

**A Study of Vocabulary Knowledge and Vocabulary Learning  
Strategies of Chinese EFL Learners**

Na Fan

A thesis submitted in fulfilment of the  
requirements for the degree of  
Doctor of Philosophy

Department of Linguistics  
Faculty of Human Sciences  
Macquarie University  
Sydney, Australia

07 April 2015



## Acknowledgements

My sincere gratitude goes to the following persons who have contributed to make this research possible. First and foremost, I would like to express my deepest appreciation to my principal supervisor, A/P Mehdi Riazi for his excellent supervision, understanding, patience and consistent encouragement throughout my Ph.D. study. He has always made himself available to clarify my doubts and provide prompt comments on my writings despite his busy schedules. It is a great opportunity for me to do my Ph.D. under his supervision and to learn from his research expertise. My thanks also go to my associate supervisor Ms. Jean Brick.

I would like to register my sincere thanks to my thesis examiners, Dr. Peter Gu, A/P David Hirsh and A/P Michael Lessard-Clouston, for their critical but extremely constructive comments on an earlier version of this thesis.

I would also like to express sincere thanks to my volunteer research participants. Their enthusiastic participation has made this thesis possible. Meanwhile, I would like to thank my friends Mr Xu Wang, Miss Li Huang, Mr Xiaohua Liu, Miss Ling Feng for their kind help in the processes of recruiting participants. Without their help, I would not be able to recruit so many EFL learners within such a short period of time.

Particular thanks go to A/P John Read for his permission to use his unpublished version of Word Associates Test. I am also greatly indebted to my colleagues at Macquarie University and some scholars I met at academic conferences who have provided support through their constructive feedback and discussions. My debt is also extended to my beloved friends who have made me enjoy my life so much during my Ph.D. study in Sydney, in particular Changpeng Huan, Susanna Widyastuti, PeiSoo Ang, Chao Han, Dr. Hua Zhong, Dr. Yupaporn Piriyaasilpa, Mr. Jalu Kuncoro Adhimy, my little friends Revive Jabeza, Jayshawn Yeo and Tylia Yeo.

I am also grateful for the financial support for my Ph.D. study at Macquarie University provided by China Scholarship Council Scholarship and International Macquarie University Research Excellence Scholarship (CSC-iMQRES Ref. No. 2011164). Finally, I would like to express my deepest gratitude to my parents for their understanding and great support.



## **Abstract**

Previous studies on vocabulary knowledge (VK) have primarily focused on EFL learners' vocabulary learning strategies (VLSs) in relation to their breadth of VK. However, the relationship between VLSs and the depth of VK has as yet been underexplored. This study sets out to examine how Chinese EFL learners' gender, discipline, proficiency levels, and VLSs might contribute to their depth, as well as their breadth of VK. A number of 419 Chinese EFL students from four universities in Mainland China took the Vocabulary Size Test (VST), and Word Associates Test (WAT), completed a VLS questionnaire, and wrote an English essay. The two vocabulary tests and the essay were adopted to assess the participants' breadth and depth of vocabulary knowledge respectively. Follow-up semi-structured interviews were conducted to seek any possible convergence and corroboration with the results drawn on the analyses of the quantitative data, or to shed more light on some of the quantitative findings where such convergence is not found. My findings showed that firstly the participants' depth of VK measured by WAT was less satisfactory in comparison to their breadth of VK measured by VST. Secondly, the participants used less frequent and less familiar words in their writings and they seem not good at using various words or phrases to express meaning. Thirdly, the participants favored repetition strategies, inferencing strategies and using dictionaries, but disfavored note-taking strategies, production-based strategies and association and imagination strategies. Fourthly, inferencing strategies made a significant and positive contribution to participants' breadth and depth of VK (VST and WAT) and productive lexical diversity, whereas repetition and social strategies made a significant but negative contribution to their breadth of VK. Self-initiated strategies significantly and positively correlated with some aspects of participants' productive VK. Finally, the overall variables (i.e. gender, discipline, proficiency levels, and VLSs) explained 13.2% of the variance in participants' breadth of VK (VST) and 13.7% of the variance in their depth of VK (WAT) while these variables explained less variance in participants' productive VK. This study may inform curriculum developers and shed light on both vocabulary teaching and learning.

## **Certificate of Originality**

I hereby certify that this work is the result of my own research and that the work has not been submitted for a higher degree to any other institution university or institution. I certify that sources of information used and the extent to which the work of others has been utilized have been indicated in the thesis.

## Table of Contents

<b>Acknowledgements .....</b>	<b>i</b>
<b>Certificate of Originality .....</b>	<b>iv</b>
<b>Table of Contents .....</b>	<b>v</b>
<b>List of Figures .....</b>	<b>ix</b>
<b>List of Tables .....</b>	<b>x</b>
<b>Chapter 1 Introduction.....</b>	<b>1</b>
<b>1.1 Background to the Current Research .....</b>	<b>1</b>
<b>1.2 Research Gap .....</b>	<b>2</b>
<b>1.3 Purpose of the Study .....</b>	<b>4</b>
1.3.1 Research Questions .....	5
<b>1.4 Theoretical Framework .....</b>	<b>6</b>
1.4.1 Richards's (1976) Vocabulary Knowledge Framework .....	7
1.4.2 Nation's (2001) Vocabulary Knowledge Framework.....	8
<b>1.5 Significance of the Study.....</b>	<b>8</b>
<b>1.6 Definitions of Key Terms and Abbreviations .....</b>	<b>9</b>
<b>1.7 Organization of the Thesis.....</b>	<b>10</b>
<b>Chapter 2 Literature Review .....</b>	<b>13</b>
<b>2.1 Studies on English as a Second Language (ESL) Vocabulary Knowledge .....</b>	<b>13</b>
2.1.1 Breadth of Vocabulary Knowledge and Assessment.....	13
2.1.2 Depth of Vocabulary Knowledge and Assessment.....	18
2.1.2.1 Conceptualizing Depth of Vocabulary Knowledge .....	18
2.1.2.2 Assessment of Depth of Vocabulary Knowledge .....	23
2.1.3 Summary .....	28
<b>2.2 Studies on Vocabulary Learning Strategies (VLSs).....</b>	<b>28</b>
2.2.1 Conceptualizing VLSs as Part of LLSs .....	29
2.2.2 Studies on Individual Vocabulary Learning Strategies .....	32
2.2.2.1 Rote Rehearsal .....	32
2.2.2.2 Mnemonic Strategies.....	33
2.2.2.3 Inferencing Strategy .....	35
2.2.2.4 Dictionary Strategies.....	36
2.2.2.5 Summary .....	38
2.2.3 Studies on Mixed Use of Vocabulary Learning Strategies .....	39
2.2.3.1 Operationalizing VLSs through VLS Questionnaires.....	39
2.2.3.2 Main Findings of the Studies on Mixed Use of VLSs .....	42
2.2.3.3 Research Methods of Studies on Mixed Use of VLSs.....	47
2.2.3.4 Summary .....	48
<b>2.3 Conclusion.....</b>	<b>49</b>
<b>Chapter 3 Quantitative Study .....</b>	<b>51</b>
<b>3.1 Research Questions .....</b>	<b>51</b>
<b>3.2 Methods .....</b>	<b>52</b>
3.2.1 Participants.....	52
3.2.2 Instruments.....	54

3.2.2.1 VLS Questionnaire .....	54
3.2.2.2 Word Associates Test (WAT).....	56
3.2.2.3 The Vocabulary Size Test (VST).....	57
3.2.2.4 Essays.....	58
3.2.3 Data Collection Procedures .....	58
3.2.4 Data Analysis Procedures .....	59
<b>3.3 Results .....</b>	<b>62</b>
<b>3.4 Summary of the Quantitative Study.....</b>	<b>88</b>
<b>Chapter 4 Qualitative Study .....</b>	<b>93</b>
<b>4.1 Research Question.....</b>	<b>93</b>
<b>4.2 Methods.....</b>	<b>93</b>
4.2.1 Participants .....	93
4.2.2 Data Collection Procedures .....	94
4.2.3 Data Analysis Procedures .....	95
<b>4.3 Results .....</b>	<b>96</b>
4.3.1 Significance of Learning English Vocabulary.....	96
4.3.2 Difficulties of Learning English Vocabulary .....	98
4.3.3 Current Vocabulary Learning Strategies .....	101
4.3.3.1 Use of Vocabulary Learning Strategies and Perceived Usefulness.....	101
4.3.3.2 Sources of the Vocabulary Learning Strategies.....	109
4.3.4 Perspectives about Vocabulary Knowledge .....	109
4.3.4.1 Perspectives about Vocabulary Size .....	109
4.3.4.2 Perspectives about Vocabulary Depth .....	110
4.3.4.3 Perspectives about the Importance of Vocabulary Size/Depth.....	111
<b>4.4 Summary of the Interview Analyses.....</b>	<b>113</b>
<b>Chapter 5 Discussion of the Quantitative and Qualitative Findings .....</b>	<b>117</b>
<b>5.1 Discussion of Quantitative and Qualitative Findings .....</b>	<b>117</b>
5.1.1 The Pattern of VK among Chinese Students .....	119
5.1.2 The Pattern of Reported VLSs among Chinese Students .....	125
5.1.3 The Relationship between Reported VLSs and VK .....	128
5.1.4 The Relationship between VK and Individual Variables (gender, discipline, and proficiency levels) .....	131
5.1.5 The Relationship between Reported VLSs and Individual Variables (gender, discipline, and proficiency levels).....	135
5.1.6 The Contribution of Overall Variables (gender, discipline, proficiency levels and VLSs) to VK.....	137
<b>5.2 Conclusion.....</b>	<b>138</b>
<b>Chapter 6 Conclusions, Implications, Limitations, and Recommendations for Further Research.....</b>	<b>143</b>
<b>6.1 Summary of the Quantitative and Qualitative Studies.....</b>	<b>143</b>
<b>6.2 Conclusions: Main Findings.....</b>	<b>144</b>
6.2.1 Main Findings Related to VK.....	144
6.2.2 Main Findings Related to Reported VLSs.....	145
6.2.3 Main Findings Related to the Relationship between VK and Reported VLSs ..	145
<b>6.3 Implications of the Study .....</b>	<b>146</b>
<b>6.4 Limitations of the Study .....</b>	<b>147</b>
<b>6.5 Recommendations for Future Research .....</b>	<b>148</b>

<b>References.....</b>	<b>151</b>
<b>Appendix A.....</b>	<b>161</b>
<b>Appendix B.....</b>	<b>167</b>
<b>Appendix C.....</b>	<b>172</b>
<b>Appendix D.....</b>	<b>177</b>
<b>Appendix E.....</b>	<b>187</b>
<b>Appendix F .....</b>	<b>188</b>
<b>Appendix G .....</b>	<b>190</b>
<b>Appendix H .....</b>	<b>191</b>



## List of Figures

Figure 3.1: The distribution of the participants' scores on VST .....	63
Figure 3.2: The distribution of the participants' scores on WAT .....	63
Figure 3.3: Scree plot.....	73
Figure 3.4: Normal P-P plot of regression standardized residuals .....	80
Figure 3.5: Scatterplots of the standard residuals .....	80
Figure 3.6: Normal P-P plot of regression standardized residuals .....	82
Figure 3.7: Scatterplots of the standard residuals.....	83
Figure 3.8: The distribution of the transformed WAT .....	83
Figure 3.9: Normal P-P plot of regression standardized residuals .....	84
Figure 3.10: Scatterplots of the standard residuals.....	84
Figure 5.1: A summary of the quantitative and qualitative findings as a frame for discussions .....	118

## List of Tables

Table 2.1: Individual VLSs in relation to language proficiency .....	38
Table 2.2: Empirical studies on mixed use of VLSs .....	43
Table 3.1: Summary of the participants' demographic information ( $N = 419$ ).....	53
Table 3.2: Descriptive statistics for the participants' VST and WAT scores ( $N = 419$ ).....	62
Table 3.3: Descriptive statistics for the lexical indices of the participants' essays as computed by Coh-Metrix ( $N = 419$ ).....	64
Table 3.4: Summary of MANOVA test for the differences in the participants' receptive vocabulary knowledge (combined VST and WAT) in terms of their gender, discipline, and proficiency levels.....	65
Table 3.5: Summary of MANOVA test for the differences in the participants' scores on VST and WAT in terms of their gender, discipline, and proficiency levels .....	65
Table 3.6: Results of the Tukey HSD test .....	66
Table 3.7: Summary of MANOVA test for the differences in the participants' productive vocabulary knowledge (combined lexical indices) in terms of their gender, discipline, and proficiency levels.....	67
Table 3.8: Summary of MANOVA test for the differences in each lexical index among the participants' proficiency levels .....	68
Table 3.9: Results of the Tukey HSD test .....	69
Table 3.10: Total variance explained .....	71
Table 3.11: Results of Monte Carlo Parallel Analysis (PA) .....	72
Table 3.12: The comparison of initial eigenvalues and PA eigenvalues .....	72
Table 3.13: Factors underlying VLS questionnaire .....	73
Table 3.14: Descriptive statistics for the seven categories of vocabulary learning strategies reportedly used by the participants ( $N = 419$ ) .....	74
Table 3.15: Summary of MANOVA test for the differences in the participants' overall reported vocabulary learning strategies in terms of their gender, discipline, and proficiency levels .....	75
Table 3.16: Summary of MANOVA test for the differences in each group of participants' reported vocabulary learning strategies in terms of their gender, discipline, and proficiency levels .....	75
Table 3.17: Results of the Tukey HSD test .....	76
Table 3.18: Pearson product-moment correlations between participants' reported vocabulary learning strategies and their receptive vocabulary knowledge (VST and WAT) ( $N = 419$ ) .....	77
Table 3.19: Pearson product-moment correlations between participants' reported vocabulary learning strategies and their productive vocabulary knowledge ( $N = 419$ ) .....	78
Table 3.20: Model summary of multiple regression of all predictor variables on VST ( $N =$ $419$ ).....	81
Table 3.21: Summary of correlation coefficients of all predictor variables on VST .....	82
Table 3.22: Model summary of multiple regression of all predictor variables on WAT ( $N =$ $419$ ).....	85

Table 3.23: Summary of correlation coefficients of all predictor variables on WAT .....	85
Table 3.24: Model summary of multiple regressions of all predictor variables on the lexical indices ( $N = 419$ ) .....	86
Table 3.25: Summary of correlation coefficients of all predictor variables on lexical diversity .....	87
Table 3.26: Summary of correlation coefficients of all predictor variables on lexical frequency .....	87
Table 3.27: Summary of correlation coefficients of all predictor variables on word meaningfulness .....	87
Table 3.28: Summary of correlation coefficients of all predictor variables on word familiarity .....	88
Table 4.1: The significance of learning English vocabulary for the four English communication skills .....	96
Table 4.2: Difficulties of learning English vocabulary .....	98
Table 4.3: The reported vocabulary learning strategies .....	101
Table 4.4: The participants' knowledge sources about VLS .....	109
Table 4.5: Perspectives about vocabulary size .....	110
Table 4.6: Perspectives about the depth of vocabulary knowledge .....	111
Table 4.7: Perspectives about the importance of vocabulary size and depth .....	112
Table 4.8: A summary of the interview analyses .....	115
Table 5.1: A summary of discussion of the participants' vocabulary knowledge .....	123
Table 5.2: The integration of quantitative and qualitative conclusions .....	139



## **Chapter 1 Introduction**

In this introductory chapter, I will first provide a background to the research presented in this thesis and the research gap in the literature. I will then explicate the purpose of the study, present the theoretical framework of this study and its significance. I will finally define the key terms and abbreviations, and describe the overall organization of the thesis.

### **1.1 Background to the Current Research**

English as a foreign language has never been so prevailing in China after China joined World Trade Organization (WTO) in 2001 and started its globalization and internationalization. Since then, English has gradually achieved the status of core subject to be taught and assessed alongside with the Chinese language and mathematics that Chinese students have to complete successfully in order to be able to enter senior high school and university. English is also a compulsory subject for all university non-English majors and they are obligated to take the college English courses for at least two years (Cheng, 2008). Usually the second-year students (sophomores) are expected to take the College English Test Band 4 (CET-4), while senior students need to take the CET-6. Passing CET is important for Chinese university students since it is a criterion when seeking a good job after graduation. Moreover, students' performance on CET is used as a benchmark for the evaluation of EFL teachers and universities.

Wilkins's (1972, p. 111) well-known saying that "without grammar very little can be conveyed, without vocabulary, nothing can be conveyed" highlights the importance of vocabulary in English language learning especially for English as a Foreign Language (EFL) students. This is while vocabulary learning is largely regarded as one of the most demanding and challenging tasks for Chinese EFL university students (H. Chen, 2001; Cheng & Wang, 2004). Anecdotally and according to my own observation, Chinese students report that insufficient lexical knowledge has been detrimental to the improvement of their language proficiency. Moreover, research findings by other researchers (e.g., Deng & Zeng, 1998; Ma, 2001) have indicated Chinese students have a lower level of vocabulary size than what has been required in the English Teaching Syllabus.

The significance of vocabulary in foreign language learning on the one hand, and the difficulties encountered by myself and the population of Chinese EFL students at large, on the other, have inspired me to conduct a systematic study exploring the status quo of Chinese EFL students' vocabulary knowledge (VK) through identifying the challenges some Chinese EFL students are facing in acquiring sufficient vocabulary knowledge, and proposing suggestions on good practices that are positive to their vocabulary learning. Previous studies in this field (e.g., Ahmed, 1989; Gu & Johnson, 1996; Schmitt, 1997) have suggested a scale of vocabulary learning strategies in terms of their possible contribution to the growth of vocabulary knowledge. As such, a hypothesis in this research is that some vocabulary learning strategies should be prioritized over others by Chinese EFL students in their pursuit of improving their English vocabulary knowledge. One principal focus of this research situates at the retrieval of such advantageous strategies to facilitate students' enhancement of vocabulary knowledge.

## **1.2 Research Gap**

Researchers in the field of vocabulary teaching and learning have generally acknowledged that vocabulary knowledge is a coin with the two sides of breadth and depth (e.g., Anderson & Freebody, 1981; Qian, 1999, 2002; Read, 1988, 1993, 1998; Schmitt, 2010). Breadth of VK (or vocabulary size) has been extensively investigated in numerous studies, which have resulted in developing vocabulary size tests for second language (L2) learners (e.g., Meara & Jones, 1988; Nation, 1990; Nation & Beglar, 2007). These studies have investigated the proper vocabulary size needed for smooth reading or listening (e.g., Hirsh & Nation, 1992; Hu & Nation, 2000; Nation, 2006), and have explored the relationship between vocabulary size and vocabulary learning strategies (e.g., Gu & Johnson, 1996). However, the investigation of vocabulary size alone hardly suffices to interrogate EFL learners' vocabulary knowledge, since vocabulary depth is equally important and worth parallel research efforts. There are three main approaches towards conceptualizing vocabulary depth, which are dimensional approach (Nation, 2001; Richards, 1976), developmental approach (Dale, 1965; Wesche & Paribakht, 1996), and association approach (Meara, 1996a). Each approach has led to distinctive testing instruments for assessing the depth of VK, such as Word Associates Test (Read, 1993, 1998), V\_Links (Meara & Wolter, 2004) and Vocabulary Knowledge Scale (Wesche & Paribakht, 1996). By comparing the vocabulary tests under the three approaches, the tests following the dimensional approach could better assess learners' depth of vocabulary

knowledge than those of the other two approaches. This point will be elaborated in Chapter 2. Nevertheless, most tests under the dimensional approach (e.g., Ishii & Schmitt, 2009; Schmitt, 1998; Schmitt & Meara, 1997) are still subject to some common problems. For instance, each test only assesses a very limited number of words or assesses limited dimensions of vocabulary knowledge. Recent developments in computational linguistics and automated text analysis have resulted in computer programs such as Coh-Metrix, which could overcome these weaknesses. Coh-Metrix provides a variety of lexical indices of the texts produced by writers such as lexical diversity, lexical frequency and word meaningfulness (Crossley, Salsbury, & McNamara, 2011). The detailed discussion of the affordances and constraints of the above mentioned tests of the depth of VK and the computational tool Coh-Metrix will be presented in Section 2.1.2.2. As we are now equipped with technological affordances to do so, it is necessary to probe more fully into students' vocabulary knowledge.

