether the participant had understood the metaphorical expressions correctly. Prior to the main reading task, the participant was given a trial session with 8 sentences to ensure familiarity with the operation and flow of the reading programme. In the middle of the experiment a short break was designed for the participant.

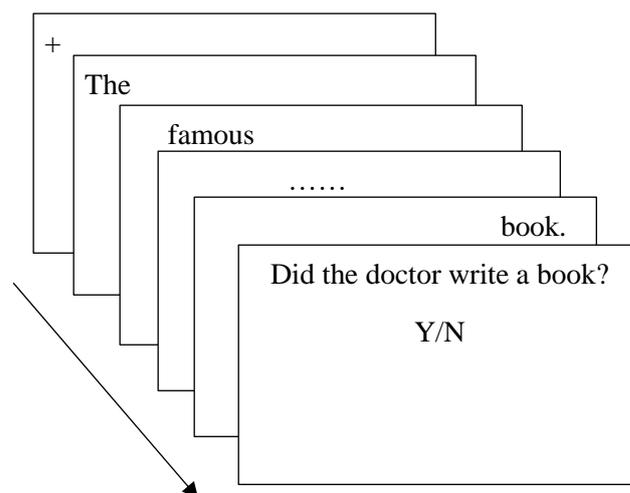


Figure 4.2 A simplified illustration of the self-paced reading task in the experiment

## 4.4 Supplementary experimental instruments

### 4.4.1 Lexical evaluation survey

The lexical evaluation survey was designed to provide a quantification of an expression's degree of metaphoricalness, and degree of markedness, as well as its relevant frequency in daily language use, and to investigate whether these factors were interrelated. The survey was inspired by Jordens and Kellerman's (1981) sorting of the meanings of 'breken' prior to investigation of meaning transferability. As discussed in Section 2.3.2, Jordens and Kellerman reduced the dimensions of word meanings to two scales, namely concreteness and coreness, and it was argued that the latter of these best represented the degree of markedness of a meaning of a polysemous word. The same dimensions of word meanings were adopted in the current thesis for the quantification of markedness, but at the same time the methodology was largely simplified.

The lexical evaluation survey was created online on Qualtrics in both Chinese and English. All the lexical items in the survey came from the acceptability judgment task and the self-paced reading task. The Chinese version contained the Chinese translation of the test sentences containing the LB, MB, LS and MS critical lexical items. Similarly, the English version contained the original test sentences containing the LB, MB, LT and MT critical lexical items. If a critical lexical item appeared in both the AJT and the SPRT, the more natural sentence (usually from the AJT) was selected. All the critical lexical items were underlined in the sentences. Each sentence was presented with three 21-point Likert scales from -10 to 10 in the form of interactive sliders: concrete/abstract, corresponding to the degree of metaphoricalness; core/peripheral, corresponding to the degree of markedness; and frequent/rare, corresponding to the degree of frequency in daily language use. An example of the layout of the survey is presented in Figure 4.3.

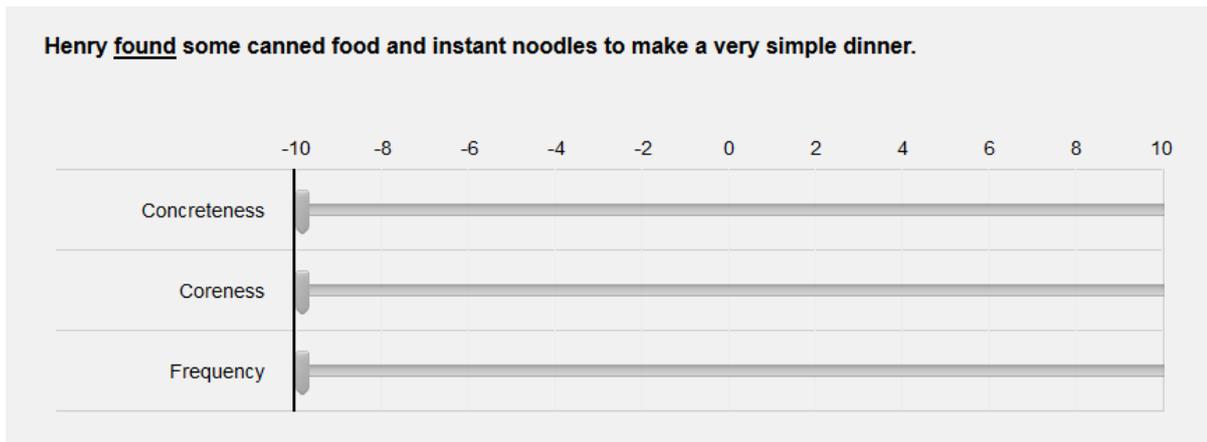


Figure 4.3 An excerpt from the online lexical evaluation survey

The lexical evaluation survey was conducted separately from the main experiments, and none of the participants in the survey declared that they had participated in the main experiments. Participants were given either the Chinese or English version, depending on their L1. In the instructions, they were asked to read each sentence in the survey, to understand the meaning of the underlined word, and to drag the sliders to the positions which they believed would most faithfully reflect the degree of concreteness/markedness/frequency of that meaning of the word.

Although, as mentioned in Section 2.3.2, the assessed degree of markedness of a meaning of a lexical item might vary between individuals, and it is possible that the participants in the lexical evaluation survey perceived the markedness of meanings differently from the participants in the main experiments, it was still decided not to ask the participants in the main experiments to complete the survey. This was to minimise the influence of the survey on the participants in the main experiments, because the appearance of the same test sentences together with the definition of ‘concreteness’ might have revealed the purpose of the main experiments and have had an impact on the participants’ performance.

#### 4.4.2 Linguistic background survey and linguistic experience questionnaire

All the Chinese learners of English participating in either of the experiments were asked to complete a linguistic background survey. The linguistic background survey mainly investigated participants’ experience of English learning, including the length of time of English learning, age of first exposure, weekly exposure to English, history of any standardised proficiency test (e.g. TOEFL, IELTS, etc.) and overseas experience in an English-speaking country. Apart from

that, basic information on each participant was also included, e.g. age, gender and university major (if applicable).

All the learner participants in the acceptability judgment task were also asked to complete a general language experience questionnaire, as a wrap-up survey for the whole task. It included questions about the participants' feeling towards English learning in general, and difficulties they had met in the acceptability judgment task. It also included questions about participants' additional linguistic exposure, that is, their learning experience and regular exposure to languages other than English. The purpose of this was to examine whether additional linguistic exposure might influence a learner's perception of psychotypology (see Xia 2017 for relevant study).

#### 4.4.3 Proficiency test

A version of the Oxford Quick Placement Test series (UCLES 2001; henceforth OQPT) was used to evaluate all learner participants' general English proficiency level; native English participants did not receive this test. The complete placement test contained 60 multiple choice questions, including 25 questions distributed over 5 cloze tests and 35 sentence completion questions. It was estimated that an advanced learner would spend less than 30 minutes completing the test, while an intermediate learner might spend around 40 minutes.

Since the OQPT was designed to provide a quick and relatively accurate estimation of a learner's general English proficiency by way of a paper-and-pen test, the questions in it mainly focus on the accurate use of (1) standard written English grammar and (2) vocabulary and common collocations. None of the questions specifically addresses metaphorical meanings of lexical items in English. Therefore, the result of the OQPT can be seen as an independent indicator of a participant's general knowledge of English, and thus can be contrasted with the focus of the current study, namely specific knowledge of metaphorical expressions.

The instructions for the OQPT series recommended that the entire test should be divided into two parts and the second part of the test should be given to the more advanced test-takers selectively after the first part was completed and marked. However, due to restrictions on the venue and time slot, both parts of the test were presented to the participants at the same time in the experiments, and they all completed the entire task at once. Although this did not fully comply with the instructions for the OQPT series, it could be regarded as an acceptable compromise between accuracy of results and temporal cost for the test. Considering that the focus of the experiment was on the AJT and the SPRT, it made it possible to keep the

proficiency test relatively simple for the participants, so that their time could be saved for the main tasks. Participants were advised to finish the test without spending too much time on it, and to leave questions blank if they did not understand.

#### 4.4.4 Psychotypology survey

The proposal of Jordens and Kellerman (1981) is considered to be one of the first discussions on the influence of psychotypology on cross-linguistic influence, and the influence of their work has lasted for more than three decades, and the influence of psychotypology is still widely investigated today. However, there are currently few experimental methods that have been created or adopted to quantify the psychotypological distance between languages among learners. Several ongoing studies on the acquisition of an additional European language by multilinguals who have already mastered several European languages (e.g. Neuser 2016 on the acquisition of English by multilingual speakers of Luxembourgish, German and French; Suhonen 2016 on the acquisition of English by bilingual speakers of Swedish and Finnish) make use of a series of Likert scales to measure the perceived similarity of different elements in several languages, including phonetic and phonological features, morphological and syntactic features, lexical features, orthographic features and general impressions. However, this paradigm may not be applicable to Chinese learners of English if they are asked to estimate the psychotypological distance between Chinese and English. There are several reasons for this: (1) the lack of multilingual exposure of these participants, especially those residing in China, may create difficulty if they are asked to compare several languages; and (2) the different orthographic forms of Chinese and English may lead the participants more to be more reluctant to discover the possible features shared by Chinese and English in terms of phonology, morphosyntax and semantics. Therefore, other possible measurements of psychotypological distance, and possibly some more impressionistic approaches, were developed in this study in order (1) to measure the psychotypological distance between languages inside and outside Europe; and (2) to capture possible individual differences in the perception of language distance.

This study has therefore made use of magnitude estimation (c.f. Bard and others 1996) to assess the psychotypological distance between pairs of languages as perceived by participants. Magnitude estimation is a psychological paradigm that is widely used in psychological studies on sense, perception and grammaticality judgment in syntax. Its mechanism works by guiding participants in estimating the perceived degree of one feature of a target object (e.g. the loudness

of a sound, the brightness of a shade, the grammaticality of a sentence, the appropriateness of a sentence in a context, etc.), using a given example as a reference. Such a mechanism matches the estimation of perceived psychotypological distance between two languages perfectly, since, as discussed in Section 2.3.1, psychotypological distance is a belief held by individuals rather than a fixed feature of a language pair, and it is likely to vary among participants. These properties of psychotypology happen to be the typical features of psychological perception that can be captured by magnitude estimation. The method is also more accurate than other similar paradigms, such as the use of a single or a series of Likert scales to give a general impression, or the use of ranking, since it makes it possible to indicate the delicate differences between two language pairs that cannot be captured by ordinal data.

All learner participants were required to finish the psychotypology survey before the main experiments. The survey in the study followed the conventional regulation of magnitude estimation. It was created online on a Chinese survey website, and each of the questions and instructions was displayed on a separate page. At the beginning of the survey, participants were introduced to the fundamental mechanism of magnitude estimation by conversion of the length of given lines to numbers with reference to a standard, reference line. They were then instructed to perform similar tasks on languages and ‘estimate the distance’ between Chinese and other languages using numbers. After two standard examples and six trial questions, the participants proceeded to the main session, in which they needed to ‘estimate the distances’ between Chinese and eleven languages that are well-known to Chinese people, including English, Tibetan, Japanese, Korean, Mongolian, Vietnamese, Thai, German, French, Spanish and Arabic. The focus of the survey was English, and other languages were used to provide a more generalised picture of a participant’s psychotypological perception. Among these languages, some were geographically close to Chinese (e.g. Korean and Thai), some orthographically close (e.g. Japanese), some typologically close (Tibetan) and some totally irrelevant typologically or geographically. The variety of languages was intended to lead participants to think about different aspects, either linguistic or extra-linguistic, of a given language, and provide a more impressionistic judgment. They were also given opportunities to justify their answers after estimating the distance between each pair of languages. At no point in the survey could the participants return to a previous question once they proceeded to the next one; thus, they could not modify their answers. These answers will be included in the final estimation of individual psychotypological distances.

#### 4.4.5 Backward digit span test

All participants in the self-paced reading experiment were instructed to complete a visual backward digit span task on an iPad before the experiment. They were asked to use an iOS app named ‘Digit Span’ (Fu 2012), which allowed them to complete a customised digit span test on the touchscreen. The participants were told that they needed to key in the numbers in reverse order on the screen; when they had completed one string correctly, the next string would be a digit longer, and if they made a mistake the next string would be a digit shorter. The app began its first cycle by displaying a 3-digit string and would jump to the next cycle if two mistakes were made consecutively. Each participant was given four cycles to complete, and the average backward digit span was calculated by the app automatically after a participant completed all the cycles. The backward digit span test was intended to reduce the influence of working memory capacity on the self-paced reading test, and the result of this test was considered as a random factor in the final data analysis.

### 4.5 Participants

#### 4.5.1 Chinese learners of English

86 Chinese learners of English were recruited for the acceptability judgment experiment, and another 81 learners were recruited for the self-paced reading experiment. They were from four major backgrounds: (1) Grade 11 (the second year of senior high school) students in a secondary school in Beijing, China; (2) non-English/linguistic major college students in major universities in Beijing, China; (3) English or linguistic major college students in major universities in Beijing, China; and (4) overseas Chinese students living in the UK, mainly at Cambridge. The ‘college students’ here include both undergraduate and graduate students. There was no explicit restriction on other socio-linguistic background factors for participants in the acceptability judgment experiment; age restrictions were applied to the self-paced reading experiment in order to reduce the influence of age on reading and reaction time, and all the participants were aged between 16 and 36 at the time of the experiment (note that age is not thought to influence judgment of degree of acceptability).

Based on their performance in the OQPT and their linguistic background, the learner participants were placed into four groups of similar sample size in each experiment: (1) the intermediate group (henceforth IN), including learners at B1 level in the CEFR system; (2) the low-advanced group (henceforth LA), including learners at B2 level; (3) the high-advanced group (henceforth HA), including learners at both C1 and C2 levels; and (4) the high-advanced

overseas group (henceforth HO), including learners that matched the proficiency level of the HA group but had also had long term exposure to a native English-speaking environment. The linguistic backgrounds of the learner groups in the two experiments are outlined below, in Tables 4.3 and 4.4. The average digit span of the participants in the reading experiment is also included in Table 4.4.

Table 4.3 Linguistic backgrounds of the 86 learner participants in the acceptability judgment experiment

		IN	LA	HA	HO
Group size		21	26	21	18
Average OQPT score (SD)		36.81 (3.44)	44.5 (1.79)	52 (2.74)	53.18 (4.52)
Average age (SD)		18;5 (3.16)	20;2 (3.70)	22;7 (2.31)	24;2 (2.96)
Gender	Male	10	9	7	6
	Female	11	17	14	12
Average years of English learning (SD)		11.86 (2.80)	11.38 (2.32)	12.81 (2.42)	15.83 (3.59)
Average age of first exposure to English (SD)		6;2 (2.15)	7;8 (3.13)	8;9 (3.14)	7;9 (2.76)
Average weekly exposure to English in hours (SD)		13.17 (17.03)	15.48 (13.79)	25 (19.89)	46 (25.50)
Number of participants in English/linguistics major		0	9	13	0
Number of participants with experience of standardised English tests		0	3	7	15
Number of participants with overseas experience		0	1 <sup>7</sup>	4 <sup>8</sup>	18 <sup>9</sup>

<sup>7</sup> The participant (28 years old) stated that she had been in the US for 36 months two years before the experiment (when she was 23). Considering that she had been living in China ever since and had not received any exposure to a native English environment since her return, she was categorised in the low advanced group according to her performance in the OQPT.

<sup>8</sup> All overseas experiences recorded had taken place at least one year before the experiment. Average length of stay was 3.5 months (SD = 3.79). Average age of first overseas experience was 16;3 years old (SD = 8.10).

<sup>9</sup> Average length of stay was 32.11 months (SD = 30.62). Average age of first overseas experience was 20;10 years old (SD = 3.70).

Table 4.4 Linguistic backgrounds of the 81 learner participants in the self-paced reading experiment

		IN	LA	HA	HO
Group size		21	22	20	18
Average OQPT score (SD)		35.47 (3.59)	44.27 (1.80)	51.45 (3.0)	51.44 (2.36)
Average age (SD)		22;2 (2.85)	21;3 (2.62)	23;2 (2.28)	23;5 (2.18)
Gender	Male	6	4	2	5
	Female	15	18	18	13
Average years of English learning (SD)		12.48 (2.34)	12.77 (2.16)	13.4 (2.85)	15.44 (3.13)
Average age of first exposure to English (SD)		9;3 (2.47)	8;2 (3.34)	9;2 (2.32)	7;1 (2.97)
Average weekly exposure to English in hours (SD)		9.07 (6.00)	12.86 (12.71)	27.95 (24.30)	49.17 (24.55)
Number of participants in English/linguistics major		7	12	14	2
Number of participants with experience of standardised English tests		2	7	8	18
Number of participants with overseas experience		1 <sup>10</sup>	1 <sup>11</sup>	4 <sup>12</sup>	18 <sup>13</sup>
Average backward digit span (SD)		6.71 (1.89)	6.89 (1.53)	6.7 (1.67)	7.47 (1.58)

Since two different groups of participants were tested in the two experiments, it was necessary to ensure that the two groups were indeed homogenous so that the results of the two experiments could be compared, even though the criteria for selection were identical. Therefore,

<sup>10</sup> The participant (27 years old) stated that he went to the UK for 1 month when he was 22.

<sup>11</sup> The participant (23 years old) stated that she had been in the UK for 12 months four years before the experiment (when she was 19). Considering that she had been living in China ever since and had not received any exposure to a native English environment since her return, she was categorised in the low advanced group according to her performance in the OQPT.

<sup>12</sup> All overseas experiences recorded had taken place at least one year before the experiment. Average length of stay was 4.75 months (SD = 5.68). Average age of first overseas experience was 20;9 years old (SD = 3.77).

<sup>13</sup> Average length of stay was 37.78 months (SD = 29.30). Average age of first overseas experience was 19;3 years old (SD = 3.04).

the proficiency of the participants in the acceptability judgment experiment (AJT) and the self-paced reading experiment (SPRT) were compared, and the results are displayed in Table 4.5. It can be seen that the participants in the two experiments did not show significant differences in terms of their sample size, sample distribution or performance in the OQPT. This indicates that the results of the two experiments can be compared, thanks to the homogeneity of the participants. By taking a combination of the two we can obtain a full picture of the processing of metaphorical expressions by Chinese learners of English in both offline and online tasks.

Table 4.5 Comparison between the participants in the acceptability judgment task and those in the self-paced reading experiment

Group	Experiment	Sample size	Average OQPT score (SD)	Mann-Whitney U test result
Intermediate	AJT (1)	21	36.81 (3.44)	U=142.43, p=0.22
	SPRT (2)	21	35.47 (3.59)	
Low-advanced	AJT (1)	26	44.5 (1.79)	U=192.29, p=0.66
	SPRT (2)	22	44.27 (1.80)	
High-advanced	AJT (1)	21	52 (2.74)	U=135.09, p=0.45
	SPRT (2)	20	51.45 (3.0)	
High-advanced overseas	AJT (1)	18	53.06 (4.41)	U=99.99, p=0.24
	SPRT (2)	18	51.44 (2.36)	

#### 4.5.2 Native speakers of English

There were 24 native British English speakers who were recruited for the acceptability judgment experiment, and another 21 native speakers were recruited for the self-paced reading experiment. All native speaker participants declared that they had no prior knowledge of Chinese at the time of the experiment. There was no explicit restriction on other socio-linguistic backgrounds for participants in the acceptability judgment experiment; an age restriction was applied to the self-paced reading experiment in order to reduce the influence of age on reading and reaction time, and all the participants were aged between 18 and 36 at the time of the experiment. The average age of the participants in the self-paced reading task was 21;1 years old (SD=2.23) and the average backward digit span was 5.48 digits (SD=1.17).

## 4.6 Procedures

### 4.6.1 Acceptability judgment experiment

All participants, except for those recruited in a secondary school, were recruited via an online advertisement; in the advertisement, they were told to send an email to the experimenter to express their interest. When the experimenter replied to the emails from the prospective participants, she assigned a participant number to each of them and attached the consent form and information about the experiment for their reference. After they had confirmed that they would like to participate in the experiment and signed the consent form, they were sent the proficiency test (learners only), the language background and linguistic experience survey (learners only) and the acceptability judgment task (learners and native speakers), all in the Microsoft Word Document (.docx) format, as well as a link to the online psychotypological survey (learners only). In each part of the experiment they were asked to enter their participant number and complete the task, and they were required to complete the tasks entirely on their own. After completion, all the .docx questionnaires were sent back to the experimenter via email, and a small remuneration was sent to the participants after the experimenter had confirmed that all parts of the experiment had been fully completed.

The learner participants in secondary school received the oral advertisement from the experimenter who was at that time accompanied by the head teacher, and after the advertisement they contacted the experimenter privately and voluntarily. Once the participants had confirmed that they would like to participate in the experiment and signed the consent form, they were organised together in a classroom and given all the paper materials, including the proficiency test, the language background and linguistic experience survey and the acceptability judgment task. The entire process was supervised by the experimenter who was accompanied by one of their senior teachers who also taught them English. When the participants had completed the paper questionnaires, they handed them back to the experimenter and took the online psychotypological survey under the supervision of the experimenter. The experimenter occasionally provided guidance on the trial session of the psychotypological survey, since it might be difficult for secondary school students to convert the distance between languages to numerical scales. After completion, all the participants received a small souvenir from the experimenter.

### 4.6.2 Reading experiment

The recruitment procedure for the reading experiment was the same as that for the acceptability judgment experiment. After the participants had confirmed that they would like

to participate in the experiment and had signed the consent form, they were asked to confirm the time of the lab session for the reading experiment. The learner participants, except for those from secondary school were sent the language background survey and the proficiency test in the Microsoft Word Document (.docx) format as well as a link to the online psychotypological survey and were asked to complete these two parts before the reading experiment. The participants from secondary school received hard copies of the language background survey and the proficiency test and were required to complete them before the reading experiment. After they had handed in the paper questionnaires, they completed the online psychotypological survey under the supervision of the experimenter, with the experimenter occasionally providing guidance.

The lab sessions for the reading experiment were conducted in a quiet classroom in Beijing and in the psycholinguistic lab in Cambridge to avoid any possible distraction. All the participants, including the learners and native speakers, underwent the same procedure in the lab session. The experimenter verified the identity of each participant to ensure the online questionnaires had been completed before the experiment, and then asked them to complete the backward digit span test. Then the experimenter briefly introduced the task to the participants, and they began the self-paced reading task. In the trial session, the experimenter supervised the participants and answered any questions regarding the operation of the programme, after which they were left alone to accomplish the main session and the experimenter only observed silently in the lab/classroom. The participants could choose whether to take a break in the middle. The experiment was completed when the participant had finished the self-paced reading task. After completion, all the participants received a small remuneration or a small souvenir from the experimenter.

## 4.7 Data adjustment and processing

### 4.7.1 Acceptability judgment score

Several problems emerged in the analysis of the raw data from the AJT, and there were two prominent ones that might influence the results of the experiment significantly: one was related to the definition and scope of ‘acceptability’ in the experiment, and the other to the variability of acceptability judgment scores between individual participants as well as between participant groups. The methods of data adjustment reported here, therefore, are intended to minimise the influence of the two problems on the results and the interpretation of the experimental data in the acceptability judgment task.

In the acceptability judgment experiment, when participants were asked to give scores regarding their ‘peers’ writing’, the item that was rated was not a single literal or metaphorical expression, but a full sentence in which the critical lexical item was included as a constituent. Using full sentences rather than individual phrases as the test item in the acceptability judgment task brought both advantages and shortcomings to the investigation of metaphorical expressions. On the one hand, a full sentence could provide the context for a metaphorical expression and help the participant to understand the intended meaning of the critical lexical item; on the other hand, a full sentence might also disguise the designated target of the experiment, namely the acceptability of the included metaphorical expression. It was possible for the learner to choose to accept the metaphorical expression in the test sentence but rate the whole sentence unacceptable because some other element in the sentence was believed unacceptable, such as tense/aspect or the use of articles.

It was therefore crucial to ensure the acceptability of metaphorical expressions was what was measured in the test. Thus, after data collection, all feedback sentences were coded and sorted. If a feedback sentence showed irrelevance to the metaphorical expression or the designed grammatical error included in the sentence, for instance, the participant focused on the tense of the sentence, or ‘corrected’ an acceptable literal counterpart or a grammatical filler, then the score for the sentence given by that participants was excluded from the data set. On average, 9.32% (450 out of 5280) of the original data points were discarded; around 10.59% (112 out of 1252) of the scores provided by native speakers were excluded from the analysis, while this percentage ranged between 5.56% (48 out of 864) and 10.52% (106 out of 1008) among different learner groups. It should be noted that some participants decided not to provide any feedback sentences and all the scores given in their answers were therefore preserved in the analysis, even if it was possible that some of these data included low scores that were irrelevant to the acceptability of the metaphorical expressions themselves. In sum, the adjustment of the AJT data was only able to maximise the reliability of the data in the current situation, while noise from judgments of other non-critical constituents in the sentence could not be fully eliminated due to the nature of the task.

A second problem with the acceptability judgment task was that participants’ preference varied when they evaluated the acceptability of sentences on the 11-point Likert scale. Such variety was observed both individually within a proficiency group and collectively between different proficiency groups. Although sample sentences and scales were provided in the instructions to show that it was recommended that use be made of the full scale (from 0 to 10) to indicate the degree of acceptability, some participants in fact made use of only part of the

scale, namely the second half (from 5 to 10), to measure the acceptability of sentences. Such discrepancies in the use and interpretation of the acceptability scale might lead to an inaccurate representation of the data when the acceptability scores are analysed collectively.

Variation in the use of the acceptability scale was also found in between-group comparisons, and this was even more profound than the individual differences within a group and might influence the analysis even more significantly. Raw scores for the degree of acceptability of each category of test items are shown below in Figure 4.4. It can be seen that both the low-advanced group and the high-advanced overseas group seemed to use the scale in a more conservative way, tending to give lower scores to all sentences; otherwise, it would be difficult to explain why advanced participants seemed to accept literal expressions less than the intermediate learners, and why overseas learners seemed to accept literal expressions less than the high-advanced group, given that the two groups had achieved the same level of proficiency in general. These results, if directly processed without adjustment, could be easily interpreted as meaning that less advanced learners show a significantly higher degree of acceptance of the literal expressions and shared metaphorical expressions than more advanced learners, which was counterintuitive and could potentially lead to false interpretations.

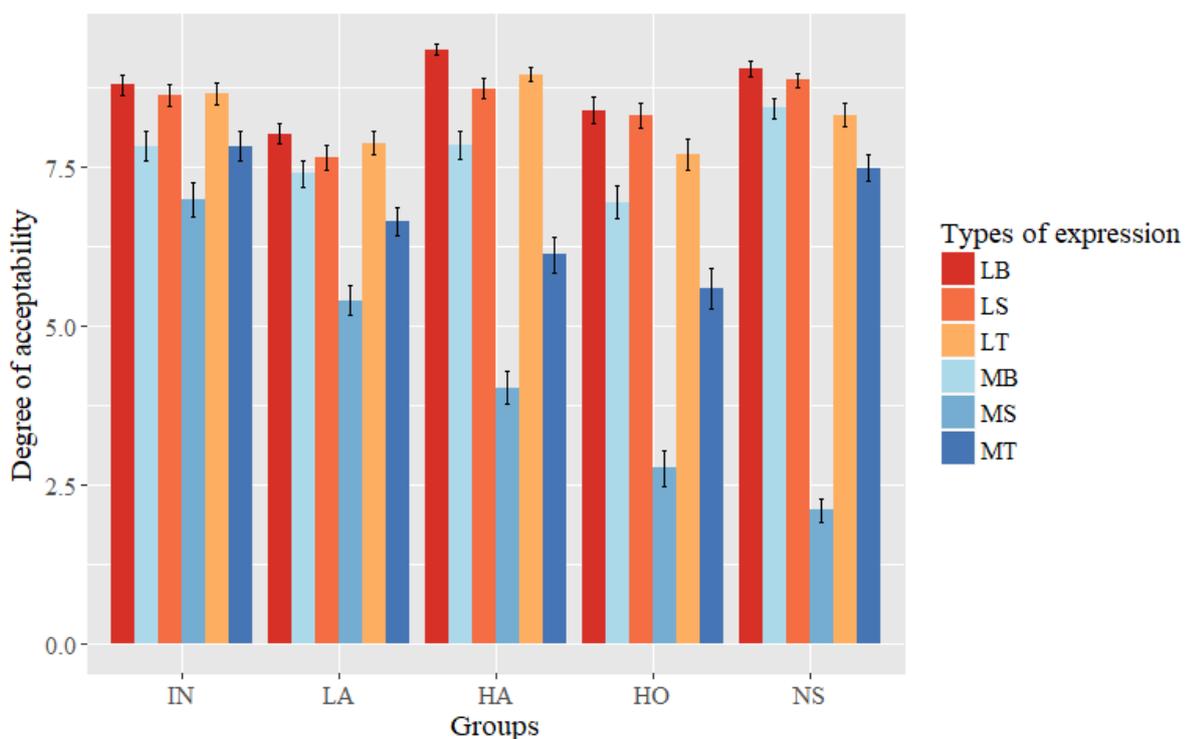


Figure 4.4 Raw acceptability scores for different categories for five groups of participants

To deal with the in-group and between-group differences, mixed-effect linear models were adopted as the main statistical tool in the analysis of the results of the acceptability judgment experiment, with individual participants being considered one of the random factors (i.e. noise factors), and individual test items being considered another random factor to avoid the problem of ‘language as a fixed effect’. At the same time, a data adjustment based on mixed-effect linear modelling was also conducted to reduce between-group differences. Within each proficiency group, a group intercept was calculated using a mixed model with three random factors but no fixed factors. The three random factors were ‘individual participants’, ‘individual critical lexical items’ and ‘individual confidence level’. The output of such a model could be seen as the group intercept of degree of acceptability of all the critical test sentences, in which the noise between individual participants and the noise between individual lexical items were maximally reduced. After that, a delta score (henceforth ‘the acceptability score’) for each valid judgment in each group was calculated by subtracting the ‘group intercept’ from the raw score. The process of acceptability score calculation could be regarded as one involving removing the lower parts of the raw scores while only maintaining the upper parts of differences. After data adjustment, a positive acceptability score indicated that a participant had accepted the given test sentence only when being compared with (1) other participants judging the same test sentence and (2) the same participants judging other test sentences. Similarly, a negative acceptability score indicated that that participant had rejected a given test sentence when comparison was made across participants or across test sentences. Previous methodological discussions of acceptability judgment show that a participant does not simply ‘accept or reject’ a sentence, i.e. the degree of acceptability of a sentence is not binary. Therefore, when interpreting the results of the acceptability judgment task, it would be problematic if we simply defined a positive acceptability score as definite acceptance. Only by comparison between groups and between categories can we see whether a group of participants shows acceptance of a category of expressions. The distribution of the acceptability scores across proficiency groups is shown in Figure 4.5. It is clearly shown that (1) the distribution of the acceptability scores resembled the upper parts of the raw scores, and (2) the problem of the higher degree of acceptance among less proficient learners disappeared.

