THE EFFECTS OF METAPHOR AWARENESS ON L2 LEARNERS' RETENTION OF ENGLISH IDIOMATIC PHRASES

BRANDON CHAI AN WENG

FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR

2019

THE EFFECTS OF METAPHOR AWARENESS ON L2 LEARNERS' RETENTION OF ENGLISH IDIOMATIC PHRASES

BRANDON CHAI AN WENG

DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ENGLISH AS A SECOND LANGUAGE

FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR

2019

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate	: Brandon Chai An Weng
Registration/Matric No	: TGB 150036
Name of Degree	: Master of English as a Second Language
Title of Dissertation	: "The Effects of Metaphor Awareness on L2 Learners' Retention of English Idiomatic Phrases"
Field of Study	: Language Acquisition

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date

Subscribed and solemnly declared before,

Witness's Signature Date

Name:

Designation:

THE EFFECTS OF METAPHOR AWARENESS ON L2 LEARNERS' RETENTION OF ENGLISH IDIOMATIC PHRASES

ABSTRACT

This study sought to examine the effects on idiom retention by raising the metaphor awareness of L2 learners of English. The results presented positive pedagogical implications for intentional (formal) vocabulary teaching. An explanatory sequential research design, with one experimental group and one control group was employed in this study, $N = 42$. A within-group analysis of the experimental method revealed that between the pre and delayed post-test levels (strict scoring), the experimental group's participants' retention of the taught idiomatic phrases had improved by 31%, although statistical significance was not achieved. A between-groups analysis comparing the experimental and conventional methods of teaching idioms showed that the overall differences between the two groups were not statistically significant as well, under both strict and lenient scoring schemes. However, qualitative analyses revealed mainly positive attitudes toward the experimental method, while mostly negative attitudes were attributed to the conventional "blind memorisation" method for learning idioms.

Keywords: metaphor awareness, L2 learner, retention, intentional vocabulary teaching, idiom

KESAN-KESAN PENINGKATKAN KESEDARAN METAFORA TERHADAP PENGEKALAN FRASA- FRASA IDIOMATIK BAHASA INGGERIS BAGI PELAJAR L2

ABSTRAK

Kajian ini dijalankan untuk mengkaji kesan-kesan pada pengekalan frasa-frasa idiomatik melalui peningkatan kesedaran metafora pelajar-pelajar L2 Bahasa Inggeris. Keputusan-keputusan yang diperolehi melalui penganalisian maklumat telah mendedahkan pelbagai implikasi positif bagi pengajaran perbendaharaan kata secara formal. Reka bentuk penyelidikan "explanatory sequential", dengan satu kumpulan eksperimental dan satu kumpulan kawalan digunakan dalam kajian ini, $N = 42$. Analisis secara dalam-kumpulan pada kaedah eksperimental telah mendedahkan bahawa antara peringkat pra-ujian dan ujian pasca-tertunda (pemarkahan analitik "ketat"), para peserta kumpulan eksperimental telah mencapai peningkatan markah sebanyak 31% pada pengekalan frasa-frasa idiomatik yang diajar. Walau bagaimanapun, kesignifikan statistik tidak tercapai. Analisis secara antara-kumpulan yang membandingkan kaedah eksperimental dan konvensional dalam pengajaran frasa idiomatik telah menunjukkan bahawa perbezaan secara keseluruhan antara kedua-dua kumpulan tidak mempunyai kesignifikan statistik, melalui skema pemarkahan analitik bercorak "ketat" dan "ringan" kedua-duanya. Walau bagaimanapun, analisis kualitatif mendedahkan persepsi positif terhadap kaedah experimental, manakala kebanyakan persepsi negatif dikaitkan dengan kaedah "penghafalan buta" yang merupakan cara konvensional dalam pembelajaran frasa-frasa idiomatik.

Kata kunci: kesedaran metafora, pelajar L2, pengekalan, pengajaran perbendaharaan kata secara formal, idiom

ACKNOWLEDGEMENTS

I thank God for these people in my life:

To my family, thank you for always being there for me.

To Dr. Patricia Riget, thank you for being such a supportive dissertation supervisor. Your keen observation and patient demeanour have made a lasting impact in both my personal and professional life.

To my friends and all the Faculty of Languages and Linguistics staff who have helped me in the grimiest of academic and administrative issues, thank you so much.

No man is an island.

Thank you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ANOVA	:	Analysis of Variance
CEFR	:	Common European Framework of Reference for Languages
CL	:	Cognitive Linguistic(s)
CLT	:	Cognitive Load Theory
CMT	:	Conceptual Metaphor Theory
CRT	:	Criterion Referenced Test
ESL	:	English as Second Language
L2	:	Second Language
LoP	:	Levels of Processing
LTM	:	Long-Term Memory
MWE	:	Multi-Word Expressions
NRT	:	Norm-Referenced Test
RQ	:	Research Question
SD	:	Standard Deviations
SLA	:	Second Language Acquisition
STM	:	Short-Term Memory
TESL	:	Teaching English as a Second Language
WM	:	Working Memory

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1 Background of Study

Idioms typically present themselves as forms of figurative expression (Nunberg and Wasow, 1994), and most idioms contain conceptual metaphors which can be categorised for pedagogical uses (Charteris-Black, 2000). A growing body of research in vocabulary development has been examining the potential of using conceptual metaphors as a teaching platform for effectively retaining learned vocabulary (Deignan et al., 1997; Boers, 2000; Beréndi et al., 2008). Considering the increasing amount of research concerning the ubiquity of figurative language in our daily lives, as first expounded extensively in Lakoff and Johnson's "Metaphors We Live By" (1980), the present study sought to investigate the effect(s) of raising the participants' metaphor awareness on their retention of idiomatic phrases, as well as to discuss the viability of phrasal vocabulary teaching ("grouping" or "chunking"). In the context of this study, this chunking method of teaching vocabulary refers to the raising of their conceptual metaphor awareness when dealing with idioms among L2 learners, as supported by researchers in the field, such as Andreou and Galantomos (2008).

The existing body of research surrounding this issue largely consists of investigations on participants from a non-teaching academic background such as Chen and Lai, (2013) and Jermoškin, (2017). The present study differs by incorporating tertiary level participants enrolled in a Teaching English as a Second Language programme (TESL). By doing so, meaningful discussions can be generated on figurative language education, and to a greater extent, vocabulary education, from the points of view of potential teachers of English. This method of raising one's metaphor awareness to better retain idioms is the Cognitive Linguistic (CL) approach to teaching idioms.

1.2 Problem Statement

"I can take part effortlessly in any conversation or discussion and have a good familiarity with idiomatic expressions and colloquialisms…"

[2001: 27, CEFR, as cited in Pérez (2017)]

It is worth noting that Littlemore and Low (2006) elaborated on the importance of metaphoric competence as a potent pedagogical aid in overall communicative competence, which is in agreement with Pérez (2017), who had argued for an increased interest in developing metaphoric competence among learners of English, so much so that the skill had already been integrated into the self- assessment scales of the CEFR (Common European Framework of Reference for Languages, 2001). As the Malaysian Ministry of Education is promoting CEFR as a key element in the "Agenda for Reform 2015-2025" educational blueprint (Darmi et al., 2017), efforts to increase the metaphoric competence among Malaysian learners should therefore receive rapturous support.

To achieve the goals set by this agenda, where metaphoric competence is concerned, increased emphasis on expanding the learners' lexicons must be equally considered as well; a wide lexicon benefits metaphoric competence, and vice versa. Learners who utilise metaphoric competence can generate clearer groupings of words, and the grouping strategy itself aids vocabulary recall (Schmitt, 1997).

The findings of Mathai et al., (2004) are consistent with various studies concerning the state of the vocabulary proficiency of Malaysian university students (e.g. Asgari and Mustapha, 2011; Che Musa et al., 2012), in that Malaysian university students generally lack English vocabulary knowledge. This problem with English vocabulary knowledge is part of a greater and more complex intertwining of issues surrounding English language education in Malaysia. Based on a review of various recent studies concerning English language teaching and learning in this country, it was reported in Che

Musa et al. (2012) that one of the underpinning issues stems from the Primary to Secondary English language education, in which these schools have been placing an over-emphasis on the learning and testing of grammatical skills over communicative competence, thus creating an exam-oriented culture. Earlier, Ming's (2003) study which investigated the issue of memorisation among English as Second Language (ESL) Learners found that Malaysian students in a public university tend to rely on memorisation when dealing with English language related tasks. Interestingly however, students who stayed off-campus were found to rely on memorisation as well, but with understanding as the underpinning goal, unlike their on-campus peers, who utilised only surface level memorisation. Hence, this has created a new breed of tertiary level students who have passed their English language requirements for entering universities, but are unable to effectively use the language in communicative contexts. This situation comes as no surprise since native-like communication is replete with figurative language.

The present study heeds the call in the exhortations made by Che Musa et. al (2012) for a more focused effort to promote independent learning among Malaysian learners in general. In addition, Folse (2006) also encouraged more research into idioms vis-a-vis vocabulary retention, after reporting that multiple-sentence completion tasks facilitated better vocabulary retention than writing original sentences. It is posited that if learners become aware of how conceptual metaphors shape language, which in turn, enrich vocabulary, then they would be better equipped to perform communicative tasks more effectively, on their own.

1.3 Purpose of this Study

English idioms appear frequently in all manner of discourse, such as conversations and in mass media, and are not merely found in informal contexts; idiomatic language can be found in academia as well (Simpson and Mendis, 2003). Hence, by explicitly dedicating time to the teaching and learning of idioms, L2 learners would potentially gain what is lacking in today's educational environment: communicative competence. As discussed earlier, too much emphasis has been and is still being placed on the learning and testing of grammatical skills, while the exploration of the dynamic nature of real world conversations and the exchanging of ideas has over the years, taken a back seat to make way for grades which do not reflect these L2 learners' actual ability to navigate the complexities of real world communication, whether in speech or writing.

The present study endeavoured to shed light on a Cognitive Linguistics (CL) inspired method of teaching idioms that could help to realign the scales in ESL education, which has been placing far too much weight on the mastering of grammatical skills, to the detriment of L2 learners of English.

1.4 Research Objectives

Based on the issues above, this study sought to examine the effect(s) of applying a CL approach in the form of raising metaphor awareness, on the participants' retention of the figurative items, as well as to compare the CL approach with a known "conventional" method of vocabulary instruction in Malaysia: memorisation.

In addition, this research endeavoured to obtain insights into the perceptions of the participants on the viability of using idioms for aiding vocabulary retention through the respective methods.

1.5 Research Questions

The research questions which guided the present study were as follows:

1. What effect does raising metaphor awareness have on learners' retention of English idiomatic phrases?
2. What are the comparative effects between the raising metaphor awareness method and the memorising method on learners' retention of English idiomatic phrases?
3. How do the students in the control and experimental groups view the use of idioms for aiding vocabulary retention?

The research questions above are mainly answered in chapter 4 of this present study. In addition, chapter 5 builds upon further discussions on the pedagogical implications of the findings in chapter 4. Briefly, the first two research questions were answered using quantitative methods, while the third research question was answered qualitatively.

1.6 Outline of the Study

Chapter 1 addressed issues of the lack of communicative competence in English among L2 learners in Malaysia. The reason for this was found to be due to the over emphasis of grammar education. To increase learners' communicative competence, the researcher of this study decided to search for answers within the paradigm of Cognitive Linguistics.

Chapter 2 presents pertinent literature concerning 3 areas of knowledge: idioms and its relationship with metaphors from the point of view of Cognitive Linguistics, Psycholinguistic work on memory and its role in learning, and vocabulary education within the scope of this study.

Chapter 3 discusses the methods employed in the collection and analyses of the data, as well as their methodologies. Subsequently, Chapter 4 illustrates the marking schemes used to score the participants' answers in the cloze tasks, as well as all the quantitative and qualitative findings.

Chapter 5 continues the discussions from Chapter 4, with emphasis on pedagogical implications. The limitations of the present study are acknowledged, and this study concludes with a recommendation for further research.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

This chapter highlights key issues related to the present study, with Conceptual Metaphor Theory (CMT) serving to anchor 3 domains pertinent to this research. The first domain concerns figurative language itself, which consists of discussions on the figurative aspect of the English language in the form of metaphors and idioms, and how idioms are typically metaphorical in nature. The second domain deals with psycholinguistic literature on memory, regarding learning. Finally, the third domain highlights how the two prior domains are related to vocabulary teaching.

2.2 Conceptual Metaphor Theory (CMT): Pervasiveness of Figurative Language

To be human is to be inherently metaphorical in cognition. The mind gives shape to that which is abstract; using conceptual systems, the mind forms meaning and reason by way of conceptual metaphors. Conceptual metaphors deal with the mapping of the target domain, that which the mind intends to understand, with a source domain, which provides metaphorically concrete images based in the experiential world, to aid in the understanding of that abstract conceptual item in the target domain. An example of this notion would be "*keep your hair on*", in which "*hair*" (source domain) represents calmness in the target domain. Essentially, the mind seeks to embody abstract thought (Lakoff & Johnson, 1999; Gibbs, 1993).

Since Lakoff and Johnson's "Metaphors We Live By" was first published in 1980, interest in metaphor studies with implications on pedagogy has been continually sustained by research across various cultural contexts (e.g. Chen and Lai, 2013; Tsaknaki, 2016). Since its inception, metaphor studies have remained popular in and outside of the field of Cognitive Linguistics (Kövecses, 2008). However, despite the convincing arguments put

forth by Lakoff and Johnson, as well as other proponents of the notion that metaphors are a part of our cognitive mechanisms, criticisms have surfaced against the "traditional" model of Conceptual Metaphor Theory (CMT), as propounded by Lakoff and Johnson in their seminal work on metaphors. According to critics of the Lakoff and Johnson CMT model, the most pertinent issue they contend with is the methodology incorporated in metaphor research. They argue that the traditional CMT model is based on intuition and selective practice rather than hard empirical evidence. According to Hoover (2016), who used a corpus driven approach to arrive at hypothetical conceptual metaphors (the bottom up approach), Lakoff and Johnson were wrong about most of the metaphors highlighted in their seminal work, including later revisions. Hoover found that conceptual metaphors like ARGUMENT IS WAR/ STRUGGLE, were erroneous, seeing as the source domains WAR and STRUGGLE do not appear to be used pervasively in the target domain. The author found that among the most commonly occurring collocations about ARGUMENT, WAR and STRUGGLE collocates do not appear as frequently as EVALUATIVE collocates. However, Kövecses, (2008) had already countered this line of argument by explaining that there is a difference in research goal between "traditional" CMT researchers and those in favour of corpus driven research: quantitative methods cannot answer qualitative questions. The researchers in favour of this "traditional top-down" approach seek to postulate conceptual metaphors from decontextualized metaphorical expressions, which can be found in dictionaries, while "bottom-up" researchers seek to systematically identify and analyse metaphorical expressions found in specific naturally occurring contexts like news articles. In other words, these two fronts help to contribute knowledge on metaphors from two distinct perspectives.

The current study incorporated the "traditional" Lakoff and Johnson's intuitive approach.

2.2.1 Idioms as Metaphorical and Figurative Language

Since it is generally agreed that figurative language plays a fundamental role in human cognition, it is in the interest of the present study to discuss how idioms are metaphorical, thereby also being figurative in nature.

A vast number of idioms contain rich figurative meanings (Gibbs, 1993). This is because the conceptual metaphors of idioms can generally be traced or hypothesised about (Boers et al., 2004). Since metaphors constitute figurative language, the terms "metaphorical" and "figurative" are used interchangeably in this study.

Before Lakoff and Johnson's seminal work on cognitive semantic metaphor theory, the prevailing notion on idioms was that they were arbitrarily contrived, or "dead". However, a new school of thought emerged, and posited that a multitude of figurative idioms are "motivated" instead (Lakoff, 1987); idioms are created through the mental process of conceptualising metaphor, which is at the heart of this dissertation. In fact, this new paradigm argued that many figurative expressions are rooted in a small number of conceptual metaphors, which are categorisable and analysable (Boers, 2001).

Since idioms can be categorised and analysed according to metaphoric themes, the current study sought to observe the effect(s) of using the raising metaphor awareness method for retaining taught idioms.

2.3 The Human Cognitive Architecture

The human cognitive architecture is the overarching platform in which psycholinguists typically carry out their investigations on memory. The term "Cognitive Architecture" was coined to refer to the way in which "cognitive structures" are organised in human cognition (Sweller, 2003). Within this human cognitive architecture, the literature surrounding learning all strongly concur that memory in general plays a crucial role in the processes of learning or problem solving (e.g. Baddeley and Hitch, 1974; Cowan, 2008; Sweller, 2016; Sweller and Chandler, 1994). The central discussions in studies about memory and learning are conventionally explored through 3 inter-related phases or processes of memory: working memory (WM), short-term memory (STM), and long-term memory (LTM).

In addition, a brief examination on how the terms "retention" and "recall" are used in the present study is given, as well as highlighting Schmidt's Noticing Hypothesis (Schmidt, 1990). These various literatures constitute the second domain of knowledge in this chapter, and are all related to studies on memory.

2.3.1 Working Memory, Short-Term Memory, and Long-Term Memory

For the purposes of this study, only the latter 2 types of memory, namely short-term memory and long-term memory were discussed in relation to the findings and discussions section. The reason why STM and LTM were more pertinent to the current study than WM, was due to the stance taken by the researcher, in that WM serves as a multicomponent system which utilises STM and LTM to solve problems (Cowan, 2008). Thus, it was beyond the scope of this study to analyse the cognitive mechanisms used by the participants to solve the language tasks. Rather, the researcher mainly endeavoured to study the extent to which the participants could retain the taught idioms, after their metaphor awareness was raised. In other words, a study on WM would entail findings

and discussions on how the participants' used their WM to manipulate their STM and/or LTM resources to solve the tasks. Whereas, the current study focused on the extent to which the participants could retain the taught input via recall.

As a brief note, there is still a lack of consensus among prominent researchers on the definition of Working Memory, as well as what model best represents memory in general (Baddeley et al., 2019). The present study uses the terms "retention" and "recall" conformingly to the "traditionally popular" multicomponent model, which asserts that there are functionally distinct roles between WM, STM, and LTM, although they are interconnected (Repovš and Baddeley, 2006).