As the domain of language teaching has been shifting from a predominantly teacher-oriented to a learner-focused orientation, a heightened emphasis has been laid on helping students take more responsibilities for meeting their own language learning needs. Scholarly efforts (Cohen, 1998; O'Malley & Chamot, 1990; Oxford, 1990) in the field of language learning strategy (LLS) over the past four decades, have resulted in the development of more specific inventories such as vocabulary learning strategies (VLSs). Research on VLSs stems from two directions of research. The first one is the research oriented towards exploring the effectiveness of individual VLSs, such as rote rehearsal (e.g., Dahlen & Caldwell-Harris, 2013), mnemonic strategies (e.g., Hulstijn, 1997; Rodriguez & Sadoski, 2000), dictionary strategies (e.g., Chan, 2012; Fraser, 1999), and inferencing strategies (e.g., Fraser, 1999; Hu & Nassaji, 2012). The second one is the research which focused on exploring the most efficient configuration of mixed use of VLSs by L2 learners (e.g., Ahmed, 1989; Catalán, 2003; Fan, 2003; Gu & Johnson, 1996; Kojic-Sabo & Lightbown, 1999; Lawson & Hogben, 1996; Sanaoui, 1992, 1995; Schmitt, 1997; Stoffer, 1995). These studies have explored various facets, including the VLSs adopted by both successful and unsuccessful language learners, how VLSs relate to the success of language learning and what factors would impact on the use of VLSs. Despite such voluminous studies in this field, the following challenges have yet received insufficient research efforts.

Firstly, previous studies on the mixed use of VLSs have detailed the relationship between VLSs and breadth of VK. However, studies investigating the contribution of VLSs as a whole to the depth of VK have been rather insufficient. Our knowledge is also limited regarding the contribution of such variables as gender, discipline, proficiency levels and VLSs to VK as a multi-dimensional construct.

Secondly, despite scholarly efforts (e.g., Fan, 2003; Gu & Johnson, 1996; Schmitt, 1997) devoted to examine and categorize VLS questionnaires on theoretical grounds or vocabulary learning processes, the validity of these classifications has not yet been fully and systematically assessed with the assistance of statistical analysis. Additionally, current VLS questionnaires have unnecessarily prioritized the strategies for learning the meaning of new words over those for learning the form of new words and how to use new words.

Thirdly, a fuller exploration of the status quo of Chinese students' VK, VLSs and their relationship, demands adopting mixed methods in its investigation. However, most previous studies premised either on quantitative or qualitative methods in this area. As Riazi and Candlin (2014) suggest, mixed-methods research (MMR), as an emerging and promising research design, could avoid the adversarial incompatibility of pure quantitative and qualitative studies and the deficiencies of each approach. Accordingly, to overcome the caveat of previous studies, the current research adopts MMR in favor of achieving a more comprehensive understanding of Chinese students' VK and VLSs.

### **1.3 Purpose of the Study**

In light of the background to the research and the research gap identified, as outlined above, and the theoretical framework, as will be discussed in the next section, the present study concentrates on investigating Chinese university students' breadth and depth of VK, paralleled with an interrogation of the relationship of their VK with several variables such as discipline, gender, English proficiency levels and VLSs use. To do so, non-English major university students at four universities in China provide the required data for the study. The data include vocabulary knowledge measured by a range of computational indices related to lexical competence along with the Vocabulary Size Test (Nation & Beglar, 2007) and the Word Associates Test (Read, 1998). Participants' use of VLSs is also be measured by a VLS questionnaire containing, in addition to VLSs, the information about participants' discipline and gender. Participants' English language proficiency level is assessed by using their CET-4 test scores.

The design of this research is, by its nature, a concurrent mixed methods triangulation study (Onwuegbuzie & Johnson, 2004; Tashakkori & Teddlie, 2003). A battery of tests such as the Vocabulary Size Test (VST), Word Associates Test (WAT), and a VLS questionnaire, and a 300-word English essay are used to collect quantitative data so as to explain VK and VLSs of Chinese EFL learners and the interplay of VK with several individual variables such as discipline, gender, English proficiency levels and VLSs. To corroborate and explain the results of the quantitative part of the study, a follow-up interview is used to elicit further information from participants regarding their approach to vocabulary learning and use. Details of the design and methodology of the research will be presented in the third and fourth chapters.

The following are the specific research questions to be answered in this study.

### **1.3.1 Research Questions**

- (1) What is the level of vocabulary knowledge of Chinese non-English major university students in terms of breadth and depth as measured by Vocabulary Size Test (VST), Word Associates Test (WAT) and the lexical indices of their essays?
- (2) Are there any significant differences in the participants' receptive vocabulary knowledge (VST and WAT) in terms of their gender, discipline, and proficiency levels?
- (3) Are there any significant differences in the participants' productive vocabulary knowledge as represented by the lexical indices of their essays in terms of their gender, discipline, and proficiency levels?
- (4) What are the major vocabulary learning strategies reportedly used by Chinese non-English major university students?
- (5) Are there any significant differences in the participants' reported vocabulary learning strategies in terms of their gender, discipline, and proficiency levels?
- (6) Is there any relationship between the participants' reported vocabulary learning strategies and their receptive vocabulary knowledge (VST and WAT)?
- (7) Is there any relationship between the participants' reported vocabulary learning strategies and their productive vocabulary knowledge as represented by the lexical indices of their essays?

- (8) What is the contribution of individual and overall variables (gender, discipline, proficiency levels and vocabulary learning strategies) to the participants' receptive vocabulary knowledge (VST and WAT)?
- (9) What is the contribution of individual and overall variables (gender, discipline, proficiency levels and vocabulary learning strategies) to the participants' productive vocabulary knowledge as represented by the lexical indices of their essays?
- (10) What are the Chinese EFL students' perspectives on vocabulary knowledge and vocabulary learning strategies, and how do they think vocabulary learning strategies affect their vocabulary knowledge?

Four key concepts (i.e. the breadth and depth of vocabulary knowledge and receptive and productive vocabulary knowledge) are particularly significant for understanding the above research questions, and therefore need a clearer explanation. Breadth of vocabulary knowledge refers to the number of words a learner knows. Depth of vocabulary knowledge refers to how well a learner knows a word. These two aspects together constitute the vocabulary knowledge. Receptive vocabulary knowledge refers to the vocabulary knowledge possessed by students when they are engaged in comprehending reading or listening texts. Productive vocabulary knowledge refers to the vocabulary knowledge possessed by the students when they are engaged in and use with speaking or writing activities. The learning of vocabulary knowledge occurs in both receptive and productive processes. In accord with these definitions, in the present study, breadth of vocabulary knowledge is measured by the Vocabulary Size Test (VST), which is also considered as a part of learners' receptive vocabulary knowledge. Depth of vocabulary knowledge is measured by the Word Associates Test (WAT) and the lexical indices of their essays, with the former test assessing learners' receptive vocabulary depth and the latter one measuring their productive vocabulary knowledge. For the sake of clarity, in the following chapters, I will use "depth of VK" to refer to participants' scores on WAT while use "productive VK" to refer to the vocabulary knowledge as represented by lexical indices of their English essays.

#### **1.4 Theoretical Framework**

This research draws on two theoretical frameworks, Richards's (1976) vocabulary knowledge framework, and Nation's (2001) vocabulary knowledge framework. The two frameworks will be discussed next.

### 1.4.1 Richards's (1976) Vocabulary Knowledge Framework

Richards's vocabulary knowledge framework considers the nature of word knowledge as a multi-dimensional construct. Based on Richards's framework, knowing a word does not just mean knowing its meaning, rather, word knowledge is by nature a multi-dimensional construct, which entails semantic, syntactic and pragmatic features of the word. Richards (1976) proposed eight assumptions and only assumption two to eight describe the construct of vocabulary knowledge, which are shown below (Assumption 2-8).

Knowing a word means knowing:

- the degree of probability of encountering that word in speech or print. For many words we also “know” the sort of words most likely to be found associated with the word.
- the limitations imposed on the use of the word according to variations of function and situation.
- the syntactic behavior associated with that word.
- the underlying form of a word and the derivations that can be made from it.
- the network of associations between that word and other words in language.
- the semantic value of a word.
- many of the different meanings associated with the word. (pp.79-83)

Assumption two in Richards's framework requires the learner to recognize that some words are more frequent and familiar than other words. For example, a word like “book” is more frequent than “manual” or “dictionary” (Richards, 1976). In addition, assumption two asserts that knowing a word means knowing other words most likely associated with the word. For example, when encountering the word “fruit” we can expect the words such as “ripe” and “green” as associates (Richards, 1976). This assumption overlaps with the lexical indices of word frequency, word familiarity and word meaningfulness in the computational tool, Coh-Metrix, which will be described later (see Section 1.6). Assumption six also emphasizes the network of association between words, for example, antonym, synonym, subordinate classification, coordinate classification and superordinate classification. This assumption overlaps with some lexical indices in Coh-Metrix, such as semantic co-referentiality and hypernymy. Assumption seven implies knowing a word means knowing the semantic value of a word which can be broken down into a basic set of minimal semantic features, which is related to the lexical index of word concreteness in

Coh-Metrix. Concrete words always have more semantic values than abstract words. Assumption eight entails that by knowing a word, learners should know different meanings associated with words; the lexical index of polysemy in Coh-Metrix. Although Richards's (1976) VK framework is relatively comprehensive, it fails to clarify the classification of receptive and productive VK.

#### **1.4.2 Nation's (2001) Vocabulary Knowledge Framework**

Nation (1990, 2001) further specified the dimensions of vocabulary knowledge and proposed three main categories: form, meaning, and use in both receptive and productive aspects. Form embraces spoken form, written form, and word parts; meaning incorporates form and meaning, concept and referents, and associations; and use encompasses grammatical function, collocations, and constraints on use. The present study adopts two VK tests (VST and WAT) to assess the participants' receptive breadth and depth of VK, while it also examines their productive VK in their English essays by using relevant lexical indices in Coh-Metrix. In addition, the VLS questionnaire used in this study is also mainly built on Nation's vocabulary knowledge framework. Vocabulary knowledge is multi-dimensional as proposed by Nation (2001). For this reason, a comprehensive questionnaire of VLSs is expected to cover the strategies for learning the form, meaning and use of words in receptive and productive processes. Form will be excluded for consideration in the present study because it was found that the length of the questionnaire will reduce the reliability of the instrument. Therefore, the VLS questionnaire is constructed based on two sections: strategies for learning the meaning of new words and strategies for learning the use of new words.

### **1.5 Significance of the Study**

This study's outcomes will make empirical and pedagogical contributions to second language research and will also inform English vocabulary teaching and learning in China.

Firstly, it provides a more balanced exploration of VK of Chinese non-English major university students by examining their breadth and depth of VK and explores the differences of VK in terms of gender, discipline and English proficiency levels.

Secondly, a more comprehensive and tailored VLS instrument is constructed based on Nation's (2001) vocabulary knowledge framework and the VLS questionnaires proposed by Gu and Johnson (1996) and Schmitt (1997). Accordingly, some items for learning how to use a word are incorporated in the present VLS questionnaire. Moreover, the underlying

dimensions of the present VLS questionnaire are explored and verified by using more sophisticated statistical procedures, factor analysis.

Thirdly, it sheds light on those VLSs that participants favor and disfavor in their English vocabulary learning, together with the differences therein in relation to the participants' gender, discipline and English proficiency levels. Therefore, the study outcomes will provide English educators in China useful insights and a better understanding of the pattern of VLSs used by Chinese non-English major university students.

Fourthly, it presents a holistic view on the contribution of VLSs, English proficiency levels, gender and discipline as a whole to participants' VK, which may offer Chinese EFL teachers and learners a better understanding of what factors could account for and facilitate English vocabulary learning.

Finally, the present study adopts the mixed methods approach which combines a relatively large-scale quantitative study with an in-depth qualitative interview study so as to achieve a possible convergence of the results from the two strands to better explain the quantitative results in light of the qualitative results. The present study demonstrates the strength of using the mixed methods approach and therefore it will make methodological contribution to second language research.

## 1.6 Definitions of Key Terms and Abbreviations

The following terms and acronyms are defined below in order to provide clarity concerning these terms within the study.

- **Word family** – a word family is composed of a headword, its inflected forms, and its closely related derived forms (Nation, 2001).
- **Lexical diversity** – refers to a variety of different words that appear in a text.
- **Lexical frequency** – refers to how frequently a word occurs in print (Graesser, McNamara, Louwerse, & Cai, 2004).
- **Word imaginability** – refers to how easy it is to trigger a mental image of a word (Graesser et al., 2004).
- **Word meaningfulness** – refers to how strong a word is associated with other words (Graesser et al., 2004).

- **Word familiarity** – refers to how familiar a word is to an adult (Graesser et al., 2004).
- **Word concreteness** – refers to how concrete (non-abstract) a word is (Graesser et al., 2004).
- **Latent Semantic Analysis (LSA) sentence adjacent** – refers to how semantically similar each sentence is to the next sentence (Graesser et al., 2004).
- **Word hypernymy** – refers to “the number of levels in a conceptual taxonomic hierarchy that is above a word” (Graesser et al., 2004, p. 198).
- **Word polysemy** – refers to the potential meanings a word has (Graesser et al., 2004).
- **L1** – an acronym for first language
- **L2** – an acronym for second language
- **NS** – an acronym for native speaker
- **NNS** – an acronym for non-native speaker
- **ESL** – an acronym for English as a second language
- **EFL** – an acronym for English as a foreign language
- **VK** – an acronym for vocabulary knowledge
- **VLS** – an acronym for vocabulary learning strategy
- **LLS** – an acronym for language learning strategy
- **CET-4** – an acronym for College English Test Band 4
- **CET-6** – an acronym for College English Test Band 6
- **WAT** – an acronym for Word Associates Test
- **VST** – an acronym for Vocabulary Size Test

## 1.7 Organization of the Thesis

This thesis is organized into 6 chapters. In addition to this introductory chapter, Chapter 2 first reviews theories and empirical research related to breadth and depth of VK. It then reviews research on both individual VLSs and mixed use of VLSs.

Chapter 3 first describes the methods adopted in the quantitative study consisting of: a) research participants; b) instruments; c) data collection procedures; and d) data analysis. It then presents the quantitative results and concludes with a brief summary.

Chapter 4 presents the methods employed in the qualitative study including the description of: a) research participants; b) data collection procedures; and c) data analysis. It reports the qualitative results in relation to four main themes of a) significance of learning English vocabulary knowledge; b) difficulties of learning English vocabulary knowledge; c) current vocabulary learning strategies; and d) perspectives about vocabulary knowledge of the participants.

Chapter 5 synthesizes and discusses both quantitative and qualitative results in the order of a) the pattern of VK among Chinese students; b) the pattern of reported VLSs among Chinese students; c) the relationship between reported VLSs and VK; d) the relationship between VK and individual variables (gender, discipline, and proficiency levels); e) the relationship between reported VLSs and individual variables (gender, discipline, and proficiency levels); and f) the contribution of overall variables (gender, discipline, proficiency levels and VLSs) to VK.

Chapter 6 concludes this research by summarizing results, explicating research limitations, and providing implications for English vocabulary learning and teaching in Chinese EFL context and for future research on VLSs.



## **Chapter 2 Literature Review**

This chapter provides the theoretical background related to vocabulary knowledge tests (VKT) and vocabulary learning strategies (VLSs). This review will provide the required background for the selection of VKT as well as the development of the VLS questionnaire to be adopted in this study. This chapter sets out to review the literature pertaining to VKTs and VLSs and the relationship between them. It will therefore examine the studies on breadth and depth of vocabulary knowledge and evaluate the associated tests and scales which have been developed and applied to measuring these two constructs. It will then focus on vocabulary learning strategies through analyzing and comparing some key empirical studies available in the literature both on individual VLSs and mixed use of VLSs.

### **2.1 Studies on English as a Second Language (ESL) Vocabulary Knowledge**

Researchers in the field of ESL vocabulary teaching and learning have generally acknowledged that the concept of vocabulary knowledge is consisted of two equally important aspects: breadth and depth (Anderson & Freebody, 1981; Qian, 1999, 2002; Read, 1988, 1993, 1998; Schmitt, 2010; Wesche & Paribakht, 1996). The review of vocabulary knowledge in this section is accordingly divided into two sections. The first section reviews the breadth of vocabulary knowledge followed by a review of the depth of vocabulary knowledge in the second section. The main points to be presented in this section (Section 2.1) will be summarized and presented in Section 2.1.3.

#### **2.1.1 Breadth of Vocabulary Knowledge and Assessment**

Breadth of vocabulary knowledge (or vocabulary size) refers to the number of words known by learners, or at least some significant aspects of the word meaning known by learners (Anderson & Freebody, 1981; Nation, 2001). The surge of research on vocabulary size in L2 began in the 1990s (Goulden, Nation, & Read, 1990; Hirsh & Nation, 1992; Laufer & Nation, 1999; Nation, 1993, 2001, 2006). In order to identify long-term vocabulary learning goals for L2 learners, some researchers explored the vocabulary size of native speakers. Their findings suggested that there were approximately 54,000 word families in English (Nation & Waring, 1997), of which well-educated native speakers

generally mastered about 17,000 (D'Anna, Zechmeister, & Hall, 1991; Goulden et al., 1990). A word family is defined as a base word, along with its inflected and derived forms that can be understood by a learner with little or no extra effort (Bauer & Nation, 1993). Schmitt (2010) found similar results, indicating that the vocabulary size of educated native speakers of English ranged from 16,000 to 20,000 word families. Other studies have focused on the vocabulary size required for different language modalities, such as reading and listening (Hu & Nation, 2000; Nation, 2006). Nation (2006) reported that as many as 2,000 and 3,000 word families were required for non-native speakers to comprehend spontaneous unscripted spoken language with a target of 95% coverage of the exchanged vocabulary. If the target coverage of the exchanged words increases to 98%, then 6,000 and 7,000 word families would be required. For reading comprehension, Nation (2006) estimated that 8,000 to 9,000 word families were required for non-native speakers to read authentic texts without trouble and with a 98% coverage.

In the meanwhile, researchers have tried to figure out the vocabulary size of L2 learners. This has resulted in the development of a number of vocabulary tests (Meara & Jones, 1988; Meara & Milton, 2003; Nation, 1983, 1990; Nation & Beglar, 2007), among which the three most influential ones are, the Vocabulary Levels Test (VLT) (Nation, 1983, 1990), Eurocentres Vocabulary Test (EVT) (Meara & Jones, 1988), and Vocabulary Size Test (VST) (Nation & Beglar, 2007). These three vocabulary tests provide ESL/EFL researchers and practitioners reliable and applicable measurements of learners' breadth of vocabulary knowledge. They are all easy to administer and mark and can test a large number of words in a relatively short time. In the following, these three tests will be presented and discussed in tandem.