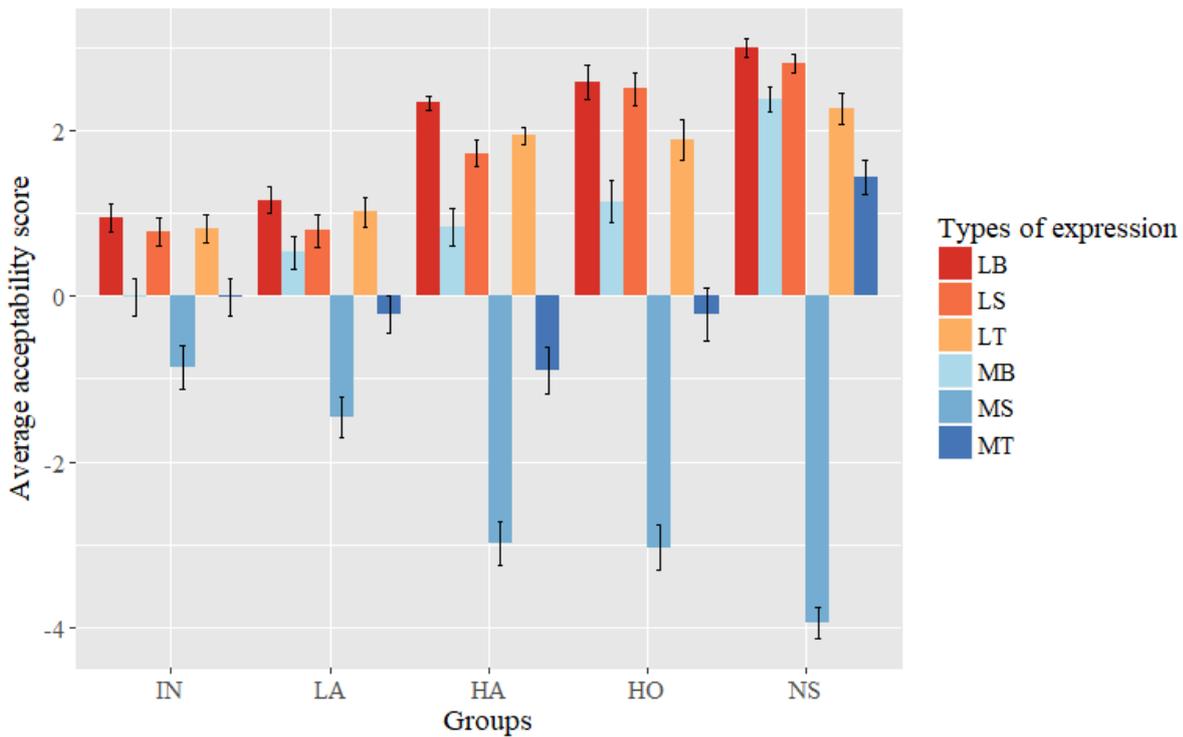


Figure 4.5 Acceptability scores of different categories of test sentence among five groups of participants

#### 4.7.2 The categorisation and coding of feedback sentences

In the sentence correction components, participants were instructed to provide a feedback sentence when they spotted an unacceptable element in the test item. The stated purpose of the feedback sentence in the instruction was to ‘help your peers to understand their problems and improve their English writing’; the intended purposes, however, were (1) to examine whether participants identified the problems in the sentences as designed by the experimenter; (2) to investigate how participants interpreted the metaphorical expressions in the sentences if they chose to paraphrase the given sentences; (3) to examine the production of alternative metaphorical expressions by comparing the feedback sentences provided by a learner and a native speaker; and (4) to survey the possible strategies learners would use when they regarded transfer from L1 to L2 as impossible. The selection, categorisation and coding of feedback sentences before the final analysis was primarily based on the four intended purposes of the feedback sentences.

In total, 2658 feedback sentences were collected across the five groups; this included correction of ungrammatical fillers, as well as paraphrasing of metaphorical expressions and other possible adjustments. It was observed that some high-advanced and native participants

provided corrections to grammatical fillers and literal counterparts that were regarded as acceptable by other native speakers. As shown in Table 4.6, more proficient participants were more willing to provide feedback sentences, regardless of the acceptability of the given test sentences.

Table 4.6 Number of feedback sentences provided by each group of participants

	IN	LA	HA	HO	NS	Total
Total sentences	323	411	617	486	821	2658
Sentences per person	15.38	15.81	29.38	27	34.21	24.16

In order to be included in the quantitative and qualitative analyses, a ‘relevant feedback sentence’ needed to meet several criteria. Correction in a relevant sentence should target the use of the metaphorical expression in the test item; thus, corrections of fillers and all literal sentences were excluded, and corrections related to other grammatical ‘problems’ in the metaphorical sentences, such as tense and aspect, were also excluded. When a participant explicitly pointed out a problematic constituent (e.g. a misused L1-specific metaphorical expression) but did not provide any sentences, the response was counted, but categorised as ‘error indicated without correction’. The distribution of these relevant sentences is listed in Table 4.7.

Table 4.7 Distribution of relevant feedback sentences provided by each group of participants across different metaphorical conditions

	IN	LA	HA	HO	NS	Total
MB feedback sentences	6	6	15	7	6	40
MS feedback sentences	20	40	86	56	118	320
MT feedback sentences	17	28	44	29	43	161
Total relevant feedback sentences	43	74	145	92	67	521
Relevant feedback sentences per person	2.05	3.19	6.90	5.11	6.96	4.74

Among all 2658 feedback sentences, 521 were relevant feedback sentences; on average, each participant provided 4.74 sentences in this experiment. In general, more proficient

participants were more willing to provide feedback sentences targeting metaphorical expressions, regardless of the availability of the expressions in English. The percentages of MS feedback sentences among total relevant feedback sentences were 46.52% in the IN group, 54.05% in the LA group, 59.31% in the HA group, 60.87% in the HO group and 70.66% in the NS group. Considering that the MS expressions were incorrect in English and should be corrected in the task, we can draw two interim conclusions from the distribution of MS feedback sentences: (1) more proficient participants tend to correct more L1-specific metaphorical expressions than less proficient participants on average; and (2) the more proficient the participants are, the more output they tend to provide.

The distribution pattern of relevant feedback sentences among different proficiency groups is also calculated, that is, whether different groups of participants provided similar percentages of relevant feedback sentences to the three metaphorical categories. There was no significant difference among the four learner groups, which meant that the learner participants performed in a qualitatively and quantitatively similar way when they provided relevant feedback sentences for different types of metaphorical expression, regardless of their proficiency. However, the pattern adopted by the learner groups was significantly different from that of the native speaker group (Pearson's  $\chi^2=10.99$ ,  $p=0.004$ ). The native speakers provided significantly fewer relevant feedback sentences for the MB and MT expressions that were acceptable in English.

#### 4.7.3 Coding and classification of level of confidence

The level of confidence was recorded using a 5-point Likert scale with each participant providing a judgment for each test item. Each participant was asked to choose from among five levels of confidence: *not at all confident*, *not very confident*, *neutral*, *confident* and *very confident*. Before the data were analysed, the five levels were transformed into numerical confidence scores from 1 to 5, with 'not at all confident' being assigned 1 and 'very confident' 5. A higher confidence score therefore meant that the participant was more confident in a judgment, and vice versa.

Although 'neutral' was given as a possible level of confidence and was intended to be used in a situation in which participants were neither confident nor unconfident, it was revealed in some informal post-experiment interviews that most participants interpreted 'neutral' as an indicator of 'lack of confidence'. Based on this piece of feedback, classification of level of confidence in the analysis used the boundary between 3 and 4 as the cut-off point: if a

participant gave a confidence score of 4 or 5, that data point was coded as ‘confident’; otherwise, the data point was coded as ‘not confident’. This coding system could be combined with AJT scores and used to analyse possible traces of cross-linguistic influence, as will be discussed in Section 5.3.

#### 4.7.4 Reading time for sentences

In the experiment, three types of data were recorded automatically by the programme: the reading time for each word, the reaction time for the comprehension questions (including reading time and answering time), and the score for the answer to the comprehension questions (0 for incorrect and 1 for correct). As a type of reaction time data, reading time data were treated in a two-step trimming method, which is the conventional method of reaction time data normalisation. The first trim was to remove the outliers in terms of reading time. A reading time of less than 100ms was discarded because it might result from an accidental press of the spacebar. For English native speakers, a reading time above 2000ms was trimmed to 2000ms; for learner participants, the threshold of trimming was 4000ms (see Papadopoulou and others 2013 for a summary of trimming criteria). The second trim was intended to ‘regularise’ the data by setting an upper boundary for all reading time data for an individual participant. This upper boundary for a set of reading time data was the average of all the reading times plus two standard deviations ( $\text{mean}+2\text{SD}$ ); all the reading times above the upper boundary were trimmed to fit within the upper boundary.

After this two-step trimming modification, the residual reading time for each word for each participant was calculated. This residual time was used in the data analysis to eliminate the effect of word length on the reading time, since the self-paced reading task relied on visual perception, with the number of letters in a word possibly affecting the reading time for each word. To calculate the residual reading time, the expected reading time was first calculated using a linear model to compare the actual reading time and the length of each word in letters. After that, the residual reading time was calculated by subtracting the expected reading time from the actual reading time. A positive residual reading time means that a participant spent more time reading the word than was expected, while a negative residual reading time means that less time was spent than expected.

Finally, the residual reading time was analysed in a segment-by-segment manner, as indicated in Table 4.3 before; that is, the total residual reading times for every three words were compared in the analysis. Although the sentences were displayed in a word-by-word fashion in

the experiment, the metaphorical expressions used in the sentences were not individual words but three-word phrases, and the reading times for the metaphorical expressions could be best represented by the total residual reading time. Moreover, the spill-over effect can continuously affect more than one word after the critical region (Jegerski 2013), and it was more precise to analyse the reading time of a larger segment for the spill-over effect. Therefore, two large segments were the focuses of the data analysis: the second segment of the whole sentence, which included the fourth to sixth words, and reflected the reading time for the whole metaphorical expression; and the third segment, i.e. the seventh to the ninth words, which showed a potential spill-over effect, as reflected in a hesitation after the metaphorical expression.

The reaction time for the comprehension questions was treated in a similar way to the reading time for individual words. However, only the second trim in the two-step trimming method was applied to the reaction time for the questions, because there is no conventional threshold for the trimming of comprehension questions (c.f. Papadopoulou and others 2013). The upper bound of the reaction time to the questions was set to the average reading time plus two standard deviations for each participant (mean+2SD). The residual reaction time was also calculated in a similar way to the residual reading time for individual words, by subtracting the linearly expected reading time for the whole question from the actual reaction time. A positive residual reaction time, therefore, means that a participant spent more time answering the question than was expected, while a negative residual reaction time indicates that less time was spent than expected.

#### 4.7.5 Psychotypological score

The responses to the psychotypology survey were standardised using the standard treatment of magnitude estimation (see Siddharthan and Katsos 2010), namely z-score standardisation; after standardisation, the z-score of the result for the estimated psychotypological distance between Chinese and English was referred to as the *psychotypological score* for this pair of languages. The conventional way of calculating z-scores of data points is as follows:

$$z\text{-score} = \frac{\text{raw score} - \text{mean}}{SD}$$

From the formula above, it can be seen that a negative psychotypological score for the Chinese-English pair indicated that a participant believed that English was relatively close to Chinese compared to the other eleven languages, at the time of the survey. This was because

the value of ‘raw score minus mean score’ was below 0, and thus the distance between the two languages in the perception of the participant was smaller than the average distance of all the given languages from Chinese. A positive psychotypological score between Chinese and English, similarly, indicated that that participant perceived English as a relatively distant language from Chinese compared to other given languages at the time of the survey.

The psychotypological distance between Chinese and English ranged from -2.25 to 3.01 among the intermediate participants, from -1.30 to 1.90 among the low-advanced participants, from -1.49 to 1.48 among the high-advanced participants, and from -0.70 to 1.78 among the high-advanced overseas participants. The perception of psychotypological distance between the two languages seemed to be more unified among more proficient participants, while less proficient participants, especially those with limited linguistic experience, were more likely to make certain radical judgments.

Table 4.8 Distribution of psychotypological distance between Chinese and English among different groups of participants

	IN	LA	HA	HO	Total
Close ( $z < 0$ )	22	23	18	10	73
Middle ( $z = 0$ )	0	1	1	1	3
Distant ( $z > 0$ )	20	24	22	25	91
Total	42	48	41	36	167

The distribution of the perception of language distance and proficiency levels is demonstrated in Table 4.8 above. 0 is used as a borderline to classify participants’ perception of language distance. Although there seemed to be a trend for more proficient learners to more likely perceive English as a distant language from Chinese, a correlation test revealed that there was no significant correlation between the English proficiency of participants and the psychotypological distance between Chinese and English.

#### 4.7.6 Lexical property scores of test items

Three different lexical property scores for each test item (either literal or metaphorical) were computed using linear mixed models after the lexical evaluation task. The participants in the evaluation task were asked to rate all the test items along three dimensions: concreteness (the

degree of metaphoricalness), coreness (the degree of markedness) and relative frequency. The words in Chinese and in English were computed separately, but all the lexical items in one language were computed together.

The three lexical property scores for each test item are listed in Appendix 3. A positive score indicates that the use of a lexical item is more concrete (literal), or more core (unmarked), or more frequent than other uses of lexical elements as perceived in everyday use. Similarly, a negative score indicates that the use of a lexical item is more abstract (metaphorical), or more peripheral (marked), or rarer as perceived in everyday use. Since the LB and MB expressions are available in both Chinese and English, two sets of lexical property scores were calculated, each representing the Chinese or English version.

It is shown from the data that the concreteness (degree of metaphoricalness) and coreness (degree of markedness) of a particular use of a lexical item are significantly correlated with each other in both Chinese and English (Pearson's  $r=0.70$ ,  $p<0.001$  for Chinese; Pearson's  $r=0.90$ ,  $p<0.001$  for English); this indicates that individual perception is likely to connect the degree of metaphoricalness to the degree of markedness. In the current study, the influence of metaphoricalness and markedness somehow overlap: if any influence on transferability results from the markedness of a use of a lexical item, it is likely that the degree of metaphoricalness also plays a role.

#### 4.8 Ethical considerations

All the experiments, including but not limited to the research methodologies mentioned above and the recruitment of minor participants (participants between 16 and 18 years old), had been approved by the Humanities and Social Sciences Research Ethics Committee by the time of the experiment. The institutions in China where the experiments were conducted were satisfied with the ethical regulation at the University, and no additional ethical review was required in China. All the participants had received sufficient information related to the experiment to give informal oral consent. They gave written consent before their participation and participated in the experiments voluntarily.

Throughout the experimental process, all participants were identified by a random but unique participant number. All the data collected were anonymised and kept strictly confidential in a secure location that could only be accessed by the author and the supervisor. Participants cannot be identified by the data collected and were not judged based on their

performance in the experiment. Personal information on the participants was not disclosed to any person who was not involved in the study.

## 5 The judgment of metaphorical expressions

### 5.1 Introduction

This chapter reports the results of the acceptability judgment task. On the one hand, the results will reflect how learners and native speakers perceived and judged metaphorical expressions when they were given sufficient time. On the other hand, more fine-grained traces of cross-linguistic influence and their association with the major influential factors listed in Chapter 2 will be further explored.

The chapter is divided into four major parts. The first part, Section 5.2, serves as an overview of the results of the acceptability judgment task, and focuses on differences in the degree of acceptability (1) between literal and metaphorical expressions; and (2) among different groups of participants. It further seeks to assess whether there is any interaction between metaphoricalness and proficiency that influences the degree of acceptability. The second part, Section 5.3, will be devoted to an analysis of the level of confidence as rated by the participants when they made the judgments, and a joint analysis of the level of confidence and the degree of acceptability of metaphorical expressions. The third part, Section 5.4, includes a discussion of transfer detected in the results, as well as how the three influencing factors of transfer (psychotypology, markedness, learners' knowledge) may affect the degree of subjective transferability of a metaphorical expression. The last major section reports the results of the sentence correction components, in which we will closely examine how learners express the intended meanings in their own words when they do not accept the metaphorical expressions given in the experiment.

### 5.2 Degree of acceptability of literal and metaphorical expressions

The result of the acceptability judgment task is thought to reflect the outcome of the acquisition of metaphorical expressions by Chinese learners of English. Particularly, we can evaluate learners' performance (i.e. whether they acquire a metaphorical expression or not) by comparing their results of acceptability judgment with the native group's. The basic logic is that, if a metaphorical expression has been acquired by a learner, it should be accepted when it appears in an appropriate sentence. By the same logic, if a metaphorical expression is not accepted by a learner even though it is used in an acceptable way, then it has not been acquired yet. At the same time, we also set up the baseline according to native speakers' performance: native speakers would accept all the metaphorical expressions that are available in English, and

would not accept all the metaphorical expressions that are not available in English. We then could covert learners' acceptability score to the outcome of acquisition of that expression by the learners: if there is a difference in degree of acceptability of an expression between native speakers and learners, then such difference can be attributed to (1) the incomplete acquisition of that expression by learners, if the expression is available in English; or (2) possible influence from learners' native language, namely Chinese, if the expression is not available in English. When the difference in the degree of acceptability of an expression between native speakers and learners disappears, it means that the learners have acquired the expression and can provide a native-like judgment. If there is a persistent difference in the degree of acceptability of an expression between native speakers and different groups of learners, then we should assume that the acquisition of that expression remains incomplete. We can deduce that that expression might be more difficult to acquire than other expressions. It should be noted that, such inferences of incomplete acquisition are not based on the sole results of learners, but a comparison of performance between learners and native speakers. As we have discussed in Section 2.3.3, it is difficult to identify whether an expression is acquired solely depending on the acceptability scores given by learners, and we will see in later analysis that a comparison of acceptability scores between learners and native speakers, or a combination of acceptability score and other evidences, can provide a better picture of what has been acquired by learners.

A general picture of the acceptability scores of literal and metaphorical expressions is given in Figure 5.1. While this figure clearly demonstrates differences in judgment among groups of participants and differences of degree of acceptability among types of expressions, point '0' on this graph should be interpreted carefully. '0' should be seen as a reference point of indeterminacy: a score far from '0' means that participants showed a clear acceptance or rejection, while a score close to '0' may mean that (1) individual participants were not sure whether to accept the particular expression; or (2) that there was a great degree of variability among participants. Therefore, the best way to interpret the data is to conduct pairwise or group comparisons, since, as discussed in Section 4.7.1, 'acceptable' and 'unacceptable' are relative rather than definite. In the current section, I adopt such pairwise comparison and do not refer to the exact scores of each type of expression given by each group. However, after the adjustment described in Section 4.7.1, point '0' can be *roughly* seen as the boundary between acceptance and rejection.

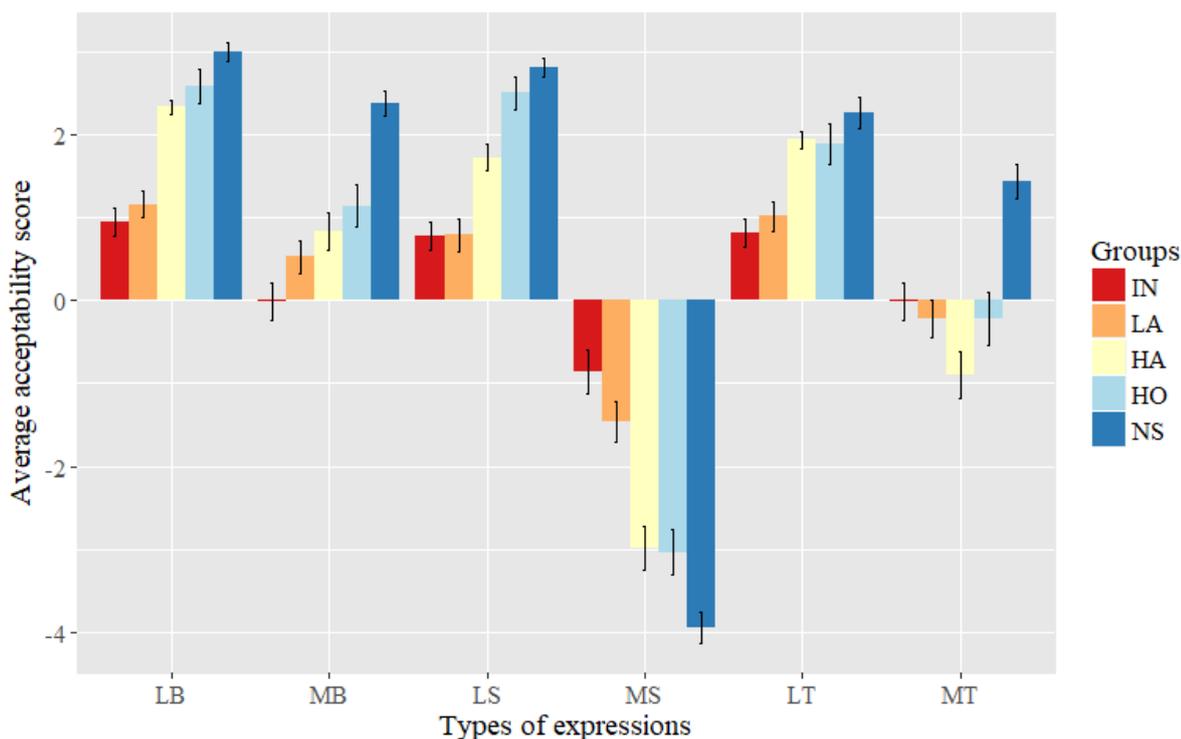


Figure 5.1 Acceptability scores for literal and metaphorical expressions by group

Overall, native speakers and learners accepted all literal expressions. Less proficient learners gave lower acceptability scores to the literal expressions compared with more proficient learners and native speakers ( $\chi^2(4)=9.56$ ,  $p=0.048$ ). Other than that, there was no significant difference observed, particularly in the case of differences of the degree of acceptability of literal expressions between more proficient learners and native speakers.

Turning to the comparison between metaphorical and literal expressions, it can be seen that all learner groups gave significantly lower acceptability scores to the MB expressions than to the LB expressions ( $\chi^2(1)=5.06$ ,  $p=0.025$  for intermediate;  $\chi^2(1)=3.86$ ,  $p=0.0496$  for low-advanced;  $\chi^2(1)=11.78$ ,  $p<0.001$  for high-advanced;  $\chi^2(1)=5.35$ ,  $p=0.021$  for overseas); such differences, however, were not seen in the native group. Compared with the native group, it has also been found that intermediate, low-advanced and high-advanced learner groups gave significantly lower acceptability scores to the MB expressions ( $\chi^2(1)=9.19$ ,  $p=0.002$  for intermediate;  $\chi^2(1)=8.24$ ,  $p=0.004$  for low-advanced;  $\chi^2(1)=11.87$ ,  $p<0.001$  for high-advanced), while the difference between native group and overseas learner group was marginal ( $\chi^2(1)=3.80$ ,  $p=0.051$ ). Although it seemed that more proficient learners generally gave higher acceptability scores to the MB expressions than less proficient groups, this difference was not significant.

The acceptability score for the LS and MS expressions clearly showed a contrast between ‘acceptable’ meanings and ‘unacceptable’ meanings, since the LS expressions are available in English but the MS expressions are not (c.f. Table 4.1). All learner groups and native speakers gave significantly lower scores to the MS expressions than to the LS expressions ( $\chi^2(1)=7.67$ ,  $p=0.005$  for intermediate;  $\chi^2(1)=10.69$ ,  $p=0.001$  for low-advanced;  $\chi^2(1)=19.30$ ,  $p<0.001$  for high-advanced;  $\chi^2(1)=21.56$ ,  $p<0.001$  for overseas). Compared to the native group, the intermediate and low-advanced groups gave significantly higher scores to the MS expressions ( $\chi^2(1)=13.94$ ,  $p<0.001$  for intermediate;  $\chi^2(1)=10.28$ ,  $p=0.001$  for low-advanced); however, the high-advanced and overseas groups did not show such significant difference from native speakers. In the case of the MS expressions, proficiency was a significant influence on the acceptability scores ( $\chi^2(3)=8.95$ ,  $p=0.030$ ): more proficient learners, such as the high-advanced and overseas groups, gave significantly lower scores to the MS expressions than less proficient learners. In particular, a distinct shift could be observed between the low-advanced group and the high-advanced group: learners classified as ‘highly advanced’ in the experiment were able to make judgments about the MS expressions similar to the native group. Such division suggests that there might be a threshold of ‘native-likeness’ for learners, with learners not reaching the threshold showing a significant difference in their judgment from native speakers of English.

Although, as mentioned in Chapter 4, both LS/MS and LT/MT pairs reflect non-transferable conditions, the case of MT expressions is of course essentially different from that of MS expressions. While MS expressions are unacceptable in the target language, the use of MT expressions is indeed acceptable to native speakers of English. The result of the comparison between the LT and MT expressions was similar to that between the LB and MB expressions: all learner groups gave significantly lower acceptability scores to the MT expressions than to the LT expressions ( $\chi^2(1)=5.70$ ,  $p=0.017$  for intermediate;  $\chi^2(1)=5.93$ ,  $p=0.015$  for low-advanced;  $\chi^2(1)=9.37$ ,  $p=0.002$  for high-advanced;  $\chi^2(1)=8.50$ ,  $p=0.004$  for overseas); again, such differences were not present in the native group. Compared with the native group, the high-advanced and overseas learner groups gave significantly lower acceptability scores to the MT expressions ( $\chi^2(1)=7.63$ ,  $p=0.006$  for high-advanced;  $\chi^2(1)=5.94$ ,  $p=0.015$  for overseas), but there was no significant difference between the intermediate group and the native group, and the difference between the low-advanced group and the native group was marginal ( $\chi^2(1)=3.83$ ,  $p=0.050$ ). Even though there were minor differences when each learner group was compared with the native group, there was no significant difference observed between all the learner groups, indicating that the proficiency of learners did not influence the acceptability scores of the LT or MT expressions significantly.

From the acceptability scores of different expressions, we can see that native speakers and Chinese learners of English gave distinct judgments depending on the type of expression. The native speaker group accepted all the expressions that are available in English, including all the literal expressions and the MB and MT expressions; at the same time, they clearly rejected the MS expressions that are not available in English. All the learner groups, however, still showed a strong preference for literal expressions when comparing acceptability scores for literal and metaphorical expressions available in English (i.e. MB and MT expressions); this was the case even for highly proficient participants with overseas experience. For the three types of metaphorical expression, learners showed a general acceptance of MB expressions and a general rejection of MS expressions, but the trend for MT expressions seems to indicate that learners with relatively high proficiency level did not accept those expressions.

If we group the given expressions as literal versus metaphorical, then we see that in the judgment task, metaphorical expressions were treated differently from literal expressions by all learners. A native speaker of English would accept a metaphorical expression that is available in English to the same extent as a literal expression; however, no matter how proficient a learner was, a lower acceptability score would be given to a metaphorical expression than to a literal expression, regardless of the availability of that metaphorical expression in the L1 or L2. Considering the association between the result of the acceptability judgment task and the outcome of acquisition of metaphorical expressions, it seems that it may be more difficult for a learner to acquire the metaphorical meaning of an expression and eventually reach a ‘native-like’ judgment. Such asymmetry between literal and metaphorical expressions might be due to the metaphoricalness, or abstractness, of a metaphorical expression; that is, as suggested by De Groot (1992) and discussed in Section 3.2.1, more abstract metaphorical meanings are generally more difficult to master than more concrete literal meanings. At the same time, this may also be due to the frequency of a metaphorical expression in a learning context. A learner may not have had sufficient exposure to a metaphorical expression, while a literal expression may have been well taught by the instructor and given sufficient input in other contexts. Furthermore, a learner may lack the ability to infer the metaphorical meanings from the literal meaning of a word and a given context (c.f. Section 3.2.3), which could widen the gap between the acquisition of literal expressions and that of metaphorical expressions.

More interestingly, learners achieved different results for metaphorical expressions when the targets of the judgment task were different. From the results for the native group, it can be seen that the judgment task involved two different targets in terms of metaphorical expressions: it is expected that one should (1) accept the metaphorical expressions available in English, and

at the same time (2) reject the metaphorical expressions that are not available in English. While the previous discussion has shown that the learner participants did not achieve the first target even when their proficiency was high, the results of the acceptability scores for the MS expressions showed that the learners achieved the second target, namely ‘learning to reject metaphorical expressions that are not available in English’. In particular, judgments made by the high-advanced and overseas groups were statistically the same as those of the native group. Such results reflected an asymmetry of treatment of metaphorical expressions: it would be easier for a learner to ‘learn to reject’ the unacceptable expressions than to ‘learn to accept’ the acceptable expressions. This may be related to strategies used in instruction in the classroom, which emphasises the correction and avoidance of ‘incorrect expressions’ in learners’ written work.

### 5.3 Judgment confidence levels and their interaction with acceptability scores

As discussed in Section 2.3.3, it is necessary to take metalinguistic feedback into consideration when aiming to discuss cross-linguistic influence and transfer strategy. Such discussion is particularly relevant for the result of an acceptability judgment task, because sometimes learners may accept (or reject) some expressions without necessarily being confident in their judgments. In this section, we will discuss how confident the participants were as they completed the acceptability judgment task, and how a combination of the acceptability scores and the confidence scores could provide more information for the analysis.

The confidence scores given by all the participants is shown below in Figure 5.2. In general, the native group was the most confident when giving scores to the given expressions, except for MS expressions. However, the native group was also more confident when judging the literal expressions than the metaphorical expressions ( $\chi^2(1)=3.16$ ,  $p=0.076$  for the LB/MB pairs;  $\chi^2(1)=8.53$ ,  $p=0.003$  for the LS/MS pairs;  $\chi^2(1)=10.40$ ,  $p=0.001$  for the LT/MT pairs). The relatively low confidence score for MS expressions for the native group should be attributed to the unfamiliarity of these expressions: None of the MS expressions is available in English and the native participants would not have encountered them. As a result, they may have difficulty inferring the meaning of the MS expressions in the task and may not be confident about their judgment.