In this study, short-term memory is defined as a temporary memory storage facility with limited capacity, for immediate retrieval (Cowan, 2008). As for long-term memory, the current study adopts the view that generally, if insufficient cognitive attention is given to a stimuli, then one would not be able to transfer it to the LTM (Kellogg et al. 1996), which can hold massive chunks of information (Atkinson and Shiffrin, 1968; Craik and Lockhart, 1972; Brady et al., 2008), and is robust against memory decay, especially for implicit memory (i.e. lacking awareness) (Mitchell, 2006). In addition, Larzabal et al. (2018) provided a case for a formidably durable explicit memory as well (which included a high attrition rate from the sample and a narrow scope for generalisable findings).

The present study was able to measure the participants' short-term retention by analysing the immediate post-test scores, because each group immediately underwent a recall task after undergoing their respective lessons. In a similar manner, the participants' long-term retention of the taught idioms was able to be measured by obtaining scores from the delayed post-tests. Subsequently, discussions relating both types of retention scores to STM and LTM respectively, were generated. The short-term retention scores were obtained by modelling after the methods of immediate testing in Boers (2000), while

studies such as Roediger III and Karpicke (2006) support the rationale for delayed post-tests being able to generate long-term retention measurements.

2.3.2 Retention and Recall

A brief mention on what constitutes "retention" and "recall" must be given, as the two terms are used throughout this study.

In various studies concerning memory testing experiments, Tversky (1973), Richardson-Klavehn and Bjork (1988), Hummel (2010) highlighted how the extent of retention of input could be measured through recall (retrieval). The current study runs along the same vein; the way in which retention was measured in this study was through analysing the recalled idiomatic phrases taught.

Based on this contextual understanding of "retention" and "recall", this study mainly sought to observe and analyse the effect(s) of raising the awareness of metaphors on the L2 learners' retention of English idioms, by measuring the extent they were able to recall the taught idioms.

2.3.3 Schmidt's Noticing Hypothesis (1990)

In terms of language education, it would seem logical for memory studies to invariably lead to the noticing hypothesis, expounded by Schmidt (1990). According to Robinson (1995), short-term memory is where noticing takes place. Essentially, there is no language learning (transfer of input into LTM) if one does not pay attention to the item(s) to be learned. Naturally, this hypothesis argues for explicit, or intentional teaching.

While it is asserted that noticing played a part in answering the cloze tests, this study did not concern itself with furthering discussions on the hypothesis. The extent to which the term "noticing" was used in the qualitative analysis of this study, was limited to postulating what might have occurred during the testing phases. The scope of the current study limited the noticing hypothesis to providing broad descriptions on the possible links between noticing, morpho-syntactical (grammar) forms, and the potential for L2 learners to better retain taught idioms, while placing an emphasis on vocabulary education.

As a deeper look into experiments concerning noticing and grammar aptitude, Izumi (2002) provided an especially interesting observation: it is not merely a matter of noticing an input well enough that predicts successful language learning, but deep processing (Craik and Lockhart, 1972) must be factored into the lesson as well. In the present study, the metaphor awareness raising method constituted one such effort to aid L2 learners in accessing deep processing, when instructed to organise the idioms according to their respective metaphor categories. Although the researcher did not instruct the participants in the present study to explicitly notice patterns among the idioms taught, insightful observations in the qualitative analysis section of this study apparently display that some degree of noticing and deep processing did take place.

2.3.4 Craik and Lockhart's Levels of Processing (1972)

Within the scope of the present study, the most pertinent issues raised using "Levels of Processing" (LoP) are concerned with "shallow" and "deep" processing. Shallow processing is confined to orthographic recognition, where learners can identify what a word looks like (with capitalisation, etc.), while the deepest LoP can be seen when learners can recall entire idioms, and use them appropriately in context.

According to Mahdavian and Kormi-Nouri (2008), it was found that when there was focused attention, deep processing, and interest in the learning item, the explicit memories of the learners were performing at their highest capabilities. Explicit memory in the context of the present study refers to the participants knowing that they would be tested on numerous occasions, and so they would naturally expend more cognitive resources to activate deep processing in solving the cloze tasks in this study.

Thus, deep processing is argued to be integral to vocabulary retention, because deep processing involves knowing the meaning of words and how to use them in the right contexts. A few observations in the qualitative findings of the present study seem to suggest that deep processing occurred.

2.4 Explicit Vocabulary Education

"...While without grammar very little can be conveyed, without vocabulary nothing can be conveyed."

[Wilkins (1972), as cited in Alqahtani (2015)]

The present study aligns itself with the greater body of research supporting efforts to reclaim the importance of vocabulary teaching from the continual magnified focus on grammar education. In addition, while incidental learning is still beneficial to vocabulary learning, it should be encouraged to act in a supplementary role, while allowing intentional vocabulary instruction to serve at the forefront of vocabulary education.

Many researchers have argued against perpetuating the common practice of over-teaching grammar, while neglecting vocabulary teaching (e.g. Elyas and Alfaki, 2014; Laufer, 1997; Swan, 2002; Ulijn and Strother, 1990). In fact, Ulijn and Strother (1990) reported that syntax plays a minor role in reading comprehension, even at advanced levels, for both native and L2 learners. Instead, lexical knowledge was found to be a

greater determinant of reading comprehension. This is not to say teaching grammar is unimportant. According to Larsen-Freeman (2001), ineffective grammar teaching occurs in the form of focusing too much on grammar rules. Instead, teachers should be aware of the contributions in corpus linguistics, in that a disproportionately high number of words occur in grammatically recurring patterns. Doing so has the added benefit of allowing students to notice how words and expressions are used in context, as decontextualised examples used in grammar instruction are not as effective (Spada and Lightbown, 1993).

On the issue of explicit (intentional) versus incidental learning of vocabulary, educationists have, for decades, held onto the idea that vocabulary is best learned through incidental reading (Waring and Nation, 2004). This means that much of a learner's vocabulary is best accumulated through personal reading sessions (Krashen, 2004). However, studies have argued against this line of thinking by stating that when learning vocabulary is the focus, intentional teaching is more effective, which leads to vocabulary retention (Smith, 2004; Schmitt, 2008). Nonetheless, incidental learning is still beneficial, even though incidental learning strategies can be error prone (Hulstijn, 1992). Furthermore, a recent literature review on "informal learning on Youtube" by Lange (2019) highlighted a few learning advantages which can be reaped from the use of Youtube in general, but the cons seem to counterbalance the potential learning benefits, which can be found under "informal learning challenges" (Lange, 2019, 8).

Since this study endeavoured to continue the argument for intentional vocabulary teaching to serve a more prominent role in vocabulary learning, the researcher replicated Boers' (2000) study on teaching vocabulary in "chunks", by using idiomatic phrases. This notion of teaching vocabulary through idioms is not new (e.g. Beréndi et al., 2008; Boers and Lindstromberg, 2008; Irujo, 1986; Lazar, 1996; Syunina et al., 2018), and more about the rationale behind this method of vocabulary instruction is presented in the sub-sections that follow, especially concerning the lexical approach (Lewis, 1993) and the Cognitive

Linguistic (CL) method of teaching vocabulary, which promotes deep processing to occur, thereby aiding vocabulary retention.

This third domain of knowledge highlights various literature concerning formal vocabulary education.

2.4.1 Communicative Competence According to the Common European Framework of Reference for Languages (CEFR)

The CEFR companion volume (2018) provides a detailed description on what constitutes communicative competence. In this study, the type of communicative competence discussed falls under the umbrella term "communicative language competences". In it, three distinct language related communicative competences are expounded: linguistic, sociolinguistic, and pragmatic. This study mainly examined the linguistic realm of communicative language competence, and within this linguistic realm, vocabulary range was specifically discussed.

Concerning the concept of "type of language" for grading vocabulary range, the CEFR has this to say, among other statements:

"... from a basic repertoire of words and phrases to a very broad lexical repertoire including idiomatic expressions and colloquialisms." (p. 132)

The designation of "C2" is the highest language proficiency scale one can obtain under the CEFR, and in regard to grading one's vocabulary range, the description of scale C2 is as follows:

"*Has a good command of a very broad lexical repertoire including idiomatic expressions and colloquialisms; shows awareness of connotative levels of meaning."* (p. 132)

Therefore, whenever "communicative competence" is used throughout this study, it concerns itself with the use of idioms in terms of vocabulary range.

2.4.2 The Second Language (L2) Learner: A Cognitive Linguistics Perspective

As a sub-field of Applied Linguistics, Second Language Acquisition (SLA) essentially concerns itself with how a person learns a second language, apart from the first language (mother tongue). According to Krashen's Monitor Theory (Krashen, 1981), there is a distinction between "learning" or "acquiring" a second language. Learning is associated with intentional teaching methods, while acquisition is linked to informal methods of language development. As such, the present study operationally used the term second language learner or "L2 learner" in short to discuss the findings, because an intentional teaching method was applied.

Many theories about how L2 learners develop their second language have been put forth, based on a variety of psychological theories of learning (Ortega, 2014). This study falls under the Cognitivist perspective of SLA, which was developed in response to the Behaviourist school of psychology (Atkinson, 2011).

As was explored in this literature review, theories examining the relationship between learning and memory serve as central tenets in the Cognitivist paradigm. Idiomatic phrases were incorporated as the linguistic element in the present study.

2.4.3 The Lexical Approach

Through a Cognitivist observation, it is argued that the lexical approach (Lewis, 1993) can effectively aid L2 Learners in English vocabulary learning. The practical idea behind the lexical approach is that educators should take advantage of the many "prefabricated" multi-word expressions (MWEs) that exist in the English language. These MWEs include collocations and idioms. Corpus linguistic data has conclusively proven that many of these MWEs occur in rather fixed grammatical patterns (Siyanova-Chanturia and Martinez, 2014), such as the inclination to wish someone "good morning" instead of

"fine morning" etc. Therefore, this highly predictable nature of MWEs is argued to help students notice recurring ways in which thoughts are expressed in English, in both form and meaning.

The notion of "grammaticalised lexis, not lexicalised grammar" is the central idea behind this approach of teaching vocabulary (Lewis et al., 1997). Grammaticalised lexis refers to lexis (single or chunks of words) being central to the production of meaningful output. Lexis is prioritised over grammar, because grammar does not contain meaning on its own. Rather, grammar plays a facilitative role in organising lexis into a coherent expression of thought. The pedagogical implication for this notion of grammaticalised lexis is that since MWEs like idioms usually occur in grammatically predictable forms, L2 learners can benefit from the lexical approach because it aids metalinguistic competence: the extent one is aware of the interplay between lexis, structure, and context to create meaningful and coherent expressions of thought (Ofte, 2014).

Besides corpus linguistic data revealing that MWEs are by and large, predictable in naturally occurring language, Miller's seminal psycholinguistic work on memory (1956) found that humans can only hold around 7 items in the short-term memory. Most pertinent to the present study is Miller's recommendation to chunk words together to process linguistic items more easily. These findings substantiate the argument for a renewed effort to teach idioms more intensively and extensively in vocabulary education, according to the lexical approach.

In essence, the lexical approach is a paradigm shift in language education. It abandons the prescriptivist stance that educators often assume in over-emphasising grammar (Thornbury, 2011) as if dispensing linguistic medicine to "sick" students. While it is acknowledged that some degree of prescriptivism complements descriptive approaches in education, the lexical approach serves to re-orient educators to a more descriptivist path, one that is closer in harmony with the field of linguistics, in that

language should be learned as it occurs (McArthur et al., 2018), without imposing one's opinions on what constitutes "proper" or "broken" language use.

The next sub-section highlights a few studies that used idioms to facilitate vocabulary retention.

2.4.4 Using Idioms to Facilitate Vocabulary Retention

Many studies have demonstrated that native-like language use is made up of highly formulaic and idiomatic expressions (Kecskes, 2007; Siyanova and Schmitt, 2008; Wood, 2002; Wray, 1999, etc.). What this entails to L2 learners is that educators should not be teaching words in isolation, but in phrases, according to appropriate contexts. This approach to vocabulary teaching is congruent with how we as humans actually store and retrieve MWEs, and that is in chunks, which aids fluency (Wood, 2002). It is therefore argued that since idiomatic phrases occur as MWEs, in which their morphosyntactic order is rather fixed, L2 learners of English can increase and retain taught vocabulary more easily, due to idioms possessing high predictability.

This notion of utilising idiomatic expressions to facilitate vocabulary retention was explored in Boers' (2000) study, in which three experiments were reported. The first experiment was most pertinent to the design of this present study. In the first experiment, 118 pupils from a Flemish secondary school, aged 16 to 17 participated in the study. They were intermediate learners of English. It was found that the experimental group performed statistically significantly better than the control group ($p < .05$). However, it must be noted that only an immediate post-test was administered. In the present study however, a pre-test and delayed post-test was included as well in order to obtain a more comprehensive observation of the retention scores.

In Beréndi et al.'s (2008) study, three experiments were carried out to investigate the viability of using idioms to activate metaphor awareness, for vocabulary teaching. While the findings were argued to have supported the notion that enhancing the awareness of conceptual metaphors and metonymies can aid the retention of figurative vocabulary, the researchers acknowledged the limitation of enlisting a low number of participants in generating the statistical data, mainly in the first and third experiments. In the first experiment, two parallel groups of 13 participants were enlisted, while the third experiment only enlisted two groups of 11 and 14 students respectively.

Yasuda (2010) carried out an investigation into the efficacy of raising the awareness of conceptual metaphors for learning idiomatic phrasal verbs. It was found that raising one's metaphoric competence could serve a complimentary role in committing unfamiliar phrasal verbs to memory. In total, 115 Japanese EFL (English as a Foreign Language) university students were enlisted in this study, with 56 participants as the control group, and 59 participants forming the experimental group.

The next two sub-sections examine the predominant method of teaching idioms (blind memorisation), and the CL method recommended in this study.

2.4.5 "Traditional" Approach to Teaching Idioms

According to Boers et al. (2004), the grammar-lexis dichotomy has traditionally dominated the conception of language. This dichotomy manifests itself in word lists containing individual words on one side, and grammatical "rules" concerning usage on the other. Idioms, as well as other MWEs, could not be fitted into this dichotomy, and as such, have been largely relegated as purely ornamental devices to be used in "spicing up" one's messages (Boers et al., 2004), as will be evident in the qualitative findings of the present study. Since idioms were considered unpredictable and uncategorizable (Chen and Lai, 2013), and thus could not be taught in a systematic manner, the only perceived method to commit idioms to memory was to blindly memorise them through simple rote (repeated) memorisation. "Blind" memorisation constitutes the "traditional" approach in the teaching of idioms (Boers, 2001), and since this approach persists to this day, the present study employs the terms "traditional" and "conventional" interchangeably, concerning this method of teaching idioms.

Idioms, though being replete in every manner of discourse in the English language, are understandably viewed as troublesome, especially so for the L2 learner and teacher. Idioms which are more metaphorically opaque tend to be much more challenging to comprehend. For this type of idiom, it is argued that one does not have to be satisfied with this traditional approach, as Boers et al. (2004) have observed that etymological elaboration (giving a historical and cultural lesson) on this "bound" type of idiom can be just as effective for opaque idioms as more "transparent" ones. In fact, a simple experiment by Dobreva (2018) found that the etymological method was effective in aiding the retention of idioms after two weeks (73% correct answers). Although, it must be said that only 15 participants were enlisted in that study, with no control group.

Since there exist studies refuting the necessity for the blind memorisation of opaque idioms, it would seem prudent to consider alternative methods of aiding retention for the more obviously metaphorically transparent idioms. The following sub-section highlights a few studies which argue that CL inspired methods are more effective than memorisation in vocabulary teaching.

2.4.6 Cognitive Linguistic Approach to Teaching Idioms: Raising Metaphor Awareness

The metaphor awareness raising method is based on the premise that the human mind perceives phenomena through metaphors. We are subconsciously mapping images from a source domain to a target domain in our minds all the time, especially when trying to understand an abstract concept. Since idioms are metaphorical in nature, the metaphor awareness raising method is argued to be suitably applied in the teaching of idioms. Simply put, idiomatic expressions reflect human patterns of cognition. There are strong pedagogical reasons for teaching metaphor explicitly. For one, because of the widely accepted notion that figurative language is pervasive in our lives, and that these metaphorical expressions often occur in collocations, idioms, and other MWEs. It is therefore imperative that educators be opportunistic about incorporating conceptual metaphor raising exercises in their language classrooms.

Armed with a confidence in CMT, researchers have reported positive outcomes in vocabulary retention. A recent study by Pan (2019) generated interesting findings and discussions that support this CL method of teaching idioms to young learners (aged 10 to 12). Overall, the author recorded a better performance from the experimental group, which incorporated a metaphor awareness raising element in the treatment, while the control group did not.

Looking further back, Gao and Meng (2010) found that vocabulary retention was higher in the experimental group (CL method) than in the control group (memorisation). However, it appeared that for the low proficiency participants in the study, it did not matter which method was employed. The authors recorded no statistically significant difference between the CL method and the memorisation method concerning vocabulary retention (p = .997).

2.5 Chapter Summary

In this chapter, studies were examined to unpack three domains of knowledge intertwined in this present study: idioms as metaphorical language, human memory and its role in learning, and intentional vocabulary teaching.

It is argued in the current study that a CL approach would be more beneficial in increasing the communicative competence of L2 learners, and that learners should not be required to carry on with the mundane act of blindly memorising idiomatic phrases. This is because raising the learners' awareness of metaphors appeals to the substantial research that exists on human memory and cognitive perception.

In the following, an extensive discussion is provided on the rationale behind the research design choices made in this study.

CHAPTER 3: METHODS AND METHODOLOGY

3.1 Introduction

An explanatory sequential research design (Creswell, 2013), with one experimental group and one control group, was applied in this study. This research design was ideal for achieving the objectives of this study, as testing for the effects of two teaching methods calls for a quantitative method of data collection. Therefore, quantitative data was used to lead the findings of this study, while the subsequent qualitative data was used to provide explanations on what might have occurred during the testing phases that lead to the findings of the quantitative data. The qualitative element supplements the quantitative aspect by offering insights concerning the pedagogical implications of using the "conventional" and experimental methods in ESL classrooms. This research design was "sequential" because it consisted of two phases: quantitative, followed by qualitative. According to Morse (1991), the designated notation for this type of sequential mixed methods research design is "QUAN → qual".

This study sought to examine and compare the effects of the approaches on the participants' retention of the learned items, to see if the experimental method would be a viable alternative to the conventional, albeit negatively perceived one.