The Vocabulary Levels Test (VLT) was developed by Paul Nation in the early 1980s and this test was designed primarily for deciding on which word frequency level learners should be given help with vocabulary learning. The VLT was available in two publications (Nation, 1983, 1990). The words on VLT were selected from the Thorndike and Lorge (1944) with reference to frequency data from the General Service List (GSL) (West, 1953) and the Computational Analysis of Present-Day American English (Kucera & Francis, 1967) using a stratified sampling procedure. The original test consists of five parts, representing five levels of word frequency in English — the 2,000-word level, 3,000-word level, 5,000-word level, the university word level, and 10,000-word level. The 2,000 and 3,000 word levels are composed of high frequency words in English, the 5,000-word level

serves as an interim or intermediate level between the high frequency level and low frequency level, and the 10,000 word level includes low frequency words. The testing words on the university word level were sampled from the University Word List (Xue & Nation, 1984) containing 836 words which are frequently appearing in university textbooks. Words in each level of the test represent all the words at that level. In a follow-up study, Schmitt, Schmitt, and Clapham (2001) replaced the university word level with the academic word level premised on the recently released, the Academic Word List (AWL) (Coxhead, 2000), a word list perceived as achieving wider coverage of academic texts than that of the UWL. The VLT adopts a form-recognition matching format and each level of word frequency embraces six sections. Each section is composed of six target words and three definitions (meanings). An example of a sample cluster of the test from 2,000 word level is provided below.

1. Original
2. Private \_\_\_\_\_ complete
3. Royal \_\_\_\_\_ first
4. Slow \_\_\_\_\_ not public
5. Sorry
6. Total

To complete the above item, the test taker will need to choose the correct word on the left side which matches the meaning of the words on the right. The test taker could get one mark for each correct matching of a word and its definition and the maximum possible score for the entire test is 90, with the maximum score for each word level being 18. If a test taker gets a score of 12 out of 18 on a word frequency level, it means that approximately one-third of the words at that level are unfamiliar to the test taker. This one third portion indicates that around 200 to 300 words in minimum are required to be mastered at that level. In addition, although some studies (e.g., Qian, 1999) have combined the frequency levels to produce a total size figure, Nation (1990) suggested that the score on each word frequency level was more useful than the total score of the test. The VLT has received considerable attention in L2 vocabulary research since it was released. This instrument has been experimented in studies by Qian (1999, 2002), Laufer (2001), and Ishii and Schmitt (2009). Qian (1999) obtained a reliability of 0.92 for VLT.

The Eurocentres Vocabulary Tests (EVT) was published by Meara and Jones (1988) for the purpose of grading students and placing them in a series of short intensive courses. This test was designed for non-native speakers of English studying in the Eurocentres schools in the United Kingdom. It contains high frequency English words sampled from the first 10,000 words of Thorndike and Lorge (1944) frequency count, and some invented words, and is divided up into a number of blocks. The first such block incorporates a representative sample of items from the first 1000 most frequent words in English; the second one embraces that from the 2000 most frequent words in English; and so on up to block 10. An example of an item in the EVT test is shown as follows, incorporating both real words and invented words.

Block A:

advisor ghastly contord implore morlorn pitiful profess stourge  
moisten discard disdain gleanse weekend boyralty partine indoores

Unlike the traditional pencil-and-paper test, the EVT is a computer-based test. The test takers are invited to indicate “Yes” or “No” for each tested word to show whether they know the tested word or not. The scoring system calculates the final result by taking into account two types of response made by the testee: “the Hit Rate (i.e. the proportion of real words that the testee thinks they know) and the False Alarm Rate (i.e. the proportion of invented words that the testee thinks they know)” (Meara, 1990, p. 3). Accordingly, an estimation of the true Hit Rate is calculated by adjusting the actual Hit Rate in the light of the False Alarm Rate (Meara, 1990; Meara & Buxton, 1987; Meara & Jones, 1988). It only takes ten minutes to complete this test.

In a recent study, Nation and Beglar (2007) developed a new instrument called the Vocabulary Size Test (VST) to comprehensively measure a learner’s overall vocabulary size from the first 1000 to the fourteenth 1000 word families of English. The target words of this test were chosen from the family-based British National Corpus (BNC) and the first 12000 word lists were produced based on the 10 million token of the spoken section of the BNC instead of the whole number of tokens in the corpus. This decision was made because, as the authors contended, the spoken ordering might better represent the order in which English learners may actually learn the words. In this test, each 1000 word frequency level contains 10 items. It is a multiple choice test and the test words are contextualized within

simple and typical situations in which the target word appears. Below is an example of one item from 1,000-word level in the VST.

They **saw** it.

- a. cut
- b. waited for
- c. looked at
- d. started

To complete this test, test takers are invited to choose the correct meaning from the four choices for the target word. Each item has one point, and the total score of this test is 140. Beglar (2010) has reported a high reliability (0.96-0.98) of VST.

Although VLT, EVT and VST serve to measuring the breadth of vocabulary knowledge, these three types of vocabulary tests differ from each other in terms of their test format, sampling sources, and the coverage of word frequency levels. In contrast to the paper and pencil format of VLT and VST, the EVT test is computer-based, which afford testees to quickly choose between “Yes” and “No” for each target item. Scoring can therefore be calculated automatically by computer and a large number of words can be tested in a short time of approximately ten minutes. Despite such advantages, there are some problems regarding this test. For example, the L1 background may have an effect on the scores of the test due to the cognate effects of the invented words. In other words, some invented words are much easier to be accepted by a group of testees because the forms of the invented words are similar to their L1 (Meara, 1990). In addition, EVT requires testees to recognize word form rather than to have a clear focus on word meaning. Therefore, this test may probably overestimate learners’ true vocabulary knowledge. L2 teachers and researchers should be cautious in applying this test.

Concerning VLT, this test seems less capable of measuring a learner’s overall vocabulary size because, as Nation (1990) illustrated, the primary purpose for designing this test has been for diagnosing on which word frequency level learners should be given help with vocabulary learning. In comparison with VLT and EVT, VST has several advantages. Firstly, the target words were chosen from the family-based BNC instead of the outdated lemma-based list advanced by Thorndike and Lorge (1944). Nation and Beglar (2007) suggested that the word family count is preferable for measuring receptive

vocabulary size, because learners beyond a minimal proficiency level have some knowledge of word affixes and word building techniques. Secondly, unlike VLT and EVT which provide a profile of vocabulary knowledge on five or ten word frequency levels, VST covers fourteen frequency levels of English words, which is more comprehensive and fine-tuned to test learners' overall vocabulary size. Finally, the meanings of some candidate answers to a testing item of VST were designed to be intimately interconnected with each other, so that the choice of the right answer could better measure testees' vocabulary knowledge than the VLT and EVT tests.

In recognition of the merits and demerits of each vocabulary test, VST was selected in the present study as the test to assess the participants' breadth of vocabulary knowledge. However, as indicated above, this test is based on the BNC lists, which are family-based instead of lemma-based, an issue which is still controversial among researchers. For instance, Schmitt and Meara (1997) and Schmitt (1998) found that language learners could not have a good mastery of the suffixes of the verbs, especially derivational suffixes, thus they may have only a partial knowledge of a word family. In this case, the family-based VST might, to some extent, overestimate the vocabulary size of learners. Therefore, the results derived from this test in Chapter 5 need to be interpreted with caution.

## **2.1.2 Depth of Vocabulary Knowledge and Assessment**

### **2.1.2.1 Conceptualizing Depth of Vocabulary Knowledge**

Depth of vocabulary knowledge is generally delimited from vocabulary size by querying how well the learner knows a word (Anderson & Freebody, 1981; Read, 1993), and it has been generally investigated in three lines of scholarly inquiry or approach: the dimensional approach (Nation, 2001; Richards, 1976), the developmental approach (Wesche & Paribakht, 1996) and the association approach (Meara, 1982).

The dimensional approach towards the investigation of the depth of vocabulary knowledge involves analyzing various aspects of vocabulary knowledge and acquisition of a word is accordingly defined as knowing all these aspects of that word. This approach draws on the L1 vocabulary assessment in which Cronbach (1942) outlined five behaviors for fully understanding a word. These five behaviors are generalization (defining the word), application (recognizing the appropriate use of the word), breadth of meaning (knowing different meanings of the word), precision of meaning (applying the word to all different situations), and availability (using the word productively). The main focus of Cronbach's

(1942) categorization is meaning and use, and aspects such as spelling and pronunciation are not included.

Later, in the field of L2 vocabulary research, Richards (1976) proposed eight aspects concerning vocabulary acquisition and vocabulary knowledge, which involves knowing various aspects of vocabulary knowledge such as words' frequency (the probability of encountering that words in speech or print), word collocation, word association (synonym, antonym, subordinate, coordinate, and superordinate classifications), limitations on use (register characteristics), syntactic pattern, derivation (the underlying form of a word and its derivations), semantic features (the semantic value) and polysemy. This framework is relatively comprehensive since it has tapped into more dimensions of vocabulary knowledge such as word frequency, word semantic features, different meanings associated with a word and its register characteristics. However, it seems that the knowledge of spelling and pronunciation of a word is still missing in Richards's relatively more comprehensive framework of vocabulary knowledge dimensions.

Following this strand of research, Nation (1990, 2001) further specified the dimensions of vocabulary knowledge that learners must know to represent their receptive and productive aspects of vocabulary knowledge. He proposed three main categories: form, meaning, and use. The form of a word embraces its spoken form, written form, and word parts. The meaning of a word incorporates form and meaning, concept and referents, and associations. Finally, use aspect encompasses grammatical function, collocations, and constraints on use (register, frequency, etc.) (Nation, 2001). Nation's framework roughly covers three aspects of the depth of vocabulary knowledge – form, semantics and stylistics in both receptive and productive domains.

Unlike Cronbach (1942) and Richards (1976), Nation (2001) adds “spoken form” and “written form” to his list and views word knowledge from the perspective of receptive and productive mastery of each different component type, which is more systematic and comprehensive. In addition, various kinds of associations are clearly illustrated in Nation's framework, such as “meronymy (the whole-part relationship, e.g., house-kitchen), troponymy (something is done in a particular manner, e.g., march, strut, traipse and amble) and entailment (engaging in one action involves engaging in the other, e.g., snore entails sleep)” (Nation, 2001, p. 54). This point was missing in Richards's (1976) framework. Nation's (2001) framework has been most widely used by researchers and it has been regarded as the most comprehensive one to date (Schmitt, 2010). However, several

dimensions of vocabulary knowledge in Richards's (1976) framework have not yet been touched upon in Nation's (2001) framework, such as word frequency, polysemy, and semantic features which are also a part of depth of vocabulary knowledge. Each framework is incomplete but complements each other in terms of some aspects of depth of vocabulary knowledge covered. Therefore, it seems better to integrate the frameworks of Richards (1976) and Nation (2001) in conceptualizing the depth of vocabulary knowledge.

The dimensional approach requires learners to master several components of vocabulary knowledge; however, these components can hardly be mastered by learners in their first acquaintance with a word. That is to say, vocabulary learning is an incremental process. As such, the developmental approach towards vocabulary knowledge and vocabulary acquisition aims to identify stages of vocabulary learning process. Various approaches have been proposed to distinguish one stage from another. In 1965, Dale, a researcher of first language, proposed a four-stage model for checking vocabulary knowledge as presented below.

Stage 1: I never saw it before.

Stage 2: I've heard of it, but I don't know what it means.

Stage 3: I recognize it in context – it has something to do with...

Stage 4: I know it.

Dale's (1965) stage model represented the general progressive process of vocabulary acquisition, and provided a rough model for subsequent researchers who developed models of vocabulary acquisition in line with this approach. Nevertheless, Dale's identification of the stages for knowing a word is incomprehensive. One main problem is that all the five stages in this model exclusively identify the receptive vocabulary knowledge at the expense of ignoring the other equally important aspect of the depth of VK – productive vocabulary knowledge.

In the field of L2 research, Paribakht and Wesche (1993) developed the Vocabulary Knowledge Scale (VKS) which was a five-category elicitation scale representing categories or degrees of vocabulary knowledge, as presented below.

1. I don't remember having seen this word before.
2. I have seen this word before, but I don't know what it means.
3. I have seen this word before, and I think it means\_\_\_\_\_.

4. I know this word. It means \_\_\_\_\_ (synonym or translation).

5. I can use this word in a sentence: \_\_\_\_\_. (Write a sentence)

Compared to Dale's stage model, VKS is more comprehensive because the latter assesses the initial productive vocabulary knowledge by requiring testees to produce a sentence. In this respect, VKS assesses both receptive and initial productive vocabulary knowledge of L2 learners. Hence, VKS seems more suitable to demonstrate the process of vocabulary learning than Dale's model.

A revised model of VKS was proposed later by Schmitt and Zimmerman (2002) with a simpler and more transparent scale, which could overcome some limitations of VKS. Schmitt and Zimmerman's (2002) scale is illustrated as follows.

A. I don't know the word.

B. I have heard or seen the word before, but am not sure of the meaning.

C. I understand the word when I hear or see it in a sentence, but I don't know how to use it in my own speaking or writing.

D. I know this word and can use it in my own speaking and writing.

In comparison with VKS, Schmitt and Zimmerman's (2002) scale is characterized with clearer and more transparent description of each stage of vocabulary knowledge development and more clearly articulate starting and ending points. To be specific, "Having no knowledge of the word" is regarded as the starting point and "can use it in speaking or writing" is the ending point (Schmitt, 2010). Schmitt however indicated that it might be better to specify stage D as either speaking or writing instead of putting them together, because in some cases, learners only can use the word in speaking or in writing.

While the above three scales roughly describe the stages of L2 word development, Henriksen (1999) attempted to depict the lexical competence of language learners and the features of the word knowledge development in a more specific and precise way. Henriksen (1999) proposed a three-dimension approach to describe the lexical competence of language learners. The three distinctive but related dimensions of lexical competence are: partial to precise knowledge (dimension 1), depth of knowledge (dimension 2), and receptive to productive use ability (dimension 3). These three dimensions also imply the continuum along which vocabulary knowledge develops. "Partial to precise knowledge" dimension means that learners' lexical development moves from vagueness to precision.

“Depth of knowledge” dimension indicates the complex nature of vocabulary knowledge and that a fuller understanding of a word requires the close examination of its various aspects of knowledge. The development of depth of vocabulary is associated with building networks between words, which is similar to the organization dimension developed by Meara (1996b). By “receptive to productive use ability” dimension, Henriksen (1999) suggested that only a limited number of words that we learn or acquire receptively would eventually be transferred into our productive knowledge. Henriksen’s (1999) three-dimension continuum model has avoided the issue of distinguishing different stages of vocabulary learning. Nevertheless, these three continuums seem to overlap with each other. For instance, “partial to precise knowledge” overlaps with “receptive to productive knowledge” and it seems difficult to distinguish them.

In general, researchers taking a developmental approach consider the process of word acquisition either as a continuum or as separate stages, usually from initial knowledge of the meaning of a word to the full mastery of it. This notwithstanding, no consensus has yet been reached among researchers as to how to define the levels or stages for word acquisition. In addition, these models view the development of lexical competence as unidirectional continuum or stages, however, as Meara (1996b) suggested they do not incorporate the issues of forgetting or attribution in the process of vocabulary acquisition. Thus Meara (1996b) argued that the process of vocabulary acquisition should be multidirectional rather than unidirectional.

The third approach was proposed by Meara (1996a) who argued that the two basic dimensions of vocabulary knowledge are vocabulary size and organization instead of vocabulary size and depth. In his framework, vocabulary size is considered as the basic dimension of lexical competence and vocabulary organization involves all kinds of associations of the word with other words (e.g., paradigmatic, syntagmatic, situational, emotional etc.). This approach is based on the psycholinguistic concept that words are stored in brain in a complex lexical network rather than as isolated units (Meara, 1996a, 1996b; Nation, 2001). There is abundant literature on word association both in L1 (e.g., Clark, 1970; Deese, 1966; Postman & Keppel, 1970), and L2 studies (e.g., Meara, 1980, 1982). One of the main findings is that native speakers have a stable pattern of word association, whereas L2 learners have a diverse and unstable network. Another finding is that, the pattern of word association produced by L2 learners develops towards native-like pattern, as their proficiency improves (Read, 1993). This approach has not yet been widely

accepted by vocabulary knowledge researchers since the concept of association is too broad a term for developing practical assessment tools. Similarly, Schmitt (2010) argued that there is still no consensus on how to best use the association approach, for instance, how to run the methodology and what kind of associations can represent mental lexicon.

Overall, the above three approaches describe depth of vocabulary knowledge from different angles, however, to some extent they are complementary and overlap with each other. The first approach, Nation (1990, 2001) and Richards (1976), conceptualized depth of vocabulary knowledge as a wide range of dimensions. One problem with this approach is that it is difficult for researchers to include all dimensions of vocabulary knowledge in one framework. The second approach, Dale (1965) and Wesche and Paribakht (1996) considered depth of vocabulary knowledge as developmental process rather than a dichotomous acquired/not acquired matter. This approach seems to face more challenges. Firstly, no consensus has been reached among researchers as to how to clearly define each stage of vocabulary learning. Secondly, it is still questionable whether the development of the stages or the continuum of word acquisition should be regarded as unidirectional or multidirectional. The third approach argued that vocabulary depth involves all kinds of associations of the word with other words (Meara, 1996a). This approach has not yet been able to create consensus among researchers either, mainly because the concept of association is too broad a term for developing practical assessment tools.

By considering the affordances and constraints of each of the three approaches, the dimensional approach seems more satisfactory in conceptualizing depth of vocabulary knowledge. Nevertheless, one demerit of the dimensional approach is that it is difficult for researchers to include all dimensions of VK in one framework. As I have illustrated, the frameworks proposed by Nation (2001) and Richards (1976) under this approach are complementary to each other. Therefore, both frameworks were deployed so as to better represent depth of vocabulary knowledge. In the next section, I will examine some empirical studies on the assessment of the depth of vocabulary knowledge based on these three approaches.

#### **2.1.2.2 Assessment of Depth of Vocabulary Knowledge**

There are a number of assessment tools that have been designed based on the three theoretical approaches to measure L2 learners' depth of vocabulary knowledge. In the developmental approach, Vocabulary Knowledge Scale (VKS) developed by Paribakht and Wesche (1993) at Ottawa University, Canada, aimed to assess vocabulary learning in the

context of academic reading activities. VKS presented a five-category elicitation scale which was believed by the test developers to present categories or degrees of vocabulary knowledge (see Section 2.1.2.1 for an explanation of VKS scale). Moreover, Paribakht and Wesche (1993) provided a scoring scale which helped researchers or teachers to score the performance of test-takers in a more reliable way. The reliability and the validity of the test have been supported by some evidence provided by the authors through their test-retest research, concurrent validity analysis, and an analysis of the correlation between the way the students rated themselves and the way their responses were scored (Wesche & Paribakht, 1996). Read (2000, p. 135) also reported that the VKS “proved to be a workable measure and seems to be sensitive to increase in vocabulary knowledge that result from reading activities”. In fact, some researchers have used this instrument in their research on assessing quality of word knowledge (e.g., Joe, 1998; Paribakht & Wesche, 1997; Vidal, 2003).

However, the VKS has its own limitations. Firstly, the intended purpose of VKS is to demonstrate the initial stages of vocabulary learning, and thus it is inappropriate for assessing the vocabulary knowledge in its general sense. Secondly, the intervals of the stages do not seem to be consistent and equivalent to allow applying the parametric statistical analysis to the scale (Schmitt, 2010). In addition, stage 5 may not indicate that the meaning of a word is known because the students often write sentences which do not indicate the clear meaning of the word (Meara, 1996b; Read, 2000; Schmitt, 2010).

Following the association approach of vocabulary knowledge, Meara and Wolter (2004) developed a software program called V\_Links to examine the associations between high-frequency words which is composed of ten items, each containing 10 words. Language learners are asked to indicate, in one minute, whether there is a connection between these words and how strong the relation is. The test words were chosen from the content words of the first 1000 words in English. In a subsequent study, Meara further revised V\_Links and later developed a new version called V\_Quint, which contains 40 items and each consists of a set of five words. Learners need to select one association within 15 seconds. The two tests are accessible from: <http://www.lognostics.co.uk/tools/index.htm>.