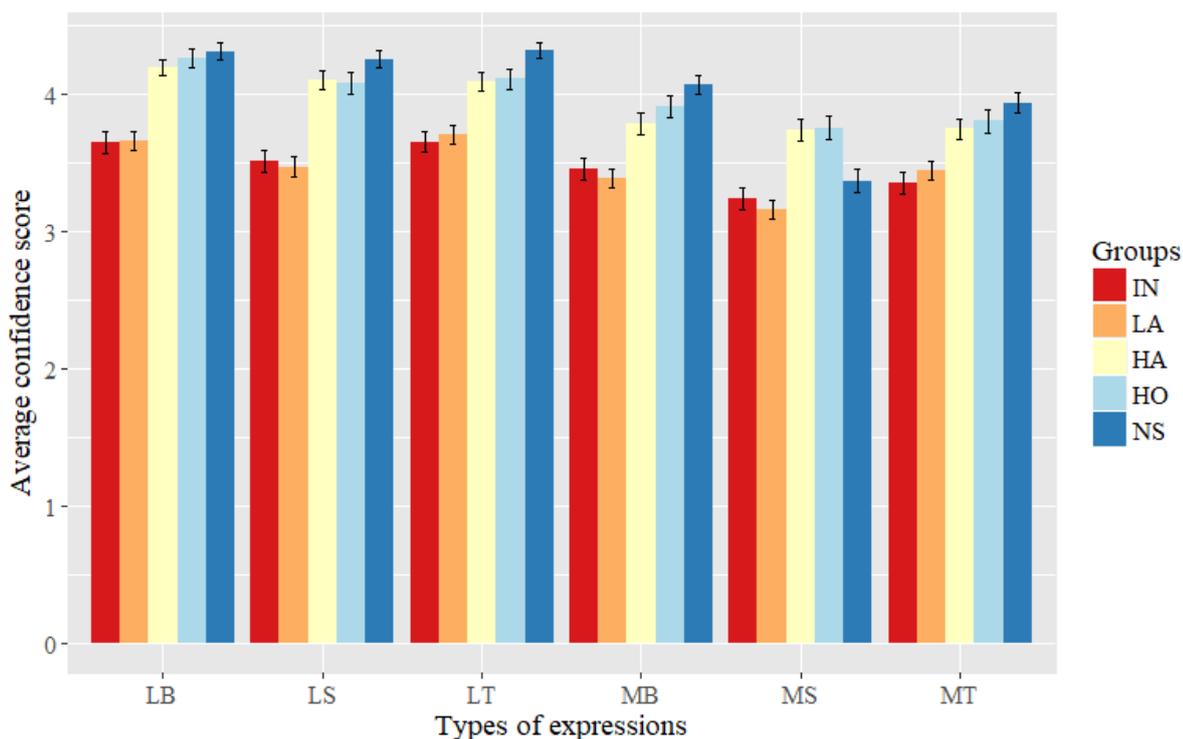


Figure 5.2 Confidence scores for the six target types for each group

The learner participants, as with the native group, were more confident when judging the literal expressions than the metaphorical expressions ( $\chi^2(1)=10.06$ ,  $p=0.002$  for the LB/MB pairs;  $\chi^2(1)=12.21$ ,  $p<0.001$  for the LS/MS pairs;  $\chi^2(1)=7.47$ ,  $p=0.006$  for the LT/MT pairs). The intermediate and low-advanced groups were less confident than other learner participants with all metaphorical expressions; in particular, the two groups were significantly less confident than other participants when scoring the MB and MT expressions ( $\chi^2(2)=10.98$ ,  $p=0.004$  for MB;  $\chi^2(2)=9.06$ ,  $p=0.011$  for MT). The confidence scores for the MB and MT expressions for high-advanced and overseas groups, on the other hand, did not significantly differ from those of the native group. As proficiency level increased, learners were significantly more confident when rating all metaphorical expressions ( $\chi^2(3)=9.56$ ,  $p=0.026$  for MB;  $\chi^2(3)=10.14$ ,  $p=0.017$  for MS;  $\chi^2(3)=7.95$ ,  $p=0.047$  for MT).

A picture of the confidence scores can help us to determine whether a learner showed indeterminacy in judging the task by investigating whether confidence in a judgment score was low. We can see from the discussion above that less proficient learners provided low acceptability scores for the MB expressions, mainly because of indeterminacy. Indeterminacy can also explain the trend of acceptability scores found for MT expressions: less proficient

learners did not truly accept the MT expressions, but showed indeterminacy, and more advanced learners made genuine rejections.

A point worth noting is that the high-advanced and overseas groups were no less confident than the native participants when they provided judgments for the MB and MT expressions; however, the scores given by the two learner groups and those given by the native speakers were significantly different. Just as the native group generally accepted the MT expressions, the high-advanced and overseas participants tended to reject the MT expressions. It could therefore be argued that, in contrast to the indeterminacy of the less proficient learners above, more proficient learners seemed to reject assertively the metaphorical expressions that they had seldom experienced in their L1, even though their intuition was entirely different from the native participants' judgments. The negative attitude towards the MT expressions shown by more proficient learners, especially in comparison to the relatively positive attitude towards MB expressions, might be largely due to the lack of availability of these metaphorical expressions in the source language. Remember that native speakers had similar confidence problems with expressions not known in their language. In the current study, the MT expressions were specific to English, while the MB expressions were shared between the language pair. The fact that none of the learner groups genuinely seemed to accept the MT expressions indicates that some language-specific use of lexical items is more difficult to acquire than the shared use of lexical items; this may be because the learners can encounter the former use only in their L2, but the latter in both languages.

Based on the interaction between the acceptability score and the level of confidence, it could easily be concluded that the learner groups displayed a 'three-fold asymmetry' in the acquisition of metaphorical expressions. Firstly, the learners, regardless of their proficiency levels, accepted the literal expressions with confidence, but did not show acceptance of the metaphorical expressions (particularly the MB and MT expressions) for which they also lacked confidence. Secondly, the learners, regardless of their proficiency levels, had successfully learned to reject metaphorical expressions that were not available in English (i.e. the MS expressions) with a relatively high level of confidence, but did not successfully acquire the metaphorical expressions available in English (i.e. the MB and MT expressions). Thirdly, most learners seemed less willing to accept the metaphorical expressions that were only available in their L2, with the exception of the overseas group. There seems to be a hierarchy of acquisition for different types of L2 word meanings: the acquisition of literal meanings being the easiest, followed by the rejection of metaphorical expressions only available in the L1, and the acquisition of shared metaphorical meanings. . The acquisition of L2-specific metaphorical

meanings is the most difficult and even highly proficient learners may not be able to achieve it; only after a period of exposure to a native English environment can learners gain some knowledge of these expressions.

## 5.4 Cross-linguistic influences on the judgment of metaphorical expressions and possible influencing factors

### 5.4.1 Identification of traces of transfer in the acceptability judgment task

Prior to analysis of how different factors can influence learners' 'strategy of transfer', or in other words, the subjective transferability of a linguistic element perceived by learners, it is crucial to identify the point at which learners transfer their knowledge from the L1 to the L2 in the acceptability judgment task. Theoretically, traces of transfer should be identified by judgments that are similar to those that would be made in the learners' L1. The cases that are most clearly identifiable as transfer are those in which the judgment does not correspond to the judgment made by most native speakers, such as accepting MS expressions and rejecting MT expressions. When the cases of possible transfer are not readily distinguishable from cases of possible acquisition, such as accepting MB expressions, we can refer to the confidence scores to see if the learners were confident about the judgments they had made.

The identification of traces of transfer followed the assumptions made in Section 4.1, and repeated below:

- A learner shows traces of transfer of an MB expression if that expression is accepted with relatively low certainty;
- A learner shows traces of transfer of an MS expression if that expression is accepted, irrespective of the level of confidence shown in accepting it;
- A learner shows traces of transfer of an MT expression if that expression is rejected, irrespective of the level of confidence.

Here the borderline between acceptance and rejection was the 0 point in the acceptability judgment result, and the confidence score marked as 'confident' was either 4 or 5 (c.f. Section 4.7.3).

In a similar fashion, traces of native-like judgment of a metaphorical expression can be also identified, following two main criteria: (1) a learner should give an acceptability score similar to that of native speakers; and (2) the learner should be confident in making the judgment. The criteria for assuming traces of native-like judgment may therefore be summarised below:

- A learner shows trace of native-like judgment of an MB expression if that expression is accepted with relatively high certainty;
- A learner shows trace of native-like judgment of an MS expression if that expression with is rejected with relatively high certainty;
- A learner shows trace of native-like judgment of an MT expression if that expression is accepted with relatively high certainty.

Finally, some learners may be in a state that amounts to neither full native-like judgment of an expression nor transfer of that expression: the expression has not been acquired at the time of experiment, and a choice is made not to transfer the expression from the L1 for various reasons. Since the reasons for transfer not being made from the L1 are not the focus of the current discussion, it is not possible to provide a more fine-grained classification of these learners, but they will instead be categorised into a single group and the phenomenon will be termed ‘transfer being blocked’.

Based on the definitions above, we can categorise all the individual judgments made by learners into three main groups: judgments that (possibly) show traces of transfer, judgments that (possibly) show traces of native-like performance, and judgments that do not belong to either of the groups mentioned above, termed ‘transfer not shown’. The criteria for the categorisation for the three types of metaphorical expression are listed below in Tables 5.1 to 5.3. It should be noted that the identification of traces of transfer is a rough, statistical estimation, rather than a case-by-case analysis, and there may be misidentification in the process. Considering that all the judgment data are analysed as groups, the accuracy of a statistical estimation is sufficient for meaningful analysis.

Table 5.1 Identification of traces of transfer in the judgments of MB expressions

	Confident (>3)	Not confident (≤3)
Acceptance (≥0)	Native-like performance	Possible positive transfer
Rejection (<0)	Transfer not shown	

Table 5.2 Identification of traces of transfer in the judgments of MS expressions

	Confident ( $>3$ )	Not confident ( $\leq 3$ )
Acceptance ( $\geq 0$ )	Possible negative transfer	
Rejection ( $< 0$ )	Native-like performance	Transfer not shown

Table 5.3 Identification of traces of transfer in the judgments of MT expressions

	Confident ( $>3$ )	Not confident ( $\leq 3$ )
Acceptance ( $\geq 0$ )	Native-like performance	Transfer not shown
Rejection ( $< 0$ )	Possible negative transfer	

By adopting the criteria for identification detailed above, we can estimate how pervasive traces of transfer can be when learners judge different types of metaphorical expression or among a particular group of learners. Table 5.4 shows the estimated percentage of traces of transfer among all the judgments given by the learner groups. A higher percentage in Table 5.4 indicates that (1) a group of learners was more likely to transfer knowledge when making the judgment; or (2) learners were more likely to transfer knowledge when judging that type of metaphorical expression.

On average, learners were most likely to transfer their knowledge when they judged the MT expressions, while they were least likely to transfer their knowledge when they judged the MB expressions. There was a clear declining trend in traces of transfer for MS expressions: traces of transfer were identified among more than half of the judgments provided by the intermediate group, while the percentage declined to only 20% when it came to both the high-advanced and overseas groups.

Table 5.4 Estimated percentage of traces of transfer of different metaphorical expressions in the learner groups

	Intermediate	Low-advanced	High-advanced	Overseas	Average
MB	27.15%	30.73%	16.44%	17.86%	23.69%
MS	58.13%	43.56%	20.37%	20.14%	36.50%
MT	35.10%	40.70%	51.55%	46.76%	43.38%
Total	40.48%	38.45%	29.85%	28.23%	34.66%

Once traces of transfer have been identified, it should be established, when traces of transfer and traces of non-transfer are compared, which aspect of ‘non-transfer’ is of greater interest, following the reasoning of Jordens and Kellerman. As shown in Table 5.1, non-transfer cases included both native-like judgment and ‘transfer being blocked’. Although it seems at first glance unreasonable to compare traces of transfer and traces of native-like judgment, for transfer always implies the absence of native-like judgment, it is theoretically feasible simply to compare cases of ‘transfer’ and ‘non-transfer’ instead of dividing ‘non-transfer’ cases into more fine-grained categories. The reason for this is that at least one factor, namely general knowledge, suggests that non-transfer includes both native-like judgments (which makes use of existent knowledge) and transfer being blocked, rather than the latter exclusively. The comparison between ‘transfer’ and ‘non-transfer’, including native-like judgment and transfer being blocked, was adopted throughout the following analysis.

#### 5.4.2 Influence of psychotypology on subjective transferability

As discussed in Section 4.7.5, the individually perceived psychological distance between Chinese and English varies among the learner participants, but the perceived psychotypological distance is not influenced by the learners’ proficiency levels. In this section, the possible influence of psychotypology on the degree of acceptability of different types of metaphorical expressions and cross-linguistic influence will be discussed. To repeat the assumption, learners who estimate the psychotypological distance between English and Chinese as shorter should transfer more knowledge from Chinese to English compared with those who estimate the distance as longer.

To examine whether the psychotypology of learners may affect the transfer strategies they use, the Mann-Whitney test was used to compare the z-scores of psychotypological distance for all participants who showed traces of transfer and those who did not show traces of transfer. If a participant provided two judgments that were identified as showing traces of cross-linguistic influence, the psychotypological distance score was counted twice in the test, since cross-linguistic influence weighed more in the overall data set. Similarly, if a participant provided one score showing traces of possible transfer and another showing traces of possible non-transfer at the same time, the psychotypological distance score would appear both in the ‘transfer’ and ‘transfer not shown’ lists. This method of calculation allowed us to account for the variability of learners: a learner may only present a certain degree of cross-linguistic influence on some critical items, and not on others. Including a data point according to its

weight in the whole data set could help to capture the delicate differences between participants and thus provide a better picture of the factors that are influential on transferability. The psychotypological distance scores were compared within each proficiency group and within each type of expression.

A significant difference in psychotypological distance scores between the intermediate learners showing traces of transfer and those not showing traces of transfer was only observed when those intermediate learners scored the MT expressions. The psychotypological distance scores of the intermediate participants who transferred their L1 knowledge to score the MT expressions were significantly smaller than for those who did not make this transfer ( $U=2095.5$ ,  $p=0.036$ ). A similar but marginal effect was also observed in the low-advanced group ( $U=4120$ ,  $p=0.099$ ), meaning the psychotypological distance scores of the low-advanced participants, who showed traces of transfer when they scored the MT expressions, were marginally lower than those of their peers who did not show traces of transfer. Such differences in psychotypological distance scores were not observed in the high-advanced and overseas groups when they scored the MT expressions. Other than that, no further significant influence from psychotypological distance scores was observed in terms of trace of transfer in other types of metaphorical expressions.

In sum, participants perceiving English as closer to Chinese showed more traces of transfer than their peers perceiving English as a language distant from Chinese. However, the result was more complicated than the primary assumption: the influence of psychotypology only affected some of the metaphorical expressions and some of the learners. Only the acceptability scores and confidence scores for the MT expressions were affected by learners' psychotypological perceptions, and only the less proficient learners were affected in that way. The results showed that less proficient learners who transferred more knowledge from their L1 to their L2 when scoring the MT expressions did perceive English as a language closer to Chinese. An effect from perception of psychotypology was not observed in (1) judgments made by more proficient learners and (2) judgments of the MB and MS expressions.

The most feasible possibilities that can be derived from the result are: (1) psychotypology may only influence judgments of metaphorical expressions at relatively early stages of acquisition, and its effect may disappear after the learners become more proficient; and (2) psychotypology may only obviously influence the judgment of a particular type of metaphorical expression, namely those expressions available in the L2 exclusively. A second possible explanation for our findings is that the learners were never exposed to these metaphorical

expressions in their L1, and they may not make accurate predictions about the availability of these expressions in their L2. Therefore, they might rely heavily on transfer to judge these expressions.

It should be noted that the current judgment task and Jordens and Kellerman's (1981) study on Dutch speakers learning English and German are essentially different in terms of their premises, so they cannot be directly compared. Jordens and Kellerman observed two groups of learners with the same L1 and different L2s, and they assumed that learners should perceive the psychotypological distances of the two pairs of languages in different ways. In the current experiment, all the learners shared the same L1 and L2, but the differences in their perceptions of the psychotypological distance between this one language pair were monitored. Compared with the study by Jordens and Kellerman, the current study involved only one target language, but at the same time possibly reflects a more diverse perception of psychotypological distance between individuals. Further investigation is needed to explore whether collectively and individually perceived differences in psychotypology can affect the judgment of linguistic elements in the same way.

#### 5.4.3 Influence of markedness of meaning on subjective transferability

As shown in Section 5.2, scores for a specific type of metaphorical expression were generally homogenous when 'metaphoricalness' was defined in a binary way. However, the degree of markedness is perceived as a continuum, and it is possible that metaphorical expressions with different degrees of markedness would show different degrees of cross-linguistic influence. This subsection focuses on the difference of degree of markedness (in Chinese) among the MB and MS metaphorical expressions and aims to discover whether degree of markedness influenced the transferability of metaphorical expressions. The degree of markedness of an expression was measured as the coreness score, as introduced in Section 4.4.1 (see Appendix 5 for detailed data): a positive coreness score means that an expression is less marked, and a negative coreness score means that an expression is marked.

Since the target of this analysis was to examine whether less marked metaphorical expressions, i.e. metaphorical expressions with a higher coreness score, were more likely to be transferred, it would be senseless to compare the degree of markedness between lexical items showing traces of transfer and lexical items not showing traces of transfer. Instead, the distribution of coreness scores within traces of transfer should be an indicator of how markedness influences transfer. If more traces of transfer are associated with a higher coreness

score, then it might be inferred that core meanings are more likely to be transferred; if more traces of transfer are associated with a lower coreness score, then this would show that peripheral meanings are more likely to be transferred. Therefore, the correlation between the extent of traces of transfer and the coreness scores was examined. Only the MB and MS expressions were included in the analysis because they may be transferred from Chinese; within each proficiency level, the MB and MS expressions were compared as one, since together they formed a longer continuum of degree of markedness and the influence of markedness would be more significant on a larger scale. The distribution of the coreness scores for the MB and MS expressions in the judgment task is shown below as Figure 5.3.



important to examine whether degree of markedness may have any influence on transferring these expressions.

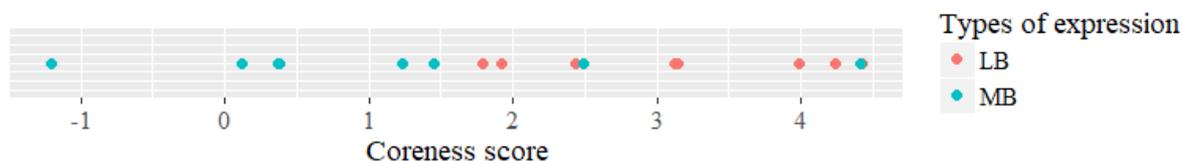


Figure 5.4 Distribution of the coreness scores of the LB and MB expressions in the AJT, as perceived by Chinese native speakers

The same statistical tool as in the previous analysis was used for this analysis. No significant correlation was found between the numbers of instances of transfer and the coreness scores for each test expression in any learner group. The reason for this lack of difference may come from the distribution of the coreness scores of these two types of expression, as illustrated in Figure 5.4. Compared with Figure 5.3, we can see that the coreness scores of most LB and MB expressions were uniformly high, with only one expression below 0. It is possible that the difference in the coreness scores was not significant enough to demonstrate their impact on the subjective transferability of a meaning.

#### 5.4.4 Influence of metaphoricalness on subjective transferability

While the previous subsection showed that the degree of markedness of an expression only had a limited influence on the transferability of metaphorical expressions, a question emerged from the results of the experiment: if markedness does not influence traces of transfer when learners judge metaphorical expressions, what is the main reason for differences in acceptability scores between LB and MB expressions? This subsection will look back and explore the possible influence of binary metaphoricalness on the subjective transferability of an expression, and to control other influential factors, the comparison between LB and MB expressions will be demonstrated.

Table 5.5 below shows the distribution of traces of transfer in LB and MB expressions across learner groups. Clear quantitative differences were shown in the distribution: compared with LB expressions, MB expressions generally showed fewer traces of transfer among learners in each group; transfer of MB expressions was also less likely to surface than that of LB

expressions when both types of expression are objectively transferable. In general, learners seemed to be less likely to transfer MB expressions than LB expressions. We might conclude on this basis that it was more difficult for them to acquire the MB expressions via transfer than the LB expressions.

Table 5.5 Distribution of traces of transfer on LB and MB expressions across learners' proficiency

	Types of expression	IN	LA	HA	HO
Possible positive transfer	LB	26.49%	24.61%	11.11%	9.70%
	MB	27.15%	30.73%	16.44%	17.86%
Native-like performance	LB	52.98%	50.79%	84.03%	80.60%
	MB	37.09%	39.06%	52.05%	55.00%
Transfer not shown	LB	20.53%	24.61%	4.86%	9.70%
	MB	35.76%	30.21%	31.51%	27.14%

A follow-up analysis showed that the degree of markedness of an expression did not significantly influence the extent of traces of transfer being blocked. Therefore, markedness did not significantly contribute to the low subjective transferability of the MB expressions, and the major possible reason for the differences shown above between LB and MB expressions may be metaphoricalness. Being metaphorical, or to be more precise, being 'non-literal', does not only indicate that an expression receives lower acceptability scores from the learners than the literal expression, but may also indicate that it is less transferable, and more importantly, learners are generally less likely to benefit from the potential for positive transfer.

#### 5.4.5 Influence of general proficiency on subjective transferability

As in Section 5.4.2, the Mann-Whitney test was used to compare the OQPT scores of all participants who presented traces of cross-linguistic influence in their judgments and those who did not. Again, the comparison of the OQPT score was performed in a weighted manner: participants who showed multiple traces of transfer were counted multiple times in the test, and participants who showed traces of both transfer and non-transfer were included in the two groups simultaneously.

A comparison of the OQPT scores reveals that the participants who presented traces of transfer when judging the MB expressions had significantly lower OQPT scores than those who

did not present traces of transfer ( $U=41764$ ,  $p=0.002$ ). The mean OQPT score achieved by the participants showing traces of transfer was 44.59 ( $SD=7.11$ ), while the mean OQPT score for the participants who did not show traces of transfer was 46.79 ( $SD=6.98$ ). It seems that less proficient learners were more likely to rely on transfer of their L1 knowledge to score the MB expressions in the task.

As for the MS expressions, the participants who showed traces of transfer had a significantly lower OQPT result than those who did not show traces of transfer ( $U=31786$ ,  $p<0.001$ ). The mean OQPT score achieved by the participants showing trace of transfer was 43.43 ( $SD=6.40$ ), while the mean OQPT score for the participants who did not show traces of transfer was 47.89 ( $SD=6.86$ ). Again, general proficiency seems to influence the transfer strategies: less proficient learners were more likely to show negative transfer when they scored the MS expressions.

However, the influence of general proficiency was not prominent on the subjective transferability of the MT expressions. Only marginal significance from the OQPT scores was observed between the participants who showed traces of transfer and those who did not ( $U=56192$ ,  $p=0.069$ ). The mean OQPT score achieved by the participants showing traces of transfer was 46.89 ( $SD=7.20$ ), and the mean OQPT score for the participants who did not was 46.04 ( $SD=6.75$ ). There was no clear difference between the two groups, while the participants presenting traces of transfer even achieved a slightly higher average OQPT score than those who did not show any traces of transfer. Furthermore, from the estimated percentage of instances of transfer listed in Table 5.4, it can also be seen that the percentage did not change drastically between different proficiency groups. Compared with the MB and MS expressions, the influence of proficiency on traces of transfer in the MT expressions was less clear. Considering that even highly proficient learners still rejected the MT expressions in general, it could be inferred that cross-linguistic influence on the MT expressions was more persistent. That is possibly because there were no correspondences to the MT expressions in the learners' L1, and the learners failed to perceive the possibility of those MT expressions.

Overall, as proficiency rose, participants gradually presented fewer traces of transfer from their L1 to their L2 in judging metaphorical expressions. While that decrease in influence was prominent among the MB and MS expressions, the MT expressions seem to be the exception, with the level of proficiency in English not influencing the transferability. This provides some evidence that the acquisition of the MT expressions may be different from that of the MB and MS expressions. The learners gradually mastered the knowledge of metaphorical expressions if there were corresponding expressions in their L1; in such cases, they were able to accept the

MB expressions available in the L2 and reject the MS expressions that were not available in the L2. However, since the MT expressions are not available in their L1, they do not have any knowledge to transfer, and remain suspicious of those expressions, which eventually leads to the rejection.

## 5.5 Learners' production of metaphorical expressions and alternatives in the sentence correction section

### 5.5.1 Four types of reaction by participants to given metaphorical expressions

In the current analysis, the relevant feedback sentences, as defined in Section 4.7.3, were only targeted at the content and the use of metaphorical expressions, and grammatical corrections were not included. All these sentences may be categorised in two different ways: (1) whether the participants provided feedback sentences as an alternative, a rejection, or a correction to the given expressions; and (2) what types of expression the participants used to produce the feedback sentences, i.e. literal or metaphorical, shared or language-specific. This subsection and the next are devoted to these two aspects of the analysis respectively, and this subsection will particularly focus on the use of feedback sentences as evidence of the rejection or correction of a given metaphorical expression.

The categorisation of reaction types involved a joint analysis of the acceptability scores of metaphorical expressions and the feedback sentences provided, resulting in one possible reaction type per item. On the one hand, a participant who accepted a given correct expression (i.e. gave a positive score to an available expression) and provided a feedback sentence at the same time, is likely to have provided an alternative expression to a given item, which was thought to be better in the context. On the other hand, if a participant gave a negative score to an available expression then any feedback sentence is more likely to have been a replacement for the test item. If a test item was designed not to be available, but the participant accepted it and presented a feedback sentence, this can be understood as a kind of tolerance to the incorrect expression, because the learner is indicating a belief that the expression was acceptable, although there could be a better way to phrase it. Rejecting an expression that was not available with a feedback sentence may be seen as a typical case of correction, especially when the instructions for the judgment task are taken into consideration. These four possible types of reaction were further categorised into two groups: The reaction types 'alternative' and 'rejection' were applied to the MB and MT expressions, while 'tolerance' and 'correction' were applied to the MS expressions.

Table 5.6 Distribution of the types of reaction across different proficiency groups

	MB		MS		MT	
	Alternative	Replacement	Tolerance	Correction	Alternative	Replacement
IN	9.30%	4.65%	9.30%	37.21%	6.98%	32.56%
LA	5.41%	2.70%	8.11%	45.95%	12.16%	25.68%
HA	2.07%	8.38%	1.38%	57.93%	0.69%	29.66%
HO	1.09%	6.52%	3.26%	57.61%	2.17%	29.35%
NS	0.60%	2.99%	2.99%	67.66%	4.19%	21.56%

Table 5.6 is a summary of the percentages of different reaction types for all participants at different proficiency levels. It can be argued that the native group clearly and exclusively concentrated on unacceptable use of the MS expressions: over 70% of the relevant feedback sentences produced by the native participants focused on the MS expressions. However, the native group occasionally produced replacement sentences for the MB and MT expressions. This was actually a good indication of individual differences of use for metaphorical expressions. Even though the test items had been verified by native English informants prior to the experiment, some participants still did not accept those expressions, and they preferred to provide their own version to express the same meanings.

The distribution of feedback sentences for the learner groups was qualitatively similar to that for the native group. As proficiency rose, there was a general trend for participants to correct the MS expressions and not to replace the MB and MT expressions with their own expressions. However, compared with the native group, the learner groups showed a more diverse focus in the sentence correction section. Learner participants produced fewer sentences to target the incorrect use of MS expressions, and they still devoted a high percentage of feedback sentences to expressions that were available in English, as seen, for example, in the higher proportion of sentences targeting MB expressions used by high-advanced learners (over 10%) and around one third of the sentences targeting the MT expressions among all the learner groups.

Although both native speakers and learners made corrections to the MB and MT expressions that were available in English, there was a quantitative difference in the results for both the acceptability judgment task and the sentence correction section. This indicates that native

participants accepted the MB and MT expressions more than any other group of learners. Such differences should be attributed to the learners' (lack of) knowledge of L2 metaphorical expressions, rather than to the participants' individual preference for metaphorical expressions.

### 5.5.2 Strategies adopted by the participants in producing feedback sentences

All feedback sentences could be assigned to four categories of strategy type: full paraphrase, near- or mis-paraphrase, alternative metaphorical expression, and error indicated without correction. Examples of the four types are given in Table 5.7 below, all extracted from the feedback sentences provided by the native participants. *Full paraphrase* means that a participant chose to rephrase the given metaphorical expression in a literal way, while the intended meaning of the sentence was fully preserved in the feedback sentence. *Near- or mis-paraphrase* is similar to full paraphrase, but the feedback sentence failed to preserve the intended meaning of the test sentence due to a misunderstanding by the participant. The third strategy, namely 'alternative metaphorical expression', indicates that the participant used another metaphorical expression or an idiom that is available in English and that is equivalent in meaning to the given expression. Finally, when participants were sometimes not sure how to express the intended meaning accurately, or failed to understand the test sentence, they may have decided not to leave any feedback sentence. In that case they would usually point out that 'this part is Chinglish', 'the sentence is not comprehensible' or 'awkward to me' with an indication of the problematic words or phrases (e.g. underlining, colouring and question marks).

Table 5.7 Examples of types of strategies adopted by participants in the sentence correction section

Strategy	Test sentence <sup>14</sup>	Relevant feedback sentence
Full paraphrase	Brian <i>ate</i> some <u>loss</u> when he started his own business.	Brian <u>suffered some loss</u> when he started his own business.
Near/Mis-paraphrase	Sally always <i>bites</i> the <u>words and phrases</u> whenever she writes an article. <sup>15</sup>	Sally always <u>cuts words and phrases</u> whenever she writes an article.
Alternative metaphorical expression	The mother <u>held a belly of gas</u> because her son failed in the exam.	The mother <u>blew a fuse</u> because her son failed in the exam.
Error indicated without correction	Sophie <u>lost her golden bowl</u> after her boss decided to shut down the company.	Sophie <u>lost her golden bowl</u> (?) after her boss decided to shut down the company.

Three of the four strategies above, i.e. full paraphrase, near paraphrase and error indicated without correction, are ‘language-neutral’; that is, they do not involve the appearance of a language-specific expression. Both the full paraphrase and the near/mis-paraphrase strategies are closely related to paraphrasing and the use of literal language; since literal expressions are generally shared between languages, these strategies are fully language neutral. The last strategy, namely ‘error indicated without correction’, does not lead to the production of any expressions, so it should be seen as language-neutral as well. The production of an alternative metaphorical expression can be either language-neutral or language-specific, depending on the resulting expression. For instance, the sentence given in Table 5.7 is a good illustration of a language-specific metaphorical expression, because ‘to blow a fuse’ is not available in Chinese, only in English. As discussed in Section 2.2, some semantic transfer may appear when a learner uses language-neutral expressions rather than L2-specific expressions. Therefore, it is important to look at the language-neutral strategy here to see whether a learner may resort to language-neutral paraphrase upon encountering a metaphorical expression.

<sup>14</sup> The words in italics are the critical lexical items that are used metaphorically in the sentences. The underlined constituents are the constituents that include the critical lexical items, and they were usually changed or rephrased in the relevant feedback sentences.

<sup>15</sup> Intended meaning: Sally always cares too much about wording whenever she writes an article.

Table 5.8 shows the levels of production of the four types of strategy commonly used in the sentence correction component by different groups of participants. The use of full paraphrase and near paraphrase, both of which are language-neutral, is pervasive among all groups. Near-paraphrase is particularly prominent in the native group, and most instances appear when participants attempt to correct the MS expressions; that is, because the native participants lack the relevant knowledge and fail to understand the meanings of those expressions. The result reveals that both learners and native speakers prefer changing a metaphorical expression to a corresponding literal expression to other strategies, and paraphrasing is used to ‘correct’ all three types of metaphorical expressions, particularly L1-specific expressions. It should be noted that the choice of strategy may be biased due to the description in the instructions: in the experiment, participants were told that they were correcting sentences for their peers, and the ‘recipients’ of the feedback sentences were English learners, so they might prefer choosing more literal and direct expressions to make the feedback sentences more accessible. Still, a preference for literal paraphrasing may indicate that literal expressions are indeed considered ‘less marked’ and metaphorical expressions ‘more marked and more difficult to comprehend’.

Table 5.8 Average number of relevant feedback sentences with different strategies provided by individual participants at different proficiency levels

	Full paraphrase	Near-/Mis-paraphrase	Alternative metaphorical expressions	Error indicated without correction
IN	1.00	0.19	0.71	0.14
LA	1.15	0.19	0.92	0.58
HA	5.10	0.10	1.33	0.38
HO	3.44	0.44	1.11	0.11
NS	3.50	0.79	2.13	0.54

At the same time, all learner groups displayed a degree of creativity in language use in the sentence correction component, particularly in proposing other metaphorical expressions to replace the given metaphors. For example, the relevant feedback sentences for one test sentence provided by learner participants are recorded and categorised in Table 5.9, and we can see that learners were able to select lexical items that they believed could accurately present the intended meaning when paraphrasing the test sentence, or derive metaphorical expressions based on the literal meanings of the lexical items. All feedback sentences except for (9) are classified as

‘replacement’ sentences using the paradigm in Section 5.5.1: this means the participants rejected the use of the given metaphorical expressions and replaced them with their own expressions.