The following page shows figure 3.1, which illustrates the overall research design of this study. The subsequent discussions within this chapter serve to unpack what is shown in figure 3.1.

Explanatory Sequential Research Design

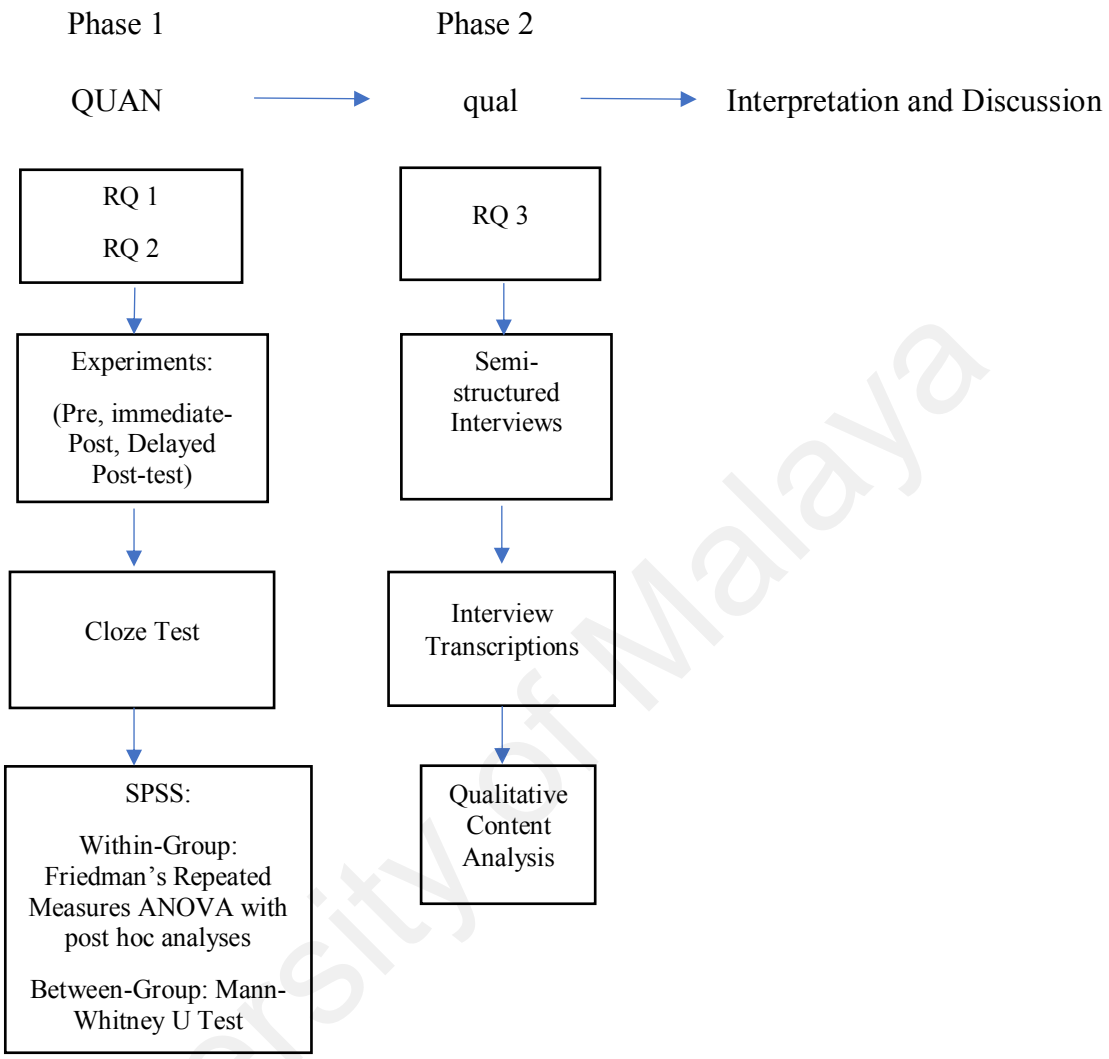

Figure 3.1: Research Design Diagram

3.2 Ethical Consideration

A permission letter to collect data from the research site was obtained and consent forms were given to all participants to obtain consent and identities were kept confidential throughout the data collection and data analysis, which includes the reporting of the data, and will remain confidential. The participants were notified that they could withdraw their participation from any point in the study. The consent form can be found in appendix J. In addition, the control group was offered the opportunity to undergo the same treatment as the experimental group after the study was completed, seeing as they may show interest in the educational implications for their own development as future English language teachers.

3.3 Quantitative Phase

The quantitative phase was the first step under this research design. The subsections that follow provide a profile of the participants enlisted in the quantitative phase, along with the instruments that were used for both control and experimental groups. The entire data collection and data analysis procedure for this phase is elaborated on as well.

3.3.1 Participants

42 Malaysian Teaching English as a Second Language (TESL) Diploma students were enlisted from a private institution of higher learning. For homogeneity, only participants who obtained a grade ranging from B- to A in their English subject for their *Sijil Pelajaran Malaysia* (SPM) certificate, were considered for the study. They are L2 English users themselves and are ethnically diverse. Only first and second year classes were chosen. The institution designated the classes which were available for data collection, and the researcher had to liaise with a supervising teaching staff. The

researcher assigned the participants randomly to each group by deciding that every class would be placed into the groups alternatingly (e.g. Class 1 to control group, Class 2 to experiment group). Altogether, the researcher collected data from four classes. Although this sampling method did not incorporate true randomisation, some degree of randomisation took place as the researcher had to assign the classes to their respective groups before meeting any of the participants. The researcher started with a set goal of 60 participants, but the attrition rate caused the total number of participants to be reduced down to 42. The attrition was caused by students not attending one of the three testing sessions.

As for the qualitative phase, ten participants were randomly selected for one-to-one interviews (five each from the control, and experiment groups). Selecting the participants at random was deemed necessary to avoid selection bias. A list of names from each group was made in Microsoft Excel, and underwent randomisation. The first 5 names were chosen for the interviews.

3.3.2 Instruments

In total, three instruments were used for the quantitative data collection phase: one "metaphor" idiom inventory, one "memorising" idiom inventory, and one cloze task sheet. These instruments can be found in appendices A, B, and C respectively. Examples of these inventories can be observed below:

Metaphor inventory (appendix A)

SHOWN EMOTION IS A FACE

Keep a straight face

Put on a brave face

Pull a face

Memorising inventory (appendix B)

Put on a brave face	Pretend you are happy about something when you are not happy.

Cloze task (appendix C)

"After all those years my father should really have known better, because he had learned from experience how she can _____ (1)."

The inventories consisted of idioms provided in McCarthy and O'Dell's *"English Idioms in Use"*, (2002). BODY PART conceptual metaphorical idioms were chosen to be included in this study as Kövecses (2000) argued that the human body is a very easily referenced source domain for metaphorical expressions, which appears widely in real contexts, and should be taught first when starting a series of lessons on conceptual metaphor among learners. The "metaphor" idiom inventory refers to the grouping of idiomatic phrases according to their respective conceptual metaphorical categories, which was used in the experimental group, while the "memorising" inventory (control group) listed down the idioms in no particular order, but had the meanings of the idioms attached to each respective one. For the current study, BODY PART idioms were expressed in the context of tense situations [modelled after Boers (2000)], who studied the effect of metaphor awareness on the retention of novel vocabulary via English idioms. This context is reflected in the cloze task sheet adapted from the ANGER cloze test by Boers (2000), and the same sheet was given to both groups across the entire quantitative data collection phase. An example of the original cloze task from Boers (2000) is provided below:

Sample Text 2:
Fill in the gaps with one word each. Sometimes there may be more than one option. In that case, try to come up with as many possibilities as you can.
Last month was my parents' wedding anniversary. A week before the anniversary my mother already suspected that my father had forgotten about it, since he hadn't asked her if she wanted to do anything special for the occasion. After all those years my father should really have known better, because he had learned from experience what a _____ (1) temper my mother has. But I felt that he deserved another lesson and I decided not to remind him of the anniversary. Days before that fatal day I could already notice my mother's anger _____ (2) up inside her. I tried to tell her to _____ (3) down, because my father might still remember after all. But he didn't, and as the day of the anniversary approached, mother was getting all _____ (4) up. On the morning of the day itself, my father left for work as usual without mentioning the wedding anniversary. Mother almost _____ (5) her lid. She had clearly reached her _____ (6) point and she kept fuming all day.

3.3.3 Data Collection Procedures

3.3.3.1 Pilot

This study enlisted the help of nine students who were taken out of the sample population, but not from the four classes provided by the supervising staff. Three participants formed the experimental group while six participated in the control group. The pilot test was useful for testing the feasibility of using the cloze task and noting on the issues that arose from administering the conventional and experimental methods of teaching idioms. Based on the participants' comments, it was found that the original inventory contained too many idioms, and so, out of the 20 idioms, the inventory was reduced to 14 idioms.

3.3.3.2 Pre-test

Both groups sat for a 10-minute pre-test in the form of a cloze exercise. The control group received two sheets: one "memorising" inventory consisting 14 BODY PART conceptual metaphorical idioms, arranged linearly by numbers, in no specific order, with the respective meaning attached to each item, and one cloze task sheet. The "memorising" inventory refers to the method of learning the idioms. Likewise, the experimental group also received two sheets: one "metaphor" inventory, and the same cloze task sheet. The "metaphor" inventory refers to the idioms which were arranged according to metaphorical themes, e.g. "FEELINGS ARE HEARTS" as in "talk to me heart-to-heart".

Both groups had five minutes to study their inventory sheets, after which, they had to put them away and complete the cloze test. 10 minutes were given for the task, and they were encouraged to write down as many contextually appropriate idioms as they could.

3.3.3.3 Treatment

After two days, each group received their respective treatment. For the control group, the researcher explained each of the idioms by referring to the respective meanings in the same idiom inventory sheet. As for the experimental group, the researcher carried out a lesson on conceptual metaphors, and explained how it was related to each of the idioms in the inventory. Both groups underwent their respective treatments for 10 minutes.

3.3.3.4 Immediate post-test

After completing their treatments, both groups had five minutes to study their respective inventories. The rest of the immediate post-testing procedure was the same as the one used in the pre-test.

3.3.3.5 Delayed post-test

After two weeks, the researcher carried out the same cloze test. This time however, both groups did not receive their idiom inventories. This was done to reduce the practice effect, and to better test their ability to recall the idioms correctly after a considerable length of time. Just as the pre-test and immediate post-test, the participants were encouraged to write down as many idioms they thought were contextually appropriate as possible. Table 3.1 illustrates the entire quantitative data collection procedure.

Table 3.1: Quantitative Data Collection Procedure

		Control Group				Treatment Group			
Day	Time (mins)	5/ 10	5	10	10	5/ 10	5	10	10
	Task	Briefing	Inventory Study	Cloze Task	Treatment	Briefing	Inventory Study	Cloze Task	Treatment
1	Pre-Test	Briefing, "memorising" inventory study, cloze test				Briefing, "metaphor" inventory study, cloze test			
3	Treatment	"Memorising" Treatment				"Metaphor" Treatment			
3	Immediate Post-Test	"Memorising" Inventory study, cloze test				"Metaphor" inventory study, cloze test			
17	Delayed Post-Test	Cloze Test				Cloze Test			

3.3.4 Data Analysis Procedures

Utilising the Statistical Package for the Social Sciences (SPSS), The within-group data were analysed via Friedman's Repeated Measures Analysis of Variance (ANOVA), along with the Wilcoxon Signed-Rank Test as the post hoc analysis where data were found to be statistically significant. The between-groups data were analysed using the Mann-Whitney U Test.

The quantitative data analysis procedure for this study was based on a specific analytical paradigm expounded in "Testing in Language Programs" (Brown, 1996). Before delving into the data, the author recommends that the teacher researcher be lucid about the educational goal(s) behind a study. Does the researcher want to analyse placement test scores, in order to know the classes in which the students should be studying? Or does the researcher want to know how successful the students were in performing a specific linguistic skill in a test? Clearly, the current study reflected the latter goal. The former type of goal falls under the umbrella term "NRT" (Norm Referenced Tests), while the latter, "CRT" (Criterion Referenced Tests). This distinction is paramount because it affects the trajectory of the analyses; the analyses must not misrepresent or misconstrue the nature of the data.

NRT type goals are best interpreted using parametric tests, as large Standard Deviations (SDs) are preferred, whereas CRT type goals favour smaller SDs. This is due to the nature of the tests: larger SDs form a normal distribution, which is useful for determining which classes students should be placed in, while smaller SDs mean that the students either experienced a meaningful or uneventful learning experience (Brown, 1996). Of course, it is in the best interest for all stakeholders that the students' scores form a negatively skewed distribution (as close to 100% as possible on each test). Therefore, since this current study intended for the participants to score as many points as possible, especially immediately after the treatment, this study employed non-

parametric methods of analyses to examine the findings. Through the non-parametric CRT lens, medians are preferred over means, and percentages over percentiles, because CRT type tests are concerned with the progress of every individual, and not between individuals. However, for the purpose of comparison in this study, a non-parametric NRT method of analysis between the two groups was required, which came in the form of the Mann-Whitney U Test. Another important reason for preferring the median is that outliers can be included. In parametric tests, which prefer means over medians, outlier scores are thrown out because they can inflate values in statistical analysis. Conversely, medians are robust against outliers. Since a largely negatively skewed distribution in the bell curve is a desirable result in CRT type statistical analysis, outlier data should not be discarded, and indeed, this study included such data.

3.3.4.1 Friedman's Repeated Measures ANOVA

By adhering mainly to the CRT domain of analytical procedures, Friedman's Repeated Measures ANOVA, the non-parametric alternative to the standard One-Way Repeated Measures ANOVA, was used for analysing the within-group data (Friedman, 1937). Subsequently, the Wilcoxon Signed-Rank Test was carried out as the post hoc analysis, only when statistically significant findings were observed in the Friedman's Test, because it analyses the differences between the three test stages (pre, immediate-post, and delayed post-test) within a group (Derrac et al., 2011; Rosner et al., 2006), whereas the Friedman's Test analyses the overall mean/ median result of the group data. The Friedman's Test was used to answer the first research question, which required an analysis of the test scores obtained from the experimental group, across three points in time, in order to generate discussions on the effect of raising the awareness of metaphors among the participants.

3.3.4.2 Wilcoxon Signed-Rank Test

The Wilcoxon Signed-Rank Test is a non-parametric test which analyses matched pair data (Woolson, 2007). In this study, the Wilcoxon Signed-Rank Test was carried out after obtaining statistically significant results from the Friedman's Test. Therefore, the Wilcoxon Test serves as a post-hoc analysis by analysing the differences between the three test stages (pre, immediate-post, and delayed post-test) within a group (Rosner et al., 2006; Derrac et al., 2011), whereas the Friedman's Test analyses the overall mean or median result of the within-group data. In other words, the Friedman's Test is used to analyse a set of interval data, while the Wilcoxon Test is used to analyse and compare the data between those intervals.

3.3.4.3 Mann-Whitney U Test

The Mann-Whitney U Test is used to compare the difference between groups, and is especially robust against outliers (Nachar, 2008). Therefore, the Mann-Whitney U Test was employed to answer the second research question, which required a comparison of the median differences between the control and experimental groups, in order to highlight which method of teaching idioms was better, and at which testing stage. In this study, the Mann-Whitney U Test analysed each respectively paired counterpart in the control and experimental groups; the scores from the control and experimental groups at the pre-test level were paired together and analysed, and the subsequent levels underwent the same analytical procedure.

The null hypothesis for the Mann-Whitney U Test is that the distributions of the two groups are equal (Nachar, 2008). According to Laerd Statistics (n.d.), what this refers to is that while the Mann-Whitney U Test makes no assumption about a normal distribution, it does however require that each paired distribution be similar in shape; one

group can possess a higher median and/or set of values than the other, but both groups should reflect more or less the same shape in distribution. If the distributions of two paired groups are different in shape, then less conclusions can be made, which are presented simply in the form of whether one group's values are higher or lower than the other in terms of mean ranks. Ideally, the Mann-Whitney U Test should report medians.

3.3.4.4 Levene's Test for Homogeneity of Variance

As such, a Levene's Test for homogeneity of variance was carried out to determine the shapes of the distributions at all levels of the test score data. For this testing purpose, the Levene's Test is a powerful and robust approach (Gastwirth et al., 2009).

When both groups were paired at each test level, it was found that the distributions of the test scores were all similar in shape, except at the delayed post-test level for both the control and experimental groups within the strict and lenient scoring paradigms (elaborated in Chapter 4). To identify the shape of distribution for each test level, one can observe the "Sig." column in Table 3.2, which represents statistical significance. Whenever $p < 0.05$, this meant that the null hypothesis of the Levene's test could not be accepted at that level, because the Levene's Test assumes that the variances of the two sampled groups are equal (Nordstokke and Zumbo, 2010). In short, the delayed post-test (both scoring schemes) was the only level at which the null hypothesis was rejected.

Therefore, it was concluded that the shapes of the paired groups' distributions at the delayed post-test level were not similar for both scoring schemes, which meant that the researcher could not report the between-groups results in medians for the delayed post-test level, irrespective of scoring paradigm, but instead reported the scores in mean ranks.

Table 3.2: Levene's Test for Homogeneity of Variance (Strict and Lenient)

Test of Homogeneity of Variance

		Levene Statistic	Sig.
Pre-test scores (strict)	Based on Median and with adjusted df	.718	.402
Immediate post-test scores (strict)	Based on Median and with adjusted df	.069	.795
Delayed post-test scores (strict)	Based on Median and with adjusted df	9.720	.004**
Pre-test scores (lenient)	Based on Median and with adjusted df	.805	.375
Immediate post-test scores (lenient)	Based on Median and with adjusted df	.261	.613
Delayed post-test scores (lenient)	Based on Median and with adjusted df	6.307	.017*

*significant at $p < .05$; **significant at $p < .005$.

At the delayed post-test level under the strict marking scheme, the variances between the two groups (control and experimental) were statistically significantly different ($p < .005$), while under the lenient marking scheme, the two groups were also found to be statistically significantly different ($p < .05$). The variances of these two statistically significant values equated to differing shapes of distribution.

As for the pre and immediate post-test scores under both marking schemes, it was found that their shapes of distribution were not statistically significant. Hence, the null hypothesis for the Levene's Test was retained; their shapes of distribution were similar at matched pairs (control and experimental groups), which allowed the researcher to report their scores in terms of medians.