There are several problems inherent in the tests of V\_Links and V\_Quint. Firstly, from my experience of carrying out the test, it involves too many kinds of associations for testees to find instead of indicating the limited and specific kinds of associations, such as limiting the types of associations to polysemy, synonymy or collocation etc. Therefore, it

may raise problems for test takers to find out the associations in a given time. Secondly, there is no feedback as to indicate whether the selected association is right or wrong. Perhaps it is for these reasons that Meara and his colleagues admitted that much more research would be required to develop a better assessment tool that can represent the association approach.

In accord with the dimensional approach, a number of researchers (e.g., Ishii & Schmitt, 2009; Read, 1993; Schmitt, 1998) have designed depth of vocabulary knowledge tests to measure various dimensions of vocabulary knowledge. Read (1993, 1998) developed the Word Associates Test (WAT) with the purpose of measuring vocabulary learning for students of English for academic purposes at university level. This test has been revised several times by Read himself (1993, 1998). The most recent version (Read, 1998) is designed to measure two aspects of depth of vocabulary knowledge: (1) word meaning, particularly polysemy and synonym, and (2) word collocation. It includes eight options within two boxes for each target word, four options in each box separately. All the target words are adjectives selected mainly from high-frequency academic vocabulary. The format of the test is shown below.

**sudden**

beautiful   quick   surprising   thirsty	change   doctor   noise   school
--	----------------------------------

To complete the above item, test takers need to select four words having a relationship with the target word “sudden” from the two boxes. The words in the left box may help explain the meaning of the target word, while the words in the right box can collocate with the target word. In this example, there are two correct answers in the left and two in the right box, but in other items of this test, there will be either one in the left and three in the right box or vice versa. This format has been used and modified by a considerable amount of researchers for assessing the depth of vocabulary knowledge in different contexts, such as advanced learners of French in Dutch universities (Greidanus & Nienhuis, 2001), Dutch elementary school students in the Netherlands (Schoonen & Verhallen, 2008), ESL university students in Canada (Qian, 1999, 2002). The reliability of WAT tests was reported to be 0.93 (Read, 1998), 0.91 (Qian, 1999), and 0.88 (Qian, 2002).

Although WAT has been widely adopted as a depth of vocabulary knowledge test, it has its own problems. One problem, put forward by Schmitt (2010), concerns the interpretation of a score of two associates and two distracters, as the two correct associates

may probably be chosen by the testees' random guessing. In addition, the target words in WAT drew entirely on adjectives, rather than other parts of speech such as nouns and adverbs. Moreover, WAT only explores three types of associations, while more associations could be involved in the test. Despite the above problems, WAT is still one of the most popular and widely adopted tests for assessing depth of vocabulary knowledge. The present study adopted WAT as an instrument for testing the participants' depth of vocabulary knowledge.

In addition, researchers such as Ishii and Schmitt (2009), Schmitt (1998), and Schmitt and Meara (1997) have designed a series of tests to explore and assess different dimensions of vocabulary knowledge. Schmitt and Meara (1997) mainly examined two types of word knowledge – word associations and grammatical suffix knowledge from both receptive and productive perspectives. A number of 20 verbs were chosen from Brown frequency list (Francis & Kucera, 1982) and a list of the most common suffixes was selected from the higher Bauer and Nation (1993) difficulty levels. A standard answer sheet was constructed by Schmitt and Meara (1997) based on native speakers' responses on the test. The L2 participants' responses were then checked against those of native speakers. If a word association provided by the test taker on the productive task was not on this standard answer sheet, but seemed native like, the appropriateness has to be judged by two native speakers. Although Schmitt and Meara (1997) made a good attempt to examine vocabulary depth, this study only includes limited dimensions of vocabulary knowledge. Following Schmitt and Meara's (1997) pioneering study in this field, Schmitt (1998) conducted another similar study investigating more dimensions of vocabulary knowledge, such as spelling, associations, grammatical information and meaning. The majority of target words were sampled from the University Word List (Xue & Nation, 1984) and a few lower-frequency words were chosen from the Brown word list (Francis & Kucera, 1982). The test of written form (spelling) deployed a 4-point rating system. The scoring system for the association measurement was based on a standard answer sheet drew from native speakers' responses on this test. Concerning the tests of word class, meanings and derivational forms, the norming lists were primarily from dictionaries. Despite this test battery tapped into more dimensions of vocabulary knowledge, it is obvious that the scoring system of this test battery is obscure and not easy to manage. In a follow-up study, Ishii and Schmitt (2009) developed a new integrated test battery of vocabulary size and depth to assess both the breadth and depth of English vocabulary of Japanese college

students. For the depth of vocabulary test, they developed three subtests, namely, TMM (Test of Multiple-Meanings), TDF (Test of Derivative Forms) and TLC (Test of Lexical Choice between Near-Synonyms). All the target words were chosen from the most frequent 2000 words of the BNC list and there were 15 to 30 target words in each test. Compared with the other existing instruments for assessing different types of word knowledge, Ishii and Schmitt's integrated test covers several types of word knowledge, including different meanings, derivative forms and the collocational or register properties of near-synonyms. Although these researchers (Ishii & Schmitt, 2009; Schmitt, 1998; Schmitt & Meara, 1997) have developed different instruments for tapping into various dimensions of vocabulary knowledge, it appears that all these tests are subject to some common problems: firstly, these vocabulary tests are relatively time-consuming and are only able to assess a very limited number of words; secondly, some scoring methods of these tests seem rather obscure, as such it is not quite practical for researchers and language teachers to implement these tests; thirdly, reliability values were rarely reported for these testing instruments; and finally, despite researchers tried to include more dimensions of vocabulary knowledge in their studies, these dimensions are only a limited part of vocabulary knowledge.

As an attempt to overcome these limitations, Crossley and his colleagues have designed and used a computational text analysis tool, called Coh-Metrix. Coh-Metrix assesses the vocabulary knowledge of text writers along with a variety of other linguistic and discoursal features of the written texts. The lexical indices of Coh-Metrix include word frequency, lexical diversity, hypernymy, polysemy, semantic co-referentiality, word meaningfulness word concreteness, word familiarity, and word imaginability. Based on these lexical indices, Crossley, Salsbury, McNamara, and Jarvis (2010) presented a model of lexical proficiency. They found that lexical diversity, word hypernymy values and word frequency explained 44% of the variance of the human evaluations of lexical proficiency in the examined writing samples. In a follow-up study, Crossley et al. (2011) concentrated on the exploration of the possibility of categorizing L2 learners' general language proficiency using lexical indices. Accordingly, 100 writing samples taken from L2 learners were analyzed using lexical indices of Coh-Metrix. The L2 writing samples were categorized into beginning, intermediate, and advanced levels based on the writers' TOEFL paper-based test (PBT) scores or TOEFL Internet-based test (iBT) scores or combined ACT Compass ESL reading and grammar tests scores. ACT Compass ESL tests aim to assess

non-native English speakers' abilities in four areas of Standard American English – Listening, Reading, Grammar/Usage, and Essay so as to place students in appropriate courses of a college. The strongest predictors of an individual's proficiency level were word imaginability, word frequency, lexical diversity, and word familiarity. In total, the indices correctly classified 70% of the texts based on proficiency level in both training and a test set of data. Moreover, Crossley and his colleagues (e.g., Crossley, Salsbury, & McNamara, 2009, 2010a) deployed Coh-Metrix to investigate the development of the value of each lexical index as L2 learners' proficiency increase. These studies showed that Coh-Metrix served as a good instrument for investigating L2 learners' lexical competence, assessing L2 lexical development, and grouping L2 learners in line with their proficiency level. Despite the capability of Coh-Metrix to assess vocabulary knowledge in an effective way, it can only assess learner's productive rather than receptive vocabulary knowledge because language learners need to first produce a text. The present study therefore used Coh-Metrix to assess the participants' depth of vocabulary knowledge (productive vocabulary knowledge).

### **2.1.3 Summary**

The above section reviewed tests and/or approaches to both breadth and depth of VK in the relevant research literature. Based on the review, this research adopted VST as an instrument to assess participants' breadth of VK. Concerning the depth of vocabulary knowledge, the dimensional approach was conceived to better conceptualize and measure the depth of VK than the other two approaches – developmental and association approaches. However, one problem with the dimensional approach is that it is difficult to include all dimensions of VK in one framework. Therefore, the present study integrated both Nation's (2001) and Richards's (1976) frameworks so as to better represent the depth of VK. In addition, given the affordances and constraints of the reviewed VK tests for depth of vocabulary knowledge, two testing instruments – WAT and the computational tool Coh-Metrix were selected to operationalize the integrated frameworks. The interconnection between these two testing instruments and the integrated frameworks has been illustrated in Section 1.4.

## **2.2 Studies on Vocabulary Learning Strategies (VLSs)**

Research on vocabulary learning strategies (VLSs) initially started with examining the effectiveness of individual VLSs such as using dictionary, inferencing, and mnemonic

strategies on vocabulary learning. From 1990s, some researchers began to investigate the mixed use of VLSs (use of VLSs in combination) by learners. Nevertheless, it still appears rather unclear regarding the way VLS is conceptualized. In the following sections, I will firstly examine how researchers have conceptualized VLSs as part of language learning strategies (LLSs), then I will review the studies on individual VLSs and the studies on the mixed use of VLSs.

### **2.2.1 Conceptualizing VLSs as Part of LLSs**

The research on vocabulary learning strategies is conducted within the context of language learning strategies. For this reason, a review of the LLSs conceptualization will provide insights into an understanding of VLSs. Accordingly, this section will firstly and briefly discuss how LLSs have been conceptualized in the literature. It will then review some prominent frameworks to categorize LLSs. Lastly, it will examine how the research on LLSs has informed the design of various VLS questionnaires.

Despite different ways of defining LLSs, there is a general consensus in the literature that the primary function of LLSs is to scaffold learners' learning of a target language. I have identified three areas of disagreement concerning the delineation of LLSs in the literature. These include whether LLSs should be defined as mental or behavioral, conscious or sub-conscious, general or specific activities.

The first area relates to whether LLSs should be defined as behaviors, mental activities or a combination of both. As will be shown in Section 2.2.3.1, studies (e.g., Sanaoui, 1992) that define LLSs as behaviors often ignore those mentally-oriented LLSs such as inferencing or monitoring; on the other hand, research (e.g., Dornyei, 2005; Macaro, 2006) that define LLSs as mental activities often neglect behavior-oriented LLSs such as using dictionaries or reading English novels. These two lines of inquiry about LLSs may be considered as a "separate" approaches (i.e. separating mental activities from language learning behaviors) in contrast to a "combined" approach. The combined approach (Gu, 2012; O'Malley & Chamot, 1990; Oxford, 1989, 2013) conceptualizes LLSs as both mental activities and behaviors. Learners' behaviors are directed by their mental activities, thus it is inappropriate to separate these two processes of learning (O'Malley & Chamot, 1990). In addition, the strength of this approach will be shown in its corresponding categorizations of LLSs and VLSs as discussed in subsequent sections.

The second area of disagreement concerns whether LLSs should be considered as conscious activities or as both conscious and subconscious activities. Some researchers (e.g., Cohen, 1998; Dornyei, 2005; Macaro, 2006; Oxford, 2013) argued that strategies must be conscious in order to be “strategic”. For instance, Cohen (1998, p. 4) described LLSs as “those processes which are consciously selected by learners and which may result in action taken to enhance the learning or use of a second or foreign language, through the storage, retention, recall, and application of information about that language”. However, this strand of research on LLSs has taken for granted that strategies are always under conscious control, and has ignored the strategies that are automated (Gu, 2012). Likewise, Bialystok (1990) suggested that learners did not need to be aware of their use of strategies and it might be inappropriate to use the concept of consciousness as a criterion to distinguish strategic from non-strategic behaviors.

The third area of disagreement is whether LLSs should be defined as general or both general and specific actions. Some researchers (e.g., Goh, 1998; Oxford, 2013) made a distinction between strategies and tactics, with the former being broad and general and the latter specific. However, this dichotomy has provided little knowledge for understanding LLSs per se because the distinction between strategies and tactics largely corresponds to that of the main categories (e.g., a group of similar strategies such as inferencing strategies) and the specific individual strategies (e.g., inferring the meaning of a word from its affixes).

Gu (2012) attempted to resolve these three issues, and defined a learning strategy as “a dynamic process with problem-solving as its central aim”. This dynamic process involves both metacognitive and cognitive components, such as the plan and the actual implementation of a strategy, implying that a learning strategy involves both mental and behavioral processes. Gu (2012, p. 340) added that “competence or performance, general or specific, controlled or automatic, learning strategies are what the learner utilizes when confronted with a learning task”. The present study adopted this definition because on the one hand it is practical by focusing on what learners actually do in learning vocabulary (or more generally learning a language) rather than on debating over the ontological nature of learning strategies per se; on the other hand, it is arguably the most comprehensive one in the literature because it integrates affordances from different approaches to learning strategies (e.g., mental vs. behavioral; conscious vs. sub-conscious; general vs. specific).

In accord with various definitions, there have also been several frameworks proposed by researchers for categorizing LLSs (e.g., Cohen, 1998; O'Malley & Chamot, 1990;

Oxford, 1990). However, I will only focus on two such categorizations proposed by O'Malley and Chamot (1990) and Oxford (1990) respectively. This is because the development of the three prominent VLS questionnaires (see Section 2.2.3.1) has been derived from these two categorizations.

Oxford (1990) established two groups of LLSs – direct and indirect strategies – with the former referring to those language learning strategies that directly involve the target language and the latter referring to those strategies that do not directly involve the target language. The two categories of strategies suggested by Oxford (1990) are presented below.

Direct strategies:

- Cognitive strategies (helping learners in the understanding and production of new language, e.g., repeating and summarizing);
- Memory strategies (facilitating learners storing and retrieving new information, e.g., grouping and using imagery);
- Compensation strategies (overcoming gaps in knowledge, e.g., inferencing or avoiding)

Indirect strategies:

- Metacognitive strategies (coordinating the learning process, e.g., planning and monitoring learning);
- Affective strategies (controlling emotions and regulating attitudes and motivations, e.g., lowering anxiety);
- Social strategies (facilitating communication with others, e.g., asking questions for clarification).

On the other hand, O'Malley and Chamot (1990) have proposed a threefold classification of LLSs, as presented below:

- Cognitive strategies (operating directly on incoming information and manipulating it in ways that enhance learning, e.g., repetition, grouping, and inferencing)
- Metacognitive strategies (planning, monitoring or evaluating the success of a learning activity, e.g., planning and self-management)

- Social/affective strategies (interacting with another person or ideational control over affect, e.g., questioning for clarification and self-reinforcement)

Although both Oxford (1990) and O'Malley and Chamot (1990) include cognitive strategies in their categorizations, they defined cognitive strategies in different ways. The cognitive strategies in O'Malley and Chamot's (1990) categorization largely correspond to the "cognitive strategies", "memory strategies", and "compensation strategies" in Oxford's (1990) categorization. In this regard, these two frameworks are similar to each other in a sense. The LLS questionnaires developed by Oxford (1990) and O'Malley and Chamot (1990) have been used widely in language teaching and learning studies.

In line with O'Malley and Chamot's (1990) categorization, Gu and Johnson (1996) categorized VLSs as metacognitive and cognitive strategies, with the former embracing selective attention and self-initiation and the latter involving inferencing, using dictionary, note-taking, rehearsal, encoding, and activation strategies. In addition, Schmitt (1997) and Fan (2003) also constructed VLS questionnaires based on the categorizations proposed by O'Malley and Chamot (1990) and Oxford (1990). These three VLS questionnaires along with others will be discussed in detail in Section 2.2.3.1. Before that, in the next section I will discuss four individual vocabulary learning strategies with a focus on their constraints in understanding language learners' use of VLSs.

## **2.2.2 Studies on Individual Vocabulary Learning Strategies**

This section discusses and evaluates four outstanding individual strategies – rote rehearsal, mnemonic strategies, inferencing strategies and using dictionaries – that are frequently studied in research literature on VLSs. The purpose of this section is twofold. On the one hand, it aims to evaluate the insights provided by each individual strategy under investigation as well as their constraints. On the other hand, the evaluation of these individual strategies serves to justify the preference of mixed use of VLSs in this research.

### **2.2.2.1 Rote Rehearsal**

Rote rehearsal refers to remembering words by using repetition strategies, such as oral repetition, written repetition and visual/silent repetition. Previous studies on rote rehearsal have indicated that using rote rehearsal could assist to learn a large number of words within a short period of time (see Gu, 2003b for a review). Some studies found that choosing the right way of rehearsing matters. Oral repetition has been reported to be more

desirable than silent repetition in remembering the meaning of a new word (e.g., Gu, 2003b; Gu & Johnson, 1996).

There are recent studies (e.g., Dahlen & Caldwell-Harris, 2013; Hummel, 2010; Lotto & De Groot, 1998; Prince, 1996; Rodriguez & Sadoski, 2000; Sagarra & Alba, 2006; Van Hell & Mahn, 1997) comparing the rote rehearsal with other VLSs. Rote rehearsal was found more helpful for vocabulary learning than semantic mapping and using translation. For instance, Hummel (2010) reported that intermediate proficiency level L2 learners performed significantly better in rote-copying condition (written repetition) compared to the two translation conditions for the short-term L2 vocabulary retention. In addition, Sagarra and Alba (2006) found that rote rehearsal was more effective than using semantic mapping for beginning L2 learners of Spanish in both short-term and long-term retention.

Rote rehearsal strategies seem to be especially more useful for low proficiency students. This viewpoint was evidenced in Prince (1996) who found that less proficient students could recall more words when they used rote rehearsal rather than learning from context. By contrast, high proficiency students seem to perform better when learning vocabulary from context than using rote rehearsal strategies. For instance, Hermann (2003) found that for ESL first-year university learners, learning vocabulary in context (e.g. learning new words in reading) was preferable than learning in isolation (e.g. learning by rote rehearsal).

#### **2.2.2.2 Mnemonic Strategies**

Mnemonic techniques refer to remembering words by using mediators such as a word, phrase or visual image. These techniques can be classified into verbal (e.g., rhyme), visual (e.g., image word form) and mixed mnemonics (e.g., the keyword strategy) (Cohen, 1987). One particular technique, keyword strategy, has been intensively studied by many researchers (e.g., Atkinson, 1975). The keyword strategy combines the phonological forms and meanings of L1 and L2 words in which the learner first finds an L1 word which sounds like the target L2 word, then creates an image combining the two concepts. Keyword strategy has been found more effective for both beginning and intermediate proficiency level L2 learners in word learning (e.g., Avila & Sadoski, 1996; Lawson & Hogben, 1998; Moore & Surber, 1992; Sagarra & Alba, 2006; Wang & Thomas, 1992). In addition, some studies (e.g., Avila & Sadoski, 1996; Sagarra & Alba, 2006) compared keyword strategy with other strategies and reported that keyword strategy was more preferable than mechanical rote learning in improving the retention of foreign words or semantic mapping. For instance, Sagarra and Alba (2006) reported that beginning learners

of Spanish who used keyword strategy performed significantly better than those who used rote rehearsal in terms of both immediate and long-term word retention. Nevertheless, contradictory results have also been documented in the literature. For instance, Van Hell and Mahn (1997) found that for beginning L2 learners, rote rehearsal was more effective than using keyword method in terms of both short-term and long-term word retention. The inconsistent results of these two studies may be explained by their different research designs. Firstly, the target words adopted in Sagarra and Alba (2006) were those words with high level of concreteness and imaginability whereas Van Hell and Mahn (1997) used half concrete words and half abstract words. As Ellis and Beaton (1993) argued, it is easier for learners to find good keywords for concrete words than abstract words. Secondly, the former study used learner-generated keywords which asked learners to find keywords by themselves, while the latter study adopted experimenter-generated keywords which are provided by researchers. This finding was consistent with Joe (1995) who found that learner-generated keyword was more helpful than experimenter-generated keyword for L2 learners' word retention. In sum, the usefulness of keyword strategy largely depend on the concreteness of keywords and the way of getting keywords (learner-generated vs. experimenter-generated keyword).