Table 5.9 All 16 relevant feedback sentences for ‘My mother said that these books would be *food* for thought’ given by different groups of participants<sup>16</sup>

Group	Strategy	Relevant feedback sentences
IN	Full paraphrase	(1) ... would be useful.
	Alternative metaphorical expression	(2) ... would be <i>supply</i> for thought.
		(3) ... would be <i>nutrition</i> for thought.
		(4) ... would be <i>sources</i> of thought.
LA	Full paraphrase	(5) ... would be good for thinking.
	Alternative metaphorical expression	(6) ... would be good <i>materials</i> to develop thinking.
		(7) ... would <i>feed</i> our thought.
		(8) ... would be the <i>fountainhead</i> of thoughts.
Error indicated without correction	(9) ... would be <b>food for thought</b> (?)	
HA	Full paraphrase	(10) ... would be a resource for thought.
		(11) These books are to human thought what food is to people.
		(12) ... would be helpful for thinking.
	Alternative metaphorical expression	(13) ... would be mental <i>nourishment</i> .
		(14) ... would be a <i>necessity</i> to thought.
HO	Full paraphrase	(15) ... would be resource of thought.
	Alternative metaphorical expression	(16) ... would be spiritual <i>nourishment</i> .

It may also be noted in the feedback sentences listed in Table 5.9 that some learner participants transfer their knowledge from the L1 into their production of an L2 sentence.

<sup>16</sup> All italicised words are lexical items used in a metaphorical sense under the MIP framework (Pragglejaz 2007). No native participant provided relevant feedback sentences for an item because as a result of generally accepting the metaphorical meaning of ‘food’.

Expressions like ‘nutrition’, ‘nourishment’ (both corresponding to *yingyang* in Chinese), ‘fountainhead’ and possibly ‘source’ (both corresponding to *yuanquan* in Chinese) are common metaphorical expressions in Chinese to describe the importance of books for studying and thinking. When learners chose these expressions to replace ‘food for thought’, they seemed to assume that these expressions were shared by Chinese and English speakers. It may also have been the case that they directly translated these expressions into English, even if they were actually less favoured in English compared to ‘food for thought’. No matter what kind of belief learners hold, we should regard the use of these alternative expressions as transfer from their L1 to the L2.

To conclude, when the participants encountered a metaphorical expression that they believed questionable, the most widely-used method was to paraphrase it using a literal sentence that conveyed the same meaning. This strategy was pervasive among learners at all proficiency levels as well as native speakers. While less proficient learners produced feedback sentences mainly by paraphrasing them literally, even more proficient learners would choose to paraphrase the shared metaphorical expressions to literal expressions. On the other hand, the fact that learner participants at all proficiency levels felt the need to paraphrase some shared metaphorical expressions indicates that the metaphorical meaning of a lexical item is more difficult to acquire even if a learner has already reached a high level of L2 proficiency. All learner participants demonstrate lexical creativity both in paraphrasing and in the construction of alternative metaphorical expressions; at the same time, a trace of transfer from their L1 could still be detected in the feedback sentences provided by the learner participants.

## 5.6 Summary

Overall, the results of the acceptability judgment task revealed that there were three types of asymmetry in the process of judging by Chinese learners of English, both between the judgment of literal expressions and that of metaphorical expressions, and between the judgment of different metaphorical expressions. Firstly, compared with literal counterparts, learners were less likely to judge the metaphorical expressions in the same way as native speakers, which indicated that the metaphorical expressions were generally less familiar to the learners than the literal expressions. Also, compared with the rejection of the metaphorical expressions that were not available in English (the MS expressions), the learners showed more difficulty in accepting the metaphorical expressions that were available in English (the MB and MT expressions), and it was more difficult for learners to reach a native-like level of judgment when accepting the

MB and MT expressions than when rejecting the MS expressions. Such differences suggested that learners' mechanism for learning to reject the MS expressions was different from the way in which they would learn to accept MB and MT expressions. Finally, the results seem to indicate that the learners were gradually able to acquire the MB and MS expressions as their proficiency level rose; however, they had persistent difficulty in acquiring the MT expressions, and even highly proficient learners were not able to make native-like judgments of the MT expressions. All three types of asymmetry demonstrate the complexity of the acquisition of metaphorical expressions in an L2, particularly if we take the learners' L1 into consideration.

In this chapter, cross-linguistic influence during the acquisition of metaphorical expressions was identified and quantified, and the three major factors influencing subjective transferability of metaphorical expressions have been examined. The influence of psychotypical distance was very limited in the current experiment: it only influenced the less proficient learners; moreover, the learners' individual psychotypical perception only influenced the transferability of the MT expressions. Influence from degree of markedness was not observed. Influence from metaphoricalness, however, was significant in subjective transferability: metaphorical expressions were generally seen as less transferable than literal expressions by learners, even if they contained the same lexical items and were shared between learners' L1 and L2. Finally, general knowledge of the L2 was the most prominent influencing factor on transferability: with a rise in proficiency level, the learners gradually showed less influence from their L1 on the L2 in their judgment of the metaphorical expressions.

In the sentence correction section, there was a general tendency for more proficient learners to provide more feedback sentences. The learners demonstrated a pervasive trend to use literal paraphrases; nevertheless, some of them also produced alternative metaphorical expressions, showing that they were able to express themselves metaphorically and creatively in their L2. Traces of transfer could also be observed in the alternative metaphorical expressions provided by the learners: some learners chose to translate the expressions from their L1 to the L2 in a word-to-word fashion, assuming that those expressions were available in both languages. It could be seen that cross-linguistic influence during the acquisition of metaphorical expressions was persistent in both perception and production. Such influences will be further discussed in Chapter 7.

## 6 The processing of metaphorical expressions

### 6.1 Introduction

This chapter reports the results of the self-paced reading task. After the discussion of the differences in the judgment of literal and metaphorical expressions by Chinese learners of English in Chapter 5, the question arises of whether such differences appear only after consideration on the part of the learner (i.e. in an offline task), or at the moment the learner reads that expression (i.e. in an online task). Although there is still debate about whether an acceptability judgment task examines a learner's implicit or explicit knowledge, it is widely recognised that a self-paced reading task, or other online tasks, examine the learner's implicit knowledge. Therefore, a comparison between the results of an offline task and those of an online task is thought to provide a better picture of learners' processing and knowledge of metaphorical expressions – irrespective of the knowledge of metaphorical expressions being more implicit or more explicit. Furthermore, as a number of proposals have been presented in Section 3.4 on the routes for retrieving metaphorical meanings in the L2, an online task was thought to provide important additional information to make it possible to decide which route(s) a learner takes when reading a metaphorical expression. It would also reveal whether learners will change their processing routes as a function of their proficiency, degree of markedness of meaning, or psychotypological perception.

The chapter is divided into three major parts. Section 6.2 presents the reading pattern for metaphorical expressions in comparison to literal expressions. Section 6.3 presents the participants' reaction to the comprehension questions after reading the metaphorical expressions in comparison to literal expressions. Section 6.4 compares the results of the reading task and the factors influencing transferability, namely psychotypology, markedness of metaphorical expression and general proficiency. Although it is difficult to observe traces of transfer in an online task, we felt it would be useful to correlate those factors and the reading time for the expressions to see if the factors may have any direct influence on the processing of metaphorical expressions.

### 6.2 Reading time for literal and metaphorical expressions

In the analyses below, we will concentrate on the reading times for segments 3 and 4, i.e., the segments following the target metaphorical (or literal) expression, and likely to show spill over of the processing of such expressions. The logic we follow for this analysis is that, if

learners show a significantly different processing pattern of literal and metaphorical expressions from native speakers, more specifically, if learners spend significantly more time processing metaphorical expressions than literal expressions while native speakers do not, then it is likely that learners use a different way to retrieve or construct metaphorical meanings. We first compare the reading times of literal versus metaphorical expressions of the same type (for example MS and LS), and then compare the reading times for metaphorical expressions of different types (MB, MT, MS). The segment-by-segment reading time pattern for the native group, used here as baseline information, is shown below in Figure 6.1. It can be observed that only the MS sentences caused a major delay: native participants spent significantly more time reading Segment 3 in all the MS sentences compared with the LS sentences ( $\chi^2(1)=5.52$ ,  $p=0.019$ ); the delay continued into the fourth segment ( $\chi^2(1)=4.40$ ,  $p=0.036$ ). Other than that, all other types of sentence, all acceptable in English, entailed a similar reading pattern for metaphorical and literal expressions. In general, the native group spent similar amounts of time reading and processing a metaphorical expression that is available in English and its corresponding literal expressions. This indicates that, if a metaphorical expression is available in English, native speakers are able to connect the metaphorical meanings directly with the word form, and the ‘literal-first’ hypothesis did not come into effect. The significant delay after reading the MS expressions showed that the native group spent additional time ‘making sense’ of these expressions, and that the effort to ‘make sense’ continued until the end of each sentence.

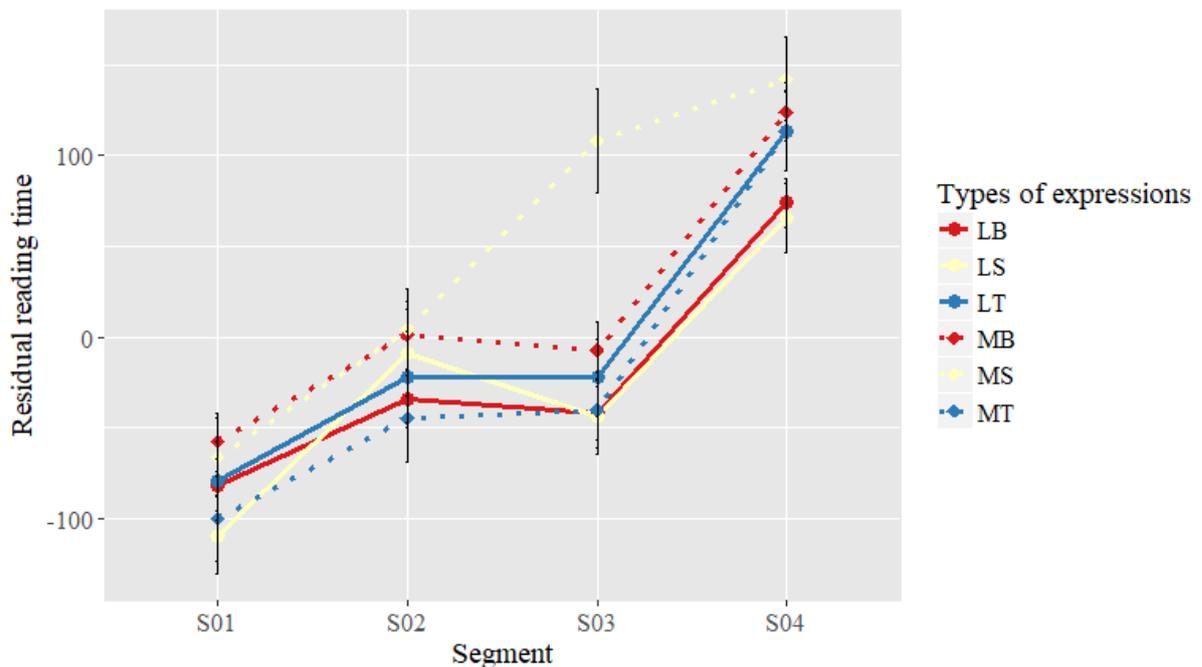


Figure 6.1 Segment-by-segment reading time pattern for the native group

It was further expected that when two metaphorical expressions were both available in English (i.e., an expression in the MB and MT class), a native speaker would spend the same amount of time reading them; however, if two metaphorical expressions were both available in Chinese (i.e. an expression in the MB and MS class), there would be a difference in reading time, since the availability of those expressions in English should also be taken into consideration. Therefore, we can make comparisons (1) between the MB and MT sentences and (2) between the MB and MS sentences. In accordance with the prediction above, native participants spent a similar amount of time reading Segment 3 of the MB and MT sentences, but they spent significantly more time reading Segment 3 in all MS sentences compared with MB sentences ( $\chi^2(1)=6.83, p=0.009$ ).

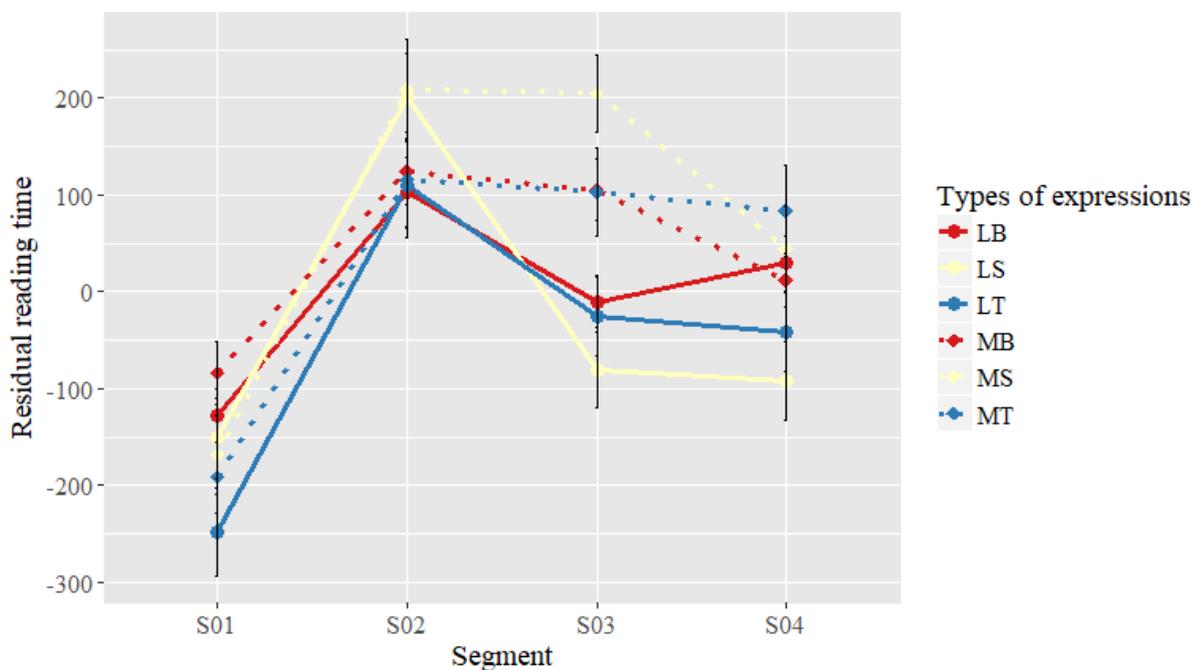


Figure 6.2 Segment-by-segment reading time pattern for the intermediate group

When comparing reading times for literal and metaphorical sentences by intermediate learners, significant differences in reading times between the literal and metaphorical sentences were observed in the Segment 3 position: the participants spent significantly more time reading the segment after all types of metaphorical expressions than after the corresponding literal expressions ( $\chi^2(1)=5.26, p=0.022$  for MB;  $\chi^2(1)=10.48, p=0.001$  for MS;  $\chi^2(1)=4.11, p=0.043$  for MT). Furthermore, the delay in processing the MS sentences also continued into the fourth segment ( $\chi^2(1)=4.02, p=0.045$ ). This suggests that the processing cost for the retrieval of all

types of metaphorical meanings continues to exist and the learners indeed showed a ‘literal-first’ preference, but only after they had read the metaphorical expression. Interestingly, although it seems that the intermediate learners showed more of a difference between LS and MS expressions than between MB and LB, and MT and LT expressions, there was no significant difference between the reading times for Segment 3 of the MS sentences and for MB sentences. Also, there was no significant difference between the reading times for Segment 3 of the MB sentences and of the MT sentences. It seems that the intermediate learners produced quantitatively similar patterns both during and after reading different metaphorical expressions.

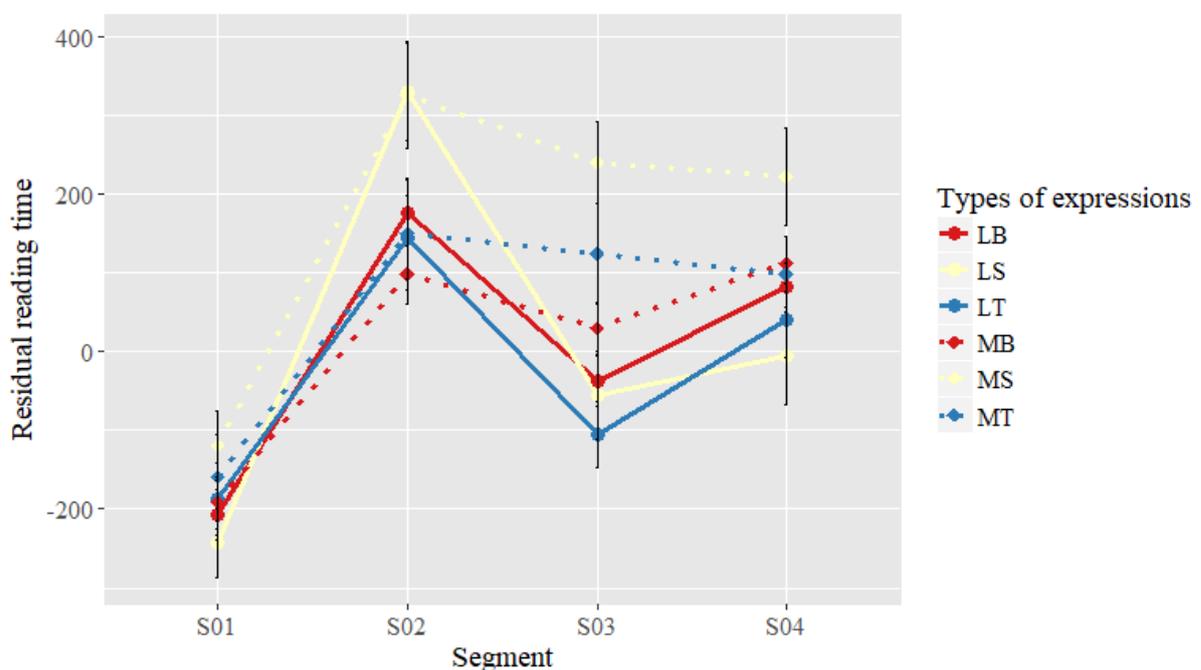


Figure 6.3 Segment-by-segment reading time pattern for the low-advanced group

The data for the low-advanced learners, as presented in Figure 6.3, again show differences in reading times between literal and metaphorical expressions at the Segment 3 position, indicating that the retrieval of metaphorical meanings might happen after the entire expression is read. The low-advanced participants spent significantly more time reading the segment after the MS expressions than that after the LS expressions ( $\chi^2(1)=5.02$ ,  $p=0.025$ ), but such differences were marginal in the fourth segment ( $\chi^2(1)=2.71$ ,  $p=0.0997$ ). The participants also spent marginally more time reading the segment after the MT expressions than the segment after the LT expressions ( $\chi^2(1)=3.54$ ,  $p=0.060$ ), and in this case the difference became more significant in the fourth segment ( $\chi^2(1)=3.97$ ,  $p=0.046$ ). There was no clear difference between

the reading patterns for the LB and MB sentences. It seems that for the low-advanced learners, both the MS and MT expressions entailed more processing effort, but this was not the case with the MB expressions. Considering the availability of the three types of metaphorical expression, it could then be argued that shared metaphorical expressions were easier for learners to process than other metaphorical expressions. However, there was no significant difference between the reading time for the segment after the MB and MT expressions; significant difference was only observed between the MB and MS expressions ( $\chi^2(1)=4.42$ ,  $p=0.035$ ). Another conclusion one might draw based on these data is that the low-advanced learners have realised the specific problem with MS sentences.

The differences at Segments 3 and 4, however, were clearly demonstrated for the high-advanced learners as shown in Figure 6.4: the participants spent significantly more time reading these two segments after an MS expression than after an LS expression ( $\chi^2(1)=8.02$ ,  $p=0.005$  for Segment 3;  $\chi^2(1)=4.26$ ,  $p=0.039$  for Segment 4), and also spent marginally more time reading Segment 3 after an MT expression than after an LT expression ( $\chi^2(1)=3.53$ ,  $p=0.060$ ). In much the same way as the low-advanced learners, the high-advanced learners only showed a significant difference in the segment after MB and MS expressions ( $\chi^2(1)=15.41$ ,  $p<0.001$ ), and not in segments after MB and MT expressions ( $\chi^2(1)=0.001$ ,  $p=0.97$ ). In general, the high-advanced learners showed similar reading time patterns to the low-advanced learners, but their hesitation with MT sentences was less evident.

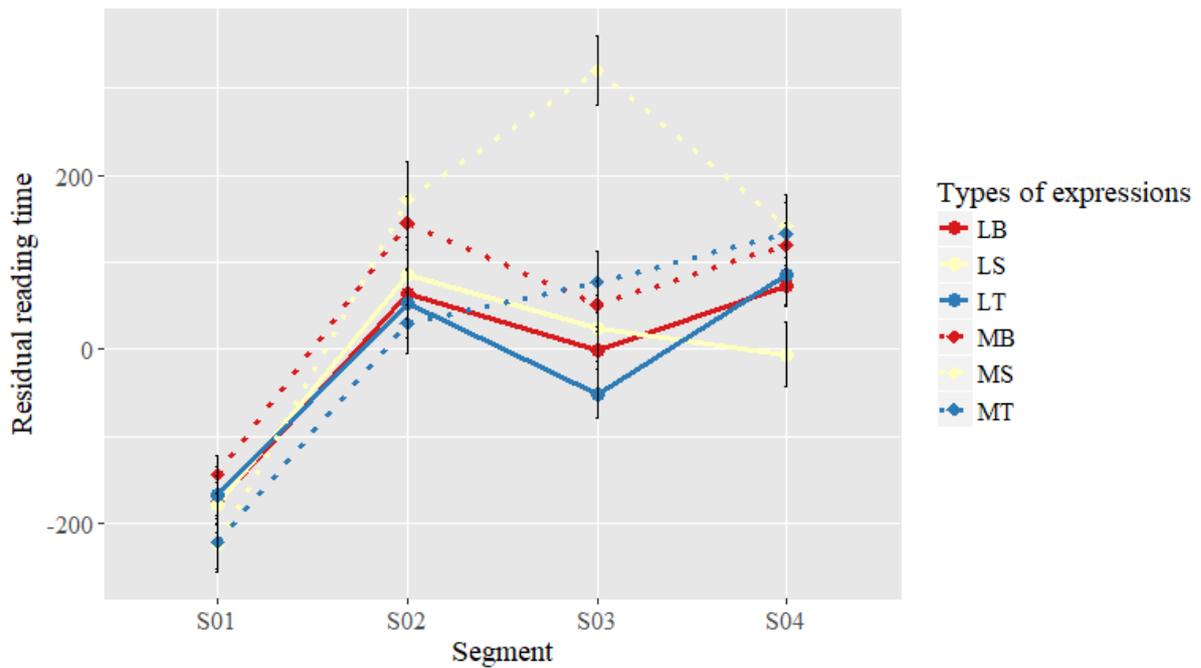


Figure 6.4 Segment-by-segment reading time pattern for the high-advanced group

Figure 6.5 shows the reading patterns for the overseas group. The only difference between the reading time for the literal sentences and for the metaphorical sentences appeared between the LS and MS sentences. The participants showed a significant hesitation after reading the MS expressions compared to the LS expressions ( $\chi^2(1)=4.44$ ,  $p=0.035$ ), while such differences disappeared at Segment 4. There was no significant difference in reading time between LB and MB, or LT and MT sentences, and thus the same reading pattern was presented as for the native speaker group. Interestingly, it was recorded that the overseas learners spent significantly less time reading the segment after MB expressions than after MT expressions ( $\chi^2(1)=5.28$ ,  $p=0.021$ ) and MS expressions ( $\chi^2(1)=16.11$ ,  $p<0.001$ ).

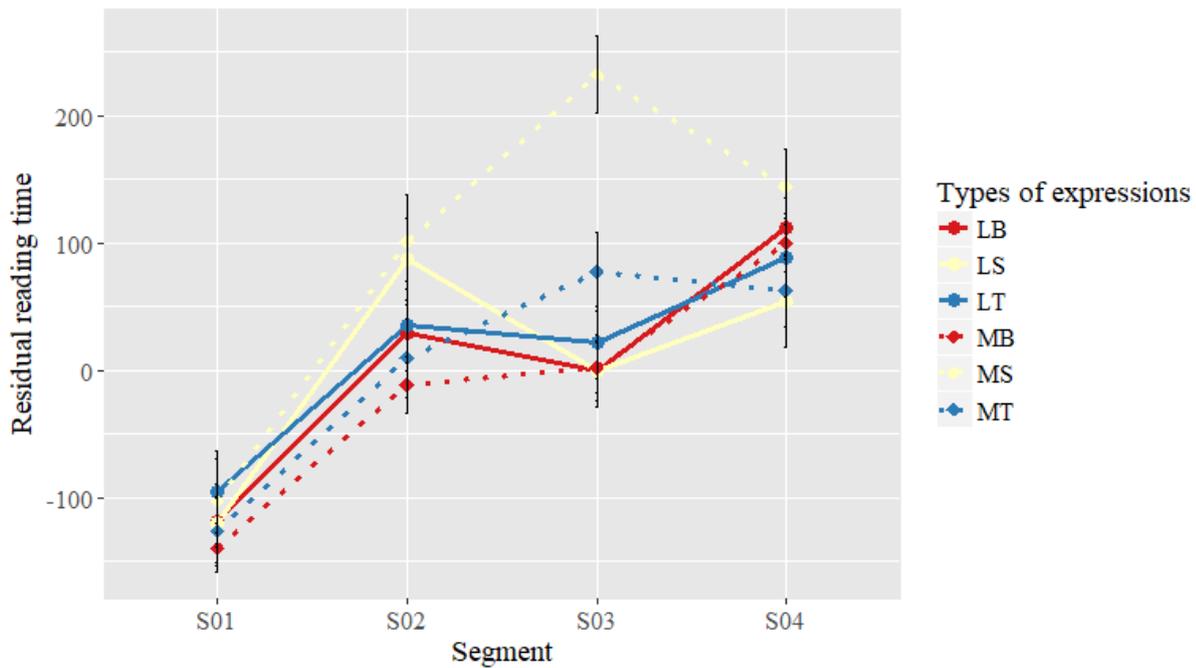


Figure 6.5 Segment-by-segment reading time pattern for the overseas group

A native-like processing pattern, as discussed by Sanders and Neville (2003), means that non-native speakers of a language demonstrate a processing pattern (e.g. reading time, ERP pattern) that is qualitatively similar to that of native speakers of the language. In the current experiment, the most prominent properties of a native-like processing pattern included the following: (1) there was no significant difference between LB and MB sentences; (2) there was no significant difference between LT and MT sentences; and (3) there *was* a significant difference at Segment 3 between LS and MS sentences, with participants showing significant hesitation at Segment 3 of MS sentences.

In terms of the reproduction of the properties of such a native-like processing pattern, there seems to be a hierarchy of difficulty. While the third property, namely a significant difference between the LS and MS sentences, was demonstrated in every learner group, the first two, especially the second property, were found only at higher proficiency levels. The intermediate group hesitated after reading the MB expressions, while the advanced learners could process the MB expressions without obvious hesitation. However, even the high-advanced participants were not able to process the MT expressions in the same way as native speakers, indicating that the MT expressions were generally more difficult for learners to process than the MB expressions. Such a hierarchy was similar to the hierarchy of difficulty displayed in the acceptability judgment task in Chapter 5: it seemed easiest to detect a lack of acceptability in

MS expressions, followed by the acceptability of MB expressions, and finally the acceptability of MT expressions.

The results also show that some learners could achieve a native-like processing pattern, although extensive native L2 input was required. We see in the results that the overseas group demonstrated a more native-like processing pattern when reading metaphorical expressions: participants in that group did not hesitate significantly when processing the MB and MT sentences, but they showed a significant delay after reading the MS expressions. While we could argue that a learner could process all the metaphorical expressions in a native-like way after (1) reaching a high proficiency level and (2) receiving sufficient exposure to the native discourse, it should be noted that there is no significant difference of residual reading time observed between the overseas group and the high-advanced group when reading Segment 3 after an MT expression. Therefore, we could not draw a firm conclusion that native exposure is the decisive factor for learners to achieve a native-like reading pattern of MT expressions.

### 6.3 Reaction time to the comprehension questions and scores for responses

The residual reaction time for the comprehension questions, as well as the scores for responses were analysed in the comparisons. Comparisons were made (1) between questions targeting the meaning of metaphorical expressions and those targeting the meaning of literal uses, and (2) between questions targeting different types of metaphorical expressions. A shorter residual reaction time indicates that participants reacted faster to questions related to that type of expression, and a higher mean score indicates that participants answered the comprehension questions more accurately.

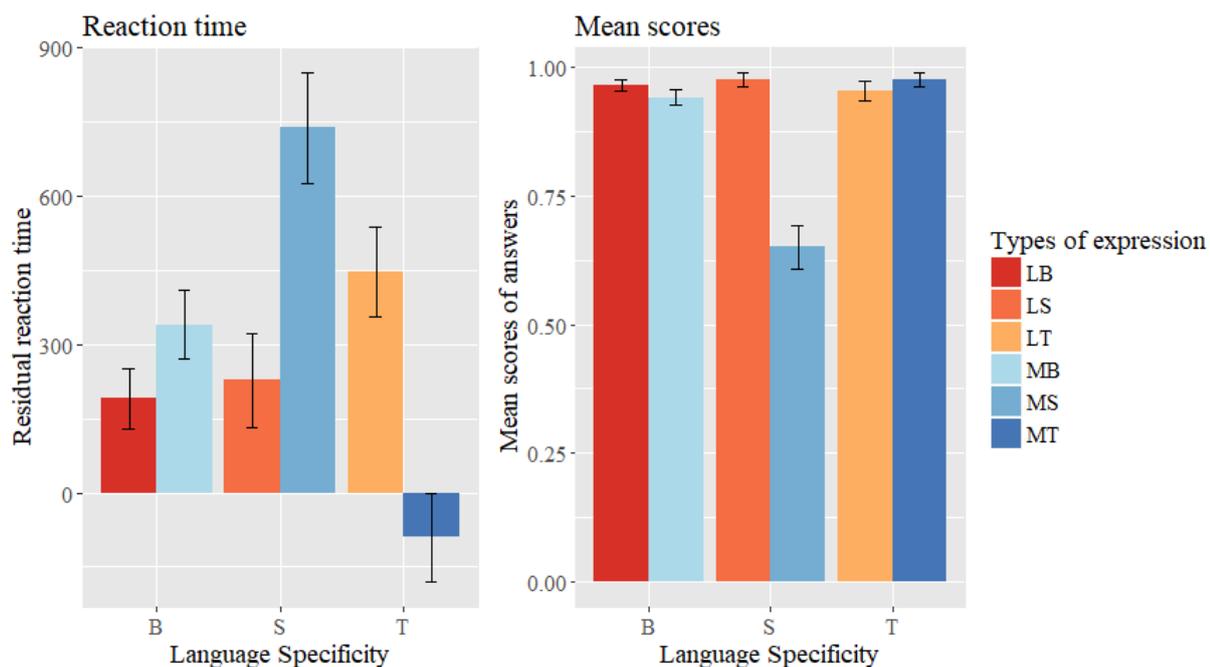


Figure 6.6 Reaction times (left) and mean scores for answers for the native group

The native group, as shown in Figure 6.6 above, did not show any difference when answering the LB and MB questions. However, significantly more time was spent answering questions related to MS sentences than those related to LS sentences ( $\chi^2(1)=4.21$ ,  $p=0.040$ ). There was also a clear difference in mean score for MS and LS questions ( $\chi^2(1)=4.05$ ,  $p=0.044$ ). The native group reacted significantly faster to questions related to MT expressions than those related to the LT questions ( $\chi^2(1)=11.96$ ,  $p<0.001$ ), but there was no difference in mean score for the two types of questions.