3.4 Qualitative Phase

The subsections that follow provide a profile of the participants enlisted in the qualitative phase, along with the interview instruments that were used for both control and experimental groups. The entire data collection and data analysis procedure for this phase is elaborated on as well.

3.4.1 Participants

The researcher only shortlisted the participants who indicated that they were willing to be interviewed in their participant's background information form (appendix I). Out of 20 participants in the control group, only 10 indicated that they were willing to be interviewed, while out of 22 participants from the experimental group, 21 stated they were willing to participate in the interview.

Their names were keyed into a Microsoft Excel sheet, and underwent randomisation to eliminate selection bias. The first five names from each group were selected to be interviewed. Each entry was designated a code, corresponding to the respective interviewee. Interviewees from the control group were labelled C1 to C5, while those from the experimental group were labelled E1 to E5. In all, 10 participants were interviewed.

3.4.2 Semi-Structured Interview Instruments

The semi-structured interview instruments consisted of two sets of interview questions: one for the control group, and one for the experimental group. The open-ended questions were designed to elicit their opinions concerning various issues related to their experiences going through the entire quantitative test phase, as well as their thoughts on merging the teaching of idioms with vocabulary teaching. These two sets of questions can be found in appendices D and E of this study. A SONY voice recorder was used to record the interviews.

3.4.3 Data Collection Procedures

The researcher made arrangements with the selected participants to meet up three days after the delayed post-test, at their university. One-to-one sessions were held one after another in vacant classrooms throughout the day, based on the participants' schedules. 10-minutes was the appointed duration for each session, but a few of the interview sessions had to exceed it due to various reasons: having to stop the interview momentarily due to noisy chatter coming from outside the classroom, the researcher having to elaborate on a few questions, and the participants requiring some extra time to process the various concepts presented to them.

After recording all 10 of the participants, the next step was to transcribe the interviews. In reality, several interview transcription conventions exist. However, it was decided that since prosody did not constitute a research objective in this present study, only an adapted corpus of grunts (Ward, 2000) was needed. For reference, it is located in appendix F. An example of an interview transcript from Participant E1 (appendix G) is provided as well.

3.4.4 Data Analysis Procedures

The Qualitative Content Analysis of the transcribed interviews was based on the tabled procedures found in Graneheim and Lundman (2004) and Bengtsson (2016). Bengtsson suggests that the analysis should be presented in the form of a table because reliability and validity can be increased through the exercise of transparently displaying the raw data, along with their condensed forms of meaning. This leads to codification for the development of sub-categories/ themes, of which finally aid the emergence of themes that can help to shed light on the studied phenomena. Therefore, the interview transcriptions were analysed using Qualitative Content Analysis to generate themes and quotes for discussions. Beginning with meaning units, the qualitative findings developed themes through the analyses of the interview recordings.

Although 10 student participants were interviewed, the data at times, had shown that a point of saturation was reached. Thus, it was not necessary to include quotes from all 5 participants from each group in every content analysis table, as the nature of the qualitative phase of this study was phenomenological, which only required the participation of at least 5 subjects in total to generate purposeful responses until a point of saturation is reached (Creswell, 1998). This meant that when the interview data offered no other additional insight to answer the objectives of this study, it was up to the discretion of the researcher to leave out the data surplus.

3.5 Chapter Summary

This chapter highlighted the overall research design of the present study, which employed an explanatory sequential design. Through this research design, quantitative findings lead the way to answer the first two research questions (RQ), which sought to observe the effects of the raising metaphor awareness method of teaching idioms to L2

learners (RQ #1), and how this method measured up with a conventional/ traditional method of teaching idioms (RQ #2). The quantitative procedures used to answer these two questions were carried out via SPSS. The Friedman's Test with the Wilcoxon Signed-Rank Test as post hoc analysis was used to analyse within-group findings for the experimental group, to observe how the participants performed over time (RQ #1), while the Mann-Whitney U Test was used for generating findings for the between-groups data, to determine which method was more effective for retaining idioms (RQ #2).

As for the qualitative phase, a transcription convention for grunts (Ward, 2000) was used. After transcribing the interview recordings, Qualitative Content Analysis (Bengtsson, 2016; Graneheim and Lundman, 2004) was carried out to generate pedagogical implications on the teaching of idioms and how it could be linked to vocabulary teaching and learning in general.

CHAPTER 4: FINDINGS AND DISCUSSIONS

4.1 Introduction

Since a sequential explanatory research design was employed in this current study, this chapter firstly reports the quantitative findings, followed by the qualitative findings. The quantitative phase is made up of two parts: within group analysis, and the between group analysis. Both of these quantitative analysis methods were carried out according to strict and lenient test scoring schemes. The qualitative phase reports on the insights of 10 of the participants (five from each group) concerning the teaching methods administered throughout the data collection phase, with emphasis on the retention of the taught idioms, and the way it is related to vocabulary teaching and learning.

4.2 Analysis of Quantitative Data

This section first illustrates the scoring procedure of the collected data, followed by the within-group data analysis. Subsequently, the between group data analysis is presented. The within-group findings were used to answer RQ #1, while RQ #2 was answered using the between-group findings. A summary is provided at the end of the within and between group data analyses.

4.2.1 Scoring Procedure

Two scoring schemes were used to award points to the participants. The first round of scoring was strict, while the second was lenient. The rationale was that having a lenient marking scheme would allow for the analysis of smaller incremental improvements (Folse, 2006). A rigid marking scheme would only serve to punish instead of motivating the participants to attempt the tasks. When the strict and lenient score results were placed side by side, salient discussions were able to be made. Table 4.1 illustrates the scoring rubric for the collected quantitative data:

Table 4.1: Strict and Lenient Scoring Schemes (Based on Oller, 1972; and Folse, 2006)

Marks Awarded	Strict	Lenient
2	Idiom in exact form, appropriately used in context.	Idiom in exact form, appropriately used in context. OR Deviation of idiom in form (syntax and/or morphology), but appropriately used in context.
1	Deviation of idiom in form (syntax and/or morphology), but appropriately used in context. OR Idiom with a synonymous meaning with exact form.	Idiom with a synonymous meaning with exact/ deviation to form (syntax and/or morphology).
None	Incorrect idiom in form, and inappropriately used in context.	Incorrect idiom in form, and inappropriately used in context.

To better illustrate this scoring procedure, examples from the cloze task (appendix C) and the scoring methods are presented below:

Examples of strict marking scheme:

"After all those years my father should really have known better, because he had learned from experience how she can shout **her** head off. But I felt that he deserved another lesson and I decided not to remind him of the anniversary." (2 points)

"After all those years my father should really have known better, because he had learned from experience how she can shout ***my*** head off. But I felt that he deserved another lesson and I decided not to remind him of the anniversary." (1 point)

Examples of lenient marking scheme:

"After all those years my father should really have known better, because he had learned from experience how she can shout ***my*** head **off**. But I felt that he deserved another lesson and I decided not to remind him of the anniversary." (2 points)

"After all those years my father should really have known better, because he had learned from experience how she can ***blow her*** head ***up***. But I felt that he deserved another lesson and I decided not to remind him of the anniversary." (1 point)

4.2.2 Within-Group Analysis: Friedman's Repeated Measures ANOVA with Post Hoc Analysis (Wilcoxon Signed-Rank Test)

A Friedman's Repeated Measures ANOVA was carried out to examine the overall progress of the raising metaphor awareness (experimental) group of participants over three points in time, based on both strict and lenient marking schemes. Subsequently, a Wilcoxon Signed-Rank Test was run when the Friedman's Test results indicated a statistically significant *p*-value. The findings under both strict and lenient scoring analyses are provided below.

4.2.2.1 Strict marking scheme

The null hypothesis for the Friedman's ANOVA was that the distribution of the scores across the 3 points in time were equal. The alpha was set at .05, with 2 degrees of freedom. Based on the strict marking scheme, there was a statistically significant difference among the three weeks, $\chi^2 = 10.88$, (2, $N = 22$), $p = .004$. Therefore, it was found that the raising metaphor awareness method of teaching idioms had caused a statistically significant variation in retention scores between the three test phases. The table below displays the findings:

Table 4.2: Friedman's Test Statistics for Experimental Group (Strict Marking)

Test Statistics[a]

N	22
Chi-Square	10.881
Df	2
Asymp. Sig.	.004
a. Friedman Test	

This statistically significant finding led to a post hoc analysis. With the alpha set at .05, the null hypothesis for the Wilcoxon Signed-Rank Test was that there was no difference in pairwise medians. However, the post hoc analysis indicated that the median immediate post-test score was statistically significantly higher than the median pre-test score, $Z = -3.29$, $p = .001$, with a large effect size ($r = .49$) (Pallant, 2007). Between the strict pre-test and immediate post-test median scores, there was an 85% increase from the pre-test score (*Mdn* = 23.22) to the immediate post-test score (*Mdn* = 42.86). Therefore, the null hypothesis is rejected. Tables 4.3 and 4.4 illustrate this analysis.

Table 4.3: Pre and Immediate Post-Test Medians for Within-Group Analysis (Experimental Group, Strict Marking)

Descriptive Statistics

		Percentiles		
	N	**25th**	**50th (Median)**	**75th**
Score Strict Pre (%)	22	13.3950	23.2150	46.4300
Score Strict Immediate Post-test (%)	22	24.1075	42.8600	60.7125

Table 4.4: Wilcoxon Signed-Rank Test for Experimental Group's Pre and Immediate Post-Tests (Strict Marking)

Test Statistics[a]

	Score Strict Immediate Post-test (%) – Score Strict Pre-test (%)
Z	-3.291[b]
Asymp. Sig. (2-tailed)	.001
a. Wilcoxon Signed-Rank Test	
b. Based on negative ranks.	

As for the strict immediate post-test and delayed post-test pairwise analysis, there was also a statistically significant finding. The null hypothesis for the Wilcoxon Signed-Rank Test was that there was no difference in pairwise medians. However, the post hoc analysis revealed that the median immediate post-test score was statistically significantly higher than the median delayed post-test score, $Z = -2.63$, $p = .009$, with a medium effect size ($r = .39$) (Pallant, 2007). Between the strict immediate post-test and delayed post-test median scores, there was a 29% decrease from the immediate post-test score (*Mdn* = 42.86) to the delayed post-test score (*Mdn* = 30.36). This result could be interpreted that there was a 29% decrease in the participants' ability to recall the idioms. This analysis can be seen in tables 4.5 and 4.6.

Table 4.5: Immediate and Delayed Post-Test Medians for Within-Group Analysis (Experimental Group, Strict Marking)

Descriptive Statistics

	N	Percentiles		
		25th	50th (Median)	75th
Score Strict Immediate Post-test (%)	22	24.1075	42.8600	60.7125
Score Strict Delayed Post-test (%)	22	7.1400	30.3550	57.1400

Table 4.6: Wilcoxon Signed-Rank Test for Experimental Group's Immediate and Delayed Post-Tests (Strict Marking)

Test Statistics[a]

	Score Strict Delayed Post-test (%) - Score Strict Immediate Post-test (%)
Z	-2.630[b]
Asymp. Sig. (2-tailed)	.009
a. Wilcoxon Signed-Rank Test	
b. Based on positive ranks.	

As before, the null hypothesis for the Wilcoxon Signed-Rank Test was that there was no difference in pairwise medians. The post hoc analysis revealed that the median strict pre-test score was not statistically significantly lower than the median delayed post-test score, $Z = -1.38$, $p = .167$, $> .05$ with a small effect size ($r = .21$). Between the strict pre-test and delayed post-test median scores, there was a 31% increase from the pre-test score ($Mdn = 23.22$) to the delayed post-test score ($Mdn = 30.36$). This result could be interpreted that the participants were able to improve their LTM retention of the idioms by 31%. Tables 4.7 and 4.8 display these findings.

Table 4.7: Pre and Delayed Post-Test Medians for Within-Group Analysis (Experimental Group, Strict Marking)

Descriptive Statistics

		Percentiles		
	N	**25th**	**50th (Median)**	**75th**
Score Strict Pre-test (%)	22	13.3950	23.2150	46.4300
Score Strict Delayed Post-test (%)	22	7.1400	30.3550	57.1400

Table 4.8: Wilcoxon Signed-Rank Test for Experimental Group's Pre and Delayed Post-Tests (Strict Marking)

Test Statistics[a]

	Score Strict Delayed Post-test (%) - Score Strict pre-test (%)
Z	-1.383[b]
Asymp. Sig. (2-tailed)	.167
a. Wilcoxon Signed-Rank Test	
b. Based on negative ranks.	

The post hoc analyses for the strict marking scheme in the experimental group show that there was an improvement in scores overall, between the strict pre-test and delayed post-test scores, but not by much. The pairwise analyses indicated that there was no statistical significance overall. However, an 85% improvement was observed after the participants had received the metaphor awareness raising lesson, at the immediate post-test stage. It could be said that an 85% increase in STM was achieved due to the experimental method. STM, and not WM is relevant in this context because Working Memory is not associated with memory storage, unlike Short Term Memory (Engle et al., 1999).

4.2.2.2 Lenient marking scheme

The null hypothesis was that the distribution of the scores across the 3 points in time were equal. The alpha was set at .05, with 2 degrees of freedom. Based on the lenient marking scheme, the participants increased their scores from the pre-test level (*Mdn* = 41.08) by 56.5% at the immediate post-test level (*Mdn* = 64.29), and subsequently decreased by 25% at the delayed post-test stage (*Mdn* = 48.22). However, there was no statistically significant difference among the three weeks, $\chi^2 = 5.67$, (2, $N = 22$), $p = .059$. Therefore, the null hypothesis was retained. A Wilcoxon Signed-Rank Test was not performed as the finding was not statistically significant. Tables 4.9 and 4.10 below display the analysis.

Table 4.9: Friedman's Test Statistics for Experimental Group (Lenient Marking)

Test Statistics[a]

N	22
Chi-Square	5.671
df	2
Asymp. Sig.	.059
a. Friedman Test	

Table 4.10: Medians for Within-Group Analysis (Experimental Group, Lenient Marking)

Report

Median		
Pre-test scores (lenient)	**Immediate post-test scores (lenient)**	**Delayed post-test scores (lenient)**
41.0750	64.2900	48.2150

4.2.3 Summary of Strict and Lenient Within-Group Analysis

According to the strict marking scheme, a statistically significant improvement was observed after the participants underwent the metaphor awareness raising treatment (85%). Between the pre and delayed post-test levels, the participants' retention of the taught idiomatic phrases had improved by 31%.

Naturally, a lenient scoring scheme meant that the participants would be obtaining a higher score in their cloze tasks, due to not being penalised for deviations in form (syntax and/or morphology) as long as their answers were contextually appropriate and the taught idioms were used, as well as being awarded points for providing different but metaphorically synonymous answers. However, the lenient marking scheme data revealed no statistically significant findings.

Nonetheless, by examining the median scores of both the strict (*Mdn* = 23.22) and lenient (*Mdn* = 41.08) pre test scores, one can observe that the participants fared much

better, in the lenient scheme. As for the immediate post-test level, the difference in median scores for the strict (*Mdn* = 42.86) and lenient (*Mdn* = 64.29) was even more pronounced. At the delayed post-test stage, the strict (*Mdn* = 30.36) and lenient (*Mdn* = 48.22) scores further showed that when contextually appropriate idioms were prioritised over form (syntax/ morphology), the participants were able to use the idioms more successfully in given contexts. Discussions on the pedagogical implications of these findings are presented in Chapter 5.

4.2.4 Between-Group Analysis

To answer the second research question, which was concerned with the comparative effects of the two methods for the retention of English idioms, the test scores of the experimental and control groups were matched according to each test phase. Overall, it was revealed that there were no statistically significant differences throughout the data. This meant that both groups performed similarly when compared side by side, irrespective of marking scheme type. Nonetheless, a breakdown is supplied.

4.2.4.1 Strict marking scheme

A Mann-Whitney U test was carried out to determine if there were statistically significant differences in test scores between the memorization group (control) and metaphor awareness raising group (experimental). The shape of distribution throughout both groups were similar, except at the delayed post-test level, according to a Levene's Test for homogeneity of variance.

At the pre-test level, test scores were not statistically significantly different between the control group (*Mdn* = 21.43) and the experimental group (*Mdn* = 23.22), $U = 187$, $p = .41$. At the immediate post-test level, test scores were also not statistically

significantly different between the control (*Mdn* = 44.65) and experimental group (*Mdn* = 42.86), $U = 197$, $p = .56$.

At the delayed post-test level, test scores were once again not statistically significantly different between the control group (*mean rank* = 20.85) and the experimental group (*mean rank* = 22.09), $U = 207$, $p = .74$. Tables 4.11, 4.12, and 4.13 show the results obtained with the strict marking scheme.

Table 4.11: Test Statistics for Between-Group Data Analysis (Strict Marking)

Test Statistics[a]

	Pre-test scores (strict)	Immediate Post-test scores (strict)	Delayed Post-test scores (strict)
Mann-Whitney U	187.000	197.000	207.000
Asymp. Sig. (2-tailed)	.405	.561	.742
a. Grouping Variable: group			

Table 4.12: Medians for Between-Group Data Analysis (Strict Marking)

Report

Median			
Group	**Pre-test scores (strict)**	**Immediate Post-test scores (strict)**	**Delayed Post-test scores (strict)**
Control	21.4300	44.6450	26.7850
Experimental	23.2150	42.8600	30.3550

Table 4.13: Ranked Data for Between-Group Data Analysis (Strict Marking)

Ranks

	Group	N	Mean Rank	Sum of Ranks
Delayed Post-test scores (strict)	**Control**	20	20.85	417.00
	Experimental	22	22.09	486.00
	Total	42		

4.2.4.2 Lenient marking scheme

Once again, distributions of the test scores throughout both groups were similar, except at the delayed post-test level, according to a Levene's Test for homogeneity of variance.

At the pre-test level, test scores were not statistically significantly different between the control group (*Mdn* = 33.93) and the experimental group (*Mdn* = 41.08), *U* = 171, *p* = .22. Despite not displaying a statistically significant difference, a question of practical significance arises, seeing as there is an apparent difference between the median scores of 41.08 (experimental group) and 33.93 (control group). However, at the immediate post-test level, concerns on practical significance do not surface, due to the differences in the scores being marginal. Naturally, there was no statistically significant difference between the control (*Mdn* = 60.72) and experimental groups at this level (*Mdn* = 64.29), *U* = 195, *p* = .53.