In addition to the above two factors, there are some other factors which could also reduce the effectiveness of keyword strategy. These factors include phonological and orthographical similarity between L1 and L2 (e.g., Wei, 2015) and teachers' ignorance of the keyword strategy (e.g., Hulstijn, 1997). For instance, a recent study conducted by Wei (2015) found that the keyword method was not as effective as using word part technique and self-strategy learning on the word translation test and word form recognition test. Wei (2015) contended that the results may be partly due to the dissimilarity between the Chinese language and English language which is difficult to find a good keyword.

In recognition of these limitations of using keyword strategy, several studies (e.g., Brown & Perry, 1991; Rodriguez & Sadoski, 2000) examined the adoption of the combined keyword with other strategies. The general finding was that the combined use of keyword strategy with other strategies such as using context or semantic strategies could produce better word recall rate than the use of keyword strategy per se.

In short, keyword strategy seems more effective for both beginning and intermediate proficiency L2 learners in vocabulary learning. However, the usefulness of keyword strategy largely depends on some factors, such as the concreteness of keywords and the

linguistic similarity between L1 and L2. In addition, the mixed use of keyword strategy with other strategies seem more desirable. As Hulstijn (1997) suggested, the keyword strategy should not be taken as a substitute for other strategies, but as a helpful supplementary strategy.

### **2.2.2.3 Inferencing Strategy**

Inferencing strategy refers to the processes involved in making guesses of the meaning of an unknown word by using available cues, such as linguistic cues, learner's general knowledge of the world, her awareness of the co-text and her relevant linguistic knowledge (Haastrup, 1991). Many studies (e.g., Fraser, 1999; Paribakht & Wesche, 1999; Parry, 1993) have shown that this strategy has been widely used by L2 learners in their reading.

The large body of work on inferencing strategy (e.g., Fraser, 1999; Haastrup, 1991; Huckin & Bloch, 1993; Hulstijn, 1992; Nation & Coady, 1988) has shown its usefulness for reading comprehension and vocabulary learning. For instance, Huckin and Bloch (1993) found that three postgraduate Chinese students studying in the United States could make 57% successful inference of 44 unknown words when reading one passage of their textbook and one journal article with readability level indices of 11 and 17 on basis of three standard readability formulas of Kincaid, Flesch, and Coleman-Liau. In addition, Fraser (1999) reported that French university students of English achieved a mean retention rate of 50% by using L1-based word identification clues to infer, with readability indices at 9.4-12.3 levels based on five readability formulas of Dale-Chall, Fog, Flesch, Flesch-Kincaid and Fry.

Despite evidence for the effectiveness of inferencing for L2 learners in reading comprehension and word learning, some studies (e.g., Hu & Nassaji, 2012; Hulstijn, 2001; Mondria & Wit-de Boer, 1991; Nassaji, 2003; Paribakht, 2005) found that the success rate of inferencing was low and even if learners could infer the correct meaning of an unknown word from context, it did not necessarily lead to word meaning retention.

Research in this area has identified a number of factors that may facilitate or inhibit word inferencing and learning from the written text, such as the nature of written texts (e.g., Frantzen, 2003; Liu & Nation, 1985; Mondria & Wit-de Boer, 1991; Nation, 2001; Paribakht & Wesche, 1999; Pulido, 2007), the selection of knowledge sources (e.g., Hu & Nassaji, 2012, 2014; Nassaji, 2003), learner differences (e.g., Bengueleil & Paribakht, 2004; Frantzen, 2003; Hammadou, 1991; Kaivanpanah & Moghaddam, 2012; Nassaji, 2006;

Paribakht & Wesche, 1997; Qian, 2005; Shen, 2010), ease of inferencing (e.g., Hu & Nassaji, 2012; Pulido, 2009), and type of words (e.g., Paribakht, 2005). As for the nature of written texts, Liu and Nation (1985) found that learners need to know at least 95% of the words in the text, which means that there is one unfamiliar word in every 20 running words in the text. Nation (2001) claimed that this was the minimum requirement for which successful inferencing happened and 98% coverage of the known words was more favorable. Thus, it is easy to understand why some empirical studies adopting higher densities of unknown words resulted in low inferencing rates. In addition, research findings also suggested that the selection of knowledge sources may impact on inferential success. For instance, Hu and Nassaji (2014) found that successful inferencers tended to use a combination of various knowledge sources, such as their linguistic, contextual and background knowledge and to frequently monitor and check their guesses. Moreover, research also suggested that learners' language proficiency level was also highly related to inferencing results. For instance, Kaivanpanah and Moghaddam (2012) found that more proficient readers could make more successful lexical inferencing than did the less proficient ones. Nassaji (2006) and Qian (2005) found that the participants' depth of vocabulary knowledge made a significant contribution to their inferential success.

In short, inferencing strategies seem to be mainly applicable in three conditions. One is that learners need to be able to know at least 95% of all vocabulary in a text; the second one requires learners to have high language proficiency; and the last one demands learners to make the right choices of knowledge sources. In other words, the effectiveness of this strategy is highly related to the individual and task differences.

#### **2.2.2.4 Dictionary Strategies**

In addition to using the contextual cues to make guesses of the meaning of an unknown word in reading, L2 learners often deploy dictionaries as an alternative strategy to discover the meaning of a new word. The adoption of dictionary use strategies affords L2 learners as a convenient means of knowing the meanings of unknown words. Most researchers (e.g., Chan, 2012; Fraser, 1999; Hulstijn, Hollander, & Greidanus, 1996; Knight, 1994; Laufer & Hadar, 1997; Luppescu & Day, 1993) have shown that dictionary use could improve reading comprehension and facilitate vocabulary retention. For instance, Luppescu and Day (1993) reported that Japanese university students who employed a bilingual dictionary scored higher on a vocabulary test than those who did not. There are occasional counterbalances to this viewpoint documented in the literature. For instance, Hulstijn

(1993) found that there was no significant difference between students who looked up many words in the dictionary and those who consulted only a few in terms of vocabulary learning. The controversies concerning the usefulness of the single strategy of dictionary use may be due to the fact that vocabulary learning strategies are often used in combination with others rather than used in isolation.

Studies examining the combined use of dictionary and other strategies have presented abundant evidence to support this point. For instance Fraser (1999) and Knight (1994) have found that the combined adoption of using dictionary strategies with inferencing strategies were more likely for L2 learners to retain the meaning of unknown words than the using of each strategy alone. To be specific, Fraser (1999) discovered that students who used a dictionary or inferred from context achieved a recall rate amounting to about 30% and 31%, respectively. However, when the students inferred and then confirmed the meaning of the word in the dictionary, their recall ascended to 50%. In sum, the findings in the research literature tend to support the positive role played by the using dictionary strategies in vocabulary learning and the combined use of dictionary with inferencing strategies would be more preferable than using each strategy alone.

However, there has been no consensus reached among researchers concerning the issue of choosing the most suitable dictionaries (e.g. monolingual, bilingual dictionaries) for L2 learners. Many English language teachers and researchers advocated that monolingual dictionaries were more desirable than bilingual dictionaries (Baxter, 1980; Folse, 2001). For instance, Folse (2001) conducted a survey and found that 37% of language teachers preferred their students to use monolingual dictionary, whereas only 5% allowed students to use bilingual dictionaries. However, the preference of monolingual dictionaries by these teachers were not well grounded on empirical evidence. The practical usefulness of monolingual dictionaries is still an issue to be debated. This point is perhaps better illustrated by studies on the bilingual dictionaries. It has been well documented in the research literature that using bilingual dictionaries could facilitate university level L2 learners' reading comprehension and vocabulary learning (e.g., Knight, 1994; Lupescu & Day, 1993). There were also studies (e.g., Chan, 2012) that have found monolingual dictionaries in some cases could cause confusion for L2 learner to understand the meaning of an unknown word.

In addition to studies on monolingual and/or bilingual dictionaries, there were studies on a third type of dictionary – bilingualized dictionary. By bilingualized dictionary, I mean

the kind of dictionary that include monolingual definitions and examples of a word and also its translation in a target language. The usefulness of the bilingualized dictionaries has largely remained unclear in the literature. Some studies (e.g., Y. Chen, 2011; Laufer & Hadar, 1997; Laufer & Kimmel, 1997) have found that the bilingualized dictionaries were preferable over bilingual and/or monolingual dictionaries for both intermediate and advanced L2 learners' vocabulary learning. However, other studies (e.g., Chan, 2011) suggested that the combined use of monolingual and bilingualized dictionaries worked better for advanced L2 learners. This was partly because the translation in bilingualized dictionary may sometimes be arbitrary and inaccurate (Chan, 2011).

In general, there has been insufficient evidence in the research literature as to which type of dictionaries works the best for L2 learners. However, one thing seems clear: monolingual dictionaries are less suitable for low proficiency students. The pedagogical implication of this point is that in choosing the right dictionary, L2 learners are not encouraged to adhere to any specific type of dictionaries, but they are invited to choose the one that can help them understand the meaning of the searched word.

### 2.2.2.5 Summary

This section has reviewed the previous studies on the four individual VLSs: rote rehearsal, mnemonic strategies, using dictionaries and inferencing strategies. Some factors, such as learners' language proficiency level, may have impacts on the effectiveness of these strategies (see the following Table 2.1). As Gu (2003b) argued, the selection of VLSs should largely depend on individual differences, learning task and learning context. Moreover, some researchers found that the combined use of VLSs (e.g., combining inferencing strategies with using dictionaries) performed better than using each individual strategy. Therefore, I will review the studies on mixed use of VLSs in the next section.

**Table 2.1: Individual VLSs in relation to language proficiency**

VLS \ Proficiency		Low	Intermediate	High
Rote rehearsal	Oral	√	√	
	Written			
	Visual			
Mnemonic	Keyword strategy	√	√	
Dictionary use	Bilingual	√	√	√
	Monolingual			√
	Bilingualized		√	√
Inferencing	Linguistic knowledge			√
	Contextual knowledge			
	Background knowledge			

### **2.2.3 Studies on Mixed Use of Vocabulary Learning Strategies**

Despite the research efforts to discover the most efficient individual strategy as shown above, vocabulary learning is always approached by learners through integrating various strategies rather than using individual strategies (Gu & Johnson, 1996). The preferable effect of using mixed VLSs in vocabulary learning has been largely established in the literature (Ahmed, 1989; Catalán, 2003; Fan, 2003; Gu, 2002; Gu, 2005; Gu & Johnson, 1996; Kojic-Sabo & Lightbown, 1999; Lawson & Hogben, 1996; Sanaoui, 1992, 1995; Schmitt, 1997; Stoffer, 1995; Takac, 2008; Tseng & Schmitt, 2008). The investigation of mixed use of VLSs is primarily conducted with the assistance of carefully designed questionnaires (e.g., Gu & Johnson, 1996). For this reason, in the following section, I will discuss some prominent questionnaires. The main findings afforded by these questionnaires will be discussed in relation to those afforded by other research methods (e.g., interviewing research participants). These different research methods will be discussed as well.

#### **2.2.3.1 Operationalizing VLSs through VLS Questionnaires**

Researchers have proposed to examine and categorize VLSs and operationalize them through questionnaires from different lenses. The first way of categorizing VLSs and constructing a questionnaire was based on the observation of learners. In this line of research, Sanaoui (1992) developed a semi-structured VLS questionnaire consisting of five aspects: 1) the amount of time spent on independent study; 2) the range of self-initiated learning activities; 3) the extent to which students make records of words; 4) the extent to which students review learnt words; 5) the extent to which they practice using those words. This categorization was based on the view that vocabulary learning strategies should be conceptualized as learners' observable behaviors. As such, it has ignored those mentally-oriented strategies, such as inferencing or monitoring (see Section 2.2.1 above).

Another kind of classification was mainly based on the categorizations of LLSs as discussed in Section 2.2.1, which conceptualizes strategies as both learners' mental activities and behaviors. In line with the categorization proposed by O'Malley and Chamot (1990), Gu and Johnson (1996) categorized VLSs as metacognitive and cognitive strategies, with the former embracing selective attention and self-initiation and the latter involving inferencing, using dictionary, note-taking, rehearsal, encoding, and activation strategies, all of which were further subcategorized. Gu and Johnson's VLS questionnaire is relatively comprehensive as it includes both mentally-oriented strategies and behavior-

based strategies. Nevertheless, there are still some caveats inherent in this VLS questionnaire. Firstly, it emphasizes strategies for learning the meaning of new words, with only one group of strategies (i.e. activation strategies) and a few in other groups for learning how to use a word. However, word knowledge is multi-dimensional involving three interconnected facets – form, meaning and use of the word in both receptive and productive learning processes (Nation, 2001). For this reason, a comprehensive questionnaire of VLSs is expected to cover the strategies for learning the form, meaning and use of words in receptive and productive processes. Secondly, Gu and Johnson (1996) used the 7-point scale from Extremely Untrue of Me (1) to Extremely True of Me (7), which seems to emphasize the extent of the correspondence between the learner and the item. Nevertheless, most of the items in the questionnaire may turn out to be ambiguous for students making the elicited response unreliable. As an example, in the item “I make use of the grammatical structure of a sentence when guessing the meaning of a new word”, students may be confused to discern whether the 7-point scale means how frequently the strategy is used or how well it is used. Therefore, it may probably result in different interpretations by the participants resulting in unreliable data.

Schmitt (1997) conceptualized vocabulary learning strategies as those factors which occur and affect the learning process. From this perspective, Schmitt (1997) categorized VLS questionnaire into two main categories: strategies for discovering a new word’s meaning (discovery strategies) and strategies for consolidating a learnt word (consolidating strategies). Premised on Oxford’s (1990) categorization of LLSs, Schmitt (1997) further categorized discovery strategies into determination and social strategies, and categorized consolidating strategies into social, memory, cognitive and metacognitive strategies. Likewise, Fan’s (2003) VLS questionnaire encompasses nine categories: management, using sources, guessing, using dictionary, repetition, association, grouping, analysis, and known words, which described the vocabulary learning process in a more specific way.

Fan’s (2003) and Schmitt’s (1997) studies have provided a new angle to the examination of VLSs by linking strategies to the dynamic processes of vocabulary learning. Despite that, the main problem of their categorizations is that the discovery and consolidating vocabulary learning processes are hardly separable but instead they constitute a continuing process. For instance, one item of Schmitt’s (1997) questionnaire is “ask classmates for meaning” for finding out the meaning of a new word. But in this discovering process through asking others, one may at the same time consolidate the

meaning(s) of the word being inquired. Additionally, Fan's (2003) and Schmitt's (1997) classifications of VLSs in relation to the vocabulary learning processes exclusively focused on the receptive vocabulary learning processes at the expense of the productive processes. Since vocabulary knowledge is learned and produced in both receptive and productive processes (Nation, 2001), it seems necessary to incorporate some VLSs in relation to the productive learning processes as well.

Most of the above reviewed VLSs classifications were developed theoretically (e.g., Fan, 2003; Schmitt, 1997) or based on observations of learners' learning behaviors (Sanaoui, 1992), but they were not subsequently subjected to statistical analyses like factor analysis. In the research literature, there are other questionnaires (e.g., Stoffer, 1995; Takac, 2008; Zhang & Li, 2011) which were developed from a large number of variables using factor analysis. As Takac (2008) developed a VLS questionnaire tailored for primary school students as opposed to tertiary level students, I focused on comparing the results of factor analysis derived from Stoffer (1995) and Zhang and Li (2011). Stoffer (1995) developed a 53-item VLS questionnaire and identified nine factors by using exploratory factor analysis (EFA). These nine factors are: 1) strategies involving authentic language use; 2) strategies involving creative activities; 3) strategies used for self-motivation; 4) strategies used to create mental linkage; 5) memory strategies; 6) visual/auditory strategies; 7) strategies involving physical action; 8) strategies used to overcome anxiety; and 9) strategies used to organize words. These nine categories involve both metacognitive strategy (factor 3), cognitive strategies (factor 1, 2, 4, 5, 6, 7, 9) and affective strategy (factor 8), which partially resembles the categorizations proposed by Oxford (1990) and O'Malley and Chamot (1990). In a recent study, Zhang and Li (2011) verified VLSs classification by using both EFA and confirmatory factor analysis (CFA). They developed a 60-item VLS questionnaire based on the LLS questionnaire of Oxford (1990) and established a 6-factor structure encompassing metacognitive, first encounter, building links, inferencing, use, and affective. Accordingly, Zhang and Li (2011) indicated that this structure was close to the three component model proposed by O'Malley and Chamot (1990), one that includes metacognitive, cognitive and affective strategies.

Stoffer (1995) and Zhang and Li (2011) have drawn three same main categories of VLS, namely metacognitive, cognitive and affective strategies. However, not all factors under the same category resemble each other. For instance, metacognitive strategies advanced by Stoffer (1995) include self-motivation, whereas Zhang and Li (2011) incorporate self-

motivation, self-monitoring and self-management. Additionally, cognitive strategies in the former study encompass strategies related to authentic language use, creative activities, mental linkage, memory strategies, visual/auditory strategy, physical action and organize words, whereas cognitive strategies in the latter study include first encounter, building links, inferencing, and use. Furthermore, affective strategies in the former study embrace strategies that are used to overcome anxiety while those in the latter study include strategies that are relevant to building learners' confidence. One possible reason for the different factors identified in the above two studies may be attributed to the different strategies investigated. The above two VLS questionnaires seem not to include sufficient coverage of vocabulary learning strategies. For instance, no inferencing strategies were mentioned in the questionnaire proposed by Stoffer (1995) while rehearsal strategies were not included in the questionnaire advanced by Zhang and Li (2011). Since the factors underlying the construct of VLS have not been clearly delimited by researchers, the statistical analyses should be applied to more comprehensive VLS questionnaires to verify its underlying factors.

In general, most VLS questionnaires predominantly emphasize the strategies for learning the meaning of new words, although Gu and Johnson (1996), Fan (2003) and Sanaoui (1992) included some strategies for learning how to use a word. However, word knowledge is multi-dimensional (Nation, 2001). For this reason, a comprehensive questionnaire of VLSs is expected to cover the strategies for learning the form, meaning and use of words in receptive and productive processes. Additionally, the design of VLS questionnaires needs a firm theoretical background as much as they need to be based on the empirical evidence to justify the categorizations.

In the next section, I will discuss the main findings afforded by theses aforementioned VLS questionnaires and other research instruments.

### **2.2.3.2 Main Findings of the Studies on Mixed Use of VLSs**

In this section, I will firstly present previous empirical studies on mixed use of VLSs in chronological order as presented in Table 2.2, then draw some general patterns concerning VLSs use from the findings of the previous studies, and finally identify gaps remaining in literature.