Compared with questions related to metaphorical expressions available in the target language (the MB and MT expressions), participants spent significantly more time answering MS questions ( $\chi^2(1)=5.38$ ,  $p=0.020$ ). However, when the mean scores for MS questions were analysed, it was found that they were significantly lower ( $\chi^2(1)=10.75$ ,  $p=0.001$ ). Although the mean score for MS questions ( $M=0.65$ ,  $SD=0.48$ ) was above chance, meaning the native participants could deduce the meaning of the MS expressions from the word form and the given sentence, it was still difficult for them to answer these questions accurately. Among the three types of metaphorical expression, the questions related to MT expressions were answered fastest: the reaction time for MT questions was significantly lower even than that for MB questions ( $\chi^2(1)=4.52$ ,  $p=0.033$ ). There was no difference in mean score for MB questions and

MT questions. It seems that native participants were best at identifying and processing the MT expressions, followed by the MB expressions.

The learner groups' reaction to the comprehension questions, including reaction time and scores for answers, formed a different pattern from the reaction time for the native group. The reactions of the intermediate group, as shown below in Figure 6.7, will be discussed first. These learners did not take more time answering questions related to metaphorical expressions than questions related to literal expressions. In particular, there was no difference in reaction time between the LS questions and MS questions, and a trend was even observed of slightly less time being spent reacting to MS questions, although no significant difference was found. While it seems that the scores of the MS questions were lower than those of the LS questions, there was, again, no significant difference between the two scores. No difference was observed, either, when the intermediate learners answered the LB/MB questions or the LT/MT questions.

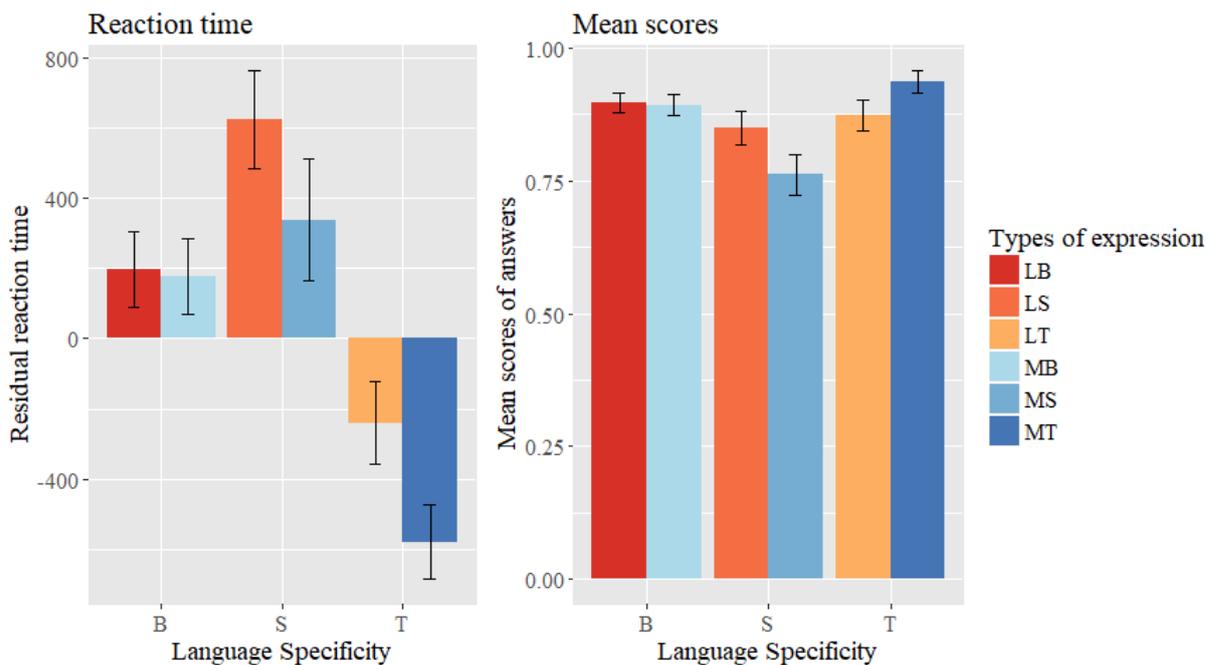


Figure 6.7 Reaction times (left) and mean scores for the answers for the intermediate group

Compared with the questions related to the metaphorical expressions available in English (the MB and MT expressions), the intermediate learners spent significantly more time answering the MS questions ( $\chi^2(1)=5.38, p=0.020$ ). Also, the scores for the MS questions were significantly lower than those for other metaphorical expressions available in English ( $\chi^2(1)=10.75, p=0.001$ ). It could possibly be inferred that, even in the case of the intermediate

learners, whose English proficiency is relatively low, there is some awareness that the MS expressions are not readily available in English when they read them.

The low-advanced group (see Figure 6.8), the high-advanced group (see Figure 6.9) and the overseas group (see Figure 6.10) displayed qualitative and quantitative similarities in terms of their reaction to the comprehension questions. No difference was observed in terms of reaction time and answers to the LB and MB questions. Participants from all three groups spent slightly more time answering the MS questions than the LS questions, and the scores for MS questions were slightly lower than for LS questions; however, such differences were not significant. All three groups also spent slightly less time answering MT questions than LT questions, but again, there was no significant difference between the two reaction times.

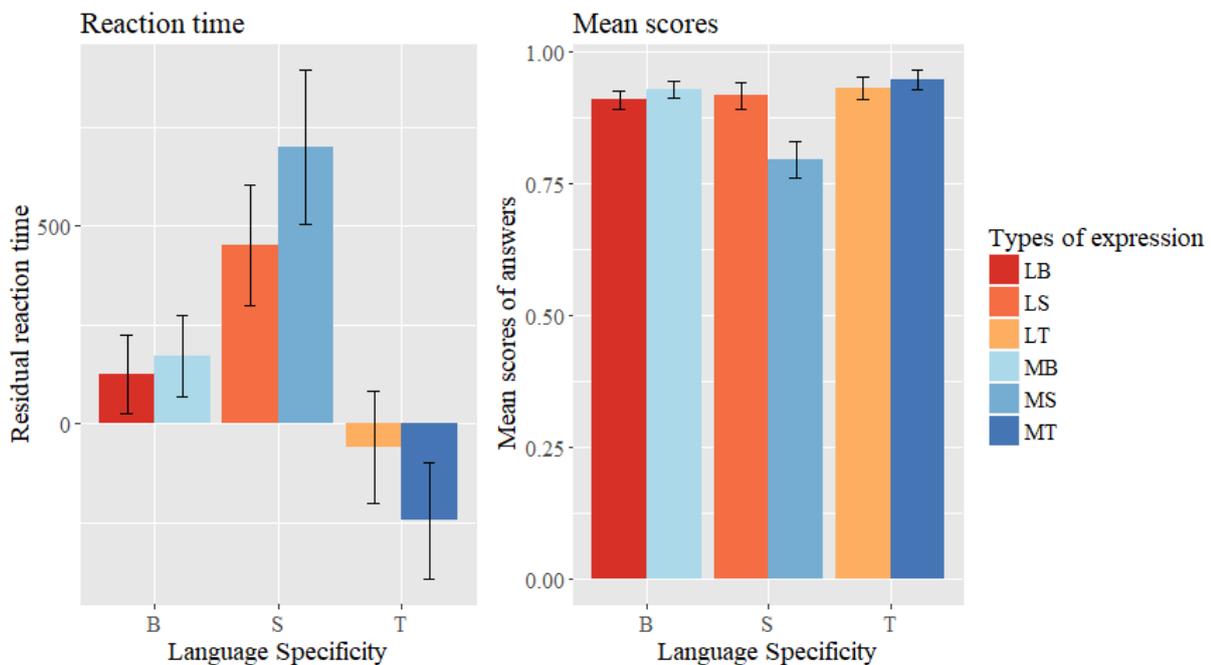


Figure 6.8 Reaction time (left) and scores for questions for the low-advanced group

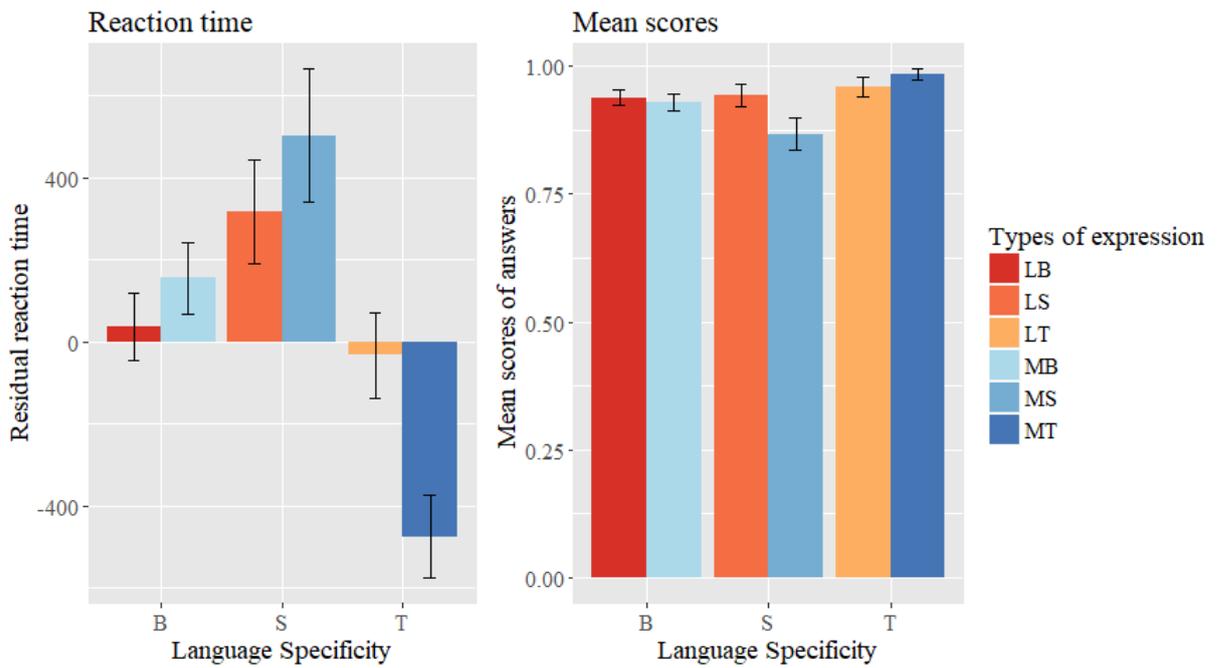


Figure 6.9 Reaction times (left) and scores for questions for the high-advanced group

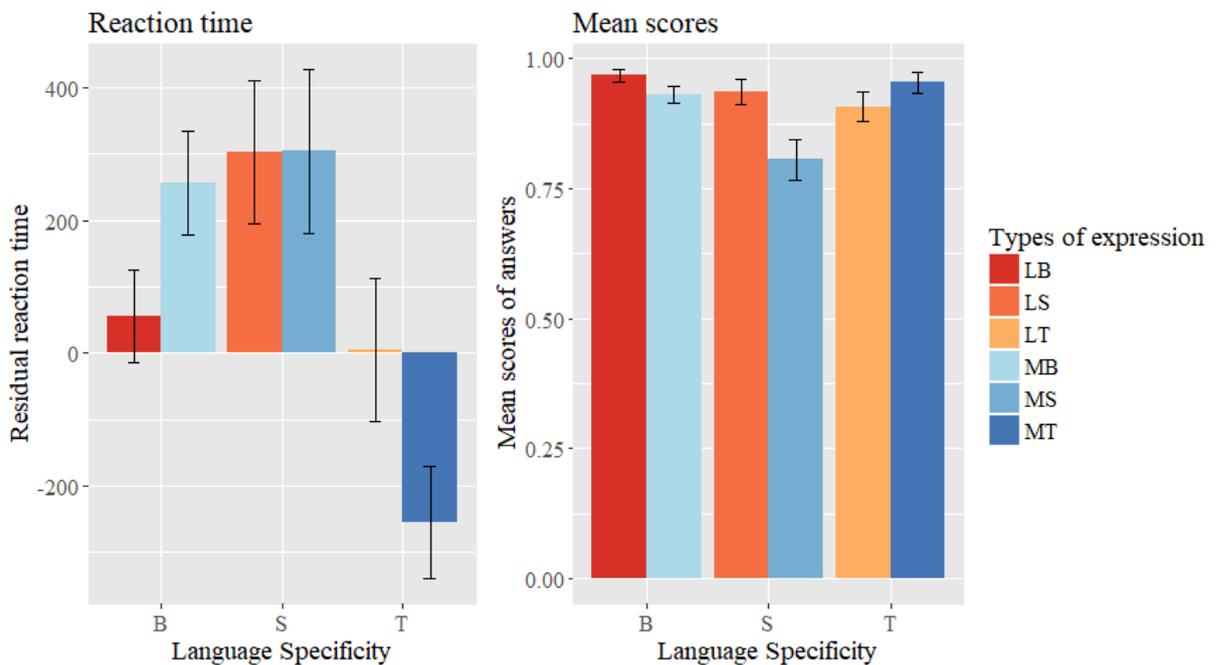


Figure 6.10 Reaction times (left) and scores for questions for the overseas group

When comparing across different types of metaphorical expression, it is found that, among the low-advanced learners and the overseas learners, the score for MS questions was significantly lower than that for MB and MT questions ( $\chi^2(1)=3.91, p=0.048$  for low-advanced;

$\chi^2(1)=4.11$ ,  $p=0.043$  for overseas). Conversely, the high-advanced learners spent marginally more time answering MS questions than MB and MT questions ( $\chi^2(1)=2.86$ ,  $p=0.091$ ), but the score for MS questions was not significantly lower than for the other two types of question. The high-advanced group spent marginally less time answering MT questions compared with MB questions ( $\chi^2(1)=3.34$ ,  $p=0.067$ ), and the difference became more significant in the overseas group ( $\chi^2(1)=7.23$ ,  $p=0.007$ ). There seems to be a trend for more proficient learners to be more likely to react faster to MT questions, which was similar to the reaction pattern of the native speaker group.

It can be concluded from these results that the learner participants performed markedly differently from the native participants, even though some of them had achieved high proficiency levels. The native participants showed significant difficulty when answering the MS questions, which can be reflected in both their reaction time and accuracy. Such difficulty was not shown among the learner participants. Although the learner participants showed clear hesitation after reading the MS expressions, they had already successfully resolved the meanings of the MS expressions by the time of answering the comprehension questions. This indicated that learners were able to build up a connection from the unacceptable L2 translation of the expressions to the L1-specific metaphorical meanings (c.f. Section 3.4), even though more time was required and learners did occasionally fail to establish the connection accurately.

There was, however, a close-to native-like performance that could be achieved by the learner participants when answering the comprehension questions. As proficiency rose, learners were gradually able to spend less time answering MT questions, while the accuracy of the answers of MT questions was also at a high level. All the learner groups showed a qualitatively similar pattern of reaction time to the LT/MT questions, and the overseas group even performed in a similar way to the native group: both spent significantly less time answering the MT questions than the LT questions. This phenomenon made an interesting contrast with the reading pattern for the MT sentences: whilst most of the learners hesitated after reading the MT expressions, they did not hesitate when answering the questions related to those expressions. The processing difficulty shown in the reading pattern did not necessarily lead to comprehension difficulty. This may suggest that the learners could actively and effectively infer the meanings of the MT expressions, and this process of inference may largely happen in the hesitation period. A further contrast can then be made between the results of the self-paced reading task and those of the acceptability judgment task reported in Chapter 5: even though learners could infer the meanings of the MT expressions, it was still difficult for them to accept this type of expression. It should be noted that the difference in results for the acceptability

judgment task and the self-paced reading task should be seen as a demonstration of general discrepancy between online and offline results rather than as a phenomenon specific to metaphorical expressions (c.f. Section 3.3).

## 6.4 Factors influencing subjective transferability and their influence on the processing of metaphorical expressions

### 6.4.1 Introduction

Previous studies (see Section 2.3 for a review) related to factors influencing subjective transferability are mostly offline studies, which means there is no known decisive evidence yet to show that the three factors proposed by Jordens and Kellerman (1981), namely psychotypology, markedness and L2 knowledge, can affect learners' performance of online tasks. In fact, neither subjective nor objective transferability of a linguistic element is defined in a way that can be used to interpret the data of an online experiment. It is not clear whether a learner using a transfer strategy would react faster or slower than another learner who has acquired that part of the knowledge. Such problems are particularly prominent in the case of MB expressions, because positive transfer and acquisition of the critical item can both lead to a similar short reaction time, and it is difficult to obtain a fine-grained boundary without the assistance of confidence levels. Therefore, it is less possible to produce a classificatory grid of cross-linguistic influence as shown in Section 5.4, and then to calculate the correlation between the presence of cross-linguistic influence and the factors that influence transferability.

Despite the fact that there is no established paradigm yet, we still would like to analyse the influence of all three influencing factors, namely psychotypology, markedness of a meaning and general knowledge of the target language, on the reading patterns for metaphorical expressions. This is to see what factors might affect learners' processing of metaphorical expressions in general. It has been observed in the previous sections that a difference in reading pattern was mainly reflected in the reading time for the segment after a metaphorical expression (i.e. Segment 3); therefore, in this section, we will look for correlations between the reading time for Segment 3 for different types of metaphorical sentences and the three possible influencing factors. Meanwhile, the reaction time for comprehension questions and the accuracy of answers for questions will also be compared to the influencing factors, to see if any of the factors might affect the comprehension of metaphorical expressions.

The influence of the three factors on the reading time for Segment 3 of different types of metaphorical sentences is predicted below:

- For the perception of psychotypology:
  - If learners perceive English as a language close to Chinese, they might expect metaphorical expressions to be the same in both Chinese and English. They should not show more hesitation in reading any type of metaphorical expression, except for MT expressions, when compared with reading literal expressions.
- For the markedness of a metaphorical meaning:
  - Learners should show more hesitation after reading metaphorical expressions that are considered more ‘marked’ than those considered less ‘marked’ in Chinese (cf. discussion of Jordens and Kellerman 1981 in Section 2.3.2). They should answer questions related to these more ‘marked’ metaphorical expressions more slowly and less accurately.
- As regards general knowledge of target language, as roughly estimated by the proficiency of individual learners:
  - If learners are more proficient in English, they might be more aware of the metaphorical expressions that are available in English. They should not show more hesitation in reading any of the types of metaphorical expressions, except MS expressions, when compared with literal expressions.

#### 6.4.2 Influence of psychotypological distance

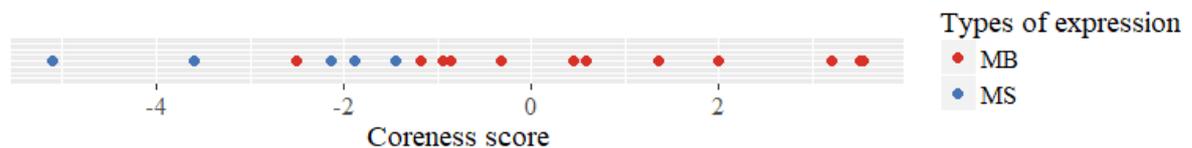
A correlation between learners’ perception of the psychotypological distance between Chinese and English and the reading time for Segment 3 of each type of metaphorical sentence was calculated to see if psychotypology had any influence on the processing of metaphorical expressions. The only statistical difference found among all the observations was that for overseas participants, those who perceived English as a language distant from Chinese showed a marginally longer hesitation after reading the MS expressions ( $\chi^2(1)=3.04$ ,  $p=0.081$ ). Other than that, no systematic trace of significant influence from psychotypological distance was observed in the reading experiment, either in terms of the reading time for Segment 3 or in terms of the reaction to comprehension questions.

The possible influence of psychotypology on the results of the self-paced reading task is drastically different from its influence on the results of the acceptability task. In particular, unlike with the results of the acceptability judgment task in Chapter 5, no influence from the perception of psychotypological distance was detected on reading time or on responses to MT

sentences. However, the absence of clear influence from psychotypical distance on the reading task cannot be used to argue against the possible influence of psychotypology on transfer strategy or the subjective transferability of a linguistic element. No matter what strategy a learner might use when processing L2 expressions, preference will always be shown for the strategy that involves the least time and effort. It is impossible for us to make a comparison of efficiency between transfer strategy and direct access in terms of individual learners' comprehension of metaphorical expressions. Therefore, we cannot use reaction time to prove that a learner is under the influence of psychotypical distance and thus prefers a transfer strategy.

#### 6.4.3 The influence of markedness of meanings

As in the discussion on Section 5.3.4 on the influence of markedness on the processing of metaphorical expressions, MB and MS expressions are discussed together because their combination leads to a longer continuum of markedness, and thus the influence of markedness will be more significant. The distribution of the coreness scores for the MB and MS expressions used in the self-paced reading task is shown below as Figure 6.11.



that the degree of markedness of a metaphorical expression in Chinese can affect the processing speed for that expression when it is presented in English. Learners generally encountered more difficulty when processing more marked expressions, which is reflected in the length of hesitation after reading the expressions. This could be seen as initial evidence from online experiments that a more marked linguistic element is less likely to be transferred by L2 learners.

Meanwhile, it was discovered in the current experiment that the degree of markedness of a metaphorical expression as perceived by English native speakers can also influence the reaction time for the comprehension questions. As we have observed in Section 6.3, there is a difference between the reaction time for MB and MT questions among the native speaker participants, even if both types of expression are available in English. This then leads us to assume that the degree of markedness of metaphorical expressions in English might influence the reaction time for the comprehension questions, and this is probably also the case for L2 learners. Figure 6.12 shows the distribution of the degree of markedness of the MB and MT expressions used in the reading task, and it may be observed that English native speakers generally perceived MT expressions as more marked than MB expressions on the scale ( $\chi^2(1)=4.35$ ,  $p=0.037$ ). It should be noted that the degree of markedness of an English metaphorical expression is unrelated to any type of cross-linguistic influence, since not all the English metaphorical expressions in the experiment are ‘transferable’.

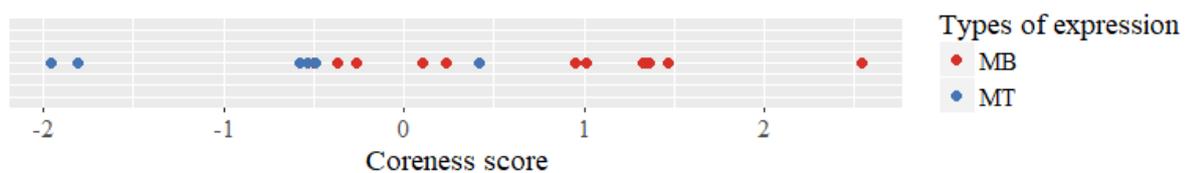


Figure 6.12 Distribution of the coreness scores for the MB and MT expressions in the SPRT, as perceived by English native speakers

The native group spent significantly less time answering the questions related to the more marked metaphorical expressions ( $\chi^2(1)=7.88$ ,  $p=0.005$ ); such influence was also observed marginally among the high-advanced participants ( $\chi^2(1)=2.93$ ,  $p=0.087$ ) and significantly among the overseas participants ( $\chi^2(1)=4.18$ ,  $p=0.041$ ). The result shows that both native speakers and highly proficient learners treated metaphorical expressions differently in terms of their degree of markedness in English. Questions related to more marked expressions were

generally answered faster by both native speakers and proficient learners than questions related to less marked expressions. This could provide some valuable information about learners' acquisition of L2 metaphorical expressions. Firstly, more proficient learners seem to know that some metaphorical expressions behave differently in their L2 when compared with other metaphorical expressions, even if, as discussed in Chapter 3, they only had limited input from the metaphorical expressions and they may not have a systematic picture of the markedness of meanings of all lexical items. Secondly and more importantly, the result may indicate that more marked metaphorical expressions are stored in a different way from less marked metaphorical expressions. This result contrasts with the prediction in Section 6.4.1, that if less marked expressions are easier to master and learners will react faster to them. However, if we compare the processing of metaphorical expressions with different degrees of markedness to the processing of other examples of figurative language (c.f. Section 3.3; see also Heredia and others 2007), we then can see that the more marked metaphorical expressions may be more likely to be memorised and processed as an entire chunk. In this way, both native speakers and learners reacted faster and more accurately to the meanings of more marked expressions, since these expressions can be directly accessed in the mental lexicon without additional semantic composition.

#### 6.4.4 Influence of general proficiency

In order to examine the impact of general proficiency on learners' processing pattern, a correlation between learners' OQPT scores and the reading time for Segment 3 of metaphorical sentences was calculated. The result shows that general proficiency did not correlate with the reading time for Segment 3 for any type of metaphorical expression. This means that more proficient learners did not hesitate more or less during or after reading any types of metaphorical expression.

This result diverges clearly from the predictions in Section 6.4.1, showing that general proficiency does not significantly influence the residual reading time for Segment 3 after each metaphorical expression. It also seems to contradict the discussion in Section 6.2, which shows that the difference in reading patterns gradually disappeared (1) between the LB and MB sentences and (2) between the LT and MT sentences, as proficiency rose.

This lack of difference can and should be analysed in terms of different types of metaphorical expression. It may be the case that learners could process the MB expressions without hesitating afterwards at a relatively earlier stage of acquisition, so the rise in proficiency

would not drastically affect the length of hesitation after reading the MB expressions. Similarly, it may also be the case that learners will experience persistent difficulty when processing the MT expressions even at a higher proficiency level, so the rise in proficiency cannot affect the length of hesitation after reading the MT expressions either. For the MS expressions, we may conclude that even less proficient learners were already sensitive to those expressions in real-time processing; therefore, there was no clear change as proficiency rose.

The reaction times for the comprehension questions and the scores for those questions were also analysed in terms of their correlation with learners' general proficiency for each type of metaphorical expression. As with the results for reading time, no significant correlation was found between learners' OQPT scores and their performance in comprehension questions, either in terms of reaction time or in terms of answer score. Such results indicate that even less proficient learners were able to answer questions related to metaphorical meanings quickly and accurately, so the rise in general proficiency level did not significantly improve their performance in this aspect.

## 6.5 Summary

The results of the self-paced reading task provided insights into the storage pattern and retrieval process for literal and metaphorical meanings in a learner's mental lexicon. In particular, it reflected a developmental trend of retrieval of metaphorical meanings between different proficiency levels, and some asymmetry. First was the asymmetry between literal and metaphorical meaning, which was prominent among less proficient learners. In this case, a learner would spend more time reading sentences with a metaphorical expression than sentences with a literal expression. If we interpret this as the learners trying to establish the metaphorical meaning of an expression taking longer than the literal meaning, it indicates that a less proficient learner may rely on the 'literal-first' strategy to understand a metaphorical sentence. Asymmetry was also displayed between metaphorical expressions available in English (i.e. the MB and MT expressions) and those that were not available in English (i.e. the MS expressions), especially when learner patterns were compared with the reading patterns of native speakers: learners were able to process the MS expressions according to a native-like pattern (i.e., with hesitation) even if they were less proficient, while they could only process the MB and MT expressions in a native-like way (with no hesitation) when they became more proficient in their L2. The final asymmetry was between the two types of metaphorical expression that were available in English: compared with the MB expressions, it took a learner more time to achieve

a native-like reading pattern for the MT expressions, again indicating that the MT expressions were more difficult to manage than other types of metaphorical expressions. The three types of asymmetry mentioned above correspond perfectly to the three-fold asymmetry reported in Chapter 5. Although the experiment in this chapter is an online one, and an online-offline discrepancy is expected in the results, the results of the two experiments still show high uniformity, which suggests that such asymmetry in the acquisition of metaphorical expressions is solid and pervasive.

It was observed in the reading pattern that all the learners experienced processing difficulty after reading the MS expressions in the test sentences. Such processing difficulty, however, did not lead to a failure of retrieval for the meanings of the MS expressions. The learner groups were still able to answer the questions related to the MS expressions both quickly and accurately, whereas native speakers of English struggled to answer questions related to the MS expressions. Such performance indicated that the MS meanings, though not available in the target language, could still be activated when the context supported that type of meaning, and a mental lexicon that integrates the word meanings of both the L1 and the L2 was practically possible. A proposal for that mental lexicon will be presented in Chapter 7.

## 7 General discussions and conclusions

### 7.1 Cross-linguistic influence and the acquisition of metaphorical expressions

The results of the experiments reported in Chapters 5 and 6 together show that Chinese learners of English made different judgments and displayed different reading patterns for metaphorical expressions compared to native speakers of English accomplishing the same task. In the judgment task, the native speakers accepted the MB and MT expressions but rejected the MS expressions. In the reading task, the native speakers read the MB and MT expressions without any processing difficulties, while they did experience processing difficulties when reading the MS expressions. They also showed difficulty answering the questions related to the meaning of the MS expressions.

Using the results of the native speakers as a reference point, we can see that Chinese learners of English performed both qualitatively and quantitatively differently from the native speakers. In the judgment task, the learners accepted the MB expressions, but to a lesser extent compared with the native speakers. They rejected the MS expressions in general; while less proficient learners rejected them to a lesser extent compared with the native speakers, more proficient learners were able to perform in a native-like way. The greatest difference was that all groups of learners seemed to reject the MT expressions, in sharp contrast to the native speakers. In the reading task, all except the intermediate learners read the MB expressions without any processing difficulty, but only high-advanced overseas learners were able to read the MT expressions without any processing difficulty. The intermediate learners showed significant processing difficulty when reading both the MB and MT expressions, and the low-advanced and high-advanced learners showed minor difficulties when reading the MT expressions. All learners showed processing difficulty when reading the MS expressions, but they had no difficulty answering the questions related to the meaning of the MS expressions after reading the entire sentences.

In the two experiments, full mastery of metaphorical expressions is indicated if native-like judgment and reading patterns are demonstrated. To show that one has acquired all three types of metaphorical expressions, one is expected to accept both the MB and MT expressions, processing them without difficulty, while rejecting the MS expressions and showing processing difficulty when reading those expressions. The discrepancy between the expected and the actual results for learner participants indicates that the learner participants have not yet fully acquired the metaphorical expressions, which leaves open the possibility for cross-linguistic influence to

affect the learners' performance. In the following, the acquisition of the three types of metaphorical expression and the effect of cross-linguistic influence will be discussed.

In the process of acquisition of MB expressions, which are shared between the L1 and L2, it is possible for a learner to transfer knowledge of MB expressions from the L1 to the L2; such transfer will not lead to any evident error in the L2. In the actual process of acquisition, as we have seen in Chapter 5, traces of positive transfer were identified in around 30% of the judgments made by the intermediate and low-advanced learners, showing that learners actively adopted transfer when they acquired the MB expressions. However, as shown in the classification of cross-linguistic influence in Section 5.4.1, the borderline between the outcome of cross-linguistic influence and the results of full mastery of knowledge is rather delicate. Since a learner shows positive transfer in a transferrable condition, the result of transfer can easily be converted to the mastery of relevant knowledge, especially if later confirmation is received from the L2 materials that the transfer is 'valid'. The effect of transfer will then disappear once the result of transfer is confirmed as 'valid' in the L2, and the learner no longer relies on transfer to judge or process the MB expressions. It is expected that the results for judgment and production will remain the same, and positive transfer will gradually 'fade out' after a learner fully acquires the metaphorical expressions. This expectation is confirmed by the results of the judgment task in Section 5.2 and the identification of cross-linguistic influence in Section 5.3: while the degree of acceptability of the MB expressions remains statistically the same across the four groups with different proficiency levels, more proficient learners show more confidence in their judgments, and the estimated percentage of instances of cross-linguistic influence on the MB expressions declines as proficiency increases. Such change is a good indication that learners gradually convert the result of positive transfer to actual knowledge of the language, and the influence of transfer, at the same time, likely 'fades out' from their performance.