At the delayed post-test level, test scores were once more not statistically significantly different between the control group (*mean rank* = 20.38) and the experimental group (*mean rank* = 22.52), *U* = 197.5, *p* = .57. Tables 4.14, 4.15, 4.16 show the results obtained through the lenient marking scheme.

Table 4.14: Test Statistics for Between-Group Data Analysis (Lenient Marking)

Test Statistics[a]

	Pre-test scores (lenient)	Immediate Post-test scores (lenient)	Delayed Post-test scores (lenient)
Mann-Whitney U	171.000	195.000	197.500
Asymp. Sig. (2-tailed)	.216	.527	.570
a. Grouping Variable: group			

Table 4.15: Medians for Between-Group Data Analysis (Lenient Marking)

Report

	Median		
Group	**Pre-test scores (lenient)**	**Immediate Post-test scores (lenient)**	**Delayed Post-test scores (lenient)**
Control	33.9250	60.7150	37.5000
Experimental	41.0750	64.2900	48.2150

Table 4.16: Ranked Data for Between-Group Data Analysis (Lenient Marking)

Ranks

	Group	N	Mean Rank	Sum of Ranks
Delayed Post-test scores (lenient)	**Control**	20	20.38	407.50
	Experimental	22	22.52	495.50
	Total	42		

4.2.5 Summary of Strict and Lenient Between-Group Analysis

While virtually no statistically significant differences were observed across the findings, it is interesting to note that at the pre-test level, the experimental group performed observably better than the control group, under the lenient marking scheme. Though not yielding statistical significance, a 7-point difference was revealed, with the experimental group performing better than the control group. This may suggest that even by simply arranging the idioms according to categories, which was then followed up by a brief explanation of their meanings, the students were able to fare better than those from the control group, who merely had their idioms listed down in no particular order, while possessing the perceived benefit of having the explanations attached to the respective idioms (refer to appendix B). The scores obtained under the lenient marking scheme seem to suggest that the participants from the experimental group were able to provide more contextually accurate answers compared to those from the control group, despite neither of the group being penalised for deviations in form (syntax/ morphology).

At the immediate post-test level, the control group managed to perform better than the experimental group, but not by a statistically significant margin. However, the experimental group exhibited a better retention of the idioms over time, as evidenced by the delayed post-test scores.

Hence, though not statistically significant, the metaphor awareness raising method seems to be somewhat more effective in helping L2 learners to retain English idiomatic phrases taught in a similarly formal setting. In terms of practical significance, the analyses suggest that the raising metaphor awareness method may be more beneficial overall than the conventional, memorizing method of teaching idioms. However, these findings do not invalidate the conventional practice of using memorisation as a method of teaching idioms. Discussions on the pedagogical implications of these findings are presented in Chapter 5.

4.3 Analysis of Qualitative Data

From a quantitative phase, we move onto a qualitative one, as is the nature of an explanatory sequential research design. Through a thorough look into the interview transcripts of 10 student participants, tabled content analyses were generated.

The tables below illustrate content analysis beginning at its base level, or the root of the theme, which presents itself in the form of "meaning units", followed by "condensed meaning units", which serve to present only the most salient information gleaned from the raw data. Next, "codes" were formed to assist in the formation of "categories" and ultimately, "themes" (Bengtsson, 2016). In other words, the qualitative analyses were carried out inductively (from the bottom up) for each test phase (pre, immediate post, and delayed post), according to the respective groups.

An analytical report is presented after each table. They serve to provide insights reflective of the results of the quantitative data. As a general point of reference, the within group data (RQ #1) showed that there was a statistically significant improvement at the immediate post-test phase (strict scoring). As for the between groups data, the analyses revealed that there were no statistically significant findings, but a question of practical significance could be raised, especially at the pre-test phase.

The qualitative results of the raising metaphor awareness (experimental) group are presented first, followed by the memorisation (control) group. The sub-headers accompanying each table indicate the focus of each qualitative analysis, and are reflective of the interview questions. The participants' experiences undergoing the three test phases are reported, because they serve as a foundation to transition into the third research question, which concerns using idioms for aiding vocabulary retention through the respective methods. Therefore, the third research question was answered at the end of both groups' qualitative content analyses. For the purpose of identification, the quotes of participants from the experimental group were attached with a label E, while those from

the control group were labelled C. An ID number was attached to each participant (e.g. participant "E1"). The quotes presented in the tables were paraphrased for brevity.

4.3.1 Qualitative Content Analysis: Experimental Group

This section discusses the qualitative findings generated via Qualitative Content Analysis for the experimental group. The discussion starts with an analysis of the interviewees' pre-test experience, followed by their immediate post-test experience, and their experience during the delayed post-test. In addition, the interviewees (potential English language teachers) were questioned on the viability of the raising metaphor awareness method for teaching idioms among school students. Subsequently, they were asked about their opinions on using idioms to retain vocabulary effectively. The final question was about using idioms to retain vocabulary through the experimental method.

4.3.1.1 Pre-test experience

Based on the content analysis of the pre-test responses in Table 4.17, it was found that all five participants in the experimental group mainly relied on just memory retrieval to solve the cloze task. It was found to be a stressful ordeal:

"...I was very blur... I'd be like, what_is this about, an- and I don't- I don't know anything about the (inaudible) and then I just... wrote what I remember (inaudible vocal fry [01.16]) ... And I just wrote uh whatever I remember I don't care about the sentences (inaudible words with laughter) ... I did my best to remember the metaphor there but once I saw the sentence and the blanks there ... I lost my (chuckling) I lost my brain ... (inaudible [2.05 to 2.06]). Oh, what did I, memorise just now, oh just write I ..."

(Excerpt from participant E1's responses on the pre-test)

Despite these challenges, the experimental group at the pre-test stage (lenient marking scheme) achieved a higher median score of 41.08 compared to the control group's 33.93. This finding may indicate that even without explicit teaching and having the meanings attached to the idioms, the participants in the experimental group received some benefit in recalling the items, through the arrangement of the idioms according to metaphoric themes.

However, the differences between these two median scores were not statistically significant (p = .22). In other words, the perceived advantage possessed by the control group did not seem to aid them as much as the raising metaphor awareness method did on the experimental group. At the pre-test stage, it seemed that the biggest issue for the experimental group concerning learning was that the task was given to them "cold" (no assistance with the comprehension of the inventory). Clearly, without the metaphor awareness raising method, the participants were stressed while completing the task, even though the idioms were arranged according to their metaphoric themes and meanings were explained orally.

Table 4.17: Experimental Group's Pre-Test Content Analysis

Theme	**Mainly relied on memory to solve a stressful task**		
Category	Cognition	Solution	Attitude
Sub- Category	Negative cognitive performance	Method for overcoming problem	Negative attitudes toward the cloze task
Codes	Feeling lost and disoriented	Reliance on memory for answers	Apathy or resentment
Condensed Meaning Unit	Very blur. Don't know anything. Lost. Hard.	Remember the metaphor. Writing remembered metaphors. Tried my best to remember. Write anything I know.	Didn't care about the sentences.
Meaning Unit	I was very blur. (E1) What is this about? (E1) I don't know anything. (E1) I lost my brain. (E1) It was kinda hard. (E2) I didn't have an introduction. (E2) I barely used the idioms before this (E3) It was hard to remember the metaphor given (E4) I was dumbstruck maybe? (E4) It was quite difficult. (E5)	I just wrote what I remember. (E1) I did my best to remember the metaphor. (E1) What did I memorise just now? Oh just write it. (E1) I just write what uh anything that I know (E3) I had to create my own idioms (due to not being able to remember them) (E5)	I didn't care about the sentences. (E1)

4.3.1.2 Immediate post-test experience

Based on the content analysis of the immediate post-test responses in Table 4.18, all five participants in the experimental group found the raising metaphor awareness method to be helpful in the task, but participant E5 felt that the task was as stressful as the first time. Naturally, memory retrieval played a part in the process of completing the task. For those who commented on how the experimental method affected their ability to recall the taught idioms, it was overall found to be beneficial in aiding the recall process:

"... it wasn't that bad because I managed to remember some of them but still there are others I can't remember... it was quite helpful ... (inaudible) I could relate to things like- like my life and (these) things (so yeah)... it (kinda) activates how like how people (inaudible) to these things... uh (inaudible [2:48 to 2:49]) it's the same for me ..."

(Excerpt from participant E5's responses on the immediate post-test)

A surge in test score performance was observed in the within group analysis. According to the strict marking scheme, there was an 85% improvement from the pre to immediate post-test. The raising metaphor awareness method can be credited for this significant result, since the participants themselves reported an improvement in cognitive ability, namely, being able to identify categories. This realisation is key to the raising metaphor awareness method because being able to notice the patterns between source and target domains can facilitate memory retrieval (Gao and Meng, 2010), which was evident here.

That being said, it appeared that participant E5 still found the task challenging, even after undergoing the raising metaphor awareness method. Nonetheless, participant

E5 still reported a positive learning experience through the method, as seen in the quotes above.

Table 4.18: Experimental Group's Immediate Post-Test Content Analysis

Theme	**Improvement of attitude towards learning the idioms which helped recall**			
Category	Learning experience	Cognition	Solution	
Sub-Category	Positive/similar learning experience	Improved cognitive ability	Method used to solve the task	Positive method used to solve the task
Codes	Feelings toward the task	A clearer understanding	Reliance on memory for answers	Beneficial method
Condensed Meaning Unit	More interest. I can do it. Same experience as the first time.	This answer is suitable for this sentence. It's easy.	Easier to remember. Memorised again.	The method is very helpful
Meaning Unit	The second time I got more interest. (E1) I think I can do it the second time. (E1) And then, it's easy actually! (E1) It was the same for me (E5)	Oh! This one is suitable for this sentence. (E1) I can see I can imagine that, this thing is much easier to understand now (E2) I actually divide them into categories like the face category, the hair category so once I divide them into categories I know what fall under bones, under face (E4) I could relate to things like my life. (E5)	I just write what I remember last time. (E1) Memorised it again. (E1) much clearer for me to understand and to remember. (E4) I managed to remember some of them but still there are others I can't remember. (E5)	It's very helpful. (E1) It's quite easy because you explained it well so I can answer it well. (E3) I think it's very helpful (E4) It was quite helpful. (E5)

4.3.1.3 Delayed post-test experience

The content analysis of the delayed post-test responses in Table 4.19 reveal that all five participants in the experimental group expressed their agreement with the raising metaphor awareness method in aiding them with recall; even though the delayed post-test was administered 2 weeks after the immediate post-test, coupled with the added difficulty of not receiving the idiom inventory, the participants still reported that the experimental method helped them with recall:

"... it's kinda hard because it was. two weeks apart... but then I just remembered where the uh trying to imagine the things the idioms. So I tried so it's still not as hard as the first task... is... because the first one I don't really understand what's going on and I- I don't really know how to imagine this thing because I don't really understand but the second when you teach me how it is and that and that the third even though it's already like... lama dah (been awhile) (chuckling) lama dah uh belajar balik benda tu (revised that thing) but... I still can recall by imagine it so it's- it's I felt the same as the second one..."

(Excerpt from participant E2's responses on the delayed post-test)

Based on the between group quantitative analysis of the delayed post-test, the experimental group fared better than the control group, albeit not yielding statistically significant results. Nonetheless, it is worth noting that participants E1, E2, and E5 mentioned that they tried to recall the idioms they had learned earlier by imagining them, which is a feature of the raising metaphor awareness method. By spending cognitive effort to visualise the source and target domains of an idiom, it is argued that it becomes easier to recall the form of the intended idiom to be used (Boers, 2004).

For instance, if one were to try to recall the idiom *grit one's teeth*, the cognitive linguistics method would be to raise the awareness of the association between *struggle* and *teeth* metaphors during the teaching/ learning phase. Once the individual has learned that the abstract concept of *struggle* can be metaphorically expressed through the concrete image of *teeth*, and since having teeth is ubiquitous to the general human experience, one would likely be able to recall the idiom with less effort.

The reasoning above could explain why the participants felt the raising metaphor awareness helped them during the delayed post-test, even though it was perceived as challenging for them.

Table 4.19: Experimental Group's Delayed Post-Test Content Analysis

Theme	**Inconsistency in cognitive performance despite conscious effort but the experimental method helped**			
Category	Cognition			
Sub-Category	Negative cognitive performance	Determination in the face of cognitive difficulties	Positive cognitive performance	Merging schemata with taught idioms
Codes	Lost and disoriented	Conscious effort	Experimental method made recall easier	Assimilation between taught and known idioms
Condensed Meaning Unit	Went blank again. Couldn't remember anything. It was difficult.	Tried recalling answers. Tried imagining.	Managed to remember certain answers. The method made a difference. Not as hard as the first task.	Know the meaning. Used own sentences.
Meaning Unit	At first I went blank again. (E1) Some of those I remember. Some of those I only remember the meanings but I don't remember the exact answers. (E1) It's kind of hard because it was two weeks apart (E2) I couldn't remember anything (E3) It was still quite difficult (E5) .	And then I tried to recall the answers that I wrote the past two weeks. (E1) But then I tried to imagine the things, the idioms. (E2) I was actually imagining why it's like this. She literally pulls her hair. (E5)	I managed to remember certain answers for the questions. Bit his finger. Talk heart to heart. Pull a face. (E1) Not as hard as the first task. (E2) The lesson made a difference (E2) The explanations (method) helped a bit (E3) It wasn't too bad (E4) The method was helpful (E5)	But I know the meaning so I just wrote what using my own sentences. (E1) I jumbled up the idioms you gave me with the ones I already know (E4)

4.3.1.4 Viability of raising metaphor awareness method for teaching idioms

Table 4.20 displays the content analysis of the responses about the viability of using the raising metaphor awareness method for teaching idioms. On this matter, all five participants from the experimental group viewed the cognitive linguistics method in a favourable light. However, participant E1 seemed to suggest that learning idioms can be a difficult task, in itself. At the same time however, participant E1 was quick to encourage the raising metaphor awareness method for teaching idioms:

"... as for me I think it's interesting but sometime (it) stress (slight chuckle) stressful because sometime I- I don't understand the idiom? and I try hard to understand the idioms but still I didn't get a clue... it helps because I can understand the sentences more i can relate with the questions (inaudible).... (so, that's it) ..."

(Excerpt from participant E1's responses on the viability of using the raising metaphor awareness method for teaching idioms)

All five participants from the experimental group seemed to be enthusiastic about the raising metaphor awareness method for teaching idioms. Despite participant E1's seeming exasperation with the ordeal of learning idioms in general, the cognitive linguistics method itself appears to be well received by the said participant. This bodes well for pedagogical discussions on this matter, and more is discussed in Chapter 5.

Table 4.20: Viability of the Raising Metaphor Awareness Method for Teaching Idioms Content Analysis (Experimental Group)

Theme	**A beneficial method when trying to learn idioms despite meeting with difficulty**		
Category	Cognition		Solution
Sub-Category	The method can present difficulty	Determination in the face of cognitive difficulties	Positive method used to solve the task
Codes	Difficult at times	Conscious effort	Beneficial method
Condensed Meaning Unit	Sometimes stressful. Sometimes don't understand the idiom. No clue.	Try hard to understand the idioms.	Interesting. It helps. Can understand the sentences more. Can relate with the questions. Enjoyable. Effective. Refreshing.
Meaning Unit	But sometimes stressful. (E1) Sometimes I don't understand the idiom. (E1) Didn't get a clue. (E1)	I try hard to understand the idioms. (E1)	I think it is interesting. (E1) It helps because I can understand the sentences more. (E1) I can relate with the questions. (E1) I think it's creative (E2) More enjoyable and effective compared to memorisation. (E2) Refreshing and helpful method. (E3) We should create awareness of metaphors. (E4) Language without metaphors is dull. (E4) Helps to create new words. (E5)

4.3.1.5 Using idioms to retain vocabulary effectively among students

The content analysis of the viability of using idioms to retain vocabulary effectively among students in Table 4.21, reveals a positive outlook on this cognitive linguistics inspired method. Overall, the participants from the experimental group opined that idioms can indeed aid vocabulary retention. However, the main challenges of this approach to vocabulary education appear to be the issues of item difficulty and the attitudes of the students toward idioms, as participant E4 explained:

"... I think uh (inaudible) metaphor using very abstract vocabulary so I think it will be it will expand their vocabulary knowledge..."

(Excerpt from participant E4's responses on the viability of using idioms to aid vocabulary retention)

Participant E4's response was particularly noteworthy as the individual made mention of the pervasiveness of the abstract quality in metaphorical language. It appears that the "abstract vocabulary" mentioned in the excerpt above refers to how there is a propensity for metaphorical language to deal with concepts such as *happiness* and *despair*. Such lexical items cannot be observed by itself, but only through vehicles such as facial expressions can they be perceived. Therefore, these lexical items are part of what the participant means by "abstract vocabulary".

Thus, students would be learning to notice these "abstract" words and could even develop a deeper understanding about how words can hold "abstract" or "concrete" meanings. A caveat must be placed however, as legitimate concerns must be dealt with in order for meaningful learning to take place: the participants suggest tackling the issue

of item difficulty as well as being aware of the students' attitudes toward using idioms in explicit vocabulary teaching.

Table 4.21: Viability of Using Idioms to Retain Vocabulary Effectively among Students Content Analysis (Experimental Group)

Theme	**A potentially effective method for vocabulary education when difficulty is calibrated well and interest is cultivated**		
Category	Attitude	Pedagogy	Reservation
Sub-Category	Positive experience	Positive method	Reservation in using idioms for vocabulary lessons
Codes	Enjoyment	Beneficial method	Issues with difficulty of idioms. Issues with students' attitudes.
Condensed Meaning Unit	Like using idioms in essay very much.	It's helpful. It will expand vocabulary knowledge	Only certain idioms I really understand. Certain idioms. Depends on the idioms. Depends on the type of students.
Meaning Unit	I very very very like using idioms in my essay. (E1)	So, I think it's helpful. (E1) It will expand their vocabulary knowledge (E4)	But only certain idioms that I really understand. (E1) Maybe there's certain idioms that we can use and maybe it depends on the idioms as well. (E1) Maybe if the idiom is easy to understand then we can use to tell the meanings. (E1) Depends on which type of students (referring to attitude towards idioms) (E5)

4.3.1.6 Using idioms to retain vocabulary via the experimental method

The prior discussions on the qualitative content analyses thus far have been largely concerned with the experiences of the participants from the experimental group, who underwent the three test phases. They were also asked to provide their insights into the viability of using the experimental method for teaching idioms in general, as well as the potential use of idioms to retain vocabulary. These analyses and discussions served to prime the participants for the actual crux of the qualitative analyses as a whole: whether the raising metaphor awareness method can be used to teach and retain chunks of vocabulary, through the explicit teaching of idioms:

"I think it's better to start when they're small because they can understand more and when they're small their mind is lim- more lim- limitless than us when we uh we're grown ... (inaudible) if you want these kids to remember well you mainly need to read books (inaudible) they have they have they have those uh idioms so so (inaudible)..."