**Table 2.2: Empirical studies on mixed use of VLSs**

Research	Participants	Instruments	Findings
Ahmed (1989)	300 Sudanese EFL university students and secondary students.	<ul style="list-style-type: none"> <li>• Think-aloud task</li> <li>• Observation</li> <li>• Interview</li> <li>• VLS questionnaire</li> </ul>	<ul style="list-style-type: none"> <li>• VLSs used by successful learners: Using a variety of strategies; Paying attention to their learning; Learning words in the context; Using semantic relationship between new and previously-learned L2 words; Using new words in real situations</li> <li>• VLSs used by unsuccessful learners: Using few strategies; Showing little awareness of how to learn new words and how to associate new words to previous knowledge</li> </ul>
Sanaoui (1992)	82 FSL and 4 ESL adult learners enrolled in continuing education courses	<ul style="list-style-type: none"> <li>• Daily written record</li> <li>• Interview</li> <li>• Semi- structured VLS questionnaire</li> <li>• Vocabulary retention test</li> </ul>	<ul style="list-style-type: none"> <li>• Structured learning approach: Creating more opportunities to learn vocabulary; Initiating extensive learning activities; Making the record of learnt words; Reviewing the words regularly; Creating opportunities to practice these words</li> <li>• Unstructured learning approach: Never or seldom take the stages of learning as mentioned in the structured approach.</li> </ul>
Stoffer (1995)	707 undergraduate foreign language learners studying French, German, Japanese, Russian, or Spanish	<ul style="list-style-type: none"> <li>• VLS questionnaire</li> </ul>	<ul style="list-style-type: none"> <li>• VLSs adopted significantly more by female students: Mental linkages; Memory; Organize words</li> </ul>
Gu and Johnson (1996)	850 Chinese non-English major university students	<ul style="list-style-type: none"> <li>• CETBAND 2</li> <li>• VLS questionnaire</li> <li>• Vocabulary size test (Goulden et al., 1990)</li> <li>• VLT at 3000 word level (Nation, 1990)</li> </ul>	<ul style="list-style-type: none"> <li>• VLSs positively correlated with vocabulary size: Self-initiation; Dictionary use; Looking-up strategies; Extracurricular time spent on English; Activation strategies; Semantic encoding</li> <li>• VLSs negatively correlated with vocabulary size: Visual repetition; Imagery encoding</li> </ul>
			<ul style="list-style-type: none"> <li>• Five different types of learners: The most successful learners were the Readers, followed in order by Active strategy users, Non-encoders, Encoders and Passive strategy users.</li> </ul>
			<ul style="list-style-type: none"> <li>• The most frequently used VLSs: Metacognitive regulation; Inferencing; Dictionary use; Note-taking, Oral repetition; Contextual encoding</li> <li>• The least frequently used VLSs: Rehearsal strategies (except for oral</li> </ul>

			repetition); Encoding strategies (except for contextual encoding); Activation strategies
Lawson and Hogben (1996)	15 participants who were enrolled in the advanced section of the first-year Italian course at a university in Australia	<ul style="list-style-type: none"> <li>• Learning task</li> <li>• Think-aloud procedure</li> <li>• Interview</li> </ul>	<ul style="list-style-type: none"> <li>• VLSs used by successful learners: Constantly and skilfully adopt VLSs; Using more than twice of VLSs than the less successful group</li> <li>• VLSs used by unsuccessful learners: Opposite to the successful group</li> </ul>
			<ul style="list-style-type: none"> <li>• VLSs positively correlated with word learning: Simple rehearsal; Appearance similarity; Sound link; Paraphrase; Mnemonic</li> <li>• VLSs negatively correlated with word learning: Spelling analysis; Simple use of context; Complex use of context</li> </ul>
Lessard-Clouston (1996)	14 adult ESL students in one TOEFL preparation class at a college in Toronto	<ul style="list-style-type: none"> <li>• A shorter and modified version of the VLS questionnaire of Sanaoui (1992)</li> <li>• A vocabulary knowledge test</li> <li>• A TOEFL test</li> </ul>	<ul style="list-style-type: none"> <li>• There are three approaches to vocabulary learning: Structured approach, Semi-structured approach and Unstructured approach.</li> <li>• Learners' vocabulary learning approach is not indicative of their vocabulary learning and language learning success.</li> </ul>
Schmitt (1997)	600 Japanese junior high school, high school, university and adults EFL students	• VLS questionnaire	<ul style="list-style-type: none"> <li>• The most frequently used VLSs: Bilingual dictionary; Guessing from context; Asking classmates for meaning; Verbal repetition; Written repetition; Studying the spelling; Saying new words aloud; Taking notes in class; Studying the sound of a word; Using word lists</li> <li>• The least frequently used VLSs: Checking for L1 cognate; Using physical action; Using cognates in study; Using semantic maps; Teachers check flash cards for accuracy</li> </ul>
Kojic-Sabo and Lightbown (1999)	47 ESL undergraduate students and 43 EFL students at the final year of preuniversity schooling	<ul style="list-style-type: none"> <li>• VLS questionnaire</li> <li>• EFL Vocabulary Tests (Meara, 1992)</li> <li>• Cloze test for assessing English proficiency</li> </ul>	<ul style="list-style-type: none"> <li>• VLSs positively related to vocabulary size and English proficiency: Independence, Time, Dictionary use and Review</li> </ul>
			<ul style="list-style-type: none"> <li>• The most frequently used VLSs: Independence; Dictionary use; Note-taking for ESL learners; Dictionary use; Note-taking; Review; Time for EFL learners</li> <li>• The least frequently used VLSs: Review; Time for ESL learners; Independence for EFL learners</li> </ul>

Gu (2002)	648 Chinese non-English major EFL university students	<ul style="list-style-type: none"> <li>• CETBAND 2</li> <li>• VLS questionnaire</li> <li>• Vocabulary size test (Goulden et al., 1990)</li> <li>• VLT at 3000 word level (Nation, 1990)</li> </ul>	<ul style="list-style-type: none"> <li>• VLSs adopted significantly more by female students: VLSs that were found to be correlated to success in Gu and Johnson (1996)</li> </ul>
			<ul style="list-style-type: none"> <li>• VLSs adopted significantly more by arts students: Words should be picked up in context; Meaning-oriented note-taking; Usage-oriented note-taking; Extracurricular time</li> <li>• VLSs adopted significantly more by science students: Words should be memorized; Using word structures</li> </ul>
Catalán (2003)	581 Spanish students learning English or Basque as a foreign language	<ul style="list-style-type: none"> <li>• Modified VLS questionnaire from Schmitt (1997)</li> </ul>	<ul style="list-style-type: none"> <li>• VLSs adopted significantly more by female students: Formal rule strategies; Input elicitation strategies; Rehearsal strategies; Planning strategies</li> <li>• VLSs adopted significantly more by male students: Imagery strategies</li> </ul>
Gu (2003a)	2 successful Chinese non-English major EFL university students	<ul style="list-style-type: none"> <li>• CETBAND 4</li> <li>• Reading task</li> <li>• Think-aloud protocol</li> <li>• Interview</li> </ul>	<ul style="list-style-type: none"> <li>• Two styles of successful learners: the freehand style and the fine brush style, corresponding to Active strategy users and Readers in Gu and Johnson (1996)</li> <li>• VLSs used by both types of successful learners: Self-initiation; Selective attention; Contextual guessing; Dictionary use; Note-taking; Skilful memorizing word lists; Using the learnt words</li> </ul>
Fan (2003)	1067 first-year ESL university students in Hong Kong	<ul style="list-style-type: none"> <li>• VLS questionnaire</li> <li>• longer version of VLT (Nation, 1990)</li> </ul>	<ul style="list-style-type: none"> <li>• VLSs used by successful learners: Management; Using sources; Inferencing; Dictionary use; Consolidating known words; Analysis strategies</li> <li>• VLSs used by unsuccessful learners: Written repetition; Association strategies</li> </ul>
			<ul style="list-style-type: none"> <li>• The most frequently used VLSs: Inferencing; Using and consolidating known words</li> <li>• The least frequently used VLSs: Management; Association; Grouping; Repetition</li> </ul>
Riazi and Alvari (2004)	40 Iranian university EFL students	<ul style="list-style-type: none"> <li>• Written self-reports</li> <li>• Classroom discussions</li> </ul>	<ul style="list-style-type: none"> <li>• The most frequently used VLSs: Rote memorization; Dictionary use; Using context</li> <li>• The least frequently used VLSs: Keyword; Word analysis</li> </ul>
Tseng and Schmitt (2008)	49 university students from Taiwan and 210 university students	<ul style="list-style-type: none"> <li>• A series of vocabulary learning scales</li> </ul>	The use of VLSs significant and positive related to learners' vocabulary knowledge which is

	from mainland China	<ul style="list-style-type: none"> <li>• A vocabulary size test (Schmitt et al., 2001)</li> <li>• Three tests of depth of vocabulary knowledge</li> </ul>	composed of both vocabulary size and depth.
--	---------------------	---	---

It may not be constructive to directly compare the results of the above reviewed research due to their adoption of different methodologies. For instance, these studies have employed different samples and instruments in their respective investigation of VLSs. However, there seem to be some points of agreement among them regarding the general patterns of VLSs use.

Firstly, a large number of studies (Ahmed, 1989; Fan, 2003; Gu, 2003a; Gu & Johnson, 1996; Lawson & Hogben, 1996; Sanaoui, 1992) found that successful learners used a wider variety of VLSs in an orchestrated manner than did less successful learners, though Lessard-Clouston (1996) did not find such a significant relationship between learners' approaches to vocabulary learning and language learning success. For instance, Gu (2003a) identified two types of successful learners as "fine brush style" and "the freehand style" (p. 96), which confirmed the two types of successful learners named active strategy users and readers in Gu and Johnson (1996). The former group was featured by their chief focus on details and mainly on textbooks, whereas the latter was characterised by their intensive attention paid to extensive reading and extracurricular materials. In addition to these differences, these two types of learners shared more similarities as they showed high levels of self-initiation and selective attention. Moreover, they used a wide range of VLSs, such as guessing from context, checking the dictionary for both comprehension and vocabulary learning, taking various notes when they felt necessary, skilful use of word list, and trying to find opportunities to use the newly learnt words.

Secondly, some researchers (e.g., Gu & Johnson, 1996; Kojic-Sabo & Lightbown, 1999) investigated the relationship between VLSs and learners' vocabulary size. Gu and Johnson (1996) found that self-initiation, dictionary use, looking-up strategies, extracurricular time spent on English, activation strategies and semantic encoding were the positive predictors of learners' vocabulary size, whereas visual repetition and imagery encoding emerged as the significant negative predictors. Similarly, Kojic-Sabo and Lightbown (1999) found that learners' independence, time, dictionary use and review positively correlated with word learning.

Thirdly, individual differences and learning context have influences on the choice and use of VLSs. Some studies (Catalán, 2003; Gu, 2002; Stoffer, 1995) found that female students used more VLSs than did male students. For instance, Gu (2002) reported that female students used significantly more VLSs that were found relevant to successful vocabulary learning by Gu and Johnson (1996). In comparison to gender, academic major was evidenced as a less potent predicting factor. Gu (2002) found that arts students adopted significantly more note-taking strategies and spent more extracurricular time than did science students, whereas the science students used significantly more word structure analysis strategies. With regard to language learning context, it seems that VLSs used by ESL learners were different from those of EFL learners. Kojic-Sabo and Lightbown (1999) found that ESL students seemed to involve in more self-initiated vocabulary learning activities, whereas EFL group adopted much more review strategies. In addition, by comparing the findings of Gu and Johnson (1996) and Fan (2003), it seems that the Chinese EFL learners in the mainland China favored more metacognitive strategies, note-taking and repetition strategies than those ESL students in Hong Kong. By contrast, the learners in Hong Kong tended to use more known words, such as revising the learnt words and using the learnt words in learning new words.

The previous research has investigated VLSs in a variety of aspects, such as, VLSs used by both successful and unsuccessful learners, the relationship between VLSs and learners' vocabulary size, the differences of VLSs use in terms of gender and discipline and VLSs used in different learning contexts. Despite fruitful findings have been found by researchers in this field, it is still unclear about the relationship between VLSs and the depth of vocabulary knowledge. Tseng and Schmitt (2008) found a significant relationship between learners' use of VLSs and their vocabulary knowledge (size and depth), however, they did not investigate which group of VLSs could better predict the depth of vocabulary knowledge. Since vocabulary knowledge is a multi-dimensional construct (Nation, 2001), it is worthwhile to examine the contribution of VLSs to the depth of vocabulary knowledge in future research so as to paint a clear picture in this respect.

### **2.2.3.3 Research Methods of Studies on Mixed Use of VLSs**

As indicated in Table 2.2, researchers have adopted different research methods in investigating VLSs employed by learners. Most of the studies (e.g., Catalán, 2003; Fan, 2003; Gu, 2002; Gu & Johnson, 1996; Kojic-Sabo & Lightbown, 1999; Lessard-Clouston, 1996; Schmitt, 1997; Stoffer, 1995) used pure quantitative approach by means of using

typical VLS questionnaires to elicit VLSs from a large number of participants. Questionnaire was regarded by Dörnyei (2003) as relatively easy to administer to a large sample and to collect a large amount of information in a short time. However, the use of questionnaire also has some limitations, for example, it can hardly provide an in-depth investigation of an issue due to the simplicity of the questions (Dörnyei, 2003). On the contrary, several studies (e.g., Gu, 2003a; Lawson & Hogben, 1996; Riazi & Alvari, 2004) have adopted pure qualitative approach by using think-aloud procedure, interview, and written self-reports to elicit learners' VLSs. Despite detailed information concerning learners' choice and use of VLSs afforded by the above qualitative studies, it is apparent that their sample size is relatively small and thus it is difficult to generalize the findings. Given the limitations of both pure quantitative and qualitative research, some researchers (e.g., Dörnyei, 2007; Riazi & Candlin, 2014) advocated for the mixed use of these two approaches. Riazi and Candlin (2014) suggested that mixed-methods research (MMR) could avoid the adversarial incompatibility of pure quantitative and qualitative studies and the deficits of each approach, which has been considered as a promising new research approach in the literature. In addition, Dörnyei (2007) also demonstrated several strengths of using MMR, such as eliminating the weakness of pure quantitative and qualitative approach, multi-level analysis of complex issues and improving validity of the findings. By examining the previous VLSs research, it can be said that only Ahmed (1989), Gu (2005) and Sanaoui (1992) combined both quantitative and qualitative approaches though not in principled ways as MMR designs require. As such, these two studies did not explicitly state their purposes of MMR (triangulation, complementarity, development, initiation, and expansion). Given the benefits of doing MMR research and the insufficient research using this approach in VLSs field, it is appealing for researchers to conduct more integrated MMR studies on VLSs to achieve clearer picture in this field.

#### **2.2.3.4 Summary**

This section has reviewed VLSs through analyzing and comparing current empirical studies on mixed use of VLSs. The main conclusions can be summarized as follows.

Firstly, although many researchers have developed VLS questionnaires based on different theoretical perspectives, it seems there are still room for modifying these questionnaires. Researchers are invited to explore more strategies for learning the form and use of a new word in both receptive and productive aspects. In addition, more empirical studies using factor analysis seem desirable to verify the previous VLSs classifications.

Secondly, studies on VLSs have investigated learners' mixed use of VLSs, such as VLSs adopted by successful and unsuccessful learners (e.g., Ahmed, 1989; Gu, 2003a), the frequency of strategy use in different language learning contexts (e.g., Fan, 2003; Kojic-Sabo & Lightbown, 1999; Schmitt, 1997), the effect of gender and discipline on VLS use (e.g., Catalán, 2003; Gu, 2002), and how VLSs relate to vocabulary size (e.g., Gu & Johnson, 1996; Kojic-Sabo & Lightbown, 1999). However, there are few studies investigating the contribution of VLSs as a whole to L2 learners' depth of vocabulary knowledge. Therefore, research efforts in this line are warranted.

Thirdly, in VLSs field, most studies were found to follow a predominantly quantitative or qualitative approach. Although two studies (Ahmed, 1989; Sanaoui, 1992) combined both quantitative and qualitative approaches, these two studies did not explicitly state their purposes, design and procedures of how the two approaches might fulfil a particular purpose in MMR. Given the advantages of MMR and the insufficient research using this approach in VLSs field, it is appealing to adopt this approach in further VLSs studies.

## **2.3 Conclusion**

In this chapter, I have provided a review of tests and/or approaches to both breadth and depth of VK in the research literature. Based on this review, the study reported in this thesis used the Vocabulary Size Test (VST) as an instrument to assess participants' breadth of vocabulary knowledge (VK). The reason for this decision was three-fold. Firstly, VST was derived from the family-based BNC rather than outdated lemma-based corpora. Secondly, it included fourteen word frequency levels, which were more comprehensive than other reviewed tests. Lastly, the test is designed in the manner that the choice of the right answers could better measure testees' vocabulary knowledge than other tests.

The depth of vocabulary knowledge is as important as its breadth. In this chapter, I argued that the dimensional approach could better conceptualize and measure the depth of VK than the other two approaches – developmental and association approaches. Nevertheless, one demerit of the dimensional approach is that it is difficult for researchers to include all dimensions of VK in one framework. As such, the frameworks proposed by Nation (2001) and Richards (1976) under this approach were deemed to be complementary. To be specific, several dimensions of vocabulary knowledge in Richards's (1976) framework have been ignored in Nation's (2001) framework, such as word frequency, polysemy, and semantic features. On the contrary, Nation's (2001) framework involves

both receptive and productive vocabulary knowledge which has not been touched upon in Richards's (1976) framework. Therefore, the present study integrated Nation's (2001) and Richards's (1976) frameworks so as to better represent the depth of VK. In recognition of the affordances and constraints of the tests designed to measure the depth of VK, two testing instruments – Word Associates Test (WAT) and the computational tool Coh-Metrix were selected to operationalize the integrated frameworks. The interconnection between these two testing instruments and the integrated frameworks has been illustrated in Section 1.3.

As for the vocabulary learning strategies (VLSs), this chapter has reviewed both individual and mixed use of VLSs. The four individual strategies that have been reviewed revealed that the usefulness of each individual strategy largely depends on learners' proficiency levels and task differences. Researchers have found that the combined use of VLSs (e.g., combining inferencing strategies with using dictionaries) worked better than using each individual strategy alone. The review of studies on mixed use of VLSs have shown several insightful findings. Firstly, the VLSs adopted by successful learners were significantly different from those of unsuccessful learners. Secondly, the factors of gender, discipline and language learning contexts impacted on VLSs use. Lastly, the use of VLSs was significantly related to learner' vocabulary size and language proficiency.

The research gap left by previous studies on VLSs is threefold. Firstly, VLS questionnaires have as yet to be more developed and modified. Secondly, the relationship between L2 learners' depth of vocabulary knowledge and use of VLSs has received scant scholarly attention. Lastly, the affordances of the MMR research have not been fully realized in the investigation of VLSs. In the following two chapters, I will address the research gap in the context of quantitative and qualitative examination of VK, VLSs and their relationship.

## **Chapter 3 Quantitative Study**

The present study seeks to investigate the vocabulary knowledge (VK) and vocabulary learning strategies (VLSs) of Chinese EFL students at the tertiary level, together with their potential interrelationship. To do so, it adopts both quantitative and qualitative studies. This chapter focuses on the quantitative aspect of the research. It firstly presents the quantitative research questions to be addressed. It then describes in detail the participants, the research instruments employed, the data collection and subsequent data analysis procedures. Finally, it presents the results drawn from the quantitative analyses.

### **3.1 Research Questions**

In order to examine the participants' VK, VLSs and their relationships, the present quantitative study proposes the following research questions to be addressed:

- (1) What is the level of vocabulary knowledge of Chinese non-English major university students in terms of breadth and depth as measured by Vocabulary Size Test (VST), Word Associates Test (WAT) and the lexical indices of their essays?
- (2) Are there any differences in the participants' receptive vocabulary knowledge (VST and WAT) in terms of their gender, discipline, and proficiency levels?
- (3) Are there any differences in the participants' productive vocabulary knowledge as represented by the lexical indices of their essays in terms of their gender, discipline, and proficiency levels?
- (4) What are the major vocabulary learning strategies reportedly used by Chinese non-English major university students?
- (5) Are there any differences in the participants' reported vocabulary learning strategies in terms of their gender, discipline, and proficiency levels?
- (6) Is there any relationship between the participants' reported vocabulary learning strategies and their receptive vocabulary knowledge (VST and WAT)?
- (7) Is there any relationship between the participants' reported vocabulary learning strategies and their productive vocabulary knowledge as represented by the lexical indices of their essays?

- (8) What is the contribution of individual and overall variables (gender, discipline, proficiency levels and vocabulary learning strategies) to the participants' receptive vocabulary knowledge (VST and WAT)?
- (9) What is the contribution of individual and overall variables (gender, discipline, proficiency levels and vocabulary learning strategies) to the participants' productive vocabulary knowledge as represented by the lexical indices of their essays?