As stated earlier, transfer of MB expressions is a form of positive transfer, and the outcome of transfer is equivalent to the outcome of full mastery of relevant knowledge. Compared with negative transfer, it is reasonable to assume that a learner is more likely to translate positive transfer into knowledge of lexical items, and such positive transfer will lead to and 'be replaced by' full mastery of knowledge even at an early stage of acquisition. Positive transfer, especially in a classroom setting, is more likely to be reinforced by (1) positive feedback from instructors and native speakers and (2) similar input received by learners. This assumption is confirmed by the results in Table 5.4: after the identification of traces of transfer, the estimated percentage of instances of transfer of the MB expressions is lowest among all three types of expressions

within each group of learner participants. Whilst the intermediate and low-advanced learners still heavily rely on transfer for their judgments of the MS and MT expressions, the percentage of instances of transfer for the MB expressions is generally lower among the two groups of participants. Considering that it cannot be the case that the MB expressions are generally seen as less transferable by learners when they begin acquisition than the MS and MT expressions, there should be an evident decline in the transfer of MB expressions at the beginning of acquisition. This may lead to two correlated arguments: (1) transfer of the MB expressions will be even more significant before a learner reaches intermediate level, and possibly at beginner level; and (2) compared with other expressions, the MB expressions are more easily acquired at an early stage of L2 acquisition, because a learner can effectively convert the content of transfer to knowledge that has been fully mastered.

Interestingly, although we have seen that more proficient learners showed more confident mastery of MB expressions, as the level of confidence was rather high in their judgments, they did not give a higher acceptability score to the MB expressions, and they never accepted the MB expressions to the same extent as the native speakers. It seems that the perception of the degree of acceptability of the MB expressions is ‘fossilised’ at a later stage of acquisition. When we use the performance of the native participants as a benchmark and suggest that the LB and MB expressions should both be highly acceptable, we can see that learners cannot reach native-like levels to eliminate the difference between the LB and MB expressions in terms of the degree of acceptability.

In contrast with the case of MB expressions, transfer of both the MS and MT expressions is negative: it will cause ‘non-nativeness’ in both reception (including judgments and reaction time) and production. However, transfers of the two types of expressions are drastically different in terms of the development of learners. The MS expressions do not exist in the learners’ L2, so transfer of those expressions should be regarded as ‘transfer to nowhere’. Any instance of transfer of the MS expressions will lead to errors related to L1-specific metaphorical expressions, especially the acceptance or production of the word-to-word translation of L1-specific metaphorical expressions. Such types of transfer can be easily suppressed by learners though. The result displayed in Table 5.4 shows that transfer of the MS expressions is rather significant when the learners are less proficient, but with a rise in the level of proficiency, the percentage of instances of transfer of the MS expressions shows a drastic decline, from nearly 60% among the intermediate learners to around 20% among the high-advanced and overseas learners. At the same time, more feedback sentences focusing on the MS expressions are found as the proficiency level of learners rises. It can be concluded that there is still a significant

impact on intermediate learners from the transfer of MS expressions, but the influence of such negative transfer is not long-lasting.

The decline in the influence of transfer on MS expressions can be attributed to several factors, most importantly increased exposure to the L2 and possibly the instruction received by learners. Since the MS expressions are not available in the learners' L2, learners will never encounter such expressions in the input, something that will become more obvious to them when their exposure to the L2 increases. Under such circumstances, learners are likely to suppress the use of the MS expressions in the L2 because they believe those expressions not to be possible. At the same time, the misuse of an MS expression can also be easily noticed by both learners and instructors. It is common for an instructor to emphasise that errors and misuses caused by so-called 'Chinglish' should be avoided (Ma 2009). We can also appreciate from such method of instruction that even if a specific MS expression is not explicitly mentioned, learner will more generally be advised by their instructors to avoid direct translation of conventionalised Chinese expressions to English in a word-for-word fashion. At a higher proficiency level, when learners are able to distinguish expressions shared between their L1 and L2 (i.e. MB expressions) and those only available in the L2 (i.e. MS expressions), they can then avoid the MS expressions but maintain the MB expressions in their use. It may then be expected that, due to both a learner's own caution and the emphasis made by the instructor, the misuse of word-for-word translation of L1-specific metaphorical expressions will significantly decrease as L2 proficiency increases.

After discussing the realisation of transfer of the MB and MS expressions, we can compare the two types of transfer, because both involve transfer of certain linguistic elements from the L1 (positively or negatively), and both become less evident as learners become more proficient. We can see that, when the learners reach intermediate level, traces of transfer of the MB become less obvious than for MS expressions. Transfer of the MB expressions has already largely disappeared after learners pass the beginner threshold, probably due to 'replacement' by actual knowledge, while transfer of the MS expressions is still clearly observable when they reach low-advanced level (corresponding to B2 in the CEFR). At a later stage, traces of transfer of MS expressions rapidly drops to the same level as for the MB expressions, and highly proficient learners are able to perform in a native-like way in terms of acceptability scores for the MS expressions and feedback sentences targeting MS expressions. These results suggest that, while positive transfer remains at a low level among learners above intermediate level, negative transfer in 'transfer to nowhere' conditions can be successfully suppressed by more proficient learners, which corresponds to our expectations. The suppression of negative transfer of the MS

expressions, however, does not sufficiently indicate that the meanings of the MS expressions cannot be accessed by advanced learners when they are forced to comprehend those expressions in their L2. We see in Chapter 6 that, by the time the learner participants had answered the comprehension questions related to the MS expressions, they had already understood those expressions and answered the questions correctly. This means that the learners were still able to access the meanings of the MS expressions when they are required to do so, suggesting a compound organisation of bilingual lexicon.

Cross-linguistic influence also affects the acquisition of MT expressions. In Chapter 4, it was predicted that cross-linguistic influence would have a negative impact on the acquisition of MT expressions; such impact was predicted to be temporary among less proficient learners, and it was expected to disappear among more proficient learners. This prediction is only partially correct, in that cross-linguistic influence does have a negative impact, but the results show that the impact is not temporary, but persistent. While native speakers generally accepted all MT expressions and processed them without significant difficulty, none of the learner groups showed full mastery of the MT expressions. Less proficient learners hesitated to accept the MT expressions, and more proficient learners entirely rejected them. All learners except those from the overseas group experienced processing difficulty when reading the MT expressions. As for traces of transfer of the MT expressions, we can see from Table 5.4 that learners at all levels of proficiency relied heavily on transfer to judge the MT expressions: more than one-third of the judgments of MT expressions made by the intermediate learners are influenced by their L1, and the proportion reached more than 50% for the high-advanced learners. The persistence of such negative influence on the judgment of MT expressions distinguishes it from the transfer of MB and MS expressions.

Here I would like to suggest that the distinction between the transfer of MT expressions and that of other metaphorical expressions is due to the difference in the content being transferred from the L1 to the L2. As we have seen before, transfer of MB and MS expressions always involves transfer of individual linguistic elements. Although transfer of MT expressions corresponds to what Kellerman refers to as ‘nothing to transfer’, there is indeed something being transferred from the L1 to the L2 in the process; what is transferred is the observation that ‘X, a linguistic element, is not possible in the L1’. This observation is part of learners’ knowledge of their L1, and when it comes to the L2, they extend that knowledge to form the belief that ‘since X is not possible in the L1, it should not be possible in the L2 either’, and thus naturally they assume that the presence of X is ‘wrong’ in the L2. Transfer of such observations means that some L2-specific metaphorical expressions are difficult to acquire, exactly as the

results of the judgment task in Section 5.2 and the results of the reading task in Section 6.2 suggest. Transfer of ‘nothing’ to MT expressions, or to be more precise, the failure of acquisition of MT expressions, shows a stronger cross-linguistic influence on acquisition than the transfer in the context of MB and MS expressions.

Whilst the cross-linguistic influence on the MB and MS expressions becomes less evident as proficiency rises, the cross-linguistic influence on the MT expressions seems to become even more significant. This is because it is more difficult to become aware of transfer of ‘the belief that there is no such linguistic element’ than transfer of ‘something’. Unless the availability of the MT expressions is explicitly pointed out in L2 exposure, learners will have little opportunity to know and use those expressions. In the case of guided acquisition, the instructors hardly have the opportunity to ‘correct’ the use of those expressions: if the students can express the intended meaning without using the MT expressions, then there is no imperative to introduce these expressions to the students. Here we might observe a deadlock in terms of the acquisition of the MT expressions: on the one hand, the lack of relevant knowledge of the MT expressions will drive the learners not to use them since they do not have other sources introducing them; on the other hand, when the learners rely on cross-linguistic influence to judge expressions, they constantly transfer ‘the belief that there is no such linguistic element’ to their L2, which leads to the rejection of the MT expressions and a lack of knowledge of those expressions. A deadlock is clearly reflected in the results of the judgment task and the sentence correction component: advanced learners confidently rejected the MT expressions (see Section 5.2), and then substituted them for other expressions that are (possibly) available in both their L1 and L2 (see Table 5.9 for examples).

One further question regarding transfer of the three types of metaphorical expression is how a learner perceives its subjective transferability. While the objective transferability of a linguistic element entirely depends on the availability of that element in a learner’s L1 and L2, learners will always hold their own perception of subjective transferability, which may differ from the actual objective transferability. We have observed that (1) a number of less proficient learners still accepted some metaphorical expressions that are not objectively transferable from the L1 to the L2, while at the same time (2) some learners chose to reject metaphorical expressions that are objectively transferable from the L1 to the L2. Both phenomena show that some learners may perceive the subjective transferability in a different way from the objective transferability of the metaphorical expressions, which requires a valid explanation.

Based on the estimated percentage of traces of transfer across the groups of learners (see Table 5.4) and the trend therein, an assumption can be made about how learners may establish their own sense of intuition to distinguish between the expressions they would like to transfer and those they do not want to transfer. As suggested in Section 2.3, subjective transferability should be seen as a continuum, and a learner can always adjust the subjective transferability of a linguistic element throughout the process of acquisition. A learner will have a relatively low L2 proficiency at the beginning of the acquisition process and may believe that every L1 expression has a similar subjective transferability. The ‘default’ degree of subjective transferability of every element should therefore be neutral on the continuum, rather than being set to ‘fully transferable’ or ‘not transferable at all’. With more understanding of the L2, adjustment will be made to the subjective transferability of each expression based on the linguistic materials the learner has been exposed to, and subjective transferability of different linguistic elements will vary at that stage, which will lead to the two observations made in the last paragraph.

Finally, it should be noted that, although (1) the metaphorical expressions are generally less acceptable and more difficult to process than the corresponding literal expressions, (2) effects of cross-linguistic influence are prominent in the judgment and reading time of metaphorical expressions, and (3) literal paraphrasing was extensively used in the sentence correction section, this does not sufficiently or necessarily indicate that the ability to use metaphorical expressions in the L2 is suppressed. The learner participants, especially the more advanced groups, are able (1) to find corresponding metaphorical expressions in the L2 to replace L1-specific metaphorical expressions; and (2) to create new metaphorical expressions in the L2 to express the intended meaning, despite the fact that cross-linguistic influence can also be observed in the metaphorical expressions they create. In particular, the percentage of alternative metaphorical expressions in the feedback sentences is similar across different groups of learners and the native group, indicating that learners are similarly productive when they apply metaphorical expressions, even if they are using their second language.

## 7.2 Factors influencing the transferability and transfer of metaphorical expressions

The current thesis explores how three major factors influence subjective transferability, as perceived by individual learners, and thus the acquisition outcomes for three types of metaphorical expressions, particularly the strategies of transfer when a learner acquires these expressions; meanwhile, the influence of the three factors, namely psychotology, markedness

and learners' knowledge, on the processing of metaphorical expressions will also be briefly discussed. While the concepts and definitions of the factors, or the so-called 'constraints on transfer', are adopted from Jordens and Kellerman (1981), I would like to emphasise that the current dissertation and the original experiments by Jordens and Kellerman are not comparable, even if they make use of the same factors in the discussions. Jordens and Kellerman, as well as other studies concerning psychotypology, compare two groups of learners who have the same L1 (Dutch in Jordens and Kellerman 1981) but different L2s (German or English in Jordens and Kellerman 1981). These studies also presume that (1) all learners with the same L1 background have the same or a similar perception of psychotypology, e.g. all the Dutch native speakers in the experiment are thought to believe that English is psychotypologically more distant than German when compared with Dutch; and (2) when other factors (e.g. the markedness of a linguistic element, or the proficiency levels of learners) are well controlled, all the learners should adopt the same transfer strategy to deal with a linguistic element, because they have the same perception of psychotypology. This view of the influence of psychotypology focuses on the difference between different pairs of L1 and L2, while the learners are treated collectively. The current dissertation, however, takes the individualist view of psychotypology: even for a single pair of L1 and L2, it is possible for individual learners to have their own perception of psychotypology. Even though we assume that psychotypology has a systematic effect across learners (i.e. it is easier for transfer to happen between languages that are perceived as psychotypologically close than between those perceived as distant), we still expect that some learners will decide to transfer their knowledge and some will not, due to individual differences in psychotypology.

The current dissertation presents several pieces of interesting evidence as well as some further hypotheses and controversies connected to the relationship between the three aforementioned factors and the cross-linguistic influence observed in the experiments. The first to be discussed is psychotypology, partly because it has been less frequently quantified in previous studies while this dissertation is among the first to measure the perception of psychotypological distance quantitatively. The psychotypological perception survey shows that, even though all the learners share the same L1 and L2 pairing, they do have varied perceptions of the psychotypological distance between Chinese and English. Although a majority of learners perceive English as a distant language from Chinese, there is a correlation in learners with more complicated linguistic experiences being more likely to perceive the distance between English and Chinese as relatively short (Xia 2017).

The impact of psychotypology on cross-linguistic influence in the course of the acquisition of metaphorical expressions is more complicated than what we have hypothesised. The most prominent issue here is that transfer is consistently observed from Chinese to English, while the two languages are generally considered psychotypologically remote. Kellerman (1983) suggests that transfer will be *suppressed* if the psychotypological distance between the L1 and L2 is too long. The conclusion of the previous section, however, shows that transfer is pervasive and sometimes persistent when a learner acquires metaphorical expressions, and as reported in Chapter 5, such transfer is hardly influenced by individual differences in psychotypological distance. In a word, transfer from Chinese to English in the context of the acquisition of metaphorical expressions survives the (generally) large psychotypological distance and is relatively stable across learners with different psychotypological perceptions. At first glance, the existence of such long-distance transfer seems to contradict Kellerman's claim, but it should be noted that 'suppress' above should be interpreted as 'making something less likely to happen', since Kellerman (as well as his colleagues) suggest subjective transferability is scalar rather than an exclusive choice between transfer or non-transfer, and psychotypology only influences the degree of subjective transferability. It is possible that on a collective level, the psychotypological distance between Chinese and English is longer than that between Dutch and English, and that at the same time transfer from Dutch to English is more likely to happen than transfer from Chinese to English. This argument, which might be the most accurate interpretation of the actual influence of psychotypology, is some distance from the possible misinterpretation: 'transfer from Chinese to English should not happen at all'.

The second issue regarding the influence of psychotypology is precisely the 'lack' of its influence in the results of the experiments. The correlation between an individual's perception of psychotypology and the extent of any traces of transfer appears only with multiple restrictions: less proficient learners who are more willing to transfer from their L1 in judging the MT expressions are found to perceive English as a language close to Chinese. In the reading task, only the highly proficient learners showed longer hesitation after reading the MS expressions when they perceived English as a language far from Chinese. Other than those two observations, the expected bidirectional correlation is largely absent in both the judgment task and the reading task. Compared with their peers, on the one hand, learners with a shorter psychotypological perception do not transfer more knowledge from their L1; on the other hand, learners who transfer more knowledge from their L1 do not necessarily have a shorter psychotypological perception. The absence of the influence of psychotypology is particularly prominent when learners judge or process MB expressions. The 'lack' of influence from

psychotypology is probably associated to a large extent with the difference between the collective view and the individualist view of psychotypology as discussed at the beginning of this section. Take the comparison between Chinese-English and Dutch-English again as an example. It may be the case that the psychotypological distance between Dutch and English is generally shorter than that between Chinese and English on a larger scale. It is likely that the shortest distance perceived between Chinese and English is still much longer than the longest distance perceived between Dutch and English, even if we take individual differences into consideration. This means the collective difference in psychotypology is more significant than the individual differences. Therefore, even though we have observed on an individual level that learners have different psychotypological perceptions, such differences may not be significant enough to trigger individual differences in the influence of psychotypology. The 'lack' of influence from psychotypology does not contradict the original proposal Jordens and Kellerman (1981), for the two studies measure and utilise the psychotypology of two languages in different ways.

The final question regarding psychotypology in this dissertation is the conditions under which the influence of psychotypology on transferability comes into effect in the acquisition of metaphorical expressions. To repeat, influence of psychotypology on learners' transfer strategies is found among less proficient learners when they judge the MT expressions. It can be observed that (1) the MT expressions are entirely unavailable in the learners' L1; (2) transfer regarding this linguistic element is negative transfer; and (3) the influence of psychotypology only affects the learners with relatively low proficiency levels. It is reasonable to assume that the influence of psychotypology on the transferability of metaphorical expressions will be more likely to happen when learners are less proficient. Since less proficient learners generally rely more on transfer in acquisition, psychotypology, as one of the major constraints, is more likely to affect their judgment of the transferability of a linguistic element, or in Jordens and Kellerman (1981), their 'strategies of transfer'. It is also feasible to assume that, in this specific area of acquisition of metaphorical expressions, psychotypology will tend to influence learners' attitude towards unknown metaphorical expressions. A learner who perceives the psychotypological distance between the L1 and L2 as greater may be more open to expressions that are not available in the L1 as a consequence of being less 'bound' by the (un)availability of these expressions in the L1. Meanwhile, a learner who perceives the L1 and L2 as two close languages would be more 'conservative' and less willing to take the risk. The second assumption could explain why the transferability of the MT expressions is affected by differences in psychotypology but the transferability of the MS expressions is not, although

both of these show negative transfer in the experiments. Finally, it should be noted that this kind of impact from psychotypology on cross-linguistic influence is exclusively summarised and analysed in the acquisition of metaphorical expressions, which are part of the domain of vocabulary and lexical meanings, while its impact may be essentially different in terms of the acquisition of other linguistic elements.

The second factor raised by Jordens and Kellerman is markedness. While markedness is mostly loosely defined in second language acquisition, this dissertation measures the degree of markedness of a metaphorical expression using a separate lexical evaluation task and uses the degree of coreness as the main measurement of markedness: a peripheral expression is more marked than a core expression. Unlike in Experiment 2 in Jordens and Kellerman (1981), this dissertation also takes the (binary) metaphoricalness of an expression into consideration, since it is doubtful whether metaphoricalness and markedness can be fully separated.

Section 5.4.3 has reported the influence of markedness on the transferability of a combination of MB and MS expressions, for both are metaphorical expressions available in the L1. The degree of markedness of a metaphorical expression does not influence the subjective transferability of that expression. Sections 5.4.3 and 5.4.4 jointly report the transferability of the LB and MB expressions, for these expressions are shared between the L1 and L2 and transferable. The results of the two subsections show that the influence of markedness in the current study is minimal. A metaphorical expression is less transferable than a literal expression, while a less marked expression is not necessarily more transferable than a more marked expression if they are shared by the L1 and L2. As suggested in Section 5.4.4, if a meaning is identified as ‘non-literal’ by a learner, it is then seen as less readily available in the L2, even if that meaning is available in both the L1 and L2, and transfer of the meaning will not lead to any error.

This outcome seems to contradict Jordens and Kellerman’s comparison between metaphoricalness and markedness. Jordens and Kellerman suggest that the subjective transferability of a meaning is not influenced by the abstractness, or metaphoricalness, of that meaning, only by its degree of markedness. Note that the LB and MB expressions are designed to have the same degree of acceptability in the L1 and the same availability in both the L1 and the L2; thus, it is expected that the two types of expression should have the same degree of objective transferability, as well as the same degree of acceptability in the L2. However, from the comparison between the LB and MB expressions in Sections 5.2 and 5.4.4, we can see that the acceptability perceived by learners and the subjective transferability of the two types of

meaning are significantly different. This result may indicate that metaphoricalness is essentially the factor influencing the subjective transferability of a meaning, which means that Jordens and Kellerman fail to distinguish the effect of metaphoricalness and the effect of markedness in their proposal. Alternatively, such contradiction may be caused by some methodological limitation in the current study, which is similar to the case of individual psychotypology. It is possible that the distinction between literal and metaphorical expressions is large enough to reflect the change of transferability effectively, while the difference in degree of markedness between the most marked expression and the least marked expression is not large enough to influence the transferability of metaphorical expressions significantly. However, this assumption cannot be tested in the current thesis, since there is no clear record of range of markedness in previous studies.

The final factor, i.e. a learner's knowledge of L2, seems to be the one with the most straightforward and consistent influence. The current study follows the tradition of previous research and measures a learner's knowledge of L2 using a separate proficiency test, i.e. the OQPT, and then compares a learner's reaction to metaphorical expressions with general proficiency level. It is shown that when learners become more proficient in their L2, they transfer less knowledge in their judgment of the MB and MS expressions, and process the MB and MT expressions with less hesitation. However, some aspects of the acquisition of metaphorical expressions remain fossilised even if a learner reaches a higher proficiency level. As with the discussion in the last section, the most significant fossilisation happens with MT expressions: even highly proficient learners rely on transfer to judge MT expressions, which creates a great obstacle for learners to achieve native-like judgments for negative transfer. The fossilisation of subjective transferability of the MT expressions leads to a reconsideration of the concept of 'knowledge' in the study of cross-linguistic influence. Knowledge of metaphorical expressions in the L2, which is crucial for the discussion of cross-linguistic influence in this dissertation, does not always develop alongside progress in L2 acquisition. Most of the factors relevant to improvement in L2 general proficiency, including the length of guided acquisition, seem only to correlate with part of the acquisition of metaphorical expressions. From a combination of the acceptability score and the estimated percentage of traces of transfer, it can be seen that the rejection of L1-specific metaphorical expressions increases as L2 proficiency rises, and a learner gradually gives up transferring unacceptable expressions from the L1. At the same time, learners do not realise what has not yet been acquired. A 'blank' is left in the mental lexicon where the L2-specific metaphorical expressions are, since there is nothing available for transfer. Even if a learner has encountered the L2-specific metaphorical meaning

of a lexical item in non-guided acquisition before the experiment or is able to infer the metaphorical meaning from the context under the time pressure, as is observed in the processing task, that learner may still not find it possible to integrate that metaphorical meaning into L2 knowledge. Without exposure to the use of those expressions frequently in daily life, as in the case of overseas students, or explicit teaching of these expressions in guided acquisition, the ‘blank’ will not automatically be filled, and will simply be left in the L2 as it is in the L1.

The gap between L2 general proficiency and improved knowledge of metaphorical expressions can provide some insights into research on both second language acquisition and language pedagogy. In investigating learners’ knowledge of L2 and the constraints of cross-linguistic influence, there are some situations in which general proficiency in the L2 cannot be a universal indicator of L2 knowledge, especially when the target of investigation is rather specific and sometimes absent from the content of language instruction. In these situations, general proficiency does not strictly match the knowledge of that target element of investigation, which may lead to a misinterpretation of the result. The lack of knowledge of metaphorical expression instruction even at a higher proficiency level indicates that the meaning and use of metaphorical expressions should be emphasised to Chinese learners of English when they acquire vocabulary. As a crucial part of daily language use, knowledge of metaphorical expressions is as important as other figurative expressions (e.g. idioms), while the meaning can be inferred from the literal meaning of a lexical item, which makes it more accessible to learners. Instructors could provide some well-formed L2-specific metaphorical expressions as a part of course material or encourage learners to derive the meaning of metaphorical expressions from the context using figurative thinking.

### 7.3 From asymmetry to symmetry: the difference between literal and metaphorical meanings of lexical items in second language acquisition and processing

While the previous two sections have focused solely on cross-linguistic influence on the acquisition and processing of metaphorical expressions, this section aims to discuss the similarities and differences between the literal and metaphorical meaning of individual lexical items in a learner’s L2 in developmental terms. As reviewed in Sections 1.2.2 and 1.2.3, if a literal meaning and a metaphorical meaning are available in language A at the same time, a native speaker of language A will accept the literal and metaphorical meanings to the same extent. If a metaphorical expression is ‘below the threshold’ in language A, the native speaker of language A will not sense the ‘metaphoricalness’ of that expression and will process the

metaphorical expression and its corresponding literal expression in the same way. Native speakers did not show any preference for the literal meanings of the lexical items, and they did not show a 'literal-first' pattern when processing the metaphorical expressions. In the experiment, the native participants accepted metaphorical expressions that were available in English (the MB and MT expressions) and their corresponding literal expressions (the LB and LT expressions) to a similar degree and showed the same pattern in processing these metaphorical expressions and their corresponding literal expressions. In a word, the judgment and processing of the literal and metaphorical expressions are symmetrical.

If we use native speakers' performance as a reference point, we can clearly see that L2 learners show a qualitative difference from native speakers. L2 learners at different levels of proficiency show a preference for literal meanings in the judgment task. At the same time, a large number of L2 learners show a 'literal-first' pattern when processing the metaphorical expressions. A learner who is given more time to consider the acceptability of a pair of literal/metaphorical expressions, will accept the literal expression without further question, but will be reluctant to accept the metaphorical expression, no matter whether it is available in the L1 or L2. Similarly, a learner, especially a less proficient one, who is asked to comprehend a pair of literal/metaphorical expressions under time pressure, will process the literal expression without any special effort, but will show hesitation after the metaphorical expression. Asymmetry is demonstrated here: a learner will show a preference for the literal use of a lexical item, while the metaphorical use is more difficult to acquire, accept and, as discussed in Section 7.2 ontransfer.

The asymmetry between the literal and metaphorical meanings of a lexical item shown among L2 learners, especially those at lower levels of proficiency, can be explained from several different perspectives. The first possibility, from a usage-based perspective, is that, overall, metaphorical expressions are used less frequently than literal expressions, so that learners have fewer opportunities to encounter, comprehend and memorise the metaphorical expressions. It has been observed in this thesis, yet not reported due to the limits of this dissertation, that if a metaphorical expression is perceived as 'frequently appearing' in English, then it is more likely to be accepted by learners. Meanwhile, the lexical evaluation task shows that in both languages the literal expressions are rated by native speakers as appearing 'significantly more frequently' than the metaphorical expressions ( $\chi^2(1)=17.80$ ,  $p<0.001$  for Chinese;  $\chi^2(1)=36.00$ ,  $p<0.001$  for English). The correlations between the acceptability and perceived frequency of a meaning show that learners have difficulty acquiring metaphorical

expressions partly because they are less frequent than literal expressions, which is likely to result in learners only having limited exposure to them.

One may argue that, even if learners receive less exposure to a metaphorical expression, they might still be able to ‘pick up’ that expression from the limited occasions on which they have encountered it. While this argument is theoretically and anecdotally possible, it may be particularly difficult for learners with certain educational settings and backgrounds. This leads us to the second probable explanation for difficulties in acquiring metaphorical expressions: learners in a classroom setting may have specific weaknesses in vocabulary acquisition, as described by Jiang (2000). To review Jiang’s observation briefly, compared with other learners who receive more natural and contextualised input, learners in a classroom setting are less able to infer the meaning of a lexical item based on the contextual information co-occurring with it. Since the metaphorical meaning of a lexical item always needs to be activated by contextual information (c.f. Section 1.2.1), a failure to utilise contextual information may lead to a failure of acquisition of metaphorical meanings, and the acquisition and retrieval of metaphorical meanings may be more difficult than the acquisition of the literal meanings of the same lexical item. This means that in a classroom setting, if a metaphorical expression is not deliberately taught, a learner will encounter some degree of difficulty when attempting to acquire an expression autonomously from occasional exposure.

Asymmetry between literal and metaphorical expressions decreases gradually as general proficiency rises, particularly in terms of real-time processing. Beginning at lower advanced level, learners no longer take more time to process the MB expressions compared with LB expressions. Although there is a persistent difficulty for less proficient learners processing the MT expressions, still, highly advanced learners with overseas experience can process MT expressions following a similar pattern to the LT expressions. The ‘literal-first’ hypothesis is no longer attested when a learner becomes more proficient, and gradually adopts a native-like processing pattern. Nevertheless, a highly proficient learner will still show a strong preference for the literal expression when asked to rate the acceptability of a pair of literal and metaphorical expressions, regardless of the availability of the metaphorical expressions in the L1. There seems to be a universal trend among learners: compared with smooth processing of metaphorical expressions, true acceptance of metaphorical expressions is to some extent delayed.

The discrepancy between preference in processing and preference in judgment may be due to the difference of nature of online tasks (e.g. self-paced reading task) and offline tasks (e.g.

untimed acceptability judgment task) (c.f. Section 3.3). However, even if we leave aside any possible task effect, we can still see that comprehending an expression and judging an expression make use of different mechanisms. A learner can understand a metaphorical expression without showing acceptance of that expression. The learner may simply infer the meaning of that expression, even under time pressure. We have seen similar performance among the native speakers of English when they read the MS expressions that are not available in English, and there is no reason to assume that learners cannot do the same in a second language. Therefore, a learner should be able to ‘understand’ the meaning of a metaphorical expression before accepting the validity of its use. If we connect this phenomenon to different types of knowledge mastered by a learner, this might constitute evidence to show the utilisation of implicit and explicit lexical knowledge: understanding meaning is more implicit, while accepting use and rating degree of acceptability rely more on explicit knowledge. Furthermore, such discrepancies also indicate that, even if a learner is able to process metaphorical expressions without significant difficulty, the metaphorical meanings are still in a secondary position in the mental lexicon, because they are less preferred when judgments of meanings are made.

A trend is observed in the processing task, especially in the reaction times for the comprehension questions, with both native speakers and learners at different levels of proficiency reacting faster to questions related to the meanings of the more marked metaphorical expressions in English. It seems that metaphorical meanings may be stored differently in the mental lexicon based on the degree of markedness. More marked meanings are stored and activated holistically with their typical collocations, and a reader can anticipate the collocations and the metaphorical meanings even before they are actually encountered; whilst less marked meanings are stored atomistically without the collocations, so that a reader should read the collocation and decide to activate the metaphorical meanings. The difference in the two types of storage and activation perfectly matches the difference between the storage of idiomatic meanings and that of literal meanings: idioms are retrieved holistically, while the literal expressions are retrieved compositionally. However, it should be noted that there is no definite borderline between the holistically stored, ‘more marked’ meanings and the atomistically stored, ‘less marked’ meanings, just as in the case of the variability of semantic transparency and idiomaticity that has been long discussed in research on idioms (see Grant and Bauer 2004; Vega-Moreno 2007 for extensive discussion). Individual native speakers or learners may store individual metaphorical expressions differently, but from a collective

perspective, more marked metaphorical expressions are more likely to be stored as a whole phrase.