(Excerpt from participant E5's responses on the viability of using idioms to retain vocabulary through the raising metaphor awareness method)

While there was an overall appreciation for the raising metaphor awareness method, the participants were concerned about its receptiveness among students. Even though this method was viewed as interesting and perhaps even better than the conventional, memorisation method, some reservations were expressed. The participants appeared to view this method as being rather advanced and maybe out of the boundaries of what Malaysian students are capable of. At the same time however, the participants were optimistic about using this method, provided the difficulty of the idioms be well calibrated. As before, the students' motivation once again was said to be crucial for a successful learning experience through this method, in terms of vocabulary education.

Participant E5's comment on this matter stood out in particular: this method should be introduced to young learners as their minds were perceived to be more malleable and agile. In addition, reading should be cultivated from a young age in order to retain a wide array of vocabulary. It would certainly be interesting to see if this method can be tailored to children as well, as most of the literature surrounding testing these ideas on vocabulary retention and metaphoric thinking consist of studies enlisting participants ranging from teenagers to adults (e.g. Boers & Demecheleer, 1998; Boers and Lindstromberg, 2008; Yasuda, 2010). Table 4.22 illustrates the analysis.

Table 4.22: Using Idioms to Retain Vocabulary via the Experimental Method (Experimental Group)

Theme	**Packaging the method and presenting it in an exciting way**		
Category	Pedagogy		
Sub-Category	Lesson Design		Task Design
Codes	Appreciation of the method	Engaging the students	Issue of difficulty
Condensed Meaning Unit	Maybe I like the method. The technique is useful.	Didn't have the interest in metaphors. Students themselves have their own interests.	Maybe a bit hard for students. Maybe boring.
Meaning Unit	Maybe as for students like me, maybe I like the method that you suggest. (E1) But still it's an interesting… (E1) This technique will really affect them really well (E2) It's much better using the technique you taught me. (E2)	… since they didn't have the interest to understand about the metaphor. (E1) But the students themselves have their own interest. (E1) Students nowadays don't like to just read and memorize they like to do more activities. (E2)	But maybe it's a bit hard for students who don't like this metaphor. (E1) Maybe they can find it boring. (E1) It might be difficult to use this method for people who don't have any interest in languages. (E4)

4.3.2 Qualitative Content Analysis: Control Group

This section discusses the qualitative findings generated via Qualitative Content Analysis for the control group. The discussion starts with an analysis of the interviewees' pre-test experience, followed by their immediate post-test experience, and their experience during the delayed post-test. In addition, the interviewees (potential English language teachers) were questioned on the viability of the memorising method for teaching idioms among school students. Subsequently, they were asked about their opinions on using idioms to retain vocabulary effectively. The final question was about using idioms to retain vocabulary through the memorisation method.

4.3.2.1 Pre-test experience

Table 4.23 displays the content analysis of the responses given by the interviewed participants from the control group, concerning their pre-test experiences. Collectively, it was not a pleasant experience. They generally felt they were given the shock and awe treatment. Only participant C3 felt that it was not too difficult; the task was viewed as challenging but due to the self-reported ability of having an active visual mind, remembering and recalling the idioms posed an adequate challenge. One response that was particularly noteworthy was provided by participant C4:

"uh quite difficult because there are some idioms that quite similar like uh pull a straight make a straight face uh and what uh is similar the the face the face is similar so I don't really remember the mea- uh the different meaning of both (the) idiom …"

(Excerpt from participant C4's responses on the pre-test)

Participant C4 appeared to view the idioms which shared similar source domains as being distractions rather than as a potentially useful mnemonic framework. Conversely, the raising metaphor awareness group arrived at precisely the opposite opinion. For the experimental group, the participants saw how the idioms could share similar source domains, while having different meanings, and that by grouping the respective idioms according to their shared source domains, recall could be improved, and indeed, the participants from the experimental group attributed their successful recall to the raising metaphor awareness method. The opposite is true here. Participant C4 viewed the list of idioms as being just that- a mere list of disassociated idioms, and not a list of idioms that could potentially be grouped together according to categories.

Participant C4's responses could be a reflection of the manner in which the control group handled the idiom inventory list, collectively. Comparatively, the experimental group performed better at the pre-test phase under the lenient marking scheme (*Mdn* = 41.08) than the control group (*Mdn* = 33.93). It appeared that when contextual appropriateness was prioritised over form, the experimental group achieved a better result. This result was not statistically significant however (p = .22), but one could argue that an 8- point difference should hardly be ignored. Even under the strict marking scheme, the experimental group (*Mdn* = 23.22) scored higher than the control group (*Mdn* = 21.43), but as one can observe, the experimental group incurred a large deduction in marks due to being penalised for deviating from the taught idioms in terms of form. Under both marking schemes, the raising metaphor awareness method seems to be more effective in teaching idioms, although a statistically significant score was not obtained.

This insight is exciting because it validates or at least warrants continual research into the explicit teaching of idioms in general, seeing as student participants like C4 had a difficult time committing the taught idioms to memory and a much harder time to recall them. Since "learning" was used in this study to mean the extent one can recall taught or

learned idioms, if one cannot recall what one has learned, then what learning has taken place? These ideas on learning are drawn from John Sweller's work on Cognitive Load Theory (1994).

Sweller and Chandler (1994) explained that problem solving skill is inextricably linked with the ability to retrieve information from the long-term memory. In the context of this current study, the problem is that many Malaysian L2 learners of English do not possess an adequate proficiency in communicating in English. Therefore, real effort should be spent on aiding learners to retain the idioms more effectively, because using idiomatic expressions improves one's communicative competence, as outlined in the CEFR [(2001, 27, CEFR, as cited in Pérez (2017)].

Table 4.23: Control Group's Pre-Test Content Analysis

Theme	**Mainly relied on memory to solve a stressful task**		
Category	Cognition	Solution	Attitude
Sub- Category	Negative cognitive performance	Method for overcoming problem	Generally negative attitude toward the cloze task
Codes	Feeling overwhelmed	Reliance on memory for answers.	Exasperation. Challenging.
Condensed Meaning Unit	Lost. Memorising was the hardest part. Understanding was the most difficult part.	Remember without any instruction. Remember the idiom first.	Shocked. Scared. Not easy but not too hard. Couldn't even understand.
Meaning Unit	I lost everything. (C1) Memorising the idioms was the hardest part. (C2) A little bit hard. (C3) Quite difficult because there are some idioms that quite similar. (C4) Understanding the idioms was the most difficult part. (C5)	I have to remember (things) without any instruction? (C1) I try to remind myself that it is uh important to remember the idiom first rather than the meaning. (C4)	Kind of shocked? (C1) I feel like very scared if I couldn't memorise all of it. (C2) I am a visual person so I think the task was not quite easy but not too hard. (C3) To remember you must understand right? I couldn't even understand some of the idioms. (C5)

4.3.2.2 Immediate post-test experience

The content analysis of the immediate post-test responses from the control group in Table 4.24 reveal that there was a marginal improvement in their attitudes toward learning the items, which can be attributed to the explanations given by the researcher:

"...okay on the second session I feel like I can do it a bit better than the first one because ah when you told me I hear it and I get the idea y-know like when you say it I-I-I can I can r-remember it by how I read it and how you told me so (it's both) ways (you know) like (listening) and memorising ..."

(Excerpt from participant C2's responses on the immediate post-test)

When asked about the experience of recalling the idioms at the immediate post-test phase, Participant C2 attributed an improved learning experience to the addition of the teaching session conducted by the researcher. Even with the perceived benefit of having the meanings of the idioms being next to each respective idiom, Participant C2 implied that explicit teaching made a difference; by listening to the researcher's elaborations and examples, the learning experience appeared to have been magnified. In contrast, the participants from the experimental group did not have any of the meanings attached to each idiom. Those in the experimental group had to rely on just listening to the researcher.

Based on the qualitative content analyses of both groups, the learning attitudes toward the tasks were similar. The only difference would seem to be the type of cognitive strategy employed, for recall. On one hand, the experimental group utilised the raising metaphor awareness method, which helped them to notice patterns and group idioms

together according to source domains. On the other, the control group had a random list of idioms to study from, but had the meanings of the idioms attached. It would appear that the conventional method for teaching idioms still holds water. In fact, while participants in the experimental group praised the raising metaphor awareness method and at the same time, criticised conventional memorisation, the findings of this study show that the memorisation method was in fact, better than the raising metaphor awareness method, at the immediate post-test stage. The control group obtained a median score of 44.65, compared to the experimental group's score of 42.86. However, the difference between these scores was not statistically significant, $p = .56$.

Table 4.24: Control Group's Immediate Post-Test Content Analysis

Theme	**Slight Improvement of attitude towards learning the idioms which helped recall**		
Category	Learning experience	Cognition	Solution
Sub-Category	Positive/ indifferent learning experience	Improved cognitive ability	Positive method used to solve the task
Codes	Feelings toward the task	A clearer understanding	Beneficial method
Condensed Meaning Unit	Same as the first time. It's better. Still quite hard.	Easier. A bit better.	Explained Showed examples
Meaning Unit	For me it's the same. (C1) I feel like i can do it a bit better than the first one. (C2) It's better than the first. (C3) Actually it's quite similar because I think it is too many for me to remember the idiom and the meaning of them. (C4)	the second time I can answer it easier because I understand what it means and I can imagine in my mind. (C1) It helped us to remember since you explained. (C3) It was better but still quite hard to remember. (C5)	Since you explained… you kind of showed some gestures and show for example pulling her own hair. (C3)

4.3.2.3 Delayed post-test experience

Based on Table 4.25, the content analysis of the delayed post-test revealed a stark difference with the experimental group's responses, matched at the same test phase. For the experimental group, the delayed post-test certainly proved to be difficult for them, however the experimental group's participants expressed that the raising metaphor awareness method helped them to recall the idioms. Participant E2 even opined that the delayed post-test felt like the immediate post-test. As the meanings of the idioms were never present on the experimental group's inventory list, it was no wonder they did not express a need to rely on it for reference.

The participants from the control group however explicitly stated that not having the opportunity to look at the idiom inventory list made it that much more difficult for them. This would suggest that they placed a reliance on the list itself. It would appear that this reliance on the reference material translated into poorer performance compared to the experimental group. Though not statistically significant, $p = .74$, the experimental group (*mean rank* = 22.09) performed better than the control group (*mean rank* = 20.85) under the strict marking scheme. Similar scores were obtained under the lenient marking scheme, with the experimental group (*mean rank* = 22.52) performing better than the control group (*mean rank* = 20.38). The real difference however, can be seen in terms of the participants' attitudes toward the respective methods in dealing with the cloze task:

"… after the last session we kind of uh I kind of forget all the idioms but uhm since the question is same like the last session it kind of tried to recall and just think logically try to put something (inaudible)… mm because we didn't have the the-the list…"

(Excerpt from Participant C3's responses on the delayed post-test)

Based on the interview responses of both groups, the experimental group evidently relied on the raising metaphor awareness method to aid their recall of the idioms. However, the control group placed their reliance on the inventory sheet itself for the same purpose. Interestingly Participant C1 implied that the idiom inventory sheet was instrumental in forming images in the mind. This response appears to affirm the nature of human cognition being one that perceives reality via mental images in the form of metaphors.

The interviewed participants from the experimental group appeared to be more optimistic about their experience during the delayed post-test. On the other hand, the delayed post-test responses from the control group did not reflect a positive outlook towards the task; while both groups fared similarly, the experimental group viewed the cognitive linguistics method positively in aiding recall, while the control group focused on lamenting their lack of opportunity to refer to the idiom inventory sheets.

Table 4.25: Control Group's Delayed Post-Test Content Analysis

Theme	**The absence of the reference material was the main reason for the negative cognitive performance**	
Category	Cognition	Role of Reference Material
Sub-Category	Negative cognitive performance	Reliance on reference material
Codes	Stressful session	The list was important
Condensed Meaning Unit	Shocked. Barely remembered anything. It was difficult.	Didn't have reference to make images. Didn't have the list. Difficult.
Meaning Unit	I kind of shocked. (C1) It was very difficult. (C2) I barely remember anything. (C2)	I didn't expect you to come back with the same thing but without the explanation. (C1) I don't have any reference to make images in my mind. (C1) It was difficult because we didn't have the list. (C3) you don't give us a time to see the (inventory list). (C4)

4.3.2.4 Viability of memorisation method for teaching idioms

The content analysis revealed in Table 4.26 shows that the attitude toward the memorisation method for teaching idioms was mostly negative. Only Participant C3 said that it was a useful method, but at the same time, the individual thought that perhaps it was due to being used to the method that led to the positive outlook. However, Participant C3 could not arrive at a conclusive answer to whether the method was beneficial for learning idioms or not. The rest, however, were clear in their responses:

"... I-I think it is not very suitable fors students because uhmm tsk y-know if you just give the the tsk what the words only? they can barely remember any of it i (grunt) except eh I mean like tsk if you put pictures or you tell stories then I think they can

remember it well because y-know they can imagine it what is it like that like (um) pull her hair macam ((like)) like that… okay (chuckles) … think by just by by doing your way the on the first session (by) not explaining anything? it is very hard for students to (take that) and tsk ya when you you know (when) ours our education (method) you just gi-uh when the test come like the third session? they can they cannot do it la (emphasis particle) they cannot do it (very well)…"

(Excerpt from Participant C4's responses on the viability of the memorising method for teaching idioms)

It was apparent that memorisation was an unsavoury method for all but one participant, and the main reason for this disdain was simply due to the perception that not everyone is good at memorising. Interestingly, Participant C5 opined that learning English should entail a dynamic pedagogy, which favours real- life communication, rather than the perceived, cold method of memorisation. This insight is noteworthy because while most of the interviewed participants did not enjoy the conventional method, the test results obtained through the memorisation method was better than that of the experimental method, at the immediate post-test phase. In addition, the control group collectively obtained a similar, albeit slightly lower score than the experimental group, at the end of the 3 test phases. To throw out memorisation in favour of the experimental method would not be prudent, as the results speak for themselves.

However, what must be considered is the L2 learners' cultivation of interest in learning idioms, especially on their own. Therefore, it does not appear that the memorisation method would be able to foster this interest in the long run, when they would no longer be required to write essays in a school or university setting.

Table 4.26: Viability of the Memorising Method for Teaching Idioms Content Analysis (Control Group)

Theme	**A conventional but generally not a beneficial method for teaching idioms**	
Category	Opinions	Explicit Teaching
Sub-Category	Mostly negative attitude toward the method	Explicit teaching of idioms in schools
Codes	Mixed attitudes	Experience of learning idioms in school
Condensed Meaning Unit	Not very suitable for students. Not everybody is good at memorising. Useful for visual learners. Not too fond of it. Not for English.	Memorised idioms in school. Sometimes gave lists. Didn't do anything with the idioms. Just explained. Encouraged to memorise idioms.
Meaning Unit	Not very suitable for students. (C2) If you give the words only they can barely remember any of it. (C2) Not for all students. (C3) I don't know whether it's a good or bad method. I kind of used to the technique. (C3) The book of articles containing idioms was useful for a visual learner like me because I can refer to it. (C3) I don't think it's good because not all students are good at remembering. (C4) Maybe i can remember the idiom but not the meaning. (C4) I'm not too fond of it. (C5) You cannot expect everybody to remember something in such a short period. (C5) I'm not a very good memoriser. (C5) Maybe for maths but not for English. (C5) English should be learned through two- way communication not memorisation. (C5)	In secondary school I also remember idioms for my essay. (C1) My teacher sometimes give the list. (C1) They don't ask me to do anything with that just explain it. (C1) Encouraged to memorise idioms. (C3)

4.3.2.5 Using idioms to retain vocabulary effectively among students

The findings from the content analysis in Table 4.27 display overlapping issues with the responses from the experimental group in Table 4.21. In the experimental group, the main issues were related to item difficulty and the learning attitudes of students. In addition to those main issues, Participant C2 advised that students need to be aware of context, and not just write down the whole idiom without checking if modifications are necessary. Participant C4 also suggested that it would not be wise to disseminate too many idioms at a given point in time. Besides criticising idioms for being too lengthy to remember, Participant C1 voiced out the seemingly arbitrary nature to certain idioms where the meanings cannot be inferred upon when observing those idioms. *Kick the bucket* serves as a good example to illustrate this legitimate frustration (Vega-Moreno, 2003). However, the pros seem to outweigh the cons when it comes to using idioms for aiding vocabulary retention:

"...like the idioms you gave us was a bit ah intermediate to a higher level right uh so maybe the ss the easier one like in the pink of health like like you have to give them yes (chuckling) you have to give them (an) idiom that is a they can use in all situations… like pink of health is used in most letters in most greetings right so it will be easier for them to remember rather than those who they uh those who they have to remember a situation then an idiom… yes because it comes in a sentence it's a package so of course they will have to remember the whole thing in order for the package to be complete…"

(Excerpt from Participant C5's responses on the viability of using idioms to retain vocabulary effectively among students)

Participant C5's responses were particularly insightful. On the issue of item difficulty, it was suggested that easier idioms be given to beginner learners of English. These easier idioms come in the form of being high frequency idioms, such as *pink of health*, which is often used in greetings and letters. Participant C5 subsequently went on to explain how students would be able to retain more words as idioms present themselves in syntactically accurate packages. This observation precisely aligns itself with the Lexical Approach (Lewis, 1993), which is concerned with learning or teaching vocabulary in chunks, or in this case, packages. Without briefing Participant C5 on this approach to vocabulary education, the individual was able to draw such a connection.

By committing whole idioms to memory, the burden of producing grammatical accuracy can be reduced. In fact, if students were to be able to potentially notice (Schmidt, 2010) the syntactic relationship between individual lexical items through learning idioms, their grammar could also very well improve, apart from knowing more words.