## 3.2 Methods

This section describes the procedures in selecting my research participants, testing instruments, data collection procedures and data analysis procedures.

### 3.2.1 Participants

In an effort to achieve a representative sample of Chinese EFL learners, this study follows the Cochran's (1977) formula to decide sample size and aligns with Bartlett, Kotrlik and Higgins's (2001) interpretation of this formula to support my selection and decision of sample size. The formula is reproduced and explained as follows:

$$no = \frac{(t)^2 * (s)^2}{(d)^2}$$

Where:

no = represents the sample size.

t = is the value for alpha value of .01 or .05 in two tail tests. This study tolerates 5% margin of error, and thus a "t" value of 1.96 is used.

s = is the estimate of standard deviation of the population, which is gained by scale points divided by number of standard deviation. The VLS questionnaire takes the form of five Likert-scale and 4 standard deviations are estimated to capture 98% of possible variables in the range. Thus, an "s" value of 1.25 is used.

d = is the acceptable margin of error for mean being estimated for categorical or continuous variable. The value arises out of multiplying numbers of scale points and acceptable margin of error. My research pertains to continuous, rather than categorical data, and therefore I get 3% for acceptable margin of error for continuous data, so the sample size can be calculated as follows when I replace the values:

$$no = \frac{(1.96)^2 * (1.25)^2}{(0.03 * 5)^2} = 266.78$$

A sample size of around 267 participants was therefore considered the minimum requirement for the present study.

Once the sample size was determined, advertisements with the description of the purpose of the research and the procedures of the data collection were posted on campus bulletin boards of four universities. A number of 455 Chinese non-English major university students from the four universities volunteered themselves to participate in the study. Two of the selected universities are located in southwest China, one in northwest and another one in the middle part of China. In all these universities English is learned and taught as a foreign language. Three of these universities are labelled as “key universities” in China (i.e. top 60), while another one is an ordinary university in China. The students were second-year (sophomore) and third-year students (junior) who had already taken the College English Test Band 4 (CET-4) (see the description of CET-4 test below) and could report their scores on this test. Although the number of volunteers was much higher than the sample size, it was decided not to exclude any of them. However, data from 36 students were excluded from the analysis as they did not complete all the data collection tasks. Therefore, the total number of valid data in the present study was 419. Table 3.1 shows the participants’ demographic information.

**Table 3.1: Summary of the participants’ demographic information (N = 419)**

Categories		<i>n</i>	Per cent (%)
Gender	Male	176	42.00%
	Female	243	58.00%
Discipline	Humanities and social sciences	220	52.51%
	Science and engineering	199	47.49%
Proficiency	Low	46	10.98%
	Intermediate	324	77.33%
	High	49	11.69%

Table 3.1 shows that 42% of the participants were male and 58% were female. These participants were broadly categorized into two groups in terms of their disciplines: 1) humanities and social sciences, and 2) science and engineering. The fields of study in

humanities and social sciences included: language and literature, education, economics, history, psychology, management, law, media and journalism, while those in science and engineering were from such disciplines as computer science, biology, civil engineering, chemical engineering, electrical engineering, and mechanical engineering. As regards participants' English language proficiency, 10.98% were low proficiency students, 77.33% belonged to intermediate proficiency group and 11.69% were students with high proficiency level. This categorization of students' proficiency level was based on the scores they obtained in the CET-4, a national English as a foreign language test developed nearly thirty years ago in China. Its purpose is to test the English proficiency of undergraduate students and postgraduate students in China. The maximum possible score for the test is 710. Testees need to obtain 426 points to get a pass. The main test incorporates three parts, namely, reading, writing and listening. The speaking test is run separately twice a year only for those students who achieved 550 points and more in the main test. Accordingly, I used this criterion to categorize the participants into three proficiency groups. The low proficiency group include those participants whose CET-4 score was lower than the cut-score of 426 in the main test. The high proficiency group include those students who scored 550, the threshold for attending the speaking test, or higher. The intermediate proficiency group include those students whose CET-4 score was 426 or higher, but less than 550 ( $426 \leq \text{intermediate group} < 550$ ).

### **3.2.2 Instruments**

Four instruments were used to collect the required quantitative data. These included an independent essay, the Vocabulary Size Test (VST), the Word Associates Test (WAT), and the vocabulary learning strategy (VLS) questionnaire. Each will be explained below.

#### **3.2.2.1 VLS Questionnaire**

The VLS questionnaire was based on the available VLS questionnaires and the relevant theories in the literature. A comprehensive review of the literature on vocabulary learning strategies was conducted and reported in Chapter 2. Accordingly, a revised VLS scale was constructed primarily based on Nation's vocabulary knowledge framework (2001). A number of items were also selected from Gu and Johnson's (1996) VLS questionnaire and from Schmitt's (1997) VLS questionnaire. There are three categories in Nation's framework: Form, Meaning and Use. "Form" embraces spoken form, written form, and word parts; "meaning" incorporates form and meaning, concept and referents, associations,

and “use” encompasses grammatical function, collocations, and constraints on use (register, frequency, etc.) (Nation, 2001). Based on Nation’s framework, the present VLS scale is composed of two sections: Strategies for learning the meaning of a new word, and strategies for learning the use of a new word. Form was excluded for consideration in the present study because it was found that the length of questionnaire will reduce the reliability of the instrument. Consequently, the VLS scale consists of two sections with 24 items in Section 1 (meaning) and another 28 items in Section 2 (use).

In the second stage, a pilot study was conducted by recruiting 160 Chinese non-English major university students. The main aims of the pilot study were: 1) to check whether the participants have any difficulties in understanding the items of the questionnaire; 2) to elicit any potential VLSs from the participants which were not included in the initial VLS questionnaire; 3) to try to get some possible feedback from the participants regarding the content and format of the VLS questionnaire, 4) to check the reliability of the questionnaire, and finally 5) to estimate the time allocation for the questionnaire. The VLS questionnaire was initially constructed in English, and it was then translated into Chinese by the researcher myself. The Chinese version of the VLS questionnaire was repeatedly reviewed and translated back into English again by a native Chinese PhD student in linguistics who was fluent in English as well. No misinterpretation was found in the Chinese version. The Chinese version of the VLS questionnaire was administered to the participants to avoid possible language barriers.

The results of the pilot study indicated that: 1) participants did not have any problem in understanding the items of the VLS questionnaire; 2) there was no suggestion for any new vocabulary learning strategy; 3) the reliability of the VLS questionnaire, using Cronbach’s alpha, was found to be 0.91, which was quite high, and 4) participants took 20 minutes to 35 minutes to complete the VLS questionnaire. Consequently, the final VLS questionnaire was prepared to be used in the main study.

The final VLS questionnaire was composed of three parts. In the first part, the participants were asked to provide their demographic information, including their gender, major, and their scores of CET-4. I relied on their reported CET-4 scores to determine their proficiency levels mainly due to three reasons. Firstly, it was not practical for all students to provide their CET-4 transcripts as some students have lost their transcripts. Secondly, the request for submitting their CET-4 transcripts would have inflicted their doubt on my promise of confidentiality of their personal information. Lastly, they have been informed

that their CET-4 scores would not impact their academic performance, and as such it is unnecessary for them to fake a different score for me. The second part was the vocabulary learning strategy scale. All items adopt a five-point Likert scale ranging from “never” to “always” to indicate the frequency of strategy use. In the third part of the questionnaire, participants were invited to write down any other strategies they thought they used but were not included in the questionnaire (see Appendix A for the English version of VLS questionnaire). As in the pilot study, the Chinese version of VLS questionnaire (see Appendix B for the Chinese version) was administered to the participants to avoid any possible language barriers.

### 3.2.2.2 Word Associates Test (WAT)

The Word Associates Test was adopted to assess the participants’ level of depth of vocabulary knowledge in English. This test has been revised several times by Read himself (e.g., 1993, 1998). The most recent version (Read, 1998), based on the concept of word association, was designed to measure two aspects of depth of vocabulary knowledge: (1) word meaning, particularly polysemy and synonym, and (2) word collocation. It includes eight options within two boxes for each target word, four options in each box. All the target words are adjectives selected mostly from Barnard’s Second and Third Thousand Word Lists (Nation, 1986). The format of the test is shown below.

#### **sudden**

beautiful   quick   surprising   thirsty	change   doctor   noise   school
--	----------------------------------

To complete the above item, test takers need to select four words from the two boxes, which relate to the target word “sudden”. The words in the left box may help explain the meaning of the target word, while the words in the right box can collocate with the target word. In the above example, there are two correct answers in the left (quick and surprising) and two in the right box (change and noise), but in other items of this test, there may be either one in the left and three in the right box or vice versa (see Appendix C for the WAT).

The WAT was initially designed to measure the depth of vocabulary knowledge of adult ESL learners in New Zealand. This test has gone through repeated piloting and refining. In one recent trial ( $N = 84$ ), the Rasch reliability coefficient of the WAT was reported to be .93 and that of the matching test was .90. There was also a high correlation ( $r = .82$ ) between the WAT and the comparison vocabulary matching test (Read, 1998).

The WAT has been modified and adopted by many researchers in their studies (Greidanus & Nienhuis, 2001; Qian, 1999, 2002). In the present study, the reliability of WAT was .87 based on Kuder-Richardson Formula 21 (K-R 21), which was relatively high.

### 3.2.2.3 The Vocabulary Size Test (VST)

The VST (Nation & Beglar, 2007) was adopted as the instrument for measuring the participants' vocabulary size. It covers fourteen levels of words, with each level consisting of 1000 words. In this test, each 1000 word frequency level contains 10 items and for each item, the testees are invited to choose one right answer that has similar meaning to the target word. The target words are chosen from the updated family-based British National Corpus (BNC) (Nation, 2006). Each item has one point, and the total score of this test is 140. There are two versions of VST – monolingual and bilingual. This research has chosen the bilingual version mainly for two reasons. On the one hand, Nguyen and Nation (2011) found that a bilingual version of VST is as valid as its monolingual version. On the other hand, a bilingual test, as Nguyen and Nation (2011) argued, demands less on participants' English skills and could avoid running the risk of perplexing students with understanding English in each item of choice. Therefore, the present study adopted the bilingual Mandarin version of VST (see Appendix D). Below is an example of one item in the VST in which the test taker must choose a correct option from the four choices given for the word "saw".

They **saw** it.

- a. cut
- b. waited for
- c. looked at
- d. started

VST has displayed a high reliability (.96- .98) as reported by Beglar (2010). Given the numerous advantages of this test and its predominant role among other vocabulary size tests (see Chapter 2), this test was adopted to measure the vocabulary size of my research participants. The reliability of the test in the present study was .70 based on K-R 21, which was acceptable as K-R 21 normally provides the conservative and minimum estimate of the reliability of a test (Alderson, Clapham, & Wall, 1995).

#### **3.2.2.4 Essays**

In order to explore the productive lexical ability of the participants, they were asked to write a timed essay of about 300 words on the topic of “Developing countries should open their doors to foreign companies to open offices and build factories. Do you agree or disagree with the above opinion?” (see Appendix E). There were two reasons for the selection of this topic. Firstly, it is a current and hot topic issue in China which has created a lot of debates, and thus it was thought the university students are both familiar with and interested in this topic. Secondly, by reviewing the topics of essays of CET-4 and IELTS, I found that most topics required an argumentative approach. For these reasons, the above topic was considered appropriate to elicit the writing samples from my research participants. In the present study, the maximum number of words of the participants’ essays was 325 words and the minimum number was 154 words. The average length of the participants’ essays was 198 words.

### **3.2.3 Data Collection Procedures**

The consent forms were distributed to the participants in advance and after their initial agreement to participate in the study. Only those who agreed and signed the consent form were invited to participate into my study. The administration of the tests was organized separately at the four universities. In order to avoid the lengthy testing time, the quantitative data collection was arranged into two sessions. In the first session, the participants completed the writing task in a classroom at their universities. Before the task, the participants were informed that this task would not influence their academic records and that their essays would be kept confidentially and used only for the purpose of this study. Participants were not allowed to use dictionaries for completing the writing task. The maximum time allowed for completing the writing task was 50 minutes, and participants finished the essay between 30-50 minutes. All the original essays were handwritten by the participants and later typed into the computer by the researcher for the purposes of computer-assisted text analysis. In this process, all errors of the participants’ essays were reserved, such as spelling and grammatical errors.

The second session of data collection was one day later in the same classrooms, and the participants were asked to complete the VLS questionnaire, VST and WAT. Presenting the tests in the same order may have some effects on research results, such as practice effect (improvement in performance due to repeated practice with a task) and fatigue effect (decline in performance as the research participant becomes tired or bored while

performing a sequence of tasks) (Cozby, 2009). Since this research does not involve repeated measurement of the same test, the practice effect is irrelevant here. In view of the maximum length of test time in my study (115 minutes maximum) and also taking into consideration other standard tests such as IELTS, TOEFL and CET which often last longer than my test time (e.g., The testing time of listening, reading and writing parts of IELTS is 150 minutes in total), the fatigue effect is perhaps not obvious and may be negligible in this study. This is further evidenced by the fact that most participants have completed the entire test, with few participants ( $n = 36$ ) leaving some sections unfinished. There was no specific time limit set for the participants as the two tests aim to measure knowledge not speed, as explained in the instructions and description of this test by Nation (2012). Before doing the VLS questionnaire and the tests, the students were told again that their responses would be kept confidential and that their scores had no effect on their academic records. Moreover, they were informed that they could ask any questions at any time in the process of testing. The time it took the participants to complete the questionnaire and the two tests was between 75-115 minutes.

### **3.2.4 Data Analysis Procedures**

The corpus of written texts produced by the participants was analyzed by using the lexical indices as performed by the computational tool Coh-Metrix. A brief description of this software is presented as follows.

Coh-Metrix is a software that could provide over 200 measures of language, cohesion, readability to analyze texts (Graesser et al., 2004). In the present study, I only used the measures that are relevant to assess vocabulary knowledge. The indices include lexical diversity as reported by Measure of Textual Lexical Diversity (MTLD) (McCarthy & Jarvis, 2010), lexical frequency as reported by Lexical Database of the Center for Lexical Information (CELEX database) (Baayen, Piepenbrock, & Gulikers, 1995), hypernymy and polysemy as reported by WordNet (Fellbaum, 1998), semantic co-referentiality as reported by Latent Semantic Analysis (LSA) (Landauer, McNamara, Dennis, & Kintsch, 2007), and word meaningfulness, word concreteness, word familiarity, and word imaginability as reported by the MRC Psycholinguistic Database (Wilson, 1988).

The outputs of the Coh-Metrix software were uploaded into SPSS 21 so that statistical analyses could be carried out. The major statistical procedures used in the analyses of the

data were: 1) descriptive statistics; 2) factor analysis; 3) one-way MANOVA; 4) correlation analysis and 5) multiple regression analysis. Each will be explained below.

### 1) Descriptive statistics

According to Pallant (2013), the descriptive statistics mainly serve three purposes: 1) providing the demographic information of the research sample; 2) checking the variables for any violation of the assumptions for some statistical techniques to be adopted to address the research questions; and 3) addressing some specific research questions.

In the present study, firstly, I used the descriptive statistics to summarize the participants' demographic information to be used in Section 3.2.1, such as the number and the percentage of students in the different categorical groups (gender, discipline and proficiency levels). Secondly, it was used to address the first research question in the present study which describes the level of vocabulary knowledge of the participants. Thirdly, the descriptive statistics were used to present the frequency of participants' reported vocabulary learning strategy use at the group-level to address the fourth research question. Finally, the descriptive statistics were also adopted to check the pattern of variation of some variables for some assumptions underlying one-way MANOVA and multiple regression analysis, such as normality and outliers.

### 2) Factor analysis

Factor analysis is widely used by researchers in the development and evaluation of the tests and scales. My studies followed two major steps in implementing FA. In the first step, an exploratory factor analysis with a pre-specified two-factor model was conducted in SPSS in order to confirm the initial structure of the VLS questionnaire. However, the result of the two-factor EFA conflicted with the initial structure of the VLS questionnaire. One possible reason for this is that I constructed the VLS questionnaire based on two available VLS questionnaires, and as such, the underlying factors did not match the original parts in the questionnaire. Consequently, the subsequent EFA was run in SPSS so as to explore any possible underlying components of the VLS questionnaire. Principle Components Analysis (PCA) was adopted in the present study as it was considered as a psychometrically sound and simpler mathematical approach as suggested by Stevens (1996). Before running the factor analysis, the suitability of the data was checked by the researcher based on two principles suggested by Kaiser (1970, 1974). On the one hand, I checked if there are some variables which highly correlate with each other ( $r > .80$ ). If yes, one of them should be

deleted since these are not two different items given the common variance between the two. On the other, I checked if there are items which their correlation with other items is low ( $r < .30$ ), such items should be deleted too because these items do not belong to the construct under investigation.

### 3) One-way MANOVA

Multivariate analysis of variance (MANOVA) is used by researchers as an extension of analysis of variance when there are more than one dependent variables. These dependent variables should be related in some way (Pallant, 2013). The results of one-way MANOVA could tell us whether there is any significant difference between the groups on the composite dependent variable and also on each of the dependent variables separately.

In the present study, I used a set of one-way MANOVA to examine any probable differences in the participants' receptive vocabulary knowledge, productive vocabulary knowledge and reported vocabulary learning strategy use in terms of their gender, discipline and proficiency levels, to address research questions two, three and five. Before conducting the MANOVA analyses, preliminary assumptions were checked for normality, univariate and multivariate outliers, linearity, homogeneity of variance-covariance matrices, equality of variance and multicollinearity.

### 4) Correlation analysis

Correlation analysis is adopted to determine the relationship between two variables in terms of the strength of the relationship and the direction (Pallant, 2013). Pearson product-moment correlation coefficients were calculated and presented in the present study to examine the relationship between participants' reported VLSs and their receptive and productive vocabulary knowledge to address research questions six and seven. Preliminary assumptions for Pearson correlation were checked to ensure that there is no violation of the assumptions of normality, linearity and homoscedasticity.

### 5) Multiple regression analysis

Standard multiple regression is used for exploring how much variance in a criterion variable could be explained by a group of predictor variables and how much unique variance in the criterion variable could be explained by each of the predictor variables separately (Pallant, 2013).

The standard multiple regression was used in the present study to examine the contribution of individual and overall variables (gender, discipline, proficiency level and vocabulary learning strategies) to participants' receptive vocabulary knowledge (VST and WAT) and productive vocabulary knowledge, as represented in participants' essays, to address research questions eight and nine. Moreover, the underlying assumptions were checked by the researcher before conducting the multiple regression analysis.

### 3.3 Results

Results of the quantitative analysis of the study are organized and presented according to each research question in the following sections.

**Research question 1:** What is the level of vocabulary knowledge of Chinese non-English major university students in terms of breadth and depth as measured by Vocabulary Size Test (VST), Word Associates Test (WAT) and the lexical indices of their essays?

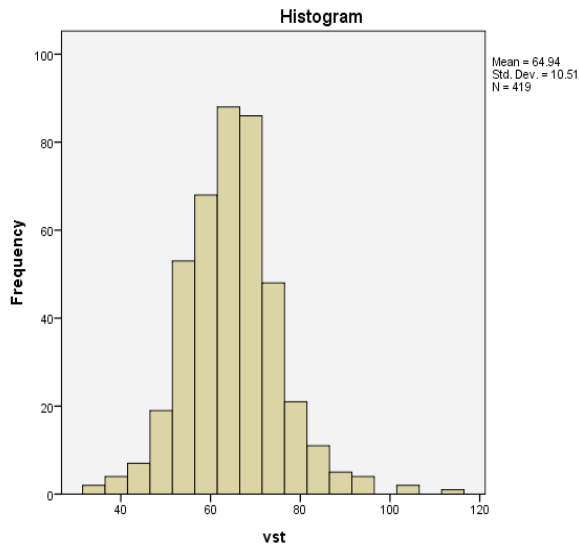
Table 3.2 presents the descriptive statistics for the participants' VST and WAT scores.