To summarise, when a lexical item has both a literal meaning and a metaphorical meaning in a learner's second language, that learner, especially in a classroom setting, will develop a preference for the literal meaning over the metaphorical meaning. The learner will show a higher degree of acceptance of the literal meaning of that lexical item, and retrieve the literal meaning first when encountering the lexical item. This is partly because the literal meaning is more frequent in the available language materials, and partly because the literal meaning will usually have been acquired first in classroom teaching. When sensing that the literal meaning does not fulfil the requirement posed by the context (e.g. the collocation of that lexical item), the learner might reject the expression, or attempt to derive the possible metaphorical meaning with some processing effort. While a preference for the literal meaning is widely observed among learners, it can be overcome when (1) a metaphorical expression is relatively frequent, so that it is or can be acquired by a learner; or (2) a learner has reached a proficiency level at which implicit knowledge of the possible metaphorical expressions is possessed. Such observation indicates that a learner is less likely to acquire the metaphorical expressions autonomously, i.e. in a non-guided environment, than in a guided environment. As with other types of figurative language, metaphorical expressions should be 'taught' rather than left to learners to 'pick up'.

#### 7.4 Implications for a framework of literal and metaphorical meanings in a bilingual mental lexicon

As our discussion on the complexity of the acquisition of metaphorical meanings is coming to an end, we will now review previous hypotheses for the bilingual lexicon and fit the metaphorical expressions into the system properly. In order to acquire a full picture of meaning activation and selection in a mental lexicon, the exact timing of meaning activation will first be discussed, and a pair of test sentences from the self-paced reading task will be used as examples, as shown in Table 7.1, which repeats Table 4.3.

Table 7.1 A pair of sample test sentences and their structure (repeated from Table 4.3)

The	famous	doctor	built	a	balcony	for	his	home	near	the	city
The	famous	doctor	built	an	argument	in	his	article	in	the	book
W01	W02	W03	W04	W05	W06	W07	W08	W09	W10	W11	W12
Segment 01			Segment 02			Segment 03			Segment 04		
Subject			Literal/metaphorical expression			Spill-over area			Wrapping-up area		

As we have seen in both the judgment task and the reading task, for a native speaker of a language, for instance English, the conventional metaphorical meanings of a lexical item behave in the same way as the literal meaning of that lexical item. Since it is generally assumed in different frameworks of the bilingual lexicon that the literal meaning of a lexical item is directly linked to the word form, the conventional metaphorical meanings should also be directly linked to the word form. A native speaker who reads a lexical item without any critical collocation or other contextual information, e.g. a native speaker reading W04 in the sentences, will activate multiple meanings of that lexical item, both literal and metaphorical. After reading the whole collocation, e.g. when reaching W06 in the sentences, the contextual information will be sufficient for that lexical item to be disambiguated, and one of the meanings will be selected and the others suppressed. Since the intended meaning is already activated before it is selected, there will not be any additional cost in searching for the intended meaning, and thus there will not be any hesitation after reading the entire literal or metaphorical expressions. If a native speaker fails to activate a metaphorical meaning, e.g. the meaning of an MS expression that is not available in English, a persistent hesitation may be shown in a later stage of processing, and a delay may also be evident when the learner is asked to interpret that expression.

A learner of English, as discussed in the last section, may have difficulty activating metaphorical meanings because such meanings may be absent from the mental lexicon. Therefore, when encountering a lexical item such as the words in W04, the learner may only activate the literal meaning, and then re-process the whole expression after reading the entire collocation (after W06) because the activated meaning does not fit in the context. This may explain why learners at different levels of proficiency often show significant or marginal hesitation after reading a metaphorical expression, for they need to suppress the literal meaning and re-activate a new meaning, either from scratch or from their mental lexicon at a slower speed.

While it takes a native speaker the same amount of time to retrieve all the possible literal and metaphorical meanings that are known in the language, different amounts of time may be spent retrieving different types of metaphorical meaning. As reported in Chapter 6, such differences mainly depend on the availability of the metaphorical meanings in a learner's L1 and L2. In general, a learner will process the shared metaphorical expressions faster than the L2-specific metaphorical expressions, and have the greatest difficulty when processing the L1-specific expressions. The reason for the difficulty in processing L1-specific expressions is that the learner speculates on the possible lack of availability of those expressions in the L2.

Table 7.2 summarises all the differences in degree of acceptability, and Table 7.3 summarises all the differences in processing cost between each pair of expressions across different proficiency groups. Dark shading indicates that there is a significant difference between the degree of acceptability or processing cost for each pair of expressions, while light shading indicates that there is a marginal difference between a pair of expressions. The abbreviation 'n.d.' means there is no difference between two types of expression when they are judged or processed by the group.

Table 7.2 Differences in degree of acceptability between literal and different metaphorical expressions

Group	LB vs. MB	LS vs. MS	LT vs. MT	MB vs. MS	MB vs. MT
Intermediate	MB < LB	MS < LS	MT < LT	MS < MB	n.d.
Low-advanced	MB < LB	MS < LS	MT < LT	MS < MB	MT < MB
High-advanced	MB < LB	MS < LS	MT < LT	MS < MB	MT < MB
Overseas	MB < LB	MS < LS	MT < LT	MS < MB	n.d.
Native	n.d.	MS < LS	n.d.	MS < MB	n.d.

Table 7.3 Differences in processing cost between literal and different metaphorical expressions

Group	LB vs. MB	LS vs. MS	LT vs. MT	MB vs. MS	MB vs. MT
Intermediate	MB > LB	MS > LS	MT > LT	n.d.	n.d.
Low-advanced	n.d.	MS > LS	MT > LT	MS > MB	n.d.
High-advanced	n.d.	MS > LS	MT > LT	MS > MB	n.d.
Overseas	n.d.	MS > LS	n.d.	MS > MB	MT > MB
Native	n.d.	MS > LS	n.d.	MS > MB	n.d.

While the intermediate group shows a clear ‘literal-first’ trend for all types of metaphorical expressions, this trend is reduced for the MS expressions and partly for the MT expressions for more proficient groups. Starting with the low-advanced group, the learners can establish a direct link from the word form to the MB meanings; however, learners should become very proficient in establishing a direct link from the word form to the MT meanings.

There seems to be a conflict between the LB and MB pair and the LT and MT pair among the low- and high-advanced learners in Table 7.3. If a learner can (1) directly access the MB meanings and (2) spend the same amount of time accessing the MB and MT meanings, then why is the processing cost for the MT expressions marginally higher than that for the LT expressions? In order to provide a feasible explanation for this conflict, we must recognise that, even within direct access or ‘literal-first’ processing, there is a difference in processing speed. For example, a learner who has acquired a metaphorical expression but is less familiar with that expression, may take more time to access the meaning of that expression compared with another learner more familiar with that expression. In the case of the low- and high-advanced learners, they may be either (1) less familiar with the MB expressions, or (2) very sophisticated in constructing the MT meanings from the LT meanings and the contexts. Both conditions might lead to the absence of differences in processing cost between MB and MT expressions. Considering the consistency of the trend in development shown in Table 7.2 and Table 7.3, both assumptions are feasible and provide possible explanations for the aforementioned phenomenon among the low- and high-advanced learners. The difference in familiarity may also explain why overseas learners show a difference in processing cost between the MT and MB expressions while there is no difference in degree of acceptability. Although the overseas learners can directly access the MT meanings from the word form, they are less familiar with

the MT expressions than with the MB expressions, so they still take longer to access the MT expressions than the MB expressions.

After summarising assumed processing costs of different expressions, the next step is to fit the differences in processing cost into a lexicon model. The storage of literal and metaphorical meanings in a monolingual lexicon is represented as Figure 7.1. In such a model, when a metaphorical meaning is available in one language, we could assume that either the metaphorical meaning of a word can be directly accessed by native speakers without any activation of literal meaning (c.f. the Direct Access view), or the metaphorical meaning and literal meaning are activated in parallel, and native speakers always go for the most salient or context-relevant meaning (c.f. Graded Salience Hypothesis). Either of the assumptions could explain the fact that native speakers do not take more time to process the conventional metaphorical meaning of a lexical item than to process the literal meaning of the same lexical item.

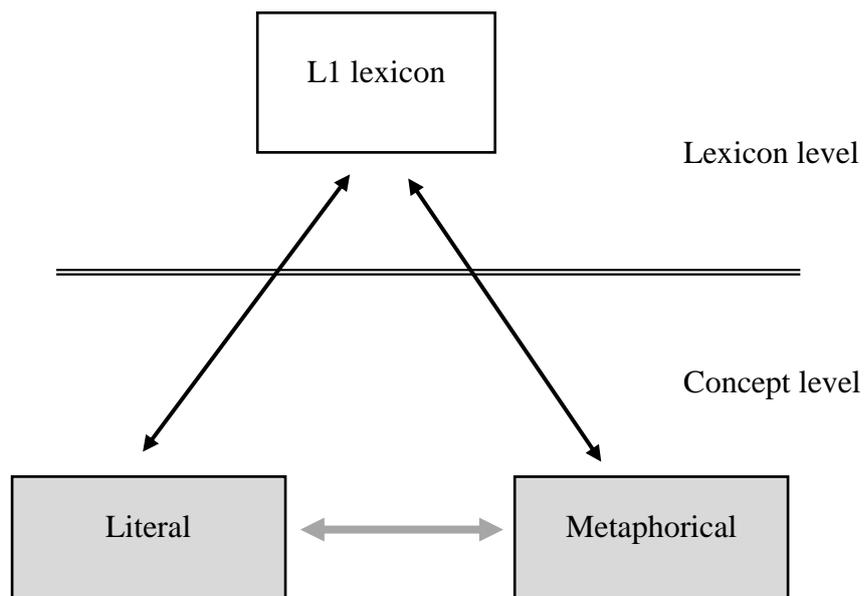


Figure 7.1 A monolingual literal and metaphorical lexicon

To take a further step, Figure 7.2, which replicates Figure 3.4, is a model of bilingual lexicon that is optimised for the acquisition of metaphorical expressions. This model, which is largely based on the Modified Hierarchical Model (Pavlenko 2009), is able to provide a developmental view of the processing cost of different metaphorical expressions, following a mandatory ‘literal-first’ principle at a lower proficiency level to symmetry between the literal and metaphorical meanings among highly proficient learners. A further elaboration of the current

model can be found in Section 3.4, while in the rest of this section, discussion will be focused on how learners at different levels establish the link from the word form to the intended metaphorical meanings.

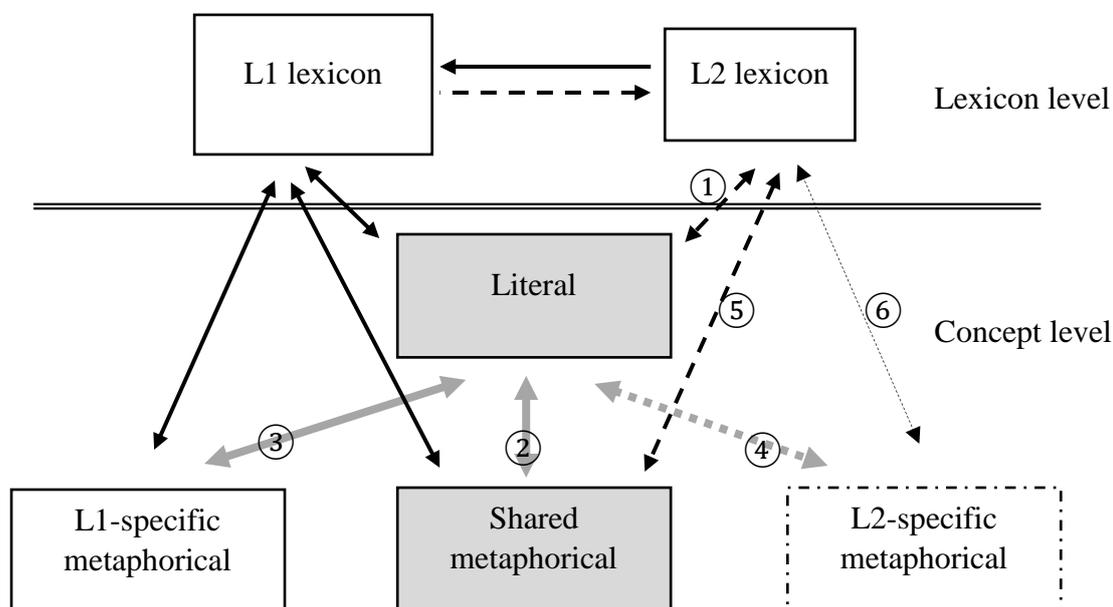


Figure 7.2 A bilingual literal and metaphorical lexicon with varied strength of links

The intermediate learners, as demonstrated in Table 7.2, generally show a smaller processing cost for the literal meanings than for all of the metaphorical meanings. An intermediate learner who always translates what is read and accesses all the concepts via the L1 lexicon, should experience the same processing cost for metaphorical expressions available in the L1 (i.e. the MB and MS expressions) and the corresponding literal expressions (i.e. the LB and LS expressions), which is not actually the case. Therefore, the more feasible proposal is that the learner arrives at the literal concept from the L2 lexicon directly, taking link ①, and then trying to generate all the metaphorical concepts from the literal concept via links ②, ③ and ④. The existence of links ②, ③ and ④ can ensure a valid explanation is provided for the additional cost for the various metaphorical expressions. This processing pattern for the intermediate learners corresponds to Bundle 3 in Table 3.5, in which all the metaphorical meanings are accessed via the literal meaning of a lexical item. This processing model is of medium complexity.

It is thought that low- and high-advanced learners cease to use link ② in Figure 7.2, because there is no additional cost for access to the shared metaphorical expressions; instead, they can directly access the literal meaning and the shared metaphorical meanings from the L2 lexicon

via link ⑤. As discussed above, they still encounter a marginal additional cost when accessing L2-specific metaphorical meanings. Two possibilities exist: they may derive L2-specific metaphorical meanings from a literal meaning, hence following links ①+④; or they may travel slowly through link ⑥ directly from the word form. Since the low-advanced learners show continuous hesitation after reading the MT expressions (c.f. Section 6.2.1), they are more likely to retrieve the L2-specific metaphorical meanings via links ①+④. Since the high-advanced learners only transit from links ①+② to link ⑤ when retrieving the shared metaphorical expressions, they may be relatively slow when passing through link ⑤, so that there will be no significant difference between the cost of links ①+④ and the cost of link ⑤. The processing patterns for the low- and high-advanced learners correspond to Bundle 4 in Table 3.5, with only the MB expressions being accessed directly.

The overseas learners and native speakers have qualitatively similar processing patterns. They directly access the shared metaphorical meanings from the word form via link ⑤, and they access the L2-specific metaphorical meanings from the word form directly via link ⑥. The only quantitative difference between the overseas learners and the native speakers is that they spend a different amount of time on link ⑥. The overseas learners may be less familiar with MT expressions since they have only just moved from links ①+④ to link ⑥; therefore, they are likely to spend more time directly accessing the L2-specific metaphorical meaning. The processing patterns of the overseas learners and native speakers generally correspond to Bundle 5 in Table 3.5, in which both the MB and MT expressions are accessed directly. The overseas learners may achieve near-native performance when processing the metaphorical expressions. It should be noted that, since the native speakers cannot have any access to link ③, they cannot effectively derive L1-specific metaphorical meanings; this explains the longer reaction time for the comprehension questions and the low accuracy of their answers for these expressions.

It can be seen from the explanation above that the strength of links ⑤ and ⑥ is crucial in the acquisition of metaphorical expressions: the acquisition of metaphorical expressions is completed only when these two links are established and, ideally, have the same strength as link ①. As proficiency level rises, link ⑤ can be gradually established and become stronger, although it is suspected that even among the most proficient learners the strength of ⑤ is still slightly lower than that of ① (c.f. the degree of acceptability of LB and MB expressions). The results of the two experiments suggest that link ⑥ is the weakest of all the word-concept links. Even if it is established, as is the case of the overseas learners, it is rather fragile compared with

the other links such as ① and ⑤. There are multiple factors that may lead to problems with link ⑥: (1) cross-linguistic influence from the L1 may suppress the establishment of link ⑥; and (2) learners, especially in a classroom setting, may not be good at making inferences for the meanings of a lexical item, and link ⑥ is therefore difficult to establish.

## 7.5 Main contributions and directions of further research

The current dissertation is among the first systematic works to explore the acquisition of conventional metaphorical expressions in a second language. It focuses exclusively on the acquisition of three types of metaphorical expression by Chinese learners of English in a classroom setting, but it is possible to develop it further as a universal paradigm to discuss the acquisition of shared, L1-specific and L2-specific metaphorical expressions.

While the current dissertation has a similar purpose to that of other studies (see García and others 2015 for a summary) on the acquisition of figurative language in a second language, it provides a more detailed picture in that field, including in the following aspects: (1) a comparison between the acquisition of literal and metaphorical meanings of a same set of lexical items; (2) an analysis of different realisations of cross-linguistic influence on metaphorical expressions; (3) an analysis of the factors influencing transferability and their impact on the acquisition and processing of metaphorical expressions, following work by Jordens and Kellerman (1981); (4) an attempt to place the metaphorical meanings of a lexical item in a bilingual mental lexicon; and (5) an attempt to explain how metaphorical expressions are accessed by L2 learners. It shows that, while cross-linguistic influence is observed in the acquisition of metaphorical expressions, and expressions with different availability are acquired according to distinct patterns, metaphorical expressions are nevertheless treated as ‘secondary’ by Chinese learners of English in a classroom setting, and these learners experience persistent difficulty in mastering metaphorical expressions. This result might provide pedagogical insights to both instructors and learners, which can be seen as an essential part of the improvement of metaphoric awareness proposed by Littlemore and Low (2006).

It has been pointed out in Section 7.2 that, although the factors influencing transferability are investigated here, the concepts involved are slightly different from Jordens and Kellerman’s original proposal, and thus any difference between the result here and that proposal may be caused by variances in concepts and the measurement of those factors. In future research, it might be possible to compare the acquisition of metaphorical expressions between learners with the same L1 and different L2s, or between learners who have acquired an L2 and an L3, to see

if the psychotypical distance collectively perceived by a group of learners between three languages has any impact on the acquisition of metaphorical expressions in general. Also, a further examination might be conducted to evaluate how much knowledge of conceptual metaphors (as per Lakoff and Johnson 1980) is involved in the acquisition of metaphorical expressions in an L2 among learners who have no theoretical background of cognitive semantics, and whether the conceptual metaphors themselves are transferable between languages. A comparison might also be drawn between child L1 and adult L2 acquisition of metaphorical expressions, to discover why highly advanced L2 learners still treat metaphorical expressions as 'secondary' while native speakers do not.

## Bibliography

- Aristotle. 1961. *Aristotle's Poetics*, Dramabook; D27, trans. by Samuel Henry Butcher (New York: Hill and Wang,), D27
- Athanasopoulos, Panos. 2015. 'Conceptual Representation in Bilinguals: The Role of Language Specificity and Conceptual Change', in *The Cambridge Handbook of Bilingual Processing*, ed. by John W. Schwieter (Cambridge: Cambridge University Press), pp. 275–92
- Bamgbose, Ayo. 1982. 'Standard Nigerian English: Issues of Identification', *The Other Tongue: English across Cultures*: 99–111
- Bard, Ellen Gurman, Dan Robertson, and Antonella Sorace. 1996. 'Magnitude Estimation of Linguistic Acceptability', *Language*, 72.1: 32–68
- Bardel, Camilla, and Christina Lindqvist. 2006. 'The Role of Proficiency and Psychotypology in Lexical Cross-Linguistic Influence. A Study of a Multilingual Learner of Italian L3' (presented at the Atti del VI Congresso di Studi dell'Associazione Italiana di Linguistica Applicata, Napoli, Italy), pp. 123–45
- Barr, Dale J., Roger Levy, Christoph Scheepers, and Harry J. Tily. 2013. 'Random Effects Structure for Confirmatory Hypothesis Testing: Keep It Maximal', *Journal of Memory and Language*, 68.3: 255–78
- Beardsley, Monroe C. 1962. 'The Metaphorical Twist', *Philosophy and Phenomenological Research*, 22.3: 293–307
- Black, Max, 1909-. 1962. *Models and Metaphors: Studies in Language and Philosophy* (Ithaca, N.Y.: Cornell University Press,)
- Blum, Shoshana, and E. A. Levenston. 1978. 'Universals of Lexical Simplification', *Language Learning*, 28.2: 399–415
- Boers, Frank, and Stuart Webb. 2015. 'Gauging the Semantic Transparency of Idioms: Do Natives and Learners See Eye to Eye?', in *Bilingual Figurative Language Processing*, ed. by Anna B. Cieślicka and Roberto R. Heredia (Cambridge), pp. 368–392

- Bortfeld, Heather. 2003. 'Comprehending Idioms Cross-Linguistically.', *Experimental Psychology*, 50.3: 217
- Bowdle, Brian F., and Dedre Gentner. 2005. 'The Career of Metaphor', *Psychological Review*, 112.1: 193–216
- Brysbaert, Marc, and Wouter Duyck. 2010. 'Is It Time to Leave behind the Revised Hierarchical Model of Bilingual Language Processing after Fifteen Years of Service?', *Bilingualism: Language and Cognition*, 13.3: 359–371
- Carston, Robyn. 2002. *Thoughts and Utterances: The Pragmatics of Explicit Communication* (Oxford: Blackwell)
- Chiappe, Dan L., and John M. Kennedy. 2001. 'Literal Bases for Metaphor and Simile', *Metaphor and Symbol*, 16.3–4: 249–76
- Cieślicka, Anna B. 2006. 'Literal Salience in On-Line Processing of Idiomatic Expressions by Second Language Learners', *Second Language Research*, 22.2: 115–44
- . 2008. 'Hemispheric Differences in Processing Salient and Nonsalient Meanings of L1 and L2 Fixed Expressions', *Lenguaje Figurado y Motivación: Una Perspectiva Desde La Fraseología*, Ed. MÀ de La Granja: 111–127
- Clahsen, Harald, and Claudia Felser. 2006. 'How Native-like Is Non-Native Language Processing?', *Trends in Cognitive Sciences*, 10.12: 564–570
- Collins, Allan M., and Elizabeth F. Loftus. 1975. 'A Spreading-Activation Theory of Semantic Processing.', *Psychological Review*, 82.6: 407
- De Angelis, Gessica. 2007. *Third Or Additional Language Acquisition (Multilingual Matters)*
- De Groot, Annette M. B. 1992. 'Bilingual Lexical Representation: A Closer Look at Conceptual Representations', *Orthography, Phonology, Morphology, and Meaning*: 389–412
- Dufour, Robert, and Judith F. Kroll. 1995. 'Matching Words to Concepts in Two Languages: A Test of the Concept Mediation Model of Bilingual Representation', *Memory & Cognition*, 23.2: 166–180

- Ellis, Rod. 2005. 'Measuring Implicit and Explicit Knowledge of a Second Language: A Psychometric Study', *Studies in Second Language Acquisition*, 27.2: 141–172
- Fauconnier, Gilles. 1994. *Mental Spaces: Aspects of Meaning Construction in Natural Language* (Cambridge University Press)
- Finkbeiner, Matthew, Kenneth Forster, Janet Nicol, and Kumiko Nakamura. 2004. 'The Role of Polysemy in Masked Semantic and Translation Priming', *Journal of Memory and Language*, 51.1: 1–22
- Frege, Gottlob. 1991. *Collected Papers on Mathematics, Logic, and Philosophy*, 1 edition (Oxford, UK ; New York, NY, USA: Wiley-Blackwell)
- Fu, Qian-Jie. 2012. *Digit Span* (Los Angeles: TigerSpeech Technology; University of California Los Angeles) <<https://itunes.apple.com/us/app/digit-span-test/id834048597?mt=8>> [accessed 13 February 2018]
- García, Omar, Anna B. Cieślicka, and Roberto R. Heredia. 2015. 'Nonliteral Language Processing and Methodological Considerations', in *Bilingual Figurative Language Processing*, ed. by Anna B. Cieślicka and Roberto R. Heredia (Cambridge), pp. 117–168
- Gibbs, Raymond W. 1984. 'Literal Meaning and Psychological Theory', *Cognitive Science*, 8.3: 275–304
- . 1995. *The Poetics of Mind: Figurative Thought, Language, and Understanding* (Cambridge England ; New York, NY, USA: Cambridge University Press)
- . 2017. 'Metaphor, Language, and Dynamical Systems', in *The Routledge Handbook of Metaphor and Language*, ed. by Elena Semino and Zsófia Demjén (New York: Routledge), pp. 56–70
- Gibbs, Raymond W., and Nandini P Nayak. 1989. 'Psycholinguistic Studies on the Syntactic Behavior of Idioms', *Cognitive Psychology*, 21.1: 100–138
- Giora, Rachel. 1997. 'Understanding Figurative and Literal Language: The Graded Salience Hypothesis', *Cognitive Linguistics*, 8.3: 183–206

- . 2002. 'Literal vs. Figurative Language: Different or Equal?', *Journal of Pragmatics*, Literal, minimal, and salient meanings, 34.4: 487–506
- Glucksberg, Sam. 2001. *Understanding Figurative Language: From Metaphors to Idioms*, Oxford Psychology Series; No. 36. (New York: Oxford University Press)
- Glucksberg, Sam, and Boaz Keysar. 1993. 'How Metaphors Work', in *Metaphor and Thought*, 2nd ed, ed. by Andrew Ortony (Cambridge: Cambridge University Press), pp. 401–24
- Gomes, Hilary, Walter Ritter, Vivien C. Tartter, Herbert G. Vaughan, and Jeffrey J. Rosen. 1997. 'Lexical Processing of Visually and Auditorily Presented Nouns and Verbs: Evidence from Reaction Time and N400 Priming Data', *Cognitive Brain Research*, 6.2: 121–134
- Grant, Lynn, and Laurie Bauer. 2004. 'Criteria for Re-Defining Idioms: Are We Barking up the Wrong Tree?', *Applied Linguistics*, 25.1: 38–61
- Grice, H. Paul. 1967. *Logic and Conversation* (Harvard)
- . 1989. *Studies in the Way of Words* (Cambridge, Massachusetts: Harvard University Press)
- Haspelmath, Martin. 2006. 'Against Markedness (and What to Replace It With)', *Journal of Linguistics*, 42.1: 25–70
- Heredia, Roberto R., Omar García, and M. R. Penecale. 2007. 'The Comprehension of Idiomatic Expressions by Spanish-English Bilinguals', in *48th Annual Meeting of the Psychonomic Society, Long Beach, CA*
- Irujo, Suzanne. 1986a. 'A Piece of Cake: Learning and Teaching Idioms', *ELT Journal*, 40.3: 236–42
- . 1986b. 'Don't Put Your Leg in Your Mouth: Transfer in the Acquisition of Idioms in a Second Language', *TESOL Quarterly*, 20.2: 287–304
- . 1993. 'Steering Clear: Avoidance in the Production of Idioms', *IRAL - International Review of Applied Linguistics in Language Teaching*, 31.3: 205–220

- Jarvis, Scott. 2009. 'Lexical Transfer', in *The Bilingual Mental Lexicon: Interdisciplinary Approaches*, ed. by Aneta Pavlenko (Bristol: Multilingual Matters), pp. 99–124
- Jegerski, Jill. 2013. 'Self-Paced Reading', in *Research Methods in Second Language Psycholinguistics*, ed. by Jill Jegerski and Bill VanPatten (New York: Routledge), pp. 20–49
- Jiang, Nan. 2000. 'Lexical Representation and Development in a Second Language', *Applied Linguistics*, 21.1: 47–77
- . 2004. 'Semantic Transfer and Its Implications for Vocabulary Teaching in a Second Language', *The Modern Language Journal*, 88.3: 416–32
- . 2007. 'Selective Integration of Linguistic Knowledge in Adult Second Language Learning', *Language Learning*, 57.1: 1–33
- Jin, Fufen. 2009. 'Third Language Acquisition of Norwegian Objects: Interlanguage Transfer or L1 Influence?', in *Third Language Acquisition and Universal Grammar*, ed. by Yan-Kit Ingrid Leung, pp. 144–61
- Jordens, Peter. 1986. 'Production Rules in Interlanguage: Evidence from Case Errors in L2 German', in *Crosslinguistic Influence in Second Language Acquisition*, ed. by Eric Kellerman and Michael Sharwood Smith (Pearson College Division), pp. 91–109
- Jordens, Peter, and Eric Kellerman. 1981. 'Investigations into the "transfer Strategy" in Second Language Learning', in *Actes Du 5e Congres de IAILA*, pp. 195–215
- Kaplan, David. 1989. 'Demonstratives', in *Themes from Kaplan*, ed. by Joseph Almog, John Perry, and Howard Wettstein (Oxford: Oxford University Press), pp. 481–563
- Keckés, Istvan. 2000. 'A Cognitive-Pragmatic Approach to Situation-Bound Utterances', *Journal of Pragmatics*, 32.5: 605–25
- Kellerman, Eric. 1979. 'Transfer and Non-Transfer: Where We Are Now', *Studies in Second Language Acquisition*, 2.01: 37–57
- . 1983. 'Now You See It, Now You Don't', *Language Transfer in Language Learning*, 54.12: 112–134

- . 1984. 'The Empirical Evidence for the Influence of the L1 in Interlanguage', *Interlanguage*: 98–122
- Kolers, Paul A. 1963. 'Interlingual Word Associations', *Journal of Verbal Learning and Verbal Behavior*, 2.4: 291–300
- Kövecses, Zoltán. 2005. *Metaphor in Culture: Universality and Variation* (Cambridge University Press)
- . 2017. 'Conceptual Metaphor Theory', in *The Routledge Handbook of Metaphor and Language*, ed. by Elena Semino and Zsófia Demjén (New York: Routledge), pp. 13–27
- Kroll, Judith F., and Annette M. B. De Groot. 1997. 'Lexical and Conceptual Memory in the Bilingual: Mapping Form to Meaning in Two Languages', in *Tutorials in Bilingualism: Psycholinguistic Perspectives*, ed. by Annette M. B. De Groot and J. F. Kroll (Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers), pp. 169–99
- Kroll, Judith F., and Erika Stewart. 1994. 'Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections Between Bilingual Memory Representations', *Journal of Memory and Language*, 33.2: 149–74
- Lado, Robert. 1957. *Linguistics Across Cultures: Applied Linguistics for Language Teachers*
- Lai, Vicky Tzuyin, Tim Curran, and Lise Menn. 2009. 'Comprehending Conventional and Novel Metaphors: An ERP Study', *Brain Research*, 1284: 145–55
- Lakoff, George. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*, New edition edition (University of Chicago Press)
- Lakoff, George, and Mark Johnson. 1980. *Metaphors We Live By* (Chicago: University Of Chicago Press)
- Lakoff, George, and Mark Turner. 1989. *More than Cool Reason: A Field Guide to Poetic Metaphor* (University of Chicago Press)
- Laufer, Batia. 2000. 'Avoidance of Idioms in a Second Language: The Effect of L1-L2 Degree of Similarity', *Studia Linguistica*, 54.2: 186–96
- Leezenberg, Michel. 2001. *Contexts of Metaphor* (Amsterdam: Elsevier)