Although the participants from the control group were not briefed about the cognitive linguistics aspect of the study, Participant C2 seemed to pick up on the idea that idioms can help learners to retain words more effectively because they can be relatable in our lives. The body part metaphors contained in the idioms used in this study were noticed and viewed as welcome features to aid recall.

Table 4.27: Viability of using idioms to retain vocabulary effectively among students Content Analysis (Control Group)

Theme	**A potentially effective method for vocabulary education when volume and length are calibrated well**		
Category	Attitude	Pedagogy	Reservation
Sub-Category	Mixed attitudes	Mostly positive method	Reservation in using idioms for vocabulary lessons
Codes	Idioms are too long. Perceived meaning may not match the form of idiom. It can help.	Idioms are too long. Using idioms for vocabulary retention is good.	Issues with form and contextual appropriateness. Issues with volume.
Condensed Meaning Unit	Don't like it. Quite long. Meaning sometimes different from the idiom. Helps to remember vocabulary.	Short and simple vocabulary is better. Good and creative. Can relate it to ourselves. For all levels. Package.	Be careful with context. Too many isn't good.
Meaning Unit	I don't like idioms because it's quite long and the meaning behind it is sometimes so different with the idioms. (C1) It can help them remember vocabulary. (C4)	better to use short and simple vocabulary. (C1) I think it is a good and creative way because we can relate it to ourselves (body parts). (C2) Suitable for intermediate to advanced levels. (C5) Use simple idioms for beginners. (C5) Yes because they have to know all the words in the package (idiom) for it to be complete. (C5)	Students need to be careful with context. Might use "pull *her* hair out" for boys. (C2) It's not good if the teacher gives too many. (C4)

4.3.2.6 Using idioms to retain vocabulary via the memorisation method

Table 4.28 displays the control group's divided sentiments on the use of idioms to retain vocabulary via the memorisation method. For those against the memorisation method, the issue of word length was highlighted again. This concern on the length of idioms was tied to the perceived challenging nature of learning idioms itself, thereby suggesting that it be most suitable for advanced learners. A concern on the practical aspect of using idioms was raised as well: if learners do not practice using the idioms in real-life settings, then they would be forgetting the idioms they learned in school. On the other hand, Participant C3 expressed that idioms which clearly contain relatable metaphors (e.g. body parts) are intrinsically grounded and connected with our daily lives, which naturally implies that memory recall would be made easier. In addition, Participant C5 explained in detail the benefits of this method:

"...like I said it was a package so of course they will remember the whole thing rather than just one word like ah pull a face they will remember pull a face and not just pull or face right so they have to remember the whole thing so in benefit they will get three words rather just (inaudible) remembering one word... uhm in my opinion.. it did help (uhm not uh) in any level it will help. because vocabulary is your like is your base if you don't have the vocab uh vocabulary in English then what would you h-h-how will you do an essay how will you speak right... so learning idioms is one way to achieve a good vocabulary even though it's simple words but those simple words will actually lead up to a whole sentence a whole essay like how can you do a good essay if you don't even have a proper base... it can like make the sentence longer (chuckling)..."

(Excerpt from Participant C5's responses on using idioms to retain vocabulary via the memorisation method)

According to Participant C5 on the matter of vocabulary retention, the main reason for approving the use of idioms to retain vocabulary through the memorisation method was because idioms are like packages. These packages contain more than one word, and the idea is that learners would be exposed to a wider variety of words through idioms; instead of learning words individually, these students can cover more ground by memorising idioms. The result would be the production of longer sentences. For Participant C5, this method of aiding vocabulary retention is suitable for all proficiency levels.

Table 4.28: Using idioms to retain vocabulary via the memorisation method (Control Group)

Theme	**Mixed attitudes toward using idioms to retain vocabulary through the memorisation method**	
Category	Pedagogy	
Sub-Category	Negative opinion about the method	Positive opinion about the method
Codes	Concerns about length, difficulty, and practical use	Relatable and suitable method for all levels to build and retain vocabulary
Condensed Meaning Unit	Idioms are too long. A lot of words in idioms. Teach words individually. For advanced level. Good method but without application, we won't remember anything. Focus on the meanings to remember idioms and the words better.	Idioms are connected (to our lives). One way to achieve good vocabulary. Students will remember the words from the packages (idioms). Benefit of 3 words rather than 1. Simple words lead up to a whole sentence. Suitable for all levels. Helpful.
Meaning Unit	Memorising idioms isn't good for me because it's too long. (C1) No because there's a lot of words in idioms. (C2) Teach the words individually. (C2) Maybe idioms is for the advanced level. (C2) It's good to memorise but if we didn't apply we won't remember anything at all after that. (C3) Instead of remembering the idioms only the teacher should stress more on the meaning of the idioms. We can remember idioms and the words better by focusing on the meaning of the idioms. (C4)	When we are mad we do make that kind of gesture pulling our face so the idiom is connected. (C3) Learning idioms is one way to achieve a good vocabulary. (C5) Since idioms are packages, the students will remember the words that form the package. (C5) They get the benefit of 3 words rather than 1. (C5) They might be simple words but they will lead up to a whole sentence. (C5) It was helpful. suitable for all levels. (C5)

4.3.3 Summary of Qualitative Data Analysis

Based on the themes generated from the responses of the experimental group participants, it was found that at the pre-test, when they were given the categorised idioms without being taught their meanings, the participants mainly relied on their ability to recall in order to solve the cloze task, which was perceived as stressful. The immediate post-test phase however saw an improvement of attitude towards learning the idioms, after learning about the raising metaphor awareness method, which helped recall. The same positive view for the raising metaphor awareness method was held at the delayed post-test phase, even though they experienced an inconsistency in cognitive performance despite putting in conscious effort to solve the cloze task.

On the issue of using the raising metaphor awareness method for teaching idioms, the experimental group participants found it to be a beneficial method when trying to learn idioms, despite encountering difficulty in understanding some of the items. Subsequently, when asked about whether idioms can actually facilitate vocabulary education, with emphasis on retention, they saw the raising metaphor awareness method as a potentially effective method for vocabulary education when difficulty is calibrated well and interest is cultivated among the learners. The participants agreed that the raising metaphor awareness method made a difference in their ability to recall the idioms, and that the method would be beneficial for the school setting. Considering all the above responses, they opined that in order for this method to be beneficial for the purpose of retaining vocabulary through learning idioms, it is important to consider the packaging of the method, as well as to present it in an exciting way.

The responses of the participants from the control group appeared to be more nuanced. Just like the experimental group at the pre-test stage, they were found to be mainly relying on memorisation to solve the task, which was also perceived as stressful. The difference here was that the control group had the meanings of the idioms alongside

each respective idiom in the inventory, while the experimental group did not throughout the 3 test phases. At the immediate post-test phase, there was a slight improvement of attitude towards learning the idioms using the memorisation method, due to the researcher explaining each of the idioms in detail. The participants agreed that the lesson helped their recall in the cloze task during this phase.

At the delayed post-test phase however, the absence of the reference material was cited as the main reason for their negative cognitive performance in the cloze task. It seemed that there was an overwhelming reliance on the physical inventory sheets, and since they were not distributed for the final test phase, the respondents were quick to voice out their collective frustration during the interview. Their experimental group counterparts however did not seem to be affected much by not receiving their idiom inventories. In fact, they met the challenge very positively due to having their metaphor awareness raised.

Concerning using memorisation for the purpose of retaining idioms, it was found to be an unpopular method. The participants saw the method as being conventional but not beneficial for the teaching of idioms. However, when asked about using idioms to facilitate vocabulary growth and retention, they opined that it is a potentially effective method for vocabulary teaching and learning when the volume of the idioms and their length (of words) are calibrated well. Overall, the participants from the control group expressed mixed attitudes toward the notion of using idioms to retain vocabulary through the memorisation method.

4.4 Chapter Summary

This chapter examined the findings from two strands of analyses. Quantitative methods yielded some preliminary findings, followed by the qualitative phase. The role of the quantitative analyses was to obtain the test scores to answer the first and second research questions, while the qualitative analyses served to provide a deeper understanding of the issue, as well as to answer the third research question. Brief summaries according to each research question are provided below:

1. What effect does raising metaphor awareness have on learners' retention of English idiomatic phrases?

2. What are the comparative effects between the raising metaphor awareness method and the memorising method on learners' retention of English idiomatic phrases?

3. How do the students view the use of idioms for aiding vocabulary retention through the respective methods?

To answer the first research question, a quantitative method using the Friedman's Analysis of Variance (ANOVA) Test was carried out to gauge how the participants from the experimental group fared overall, beginning from the pre-test, then moving on to the immediate post-test, and finally to the delayed post-test. Two sets of scores were obtained: one being from the strict marking scheme and the other, through the lenient marking scheme. The purpose of having two marking schemes was to offer a wider variety of data to be observed. The test scores obtained through the strict marking scheme reflected a focus on form and meaning (contextual appropriateness), while test scores from the lenient marking scheme prioritised meaning only. Doing so allowed the researcher to generate a more detailed discussion on pedagogical implications (Chapter 5). Overall, the

raising metaphor awareness method was proven to yield statistically significant results under the strict marking scheme (p = .004). From the pre-test to the delayed post-test phase, the participants marked according to the strict marking scheme had achieved a 31% improvement in test scores. The lenient marking scheme results yielded no statistical significance, which meant that the scores obtained across the 3 test phases were similar.

The second research question was concerned with how the raising metaphor awareness method measured up with the memorising method, a conventional method of teaching idioms. To achieve this, a between groups analysis using the Mann Whitney U Test was employed. Surprisingly, the differences in test scores between the two groups, matched at every test interval, yielded no statistical significance. In other words, the two paired groups performed similarly throughout the 3 test phases. However, a question of practical significance was raised and elaborated on, seeing as an 8-point difference was observed between the two groups matched at the pre-test phase.

As for the third research question, the researcher presumed that it would not have been beneficial to ask that question directly. This was because three ideas were present: idioms, vocabulary retention, and the memorising method/ raising metaphor awareness method. As such, the researcher ultimately approached the third research question by first introducing and engaging the interviewed participants with each concept, and how they were interlinked. Doing so had the added benefit of generating rich responses on the entire testing process.

In sum, the experimental group's responses collectively reflected a positive attitude toward the experimental method, despite a number of pedagogical concerns. As for the control group, the respondents were divided on whether the memorising method could be used for idioms in the pursuit of greater vocabulary retention. These pedagogical issues are discussed in the following chapter.

CHAPTER 5: CONCLUSION

5.1 Introduction

The main purpose of this chapter is to discuss the pedagogical implications of this study based on the quantitative and qualitative analyses of the data. Although the purpose of the third research question was to guide such a discussion, and has been examined at length in the previous chapter, a more consolidated effort was spent in this chapter to elaborate on the pedagogical implications on vocabulary teaching in particular, which include potential methods to aid vocabulary retention.

The first two discussions are presented in a sequential pattern, whereby the first discussion builds upon the next. The discussions first start with a look at the rationale for the explicit (intentional) teaching of idioms to aid vocabulary retention. Next, a practical method is offered to aid this pedagogical endeavour, in the form of organising the idioms according to metaphorical themes. In addition, the pedagogical implications of continuing the use of the conventional method of memorising idioms are also discussed.

Subsequently, the researcher provides an acknowledgement of limitations faced in this study, with explanations. The chapter ends with further research recommendations.

5.2 Pedagogical Implications

The central reasoning behind progressing from merely discussing the effects of raising metaphor awareness on L2 learners' retention of English idioms, to seeing how it was related to vocabulary education, was that idiomatic phrases serve as syntactically "fixed" packages of meaningful utterances. Idioms are packages in a sense that they are used in syntagmatically and paradigmatically fixed forms (Sprenger, 2003). In other words, the choice of lexical items and their syntactical relationships between each other

(form) contained in a given idiomatic phrase are not likely to change. It is this "fixed" nature of idioms that will be focused on concerning its pedagogical implications, in light of the analysed data.

5.2.1 The Rationale Behind Explicit Teaching of Idioms in Vocabulary Education

The CEFR (2018) has expounded on the importance of raising communicative competence. To achieve this end, learners' vocabulary range has to be increased. However, focusing on increasing one's vocabulary range would not be complete without employing effective methods to retain the taught vocabulary.

It is suggested that the intentional teaching of idioms would serve this endeavour well. This is due to the fact that since idioms occur mainly in fixed forms, educators can help learners to notice (Schmidt, 2010) the recurring metaphoric themes and morphosyntactic patterns shared among the vast number of English idiomatic expressions, which may then lead to the transferring of the input from the short term memory to the long term memory. In other words, by explicitly teaching idioms in classrooms, L2 learners of English would be increasing their vocabulary range through the learning of idioms, which has the added benefit of increasing retention.

The following discussions highlight the practical pedagogical implications of the two studied methods in this research: the raising metaphor awareness method and the memorising method.

5.2.2 Raising Metaphor Awareness: Organising Idioms According to Metaphorical Themes

Based on the interview responses from both experimental and control groups, the explicit teaching of idioms was more readily supported by the experimental group participants than those from the control group.

The interviewees from the experimental group credited the raising metaphor awareness method as being beneficial for the aim of retaining taught idioms, as well as for vocabulary. It was cited as an interesting and creative method, and a few participants repeatedly mentioned how they could imagine the idioms in their minds, which facilitated their understanding and recall of the idioms. Participant E2 also posited that the raising metaphor awareness method was more enjoyable than the memorising method, which was encouraged in secondary school.

In Cognitive Load Theory (CLT), Sweller (2008) discusses the human cognitive architecture, which mainly relies on the machinations of memory for learning. Imagination plays a very beneficial role in transferring input related to procedures or concepts from the short term memory to the long term memory. In fact, experiments utilising CLT methods have shown that when compared to conventional methods of learning, namely by simply studying the learning material, imagining the information appears to be the more beneficial technique for rapidly transferring input into the long-term memory (Leahy and Sweller, 2008). It would appear that not only is the raising metaphor awareness method of learning idioms perceived as a more enjoyable and effective alternative to memorisation on the part of the interviewees, the literature concerning learning from the cognitive science perspective appears to affirm the efficacy of the raising metaphor awareness method for teaching idioms, seeing as the method incorporates the imagining of conceptual metaphors.

However, this present study was not able to find statistically significant differences for the between-group findings. Even though the experimental group performed better than the control group at the end of the three test phases, there was no statistically significant difference in test scores. That being said, it was reported that the experimental group interviewees' collective attitude towards the raising metaphor awareness method was a positive one, whereas half of the control group's responses reflected a negative attitude towards the memorising method.

One element of the raising metaphor awareness method that was appreciated by the experimental group participants, was that the idioms were organised according to metaphorical themes. As such, educators wishing to implement this method into teaching idioms, or even extending the method by using idioms for vocabulary teaching, may consider replicating the format of the experimental group's idiom inventory (appendix A).

5.2.3 Memorisation: Attaching Meanings to Respective Idiomatic Phrases

The memorisation method used in the present study for the control group reflected a current conventional method for teaching idioms in Malaysian secondary schools, as affirmed by the control group interviewees (TESL students).

Based on the findings of this study, there was no statistically significant difference in test scores between the two groups, matched at each testing phase. The qualitative data however painted a more nuanced picture concerning the memorising method: half of the interviewees from the control group did not view the memorising method as beneficial for learning idioms. However, they liked the idea of using idioms to learn vocabulary. One participant even perceived idioms as being made up of chunks or packages of lexical items without the researcher prompting the individual.

One of the biggest difficulties they had with the memorising method was that there were too many idioms (14 in total). Interestingly, this sentiment was not present in the experimental group's responses. One possible explanation could be that by organising the same 14 idioms according to metaphorical themes without including the explanations, the experimental group's idiom inventory did not overwhelm them with information. The converse was true for the control group as their idiom inventory listed down the 14 idioms in no particular order, while having their respective meanings attached to them (appendix B). CLT explanations may affirm this observation because working memory is not able to hold more than nine elements of information at any given time (G. Cooper, 1998). It appears that the experimental method reduced cognitive load, despite employing the same number of idioms in the control group.

Therefore, while the memorisation method was not proven to be inferior to the raising metaphor awareness method for retaining taught idioms, a few negative perceptions were encountered.

5.3 Limitations

Two limitations are acknowledged in this present study: the lack of a test-retest reliability for internal consistency and a possible statistical interference from a practice effect due to the nature of a within-group design. However, rationales are provided to illustrate the unique circumstances attached to the context of this study.

5.3.1 Absence of Test-Retest Reliability for Internal Consistency

According to Brown, (1996) the goal of carrying out a reliability test for a task would be that when the task is administered to a similar group of participants, the results should be similar, which reflects the NRT type of goal. A reliability test requires the production of variances in scores to determine the feasibility of the test being used across

a wide population of students. This was not applicable to the current study because CRT type tasks are not developed for this reason. However, generalizations can still be posited, albeit to the extent of similar populations to the sample in this current study.

However, there exists reliability testing strategies that do not need a high standard deviation. Brown (1989) expounds on the alternatives to test the reliability of tasks: threshold loss agreement, squared-error loss agreement, and domain score dependability. These methods were specifically designed to test the reliability of CRT type tasks. That being said, none of those methods could be used in this study because all of them hinge on the prerequisite that the scoring method be dichotomous (right/ wrong), to which the researcher applied a weighted scoring system instead in this present study. The rationale was that it was worthwhile to examine the degrees of correctness or incorrectness, based on the parameters which determined full, partial, or no marks.

5.3.2 The Practice Effect

A discussion on the practice effect cannot be avoided when employing a within-group study design. The reason for this is due to the sequential or repetitive nature of this type of research design, in which participants can "carry-over" habits from prior tests (Altermatt, 2014). This would lead to the undesirable outcome of conflating statistical data. However, there are contexts which permit the presence of the practice effect. One such context is when a researcher intentionally wants to study the effects of repeated practice of taught material. The current study serves to illustrate this point.

One of the goals in this study was to examine the extent of the participants' ability to recall the same taught idioms over a period of time. The pre-test offered findings which show that when the participants of both groups were given their respective idiom inventories with only a simple instruction to remember the idioms, neither of the groups provided satisfying results in test performance. In fact, in both groups, the participants

overwhelmingly viewed their pre-test experiences in a negative manner. In the immediate post-test, the findings revealed that when detailed instruction was provided shortly before the test, both groups exhibited much greater median scores. In the delayed post-test, the participants endured a two-week drought of intentional contact with the idiom inventories and the cloze task. When interviewed, a few participants reported that during those two weeks, they were pre-occupied with other matters and hardly reflected on the two test phases.

The above descriptions show the particular contexts in which the participants operated in, along with the researcher's goal of observing how through repeated and interrupted practice affected their test performances. Thus, the deliberate inclusion of this practicing element should not detract the pertinent findings of this study (Greenwald, 1976).

However, the use of the same unedited cloze task for both groups over the 3 test phases is acknowledged. To reduce the risk of conflating statistical significance, the participants did not know that the same cloze task would be administered over the three points in time.

In addition, corrective feedback was not provided in any of the test phases, so that the participants would have less incentive to store the phrases and sentences from the cloze task itself in their long-term memory.

5.4 Recommendation for Future Research

The present study was conducted on the premise that L2 learners of English in Malaysia generally do not possess the vocabulary range to communicate competently, according to the CEFR (2001), which is the current English language educational paradigm held by the Malaysian Ministry of Education. To solve this issue, educators should devote more time and effort into conducting vocabulary-centred lessons, by way of teaching idioms, according to the lexical approach.

While the 42 participants included in this study cannot be reflective of Malaysian learners in general, this study does indeed paint a positive picture for the pedagogical implications of incorporating the raising metaphor awareness method of teaching idioms, to aid vocabulary teaching.

As noted in Chapter 2, etymological elaboration has continually challenged the necessity for rote memorisation. Therefore, a replication study on whether truly opaque idioms like "kick the bucket" can be best taught by memorisation compared to etymological elaboration should be considered, as such opaque idioms would not be able to be categorised metaphorically. The rationale for not dismissing seemingly arbitrary idioms is that the CEFR (2001, 110) requires that learners possess a range of such opaque idioms. Previous research has shown that syntactically bound idioms are more easily learned than syntactically flexible ones (T. C. Cooper, 1998). This means that idioms like "kick the bucket" would be more easily learned despite being very opaque. Hence, such a study could provide more pertinent pedagogical implications for the teaching of such opaque idioms.

REFERENCES

Alqahtani, M. (2015). The importance of vocabulary in language learning and how to be taught. *International Journal of Teaching and Education, 3*(3), 21-34.

Altermatt, B. (2014). Threats to internal validity for within-subjects designs. Retrieved December 2, 2019, from https://vault.hanover.edu/~altermattw/courses/220/readings/Within_subjects.pdf

Andreou, G., & Galantomos, L. (2008). Designing a conceptual syllabus for teaching metaphors and idioms in a foreign language context. *Porta Linguarum: Revista Internacional de Didáctica de las Lenguas Extranjeras,* (9), 69-78.

Asgari, A., & Mustapha, G. B. (2011). The type of vocabulary learning strategies used by ESL students in University Putra Malaysia. *English Language Teaching, 4*(2), 84-90.

Atkinson, D. (2011). *Alternative approaches to second language acquisition*: Taylor & Francis.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *Psychology of Learning and Motivation,* (2), 89-195: Elsevier.

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of learning and motivation*, (8), 47-89.

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2019). From short-term store to multicomponent working memory: the role of the modal model. *Memory & cognition, 47*(4), 575-588.

Bengtsson, M. (2016). How to plan and perform a qualitative study using content analysis. *NursingPlus Open, 2*, 8-14.

Beréndi, M., Csábi, S., & Kovecses, Z. (2008). Using conceptual metaphors and metonymies in vocabulary teaching. *Applications of Cognitive Linguistics, 6*, 65.

Boers, F. (2000). Metaphor awareness and vocabulary retention. *Applied Linguistics, 21*(4), 553-571.

Boers, F. (2001). Remembering figurative idioms by hypothesising about their origin. *Prospect*, 16(3), 35-43.

Boers, F. (2004). Expanding learners' vocabulary through metaphor awareness: What expansion, what learners, what vocabulary. *Cognitive linguistics, Second Language Acquisition, and Foreign Language Teaching*. Berlin: Mouton de Gruyter.

Boers, F., Demecheleer, M., & Eyckmans, J. (2004). Cross-cultural variation as a variable in comprehending and remembering figurative idioms. *European Journal of English Studies, 8*(3), 375-388.

Boers, F., & Lindstromberg, S. (2008). Opening chapter: How cognitive linguistics can foster effective vocabulary teaching. *Cognitive linguistic approaches to teaching vocabulary and phraseology*, 01-61.

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences, 105*(38), 14325-14329.

Brown, J. (1989). Criterion-referenced test reliability. *University of Hawai'i Working Papers in ESL, 1*, 79-113.

Brown, J. D. (1996). *Testing in language programs.* Upper Saddle River, NJ: Prentice Hall.

Charteris-Black, J. (2000). Metaphor and vocabulary teaching in ESP economics. *English for Specific Purposes, 19*(2), 149-165.

Che Musa, N., Lie, K. Y., & Azman, H. (2012). Exploring English language learning and teaching in Malaysia. *GEMA Online Journal of Language Studies, 12*(1), 35-51.

Chen, Y.-c., & Lai, H.-l. (2013). Teaching English idioms as metaphors through cognitive-oriented methods: A case in an EFL writing class. *English Language Teaching, 6*(6), 13.

Cooper, G. (1998). Research into cognitive load theory and instructional design at UNSW: Citeseer, 01-33.

Cooper, T. C. (1998). Teaching idioms. *Foreign language annals, 31*(2), 255-266.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in brain research, 169*, 323-338.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of verbal learning and verbal behavior, 11*(6), 671-684.

Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five tradition*. Thousand Oaks, CA: Sage.

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches (2nd ed.)*. Thousand Oaks, CA: Sage publications.

Darmi, R., Saad, N. S. M., Abdullah, N., Behak, F. P., Zakaria, Z. A., & Adnan, J. N. I. (2017). Teachers' views on students' performance in English language proficiency courses via CEFR descriptors. *IJAEDU-International E-Journal of Advances in Education, 3*(8), 363-370.

de Europa, C. (2018). Common European Framework of Reference for Languages: learning, teaching, assessment companion volume with new descriptors. *Estrasburgo: Consejo de Europa.* Retrieved December 2, 2019, from https://rm. coe. int/cefr-companion-volume-with-new-descriptors-2018/1680787989.

Deignan, A., Gabryś, D., & Solska, A. (1997). Teaching English metaphors using cross-linguistic awareness-raising activities. *ELT journal, 51*(4), 352-360.

Derrac, J., García, S., Molina, D., & Herrera, F. (2011). A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. *Swarm and Evolutionary Computation, 1*(1), 3-18.

Council of Europe. Council for Cultural Co-operation. Education Committee. Modern Languages Division. (2001). *Common European Framework of Reference for Languages: learning, teaching, assessment*. Cambridge University Press.

Dobreva, A. (2018). Etymology of Idioms and Their Teaching. *E-Newsletter*, 42.

Elyas, T., & Alfaki, I. (2014). Teaching Vocabulary: The relationship between techniques of teaching and strategies of learning new vocabulary items. *English language teaching, 7*(10), 40-56.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of experimental psychology: General,* 128(3), 309.

Folse, K. S. (2006). The effect of type of written exercise on L2 vocabulary retention. *TESOL quarterly, 40*(2), 273-293.

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association, 32*(200), 675-701.

Gao, L.-q., & Meng, G.-h. (2010). A study on the effect of metaphor awareness raising on Chinese EFL learners' vocabulary acquisition and retention. *Canadian Social Science, 6*(2), 110-124.

Gastwirth, J. L., Gel, Y. R., & Miao, W. (2009). The impact of Levene's test of equality of variances on statistical theory and practice. *Statistical Science, 24*(3), 343-360.

Gibbs, R. W. (1993). Why idioms are not dead metaphors. *Idioms: Processing, Structure, and Interpretation*, 57-77.

Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse education today, 24*(2), 105-112.

Greenwald, A. G. (1976). Within-subjects designs: To use or not to use? *Psychological Bulletin, 83*(2), 314.

Hoover, D. L. (2016). Metaphors we may not live by. *International Journal of Literary Linguistics, 5*(1).

Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. *Vocabulary and Applied Linguistics*, 113-125.

Hummel, K. M. (2010). Translation and short-term L2 vocabulary retention: Hindrance or help? *Language teaching research, 14*(1), 61-74.

Irujo, S. (1986). A piece of cake: learning and teaching idioms. *ELT journal, 40*(3), 236-242.

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in second language acquisition, 24*(4), 541-577.

Jermoškin, S. (2017). *The cognitive linguistic approach to teaching idioms: a case study with Estonian secondary school EFL students.* University of Tartu.

Jr, J. W. O. (1972). Scoring methods and difficulty levels for cloze tests of proficiency in English as a second language. *The Modern Language Journal, 56*(3), 151-158.

Kecskes, I. (2007). Formulaic language in English lingua franca. *Explorations in pragmatics: Linguistic, cognitive and intercultural aspects, 1*, 191-218.

Kellogg, R. T., Newcombe, C., Kammer, D., & Schmitt, K. (1996). Attention in direct and indirect memory tasks with short-and long-term probes. *The American journal of psychology*, 205-217.

Kövecses, Z. (2000). *A cognitive linguistic view of learning idioms in an FLT context*: LAUD, 01-33.

Kövecses, Z. (2008). Conceptual metaphor theory: Some criticisms and alternative proposals. *Annual review of cognitive linguistics, 6*(1), 168-184.

Krashen, S. D. (1981). *Second language acquisition and second language learning*: Oxford University Press.

Krashen, S. D. (2004). *The power of reading: Insights from the research* (2nd ed.). Westport, CT: Libraries Unlimited.

Lakoff, G. (1987). The death of dead metaphor. *Metaphor and symbol, 2*(2), 143-147.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.

Lange, P. G. (2019). Informal learning on YouTube. *The International Encyclopedia of Media Literacy*, 01-11.

Larsen-Freeman, D. (2001). Teaching grammar. *Teaching English as a second or foreign language, 3*, 251-266.

Larzabal, C., Tramoni, E., Muratot, S., Thorpe, S. J., & Barbeau, E. J. (2018). Extremely long-term memory and familiarity after 12 years. *Cognition, 170*, 254-262.

Laufer, B. (1997). The lexical plight in second language reading: Words You Don't Know, words You Think You Know, and Words You Can't Guess. *Second Language Vocabulary Acquisition*.

Lazar, G. (1996). Using figurative language to expand students' vocabulary. *ELT Journal, 50*(1), 43-51.

Leahy, W., & Sweller, J. (2008). The imagination effect increases with an increased intrinsic cognitive load. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 22*(2), 273-283.

Lewis, M. (1993). *The Lexical Approach: The State of ELT and a Way Forward*. Language Teaching Publications.

Lewis, M., Gough, C., Martínez, R., Powell, M., Marks, J., Woolard, G. C., & Ribisch, K. H. (1997). *Implementing the Lexical Approach: Putting Theory into Practice*. Language Teaching Publications.

Littlemore, J., & Low, G. (2006). Metaphoric competence, second language learning, and communicative language ability. *Applied Linguistics, 27*(2), 268-294.

Mahdavian, A., & Kormi-Nouri, R. (2008). Effects of attention and levels of processing on explicit and implicit memory function with interesting and uninteresting tasks in university students. *Journal of Applied Sciences, 8*(6), 1055-1060.

Mathai, E. J., Jamian, L. S., & Nair, S. (2004). Assessing Malaysian university students' English vocabulary knowledge. In W. Khanittanan, & P. Sidwell, (Eds.), *SEALSXIV 2004 Conference*: *papers from the 14th meeting of the Southeast Asian Linguistics Society* (pp.219-237). Canberra: Pacific Linguistics, 2008. Retrieved from http://www.academia.edu/download/13120524/seals-14-1.pdf#page=227

McArthur, T., Lam-McArthur, J., & Fontaine, L. (2018). *Oxford Companion to The English Language*. Oxford University Press.

McCarthy, M., & O'Dell, F. (2002). *English Idioms*. Cambridge.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81.

Ming, T. S. (2003). Investigating the 'problem' of memorisation among Malaysian English as Second Language (ESL) learners. Retrieved December 2, 2019, from https://www.researchgate.net/profile/Siew_Ming_Thang/publication/268435195_INVESTIGATING_THE_'PROBLEM'_OF_MEMORISATION_AMONG_MALAYSIAN_ENGLISH_AS_SECOND_LANGUAGE_ESL_LEARNERS/links/54b7e7980cf2c27adc476f35/INVESTIGATING-THE-PROBLEM-OF-MEMORISATION-AMONG-MALAYSIAN-ENGLISH-AS-SECOND-LANGUAGE-ESL-LEARNERS.pdf

Mitchell, D. B. (2006). Nonconscious priming after 17 years: Invulnerable implicit memory? *Psychological Science, 17*(11), 925-929.

Morse, J. M. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing Research*, *40*(2), 120-123.

Nachar, N. (2008). The Mann-Whitney U: A test for assessing whether two independent samples come from the same distribution. *Tutorials in quantitative Methods for Psychology, 4*(1), 13-20.

Nordstokke, D. W., & Zumbo, B. D. (2010). A new nonparametric Levene test for equal variances. *Psicologica: International Journal of Methodology and Experimental Psychology, 31*(2), 401-430.

Nunberg, G., Sag, I. A., & Wasow, T. (1994). Idioms. *Language, 70*(3), 491-538.

Ofte, I. (2014). English academic writing proficiency in higher education: Facilitating the transition from metalinguistic awareness to metalinguistic competence. *Acta Didactica Norge, 8*(2), 17.

Ortega, L. (2014). Second language learning explained? SLA across 10 contemporary theories. In B. VanPatten & J. Williams (Eds.), *Theories in Second Language Acquisition: An Introduction*. (pp. 245-272). Routledge.

Pallant, J. (2007). *SPSS Survival Manual* (3rd ed.). McGrath Hill.

Pan, M. X. (2019). The effectiveness of the Conceptual Metaphor Approach to English idiom acquisition by young Chinese learners. *Metaphor and the Social World, 9*(1), 59-82.

Pérez, R. G. (2017). Teaching Conceptual Metaphors to EFL Learners in the European Space of Higher Education. *European Journal of Applied Linguistics, 5*(1), 87-114.

Repovš, G., & Baddeley, A. (2006). The multi-component model of working memory: Explorations in experimental cognitive psychology. *Neuroscience, 139*(1), 5-21.

Richardson-Klavehn, A., & Bjork, R. A. (1988). Measures of memory. *Annual Review of Psychology, 39*(1), 475-543.

Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning, 45*(2), 283-331.

Roediger III, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*(3), 249-255.

Rosner, B., Glynn, R. J., & Lee, M. L. T. (2006). The Wilcoxon signed rank test for paired comparisons of clustered data. *Biometrics, 62*(1), 185-192.

Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker (Eds.). *Proceedings of CLaSIC 2010, Singapore, December 2-4* (pp. 721-737). Singapore: National University of Singapore, Centre for Language Studies.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158.

Schmitt, N. (1997). *Vocabulary Learning Strategies. Vocabulary: Description, Acquisition and Pedagogy* (pp.199-228). Cambridge University Press.

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research, 12*(3), 329-363.

Simpson, R., & Mendis, D. (2003). A corpus-based study of idioms in academic speech. *TESOL Quarterly*, 37(3), 419-441.

Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *Canadian Modern Language Review, 64*(3), 429-458.

Siyanova-Chanturia, A., & Martinez, R. (2014). The idiom principle revisited. *Applied Linguistics, 36*(5), 549-569.

Smith, B. (2004). Computer-mediated negotiated interaction and lexical acquisition. *Studies in Second Language Acquisition, 26*(3), 365-398.

Spada, N., & Lightbown, P. M. (1993). Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition, 15*(2), 205-224.

Sprenger, S. A. (2003). *Fixed Expressions and the Production of Idioms.* Radboud University.

Statistics, L. ((n.d.)). Mann-Whitney U Test using SPSS Statistics. Retrieved December 2, 2019, from https://statistics.laerd.com/spss-tutorials/mann-whitney-u-test-using-spss-statistics.php

Swan, M. (2002). Seven bad reasons for teaching grammar–and two good ones. *Methodology in Language Teaching: An Anthology of Current Practice*, 148-152.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*(4), 295-312.

Sweller, J. (2003). Evolution of human cognitive architecture. *Psychology of Learning and Motivation, 43*, 216-266.

Sweller, J. (2008). Human cognitive architecture. *Handbook of Research on Educational Communications and Technology* (pp. 369-381). Routledge.

Sweller, J. (2016). Working memory, long-term memory, and instructional design. *Journal of Applied Research in Memory and Cognition, 5*(4), 360-367.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*(3), 185-233.

Syunina, A. S., Yarmakeev, I. E., Pimenova, T. S., & Abdrafikova, A. R. (2018). Killing More than Two Birds with One Stone: Teaching Topical Vocabulary through Idioms. *Cypriot Journal of Educational Sciences, 13*(4), 604-612.

Thornbury, S. (2011). P is for prescriptive. Retrieved December 2, 2019, from https://scottthornbury.wordpress.com/2011/10/02/p-is-for-prescriptive/

Tsaknaki, O. (2016). Cooking verbs and metaphor. *Contrastive study of Greek and French.* 21st International Symposium on Theoretical and Applied Linguistics (ISTAL 21), 458-472.

Tversky, B. (1973). Encoding processes in recognition and recall. *Cognitive Psychology, 5*(3), 275-287.

Ulijn, J. M., & Strother, J. B. (1990). The effect of syntactic simplification on reading EST texts as L1 and L2. *Journal of Research in Reading, 13*(1), 38-54.

Vega-Moreno, R. E. (2003). Relevance Theory and the construction of idiom meaning. *UCL Work Pap Ling, 15*, 83-104.

Ward, N. (2000). Issues in the transcription of English conversational grunts. *Proceedings of the 1st SIGdial Workshop on Discourse and Dialogue.* (pp. 29-35). Hong Kong: Association for Computational Linguistics.

Waring, R., & Nation, I. (2004). Second language reading and incidental vocabulary learning. *Angles on the English speaking world, 4*, 97-110.

Wood, D. (2002). Formulaic language acquisition and production: Implications for teaching. *TESL Canada Journal*, 01-15.

Woolson, R. (2007). Wilcoxon signed-rank test. *Wiley Encyclopedia of Clinical Trials*, 01-3.

Wray, A. (1999). Formulaic language in learners and native speakers. *Language Teaching, 32*(4), 213-231.

Yasuda, S. (2010). Learning phrasal verbs through conceptual metaphors: A case of Japanese EFL learners. *TESOL Quarterly, 44*(2), 250-273.