- Leung, Yan-Kit Ingrid. 2005. 'L2 vs. L3 Initial State: A Comparative Study of the Acquisition of French DPs by Vietnamese Monolinguals and Cantonese–English Bilinguals', *Bilingualism: Language and Cognition*, 8.1: 39–61
- Libben, Maya R., and Debra A. Titone. 2008. 'The Multidetermined Nature of Idiom Processing', *Memory & Cognition*, 36.6: 1103–1121
- Link, Perry. 2013. *An Anatomy of Chinese: Rhythm, Metaphor, Politics* (Cambridge, Massachusetts: Harvard University Press)
- Liontas, John. 2002. 'Context and Idiom Understanding in Second Languages', *EUROSLA Yearbook*, 2.1: 155–85
- Littlemore, Jeannette. 2010. 'Metaphoric Competence in the First and Second Language', *Converging Evidence in Language and Communication Research (CELCR)*: 293
- Littlemore, Jeannette, and Graham Low. 2006. *Figurative Thinking and Foreign Language Learning* (Palgrave Macmillan)
- Liu, Dilin. 2002. *Metaphor, Culture, and Worldview: The Case of American English and the Chinese Language* (University Press of America)
- Ma, Qing. 2009. *Second Language Vocabulary Acquisition* (Peter Lang)
- Matlock, Teenie, and Roberto R. Heredia. 2002. 'Understanding Phrasal Verbs in Monolinguals and Bilinguals', *Advances in Psychology, Bilingual Sentence Processing*, 134: 251–74
- Meisel, Jürgen M. 2011. *First and Second Language Acquisition: Parallels and Differences* (Cambridge: Cambridge University Press)
- Miller, George A. 1969. 'A Psychological Method to Investigate Verbal Concepts', *Journal of Mathematical Psychology*, 6.2: 169–191
- Mitchell, Don C. 2004. 'On-Line Methods in Language Processing: Introduction and Historical Review', *The On-Line Study of Sentence Comprehension*: 15–32
- Nacey, Susan. 2014. *Metaphors in Learner English* (Amsterdam: John Benjamins Publishing)
- Neuser, Hannah. 2016. 'Investigating Distribution Patterns of Types of Transfer across Source Languages' (presented at the EuroSLA 26, Jyväskylä, Finland)

- Odlin, Terence. 1989. *Language Transfer: Cross-Linguistic Influence in Language Learning* (Cambridge University Press)
- Papadopoulou, Despina, Ianthi Tsimpli, and Nikos Amvrazis. 2013. 'Self-Paced Listening', in *Research Methods in Second Language Psycholinguistics*, ed. by Jill Jegerski and Bill VanPatten (New York: Routledge), pp. 50–68
- Pavlenko, Aneta. 2009. 'Conceptual Representation in the Bilingual Lexicon and Second Language Vocabulary Learning', in *The Bilingual Mental Lexicon: Interdisciplinary Approaches*, ed. by Aneta Pavlenko (Bristol: Multilingual Matters), pp. 125–160
- Pragglejaz Group. 2007. 'MIP: A Method for Identifying Metaphorically Used Words in Discourse', *Metaphor and Symbol*, 22.1: 1–39
- Quillan, M. R. 1966. *Semantic Memory* (Cambridge, Massachusetts: Bolt Beranek and Newman Inc)
- Recanati, François. 2004. *Literal Meaning* (Cambridge: Cambridge University Press)
- Ringbom, Hakan. 1978. 'What Differences Are There between Finns and Swedish-Speaking Finns Learning English?', *Papers and Studies in Contrastive Linguistics Poznan*, 8: 133–145
- . 2006. 'The Importance of Different Types of Similarity in Transfer Studies', in *Cross-Linguistic Influences in the Second Language Lexicon* (Clevedon: Multilingual Matters), pp. 36–45
- Rothman, Jason. 2013. 'Cognitive Economy, Non-Redundancy and Typological Primacy in L3 Acquisition', in *Romance Languages and Linguistic Theory 2011: Selected Papers from 'Going Romance' Utrecht 2011*, ed. by Sergio Baauw, Frank Drijkoningen, Luisa Meroni, and Manuela Pinto (Amsterdam: John Benjamins Publishing), pp. 217–47
- Rothman, Jason, Michael Iverson, and Tiffany Judy. 2011. 'L3 Syntactic Transfer Selectivity and Typological Determinacy: The Typological Primacy Model', *Second Language Research*, 27.1: 107–27
- Rutherford, William E. 1982. 'Markedness in Second Language Acquisition<sup>1</sup>', *Language Learning*, 32.1: 85–108

- Sanders, Lisa D., and Helen J. Neville. 2003. 'An ERP Study of Continuous Speech Processing: I. Segmentation, Semantics, and Syntax in Native Speakers', *Cognitive Brain Research*, 15.3: 228–240
- Sanford, Anthony J., Patrick Sturt, Linda Moxey, Lorna Morrow, and Cathy Emmott. 2004. 'Production and Comprehension Measures in Assessing Plural Object Formation', *On-Line Sentence Processing: ERPS, Eye Movements and Beyond*: 151–168
- Siddharthan, Advaith, and Napoleon Katsos. 2010. 'Reformulating Discourse Connectives for Non-Expert Readers' (presented at the Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Stroudsburg, PA, USA: Association for Computational Linguistics), pp. 1002–1010
- Singleton, David. 2016. 'Cross-Lexical Interaction and the Structure of the Mental Lexicon', in *New Perspectives on Transfer in Second Language Learning*, ed. by Liming Yu and Terence Odlin (Bristol: Multilingual Matters), pp. 51–62
- Siyanova-Chanturia, Anna, Kathy Conklin, and Norbert Schmitt. 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
- Sjöholm, Kaj. 1976. 'A Comparison of the Test Results in Grammar and Vocabulary between Finnish-and Swedish-Speaking Applicants for English', *Errors Made by Finns and Swedish-Speaking Finns in the Learning of English. Abo, Finland: Department of English, Abo Akademi.* (ERIC Document Reproduction Service No. ED122628)
- Spolsky, Bernard. 1985. 'Formulating a Theory of Second Language Learning', *Studies in Second Language Acquisition*, 7.3: 269–288
- Stanley, Jason. 2002. 'Nominal Restriction', in *Logical Form and Language*, ed. by Georg Peter and Gerhard Preyer (Oxford University Press), pp. 365–390
- Stanley, Jason, and Zoltán Gendler Szabó. 2000. 'On Quantifier Domain Restriction', *Mind and Language*, 15.2&3: 219–61
- Stern, Josef. 2000. *Metaphor in Context* (Cambridge, Massachusetts: MIT Press)

- Suhonen, Lari-Valtteri. 2016. 'Psychotypology in Conceptual Cross-Linguistic Influence: Data from Finland-Swedish Functional Bilingual Learners of English' (presented at the EuroSLA 26, Jyväskylä, Finland)
- Sweetser, Eve. 1990. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure* (Cambridge: Cambridge University Press)
- Titone, Debra, Georgie Columbus, Veronica Whitford, Julie Mercier, and Maya Libben. 2015. 'Contrasting Bilingual and Monolingual Idiom Processing', in *Bilingual Figurative Language Processing*, ed. by Anna B. Cieřlicka and Roberto R. Heredia, pp. 171–207
- Traugott, Elizabeth. 2004. 'Historical Pragmatics', in *Handbook of Pragmatics*, ed. by L. Horn and G. Ward (Oxford: Blackwell), pp. 538–61
- Türker, Ebru. 2016a. 'Idiom Acquisition by Second Language Learners: The Influence of Cross-Linguistic Similarity and Context', *The Language Learning Journal*: 1–12
- . 2016b. 'The Role of L1 Conceptual and Linguistic Knowledge and Frequency in the Acquisition of L2 Metaphorical Expressions', *Second Language Research*, 32.1: 25–48
- UCLES. 2001. *Oxford Quick Placement Test* (Oxford: Oxford University Press)
- Utsumi, Akira. 2007. 'Interpretive Diversity Explains Metaphor–Simile Distinction', *Metaphor and Symbol*, 22.4: 291–312
- Van Hell, Janet G., and Annette M. B. De Groot. 1998. 'Conceptual Representation in Bilingual Memory: Effects of Concreteness and Cognate Status in Word Association', *Bilingualism: Language and Cognition*, 1.03: 193–211
- Vega-Moreno, Rosa Elena. 2007. *Creativity and Convention: The Pragmatics of Everyday Figurative Speech* (John Benjamins Publishing)
- Wehmeier, Sally. 2000. 'Attack', *Oxford Advanced Learner's Dictionary* (Oxford; New York: Oxford University Press) [accessed 12 February 2017]
- Westergaard, Marit, Natalia Mitrofanova, Roksolana Mykhaylyk, and Yulia Rodina. 2016. 'Crosslinguistic Influence in the Acquisition of a Third Language: The Linguistic Proximity Model', *International Journal of Bilingualism*: 1367006916648859

- Wittgenstein, Ludwig. 1953. *Philosophical Investigation*, trans. by G. Anscombe (Oxford: Blackwell)
- Wong, Sau-Ling Cynthia. 1983. 'Overproduction, under-Lexicalization and Unidiomatic Usage in the Make Causatives of Chinese Speakers: A Case for Flexibility in Interlanguage Analysis', *Language Learning and Communication*: 151–163
- Xia, Mengying. 2015a. 'Layered Context in the Interpretation of Metaphor' (presented at the The 14th International Pragmatics Conference, Antwerp, Belgium)
- . 2015b. 'Towards an Indexicalist Semantic View of Metaphor: Theory, Strength and Limitation' (Cambridge)
- . 2017. 'Psychotypology of Chinese Learners of English and Its Influence on the Acquisition of Metaphorical Expressions: An Offline Study', *Cambridge Occasional Papers in Linguistics*: 237–55
- Xia, Violet, and Sally Andrews. 2015. 'Masked Translation Priming Asymmetry in Chinese-English Bilinguals: Making Sense of the Sense Model', *The Quarterly Journal of Experimental Psychology*, 68.2: 294–325
- Yu, Ning. 1998. *The Contemporary Theory of Metaphor: A Perspective from Chinese* (John Benjamins Publishing)
- Zhang, Di. 2008. 'A Study of Cross-Linguistic Transfer of Metaphor in SLA' (unpublished Master of Arts, Dalian: Liaoning Normal University)

## Appendix 1: Instruction and sentences used in the acceptability judgment task

Instruction for the acceptability judgment task:

Please imagine the following scene: your school/university now has an English writing club and you are an active member in it. Every week you need to write your own compositions and hand them to your teachers/advisors, and at the same time you will get some paragraphs from other students in the writing club, and you can point out the problems in their writing. The whole procedure is anonymous, so you will not know whose composition you are reading and marking, and the author of the composition will not know who the marker is either.

Please feel free to point out any part in the sentences that make you feel awkward or uncomfortable, and describe your feeling of acceptability by giving a score between 0 and 10 (0 for not acceptable at all, and 10 for totally acceptable; if you are not sure how to do it, you can see the examples below). If you believe that there is a better sentence, you can always write it down in the blank below the score, so your peers will know more about the mistakes they have made. After doing so, please indicate how confident you are when you mark and correct the sentences. Please do not use a dictionary and just follow your intuition, and make the judgment as quickly as possible.

The underlined lexical item is the designed critical lexical item in the sentence. Sentences are arranged in the alphabetical order of the critical lexical items.

MB (metaphorical-both) sentences:

He attacked every weak point in my argument.

I admire him because he is the father of modern biology.

The sudden death of his mother hit him hard and he cried for a long time.

Don't worry. There is always a market for good plans.

When he was young, the seeds of his great thoughts have been already planted.

When Judy came back after four years, she finally felt the warmth of her family again.

If she focuses on reading instead of wasting her time, she will be a better student.

He is very knowledgeable and I never won an argument against him.

LB (literal-both) sentences:

When everyone was asleep, the enemies attacked the village.

Lisa was happy when she got her birthday gift from her father.  
The car hit me hard in the accident and my leg was broken.  
I really enjoy taking a walk in the flower market in the morning.  
Growing vegetables from seeds may take some time, but it can save a lot of money.  
I sat back, enjoying the warmth of the sun on a Sunday afternoon.  
He wasted a lot of money on this useless machine.  
It is unlikely that the United States will win the Football World Cup this time.

MS (metaphorical-source) sentences:

Sally always bites the words and phrases whenever she writes an article.  
Sophie lost her golden bowl after her boss decided to shut down the company.  
Martin broke the crime and received an award from the police.  
He ate some loss when he started his own business.  
The teacher forced the fire in her heart down, and didn't let it break out.  
The mother held a belly of gas because her son failed in the exam.  
This incident might influence the fame of our brand, so we'd better handle it at a low pitch.  
Justin often sings an opposite tune when we discuss the homework.

LS (literal-source) sentences:

Sometimes I bite my nails when I become very nervous.  
The shiny golden bowl in the cupboard belongs to my elder brother.  
Laura found that a stranger broke her door and entered her room.  
My sister is very picky and always refuses to eat vegetables.  
Now some modern cars can use both gas and electricity as power sources.  
The brave boy saved his dog and helped put out the fire.  
When people want to sound more attractive, they will choose to speak at a low pitch.  
I heard someone singing a beautiful song in the corridor.

MT (metaphorical-target) sentences:

The idea is too stupid. He won't buy it.  
They exploded his last theory, so he felt really disappointed.  
My mother said that these books would be food for thought.  
The first two paragraphs prepare the ground for the main argument of the article.  
I see what your problem is, and I believe that I can solve it in a moment.  
I tried really hard, but just couldn't swallow his claim.

What he said left a really bad taste in my mouth.

Clare has a wealth of ideas and she always makes successful plans.

LT (literal-target) sentences:

He bought an old wooden clock on the Internet.

Someone just exploded a balloon suddenly and the loud sound scared everyone.

We cannot survive for long without enough food and drink.

Bob found his wallet lying on the ground near his bike.

I can see the boy playing with his classmates in the garden.

You can put the pills in your mouth and swallow them.

I lost my sense of taste because of the heavy cold, but I believe I will be better soon.

Peter inherited his family's wealth after the death of his father.

Fillers:

The best singer in this competition will receive a car from the host.

The toy became popular of children who enjoyed dressing it up.

The house shook so hard that she lost her footing and sat down upon the floor.

He received a lot of inspiration to the beautiful landscape.

The door was open, and she found that it led up a small passage.

She could not even get her head through the doorway.

He looked at her together a hurt expression.

In my surprise, we arrived at the airport on time despite of the heavy traffic.

Once upon a time, there was a princess living in a grand castle.

There was nothing in the glass table except a tiny golden key.

I dropped the key somewhere in here.

If you observe carefully, you will find that the shadow always moves with the sun.

I spent the whole afternoon watching a film on Chinese art.

A girl was thrown out of a motor vehicle and was killed in the accident.

He was only one between many who needed help.

I left a message for you towards your secretary.

She divided the fruit she bought equally in her four children.

My sister helped me move the books onto the second shelf.

He threw himself down into an armchair and smiled at me.

This book tells us a lot in the history of English drama.

On her return to the United States, Handler improved the design of the doll.

In order to save energy, please turn the computer out when you finish using it.

I introduced the girls to each other and they became friends quickly.

Hearing the call, Emma dropped her work and came after the door.

July is the hottest month in my hometown, and we all go picnic at that time.

With time passing by, I learn more about the secrets of cats.

## Appendix 2: Instruction, sentences and questions used in the self-paced reading task

Instruction for the self-paced reading task:

In this experiment, you will read some English sentences and provide judgments according to the contents of the sentences. Before each sentence, a "+" mark will appear on the screen to indicate the position of the first word of that sentence.

The first word of that sentence will appear after you press SPACEBAR, and the next word will appear when you press SPACEBAR again, while the previous word will disappear. You will read all the sentences in this way. Please be as natural as you could when reading, and understand the content of each sentence as quickly as possible.

After finishing a sentence, you will see a question, and please use keyboard to provide your judgment to that question according to the sentence. The key with a "Y" sticker means "Yes", and the one with an "N" sticker means "No". Please respond to all questions as quickly as possible.

The underlined lexical item is the designed critical lexical item in the sentence. The comprehension questions are not included in the appendix. Sentences are arranged in the alphabetical order of the critical lexical items. A comprehension question (CQ) is attached after each critical sentence.

MB (metaphorical-both) sentences:

The famous doctor built an argument in his article in the book.

CQ: Did the doctor only describe some facts in his article?

The pretty girl changed her mind before she went to the party.

CQ: Did the girl have a different idea when she went to the party?

The honest woman cleared her thoughts and told her story to people.

CQ: Did the woman refuse to talk about her story?

The busy musician found some time to take a very quick rest.

CQ: Did the musician have no time to rest?

The wise gentleman gave an idea to the students in the classroom.

CQ: Did the gentleman ask his student to provide an idea?

The little boy grabbed the chance and ran into the house quickly.

CQ: Did the boy use his chance effectively?

The happy lady lit the passion in the mind of her daughter.

CQ: Did the lady have no influence on her daughter?

The busy musician played a part in the creation of the song.

CQ: Did the musician contribute to the song?

The famous doctor saved his time by making a very compact schedule.

CQ: Did the doctor waste a lot of time?

The honest woman showed the solution to the question from the professor.

CQ: Did the woman explain the solution to a question?

The pretty girl spent some time on computer games with her friends.

CQ: Did the girl play computer for some time?

The young man won Anne's heart by writing a very sweet poem.

CQ: Did the man attract Anne with his own poem?

LB (literal-both) sentences:

The famous doctor built a balcony for his home near the city.

CQ: Did the doctor construct a balcony?

The pretty girl changed her clothes before she went to the party.

CQ: Did the girl wear different clothes when she went to the party?

The honest woman cleared her room and put everything into her suitcase.

CQ: Did the woman leave something in her room?

The busy musician found some food to make a very simple dinner.

CQ: Did the musician get some food?

The wise gentleman gave some money to poor people on the street.

CQ: Did the gentleman send food to poor people?

The little boy grabbed the knife and killed the crazy dog quickly.

CQ: Did the boy killed the dog with a knife?

The happy lady lit the fireplace in the bedroom of her house.

CQ: Did the lady start a fire in the fireplace?

The busy musician played the drums as a kid for several years.

CQ: Did the musician learn violin as a kid?

The famous doctor saved his wife by performing a very important procedure.

CQ: Did the doctor cure his wife effectively?

The honest woman showed a video about the story of her family.

CQ: Did the woman play a song about her family?

The pretty girl spent some money on new clothes for her dolls.

CQ: Did the girl use money to buy snacks?

The young man won that prize by writing a very sweet poem.

CQ: Did the man get that prize by composing a song?

MS (metaphorical-source) sentences:

The poor teacher bit her words when she was writing an article.

CQ: Did the teacher choose her words carefully in her writing?

The young man broke the crime and received an award from police.

CQ: Did the man fail to solve the crime?

The naughty student chewed his tongue while his classmates discussed their homework.

CQ: Did the student keep quiet when his classmates were working?

The rich manager ate some loss when he started his own business.

CQ: Did the manager get profit at the beginning of his business?

The old chef fried some stocks when he was not at work.

CQ: Did the chef buy and sell some stocks?

The naughty student raised his heart when he walked on the ice.

CQ: Did the student feel anxious when he walked on the ice?

LS (literal-source) sentences:

The poor teacher bit her nails when she was worried about something.

CQ: Did the teacher paint her nails when she was nervous?

The young man broke the door and brought the sleeping baby outside.

CQ: Did the man open the window of that room?

The naughty student chewed his gum while his classmates discussed their homework.

CQ: Did the student have some chewing gum?

The rich manager ate some fruits because they were healthy for him.

CQ: Did the manager have fruits in his diet?

The old chef fried some rice when he was not at work.

CQ: Did the chef cooked some rice?

The naughty student raised his hand when the teacher asked a question.

CQ: Did the student stay still when his teacher asked a question?

MT (metaphorical-target) sentences:

The happy lady bought an idea because it sounded clever to her.

CQ: Did the lady reject the idea?

The rich manager drove his business all the way to the top.

CQ: Did the manager run a successful business?

The poor teacher lost her temper and shouted at her students angrily.

CQ: Did the teacher keep calm to her students?

The young man paid some attention to the question from his colleagues.

CQ: Did the man ignore the question from his colleagues?

The wise gentleman saw the problem and gave a piece of advice.

CQ: Did the gentleman understand the problem correctly?

The naughty student spilled the secret when he entered the meeting room.

CQ: Did the student tell the secret to other people?

LT (literal-target) sentences:

The happy lady bought a wallet because it looked lovely to her.

CQ: Did the lady sell a wallet?

The rich manager drove his motorbike all the way to his house.

CQ: Did the manager take a taxi to home?

The poor teacher lost her wallet and was locked outside her apartment.

CQ: Did the teacher fail to open the door?

The young man paid some money to the owner of the store.

CQ: Did the man give some money to the shopkeeper?

The wise gentleman saw a schoolgirl and gave a book to her.

CQ: Did the gentleman meet a girl?

The naughty student spilled the coffee when he entered the meeting room.

CQ: Did the student drink the coffee in the meeting room?

Fillers:

The busy musician started his practice around half past seven every morning.

The busy musician went to sleep around half past eleven every night.

The famous doctor included a graph in an article in his book.

The famous doctor recorded a case in an article in his book.

The famous doctor took some photos in a village near the city.

The famous doctor decorated his office by hanging photos on the walls.

The happy lady painted the walls when she lived in the house.

The happy lady chose that plan because it sounded smart to her.

The happy lady owned the book when she was only a child.

The happy lady accepted the offer because it sounded great to her.

The honest woman told a story about the life of her mother.

The honest woman stopped a car when she walked to the park.

The honest woman closed her door and got back to work again.

The honest woman caught a bird and showed it to her friends.

The little boy greeted his neighbour when he walked to the park.

The little boy joined the club and was amazed by the events.

The little boy visited the castle and was amazed by the treasure.

The little boy watched the show and felt very excited about it.

The little boy received a present from an old friend of his.

The little boy read the book and felt very happy about it.

The little boy sang a song when he walked to the park.

The little boy met his classmate and had a chat with her.

The naughty student took a nap while his classmates discussed their homework.

The naughty student drew a picture while his classmates discussed their homework.

The old chef kept the fish and showed it to his friends.

The old chef kept the money and refused to give it away.

The poor teacher dropped her phone and cannot call her best friend.

The poor teacher prepared the paper carefully before she made the booklets.

The poor teacher forgot her umbrella and was wet in the rain.

The poor teacher prepared the photos carefully before she made the poster.

The pretty girl read some books before she went to the party.

The pretty girl earned some money by selling her handmade hats online.

The pretty girl wrote some articles about studying in a foreign country.

The pretty girl met her friend and had a chat with him.

The rich manager did some exercise because it was healthy for him.

The rich manager watched some movies when he was not at work.

The rich manager made some decisions when he started his own business.

The rich manager received the ring and wore it on his finger.

The wise gentleman sent a postcard to a student in that city.

The wise gentleman drank some tea with his friends in his house.

The wise gentleman read a book about the history of a country.

The wise gentleman sold his book to a student in his class.

The young man chose the suitcase and carried it to his house.  
The young man called the doctors and described the problem to them.  
The young man got a knife and opened the huge box carefully.  
The young man met some friends and played the guitar with them.  
The young man borrowed the camera and took it to his house.  
The young man had some friends and always hang out with them.  
The busy musician helped his colleagues as the leader in his team.  
The busy musician started his career as a teen for twenty years.  
The famous doctor organised his work by making a very compact schedule.  
The famous doctor treated his patient by performing a very effective procedure.  
The famous doctor recorded some videos in a village near the city.  
The famous doctor decorated his bedroom by hanging pictures on the walls.  
The happy lady cut the lawn when she lived in the house.  
The happy lady chose that handbag because it looked pretty to her.  
The happy lady owned a cat when she was only a child.  
The happy lady accepted the gift because it looked beautiful to her.  
The honest woman closed the window and got back to sleep again.  
The honest woman wrote a story about the life of her mother.  
The honest woman caught a bug and showed it to her friends.  
The honest woman stopped a man when she walked to the park.  
The little boy received a question from a new classmate of his.  
The little boy joined the team and was amazed by his teammates.  
The little boy visited the museum and was amazed by the exhibits.  
The little boy completed the homework and felt very proud of it.  
The little boy greeted a man when he walked to the park.  
The little boy sang a song when he did his math homework.  
The little boy met his neighbour and had a chat with her.  
The little boy watched the show and felt very happy about it.  
The naughty student played his game while his classmates discussed their homework.  
The naughty student wrote a joke while his classmates discussed their homework.  
The old chef held the insect and showed it to his friends.  
The old chef kept the secret and refused to tell other people.  
The poor teacher dropped her watch and lost track of the time.

The poor teacher prepared the materials carefully before she wrote the essay.  
The poor teacher forgot her glasses and cannot read the books clearly.  
The poor teacher prepared the ingredients carefully before she baked the cookies.  
The pretty girl took some photos before she went to the party.  
The pretty girl met her teacher and had a chat with her.  
The pretty girl made some effort when selling her handmade hats online.  
The pretty girl wrote some novels about studying in a foreign country.  
The rich manager enjoyed the sunshine when he was living in Greece.  
The rich manager watched some musicals when he was not at work.  
The rich manager enjoyed the coffee when he was in the café.  
The rich manager received the watch and wore it on his wrist.  
The wise gentleman drank some wine with his colleagues in his house.  
The wise gentleman sent a letter to a friend in that city.  
The wise gentleman read a magazine about the history of a country.  
The wise gentleman sold his pen to a friend in that city.  
The young man took the apple and cut it into two halves.  
The young man climbed a tree and took a scared cat down.  
The young man got a flashlight and walked into the dark cave.  
The young man chose the orange and cut it into two halves.  
The young man met some friends and went to have a drink.  
The young man had some friends and always sent messages to them.

### Appendix 3: Lexical property scores for the critical lexical items

Lexical property scores rated by Chinese native speakers:

Lexical item	Literal			Metaphorical		
	Concreteness	Coreness	Perceived frequency	Concreteness	Coreness	Perceived frequency
attack	5.82633	4.4274	3.891	-6.9819	0.117	4.5026
bite	5.8988	2.087	5.943	-7.81165	-1.884	-1.1004
bowl	6.0092	2.39	4.7666	-6.8566	-1.4369	6.7192
break	-0.16993	-0.6552	1.984	-5.10876	-1.4458	3.9426
build	5.3495	3.3191	6.031	-5.23283	1.3573	3.0036
change	5.57523	1.195	5.418	-5.86484	3.2018	4.9436
chew	6.0057	3.2796	6.209	-4.91618	-3.6048	-0.1124
clear	5.6995	2.0282	5.595	-6.53022	-0.9401	2.4036
eat	5.38031	3.2181	6.517	-8.78353	-2.5093	6.6968
father	6.04466	1.9245	6.164	-5.58029	-1.2006	4.6066
find	4.05575	0.4868	5.907	-5.85947	-0.8592	6.3327
fire	6.23508	2.7528	6.056	-8.21035	-2.3974	7.0574
fry	6.02597	2.0764	5.727	-6.45958	-5.1048	5.3813
gas	5.12229	-0.7983	2.474	-8.67792	-0.6749	6.6406
give	5.40671	3.1062	6.434	-3.20561	0.4415	6.7262
grab	6.41508	1.6994	4.0841	-7.94997	-2.5072	5.8297
hit	5.12113	3.1471	6.069	-7.61792	1.2321	6.675142
light	6.08868	3.6803	4.1991	-8.29239	-1.1863	2.5787
market	6.21625	1.7934	6.063	-6.49971	0.3762	4.7016
pitch	5.72485	1.6842	1.203	-7.39737	-3.4429	3.6176
play	5.85573	3.4017	5.135	-1.79351	2.0013	4.9402
raise	6.28033	2.2706	5.772	-7.63044	-2.1414	2.6486
save	-0.40462	3.4193	6.65	-0.85354	3.5412	6.2116
seed	6.37895	3.1227	5.005	-7.49968	0.3855	3.6076
show	4.54543	2.2854	5.001	4.4079	3.4992	4.3366
sing	5.54837	3.7592	6.253	-6.84813	-4.0898	3.1596
spend	-1.29724	-0.5261	7.414	-4.78579	-0.3153	7.356
warmth	1.09021	4.2433	5.882	-7.99794	1.4557	6.683
waste	0.67763	2.4315	6.368	-3.68628	2.4929	7.2043
win (1)	1.8292	3.9843	6.009	1.62894	4.4171	6.2166
win (2)	1.8292	3.9843	6.009	-6.92766	0.5862	4.2896

Footnote: The metaphorical meanings of ‘win’ in the acceptability judgment task and the self-paced reading task are different. Therefore, two different metaphorical versions of ‘win’ were examined in the lexical evaluation task, (1) for the AJT version and (2) for the SPRT version. The literal meaning of ‘win’ in the two tasks are the same.

Lexical property scores rated by English native speakers:

Lexical item	Literal			Metaphorical		
	Concreteness	Coreness	Perceived frequency	Concreteness	Coreness	Perceived frequency
attack	6.96222	4.79165	4.985	-2.23299	0.58693	2.996
build	6.91029	5.39747	4.89668	-4.62748	1.32336	-0.09032
buy	6.33739	4.76017	5.987	-4.48432	-1.80827	2.71
change	4.93221	2.88143	5.04158	-4.07764	2.54343	5.98158
clear	5.39275	3.27493	2.993	-5.09954	1.33876	1.5079
drive	6.70816	4.28102	5.6482	-4.83117	0.41525	-0.4238
explode	7.08368	5.11843	1.656	-5.64038	0.32546	-4.691
father	6.80744	4.7553	6.012	-4.92316	0.98474	1.421
find	6.62536	3.33527	4.976481	-3.7023	-0.36685	4.877481
food	6.39517	5.07086	5.9007	-5.38246	-0.32865	0.5127
give	6.06869	4.89809	5.8439	-1.65167	1.4654	4.2022
grab	6.81795	4.27247	4.94641	-5.71287	1.36528	0.19141
ground	6.1521	4.69314	5.4369	-5.06949	-0.36672	0.1699
hit	6.59071	4.12384	4.5606	-4.43679	-0.66861	3.8786
light	5.5639	4.48074	4.4938	-5.13998	-0.49967	-1.8082
lose	6.17639	4.71474	5.1463	-4.05251	-0.49139	4.9553
market	7.2007	3.96919	3.533	-3.67078	1.11089	1.8747
pay	6.17507	4.94483	5.5897	-3.86152	-1.95548	5.1707
play	5.32397	2.78393	5.5181	-0.67382	-0.26376	4.2395
save	4.85798	4.21603	4.0003	-2.54162	0.23236	5.0043
see	5.99858	5.11721	5.4772	-2.56472	-0.57465	5.8902
seed	7.26521	5.00998	3.927	-6.14803	1.11321	0.548
show	4.44655	4.6286	5.3984	3.4668	1.01337	4.6494
spend	5.93701	4.47464	5.5984	-2.48323	0.1064	5.8764
spill	6.75773	4.83644	5.0007	-5.08138	-0.53396	0.5307
swallow	6.98839	5.01018	5.3297	-5.88035	-1.2625	0.2547
taste	4.48032	4.49854	3.813	-5.15986	-0.25007	1.0436
warmth	4.93315	5.71586	4.1643	-5.34929	0.12458	2.10163
waste	1.3769	3.41856	5.928	0.20119	0.86042	5.989
wealth	4.64058	4.85016	3.813	-4.95126	1.04848	1.4649
win (1)	4.2081	4.04553	6.088	-0.4092	2.79053	4.3918
win (2)	4.2081	4.04553	6.088	-3.78974	0.95219	1.522

Footnote: The metaphorical meanings of ‘win’ in the acceptability judgment task and the self-paced reading task are different. Therefore, two different metaphorical versions of ‘win’ were examined in the lexical evaluation task, (1) for the AJT version and (2) for the SPRT version. The literal meaning of ‘win’ in the two tasks are the same.