**Table 3.2: Descriptive statistics for the participants' VST and WAT scores ( $N = 419$ )**

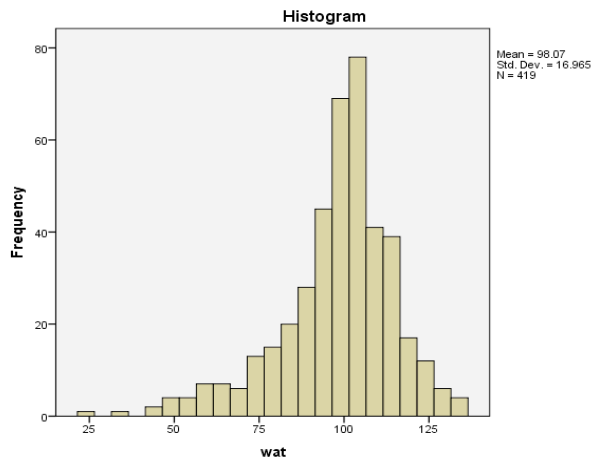
Test	Total possible Score	Minimum	Maximum	<i>M</i>	<i>SD</i>
VST	140.00	34.00	116.00	64.94	10.51
WAT	160.00	24.00	136.00	98.07	16.97

The descriptive results show that the mean value of VST was 64.94. This value should be multiplied by 100 to obtain the average vocabulary size of the participants, which was 6,494 English word families. As for the participants' depth of vocabulary knowledge, the average score of WAT was 98.07, amounting to 61.29 percent of the maximum possible score of WAT.

The distributions of the participants' scores on the VST and WAT, as representations of participants' breadth and depth of English VK, are shown in Figure 3.1 and Figure 3.2.



**Figure 3.1: The distribution of the participants' scores on VST**



**Figure 3.2: The distribution of the participants' scores on WAT**

While the distribution of the scores for VST represents a rather leptokurtic distribution, the scores on the WAT represents a negatively-skewed distribution. Further discussion of the figures will be presented in the Section 5.1.1.

Table 3.3 shows the descriptive statistics for the lexical indices drawn from participants' essays through computational tool Coh-Metrix. I will discuss these figures in the Section 5.1.1.

**Table 3.3: Descriptive statistics for the lexical indices of the participants' essays as computed by Coh-Metrix ( $N = 419$ )**

Lexical indices	Minimum	Maximum	<i>M</i>	<i>SD</i>
Lexical diversity (MTLD all words)	31.42	141.49	67.68	18.64
Lexical frequency (CELEX mean for content words)	1.88	2.86	2.47	0.11
Lexical frequency (CELEX mean for all words)	2.73	3.40	3.03	0.09
Hypernymy (nouns and verbs)	1.16	2.69	1.78	0.21
Polysemy (content words)	3.42	6.07	4.65	0.51
Semantic co-referentiality (LSA, adjacent sentences)	0.04	0.60	0.20	0.09
Word meaningfulness (content words)	383.97	504.69	443.23	15.14
Word concreteness (content words)	323.30	443.89	376.95	18.74
Word imaginability (content words)	360.60	441.18	402.77	14.52
Word familiarity (content words)	552.64	597.67	577.23	5.31

**Research question 2:** Are there any differences in the participants' receptive vocabulary knowledge (VST and WAT) in terms of their gender, discipline, and proficiency levels?

A set of one-way between-groups multivariate analysis of variance (MANOVA) were performed to investigate any probable differences in the participants' receptive vocabulary knowledge, as presented by their scores on VST and WAT, in terms of their gender, broad disciplines and their overall English language proficiency. The two dependent variables were students' scores on VST and WAT, and the independent variables were gender, discipline, and proficiency levels. Before proceeding with the main MANOVA analyses, preliminary assumptions of the test were checked for normality, univariate and multivariate outliers, linearity, homogeneity of variance-covariance matrices, equality of variance and multicollinearity. Six multivariate outliers for the dependent variables (ID = 274, ID = 117, ID = 316, ID = 331, ID = 176, ID = 86) were identified by comparing Mahalanobis distance values with the designated critical values, and were thus excluded from the analysis. A full list of the critical values can be obtained from any statistics text (e.g., Tabachnick & Fidell, 2013, Table C.4). No other serious violations were detected as related to other assumptions for MANOVA.

Table 3.4 indicates that there was a statistically significant difference between the male and female students on the combined dependent variables: their receptive vocabulary knowledge (combined VST and WAT) ( $F(2, 410) = 3.46, p = .032^*$ ; Wilks' Lambda = .98; partial eta squared = .017). The results also show statistically significant differences between the students in the two broad disciplines (humanities and social sciences and science and engineering) ( $F(2, 410) = 6.06, p = .003^*$ ; Wilks' Lambda = .97; partial eta squared = .029), and among the proficiency groups ( $F(4, 818) = 15.85, p = .000^{**}$ ; Wilks' Lambda = .86; partial eta squared = .072) in terms of their receptive vocabulary knowledge.

**Table 3.4: Summary of MANOVA test for the differences in the participants' receptive vocabulary knowledge (combined VST and WAT) in terms of their gender, discipline, and proficiency levels**

Effect	Wilks' Lambda	<i>F</i>	Hypothesis <i>df</i>	Error <i>df</i>	<i>p</i>	Partial Eta Squared
Gender	0.98	3.46	2	410	0.032*	0.017
Discipline	0.97	6.06	2	410	0.003*	0.029
Proficiency	0.86	15.85	4	818	0.000**	0.072

\* $p < .05$ , \*\* $p < .001$

When the results for the two dependent variables were examined separately (see Table 3.5), the only significant difference found between the male and female students, by using a Bonferroni adjusted alpha level of .025 (.05 were divided by the number of dependent variables), was their scores of VST ( $F(1, 411) = 6.91$ ,  $p = .009^*$ , partial eta squared = .017). The results of the mean scores show that the vocabulary size of the female students ( $M = 65.71$ ,  $SD = .63$ ) were slightly larger than that of the male students ( $M = 63.13$ ,  $SD = .75$ ). In addition, the humanities and social sciences students differed significantly from the science and engineering students both in their scores of VST ( $F(1, 411) = 7.81$ ,  $p = .005^*$ , partial eta squared = .019), and their scores of WAT ( $F(1, 411) = 8.67$ ,  $p = .003^*$ , partial eta squared = .021). The vocabulary size of the science and engineering students ( $M = 66.07$ ,  $SD = 9.64$ ) were larger than that of the humanities and social sciences students ( $M = 63.37$ ,  $SD = 9.96$ ). Moreover, in terms of vocabulary depth, the science and engineering students ( $M = 100.86$ ,  $SD = 14.23$ ) also outperformed the humanities and social sciences students ( $M = 96.25$ ,  $SD = 17.23$ ). Finally, significant differences were also detected among the three proficiency groups both in their VST scores ( $F(2, 410) = 20.95$ ,  $p = .000^{**}$ , partial eta squared = .093), and their WAT scores ( $F(2, 410) = 21.89$ ,  $p = .000^{**}$ , partial eta squared = .096).

**Table 3.5: Summary of MANOVA test for the differences in the participants' scores on VST and WAT in terms of their gender, discipline, and proficiency levels**

Source	Dependent Variable	Type III SS	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>	Partial Eta Squared
Gender	VST	665.91	1	665.91	6.91	0.009*	0.017
	WAT	311.97	1	311.97	1.21	0.271	0.003
Discipline	VST	751.21	1	751.21	7.81	0.005*	0.019
	WAT	2190.27	1	2190.27	8.67	0.003*	0.021
Proficiency	VST	3737.50	2	1868.75	20.95	0.000**	0.093
	WAT	10229.18	2	5114.59	21.89	0.000**	0.096

\* $p < .025$ , \*\* $p < .001$

In order to identify where exactly the significant differences lie among the three proficiency groups, the post-hoc comparisons using Tukey HSD test was used. The results (see Table 3.6) indicate the significant differences existed among the three proficiency groups in both VST and WAT scores. An inspection of the mean scores shows that the VST of the high proficiency students ( $M = 71.83$ ,  $SD = 10.44$ ) was larger than that of the intermediate proficiency group students ( $M = 64.29$ ,  $SD = 9.17$ ) and that of the low proficiency group students ( $M = 59.44$ ,  $SD = 10.24$ ). Moreover, the high proficiency students ( $M = 109.44$ ,  $SD = 16.21$ ) performed better in their scores of WAT than the intermediate proficiency students ( $M = 98.15$ ,  $SD = 14.65$ ), and the low proficiency group students ( $M = 88.56$ ,  $SD = 18.46$ ).

**Table 3.6: Results of the Tukey HSD test**

Dependent Variable	(I) Proficiency	(J) Proficiency	Mean Difference (I-J)	SD	p	95% Confidence Interval	
						Lower Bound	Upper Bound
VST	Low	Intermediate	-4.85	1.50	0.004*	-8.38	-1.31
		High	-12.39	1.96	0.000**	-17.00	-7.78
	Intermediate	Low	4.85	1.50	0.004*	1.31	8.38
		High	-7.54	1.46	0.000**	-10.98	-4.10
	High	Low	12.39	1.96	0.000**	7.78	17.00
WAT	Low	Intermediate	7.54	1.46	0.000**	4.10	10.98
		High	-9.59	2.43	0.000**	-15.32	-3.87
	Intermediate	Low	-20.88	3.17	0.000**	-28.34	-13.42
		High	9.59	2.43	0.000**	3.87	15.32
	High	Low	-11.29	2.37	0.000**	-16.86	-5.73
		Intermediate	20.88	3.17	0.000**	13.42	28.34
		Intermediate	11.29	2.37	0.000**	5.73	16.86

\* $p < .025$ , \*\* $p < .001$

**Research question 3:** Are there any differences in the participants' productive vocabulary knowledge as represented by the lexical indices of their essays in terms of their gender, discipline, and proficiency levels?

To answer this question, a set of one-way between groups MANOVA tests were run with the lexical indices (lexical diversity, lexical frequency, hypernymy, polysemy, semantic co-referentiality, meaningfulness, concreteness, imagibility, familiarity) as the dependent variables and participants' gender, discipline and level of proficiency as the independent variables. Before running the test, however, it was necessary to check the

underlying assumptions including normality, univariate and multivariate outliers, linearity, homogeneity of variance-covariance matrices, equality of variance and multicollinearity for the variables. Six multivariate outliers for the dependent variables (ID = 323, ID = 204, ID = 234, ID = 213, ID = 243, ID = 380) were detected and excluded from the analysis. In addition, the correlation between two of the dependent variables, concreteness and imaginability, was very high (i.e. .929) exceeding .80. Thus, the assumption of multicollinearity was violated for these two variables. Therefore, one of the two variables, imaginability, was excluded from analysis because it showed low correlations with the other two dependent variables polysemy and familiarity, which on the other hand slightly violated the assumption of moderate correlations between dependent variables.

Table 3.7 indicates that there was a statistically significant difference among the proficiency groups on the combined dependent variable ( $F(16, 806) = 4.52, p = .000^{**}$ ; Wilks' Lambda = .84; partial eta squared = .082). However, the results also show that there was no statistically significant difference between the students in the two broad disciplines (humanities and social sciences and science and engineering) ( $F(8, 404) = .50, p = .855$ ; Wilks' Lambda = .99; partial eta squared = .010), and between the male and female students ( $F(8, 404) = 1.09, p = .373$ ; Wilks' Lambda = .98; partial eta squared = .021).

**Table 3.7: Summary of MANOVA test for the differences in the participants' productive vocabulary knowledge (combined lexical indices) in terms of their gender, discipline, and proficiency levels**

Effect	Wilks' Lambda	$F$	Hypothesis $df$	Error $df$	$p$	Partial Eta Squared
Proficiency	0.84	4.52b	16	806	0.000**	0.082
Discipline	0.99	0.50b	8	404	0.855	0.010
Gender	0.98	1.09b	8	404	0.373	0.021

\*\* $p < .001$

Table 3.8 shows the results when the eight dependent variables were examined separately. By using a Bonferroni adjusted alpha value of .006, the statistical significant differences were detected among the proficiency groups in terms of lexical diversity ( $F = 23.86, p = .000^{**}$ , partial eta squared = .104), lexical frequency ( $F = 13.63, p = .000^{**}$ , partial eta squared = .062), lexical familiarity ( $F = 12.46, p = .000^{**}$ , partial eta squared = .057) and lexical meaningfulness ( $F = 5.38, p = .005^{*}$ , partial eta squared = .026).

**Table 3.8: Summary of MANOVA test for the differences in each lexical index among the participants' proficiency levels**

Source	Dependent Variable	Type III Sum of Squares	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>	Partial Eta Squared
Proficiency	Lexical diversity	14425.32	2	7212.66	23.86	0.000**	0.104
	Lexical frequency	0.30	2	0.15	13.63	0.000**	0.062
	Hypernymy	0.03	2	0.02	0.35	0.708	0.002
	Polysemy	0.20	2	0.10	0.39	0.679	0.002
	Semantic co-referentiality	0.06	2	0.03	4.50	0.012	0.021
	Word meaningfulness	2244.65	2	1122.33	5.38	0.005*	0.026
	Word concreteness	2200.30	2	1100.15	3.15	0.044	0.015
	Word familiarity	600.29	2	300.14	12.46	0.000**	0.057

\* $p < .006$ , \*\* $p < .001$

In order to identify where exactly the significant differences lie among the three proficiency groups, the post-hoc comparisons using Tukey HSD tests were used. The results (see Table 3.9) reveal that the high proficiency group was significantly different from both the intermediate proficiency group and the low proficiency group in terms of lexical diversity, lexical frequency and lexical familiarity. However, the high proficiency group was only significantly different from the low proficiency group in terms of lexical meaningfulness. An inspection of the mean values reveals that the mean value of lexical diversity of the high proficiency group ( $M = 80.48$ ,  $SD = 18.16$ ) was larger than that of the intermediate proficiency group ( $M = 65.68$ ,  $SD = 15.77$ ) and the low proficiency group ( $M = 59.00$ ,  $SD = 12.73$ ). In addition, the mean value of lexical frequency of the high proficiency group ( $M = 2.42$ ,  $SD = .09$ ) was smaller than that of the intermediate proficiency group ( $M = 2.48$ ,  $SD = .10$ ) and the low proficiency group ( $M = 2.50$ ,  $SD = .09$ ). Moreover, the mean value of lexical familiarity of the high proficiency group ( $M = 574.91$ ,  $SD = 4.76$ ) was also smaller than that of the intermediate proficiency group ( $M = 577.63$ ,  $SD = 4.65$ ) and the low proficiency group ( $M = 578.64$ ,  $SD = 4.52$ ). Finally, the mean value of lexical meaningfulness of the high proficiency group ( $M = 437.93$ ,  $SD = 12.01$ ) was smaller than that of the intermediate proficiency group ( $M = 443.32$ ,  $SD = 13.78$ ) and the low proficiency group ( $M = 445.73$ ,  $SD = 13.58$ ).

**Table 3.9: Results of the Tukey HSD test**

Dependent variable	(I) Proficiency	(J) Proficiency	Mean Difference (I-J)	SE	p	95% Confidence Interval	
						Lower Bound	Upper Bound
Lexical diversity	Low	Intermediate	-7.72	2.77	0.015	-14.23	-1.21
		High	-23.64	3.61	0.000**	-32.12	-15.15
	Intermediate	Low	7.72	2.77	0.015	1.21	14.23
		High	-15.92	2.69	0.000**	-22.25	-9.59
	High	Low	23.64	3.61	0.000**	15.15	32.12
		Intermediate	15.92	2.69	0.000**	9.59	22.25
Lexical frequency	Low	Intermediate	0.03	0.02	0.260	-0.01	0.07
		High	0.10	0.02	0.000**	0.05	0.15
	Intermediate	Low	-0.03	0.02	0.260	-0.07	0.01
		High	0.08	0.02	0.000**	0.04	0.11
	High	Low	-0.10	0.02	0.000**	-0.15	-0.05
		Intermediate	-0.08	0.02	0.000**	-0.11	-0.04
Word familiarity	Low	Intermediate	1.05	0.78	0.374	-0.79	2.88
		High	4.54	1.02	0.000**	2.15	6.94
	Intermediate	Low	-1.05	0.78	0.374	-2.88	0.79
		High	3.50	0.76	0.000**	1.71	5.28
	High	Low	-4.54	1.02	0.000**	-6.94	-2.15
		Intermediate	-3.50	0.76	0.000**	-5.28	-1.71
Word meaningfulness	Low	Intermediate	3.42	2.30	0.299	-2.00	8.83
		High	9.48	3.00	0.005*	2.43	16.53
	Intermediate	Low	-3.42	2.30	0.299	-8.83	2.00
		High	6.07	2.24	0.019	0.81	11.33
	High	Low	-9.48	3.00	0.005*	-16.53	-2.43
		Intermediate	-6.07	2.24	0.019	-11.33	-0.81

\* $p < .006$ , \*\* $p < .001$

**Research question 4:** What are the major vocabulary learning strategies reportedly used by Chinese students?

The distribution of the frequency of vocabulary learning strategy use reported by the participants was presented in Appendix F. The descriptive statistics show that the mean values of the strategy items in Section one, strategies for learning the meaning of a new word, ranged from 1.61 to 3.93, while in Section two, strategies for learning the use of a new word ranged from 2.28 to 4.00. The overall mean for the former was 3.01 (3 = sometimes, 4 = often) and the overall mean for the latter was 2.91 (2 = seldom, 3 = sometimes). The results of the descriptive analyses also reveal that the most commonly

used strategy reported by the participants was “when I don’t know how to use a word, I look it up in the dictionary” ( $M = 4.00$ ,  $SD = .99$ ). While the least commonly used strategy reported by the participants was “I try to group new words together within a storyline to remember the meanings of new words (e.g., I make up a story by using the new words)” ( $M = 1.61$ ,  $SD = .86$ ).

### **Exploratory factor analysis**

An exploratory factor analysis with a pre-specified two-factor model was adopted to validate the construct of the VLS questionnaire. The 52 items of the VLS questionnaire were subjected to principal components analysis (PCA) using SPSS version 21. Prior to performing PCA, the suitability of data for factor analysis was assessed. Kaiser (1970, 1974) suggested two principles for checking the correlation matrix. The first is to find out if there are some variables which highly correlate with each other ( $r > .80$ ). If yes, one of them should be deleted since these are not two different items given the common variance between the two. Second, if there are items which their correlation with other items is low ( $r < .30$ ), such items should be deleted too because these items do not belong to the construct under investigation. Therefore, three items (q1.7, q1.13, q1.22) of the VLS questionnaire were removed because their correlation with other items was low ( $r < .30$ ). After removing the above three items, the two-factor exploratory factor analysis was run. The Kaiser-Meyer-Olkin value was .879, exceeding the recommended value of .6 (Kaiser, 1970, 1974), and Bartlett’s Test of Sphericity (M. S. Bartlett, 1954) reached statistical significance, supporting the factorability of the correlation matrix.

The two-factor solution explained only a total of 28.49% of the variance which seems a little bit low. Factor 1 contributes 22.10% and Factor 2 contributes 6.40%, respectively. To aid in the interpretation of these two components, Oblimin rotation was used as this rotation method could result in the valid solution regardless of the correlation between the two components (Pallant, 2013). The rotated solution reveals the presence of simple structure (Thurstone, 1947), with both components showing a number of strong loadings and most variables loading substantially on only one component. It shows that the strategies related to deep cognitive processing loaded strongly on Component 1, such as production-based strategies, self-initiated strategies, association and imagination strategies, while some mechanical strategies loaded strongly on Component 2, such as repetition strategies, dictionary strategies and note-taking strategies. However, this interpretation of the solution was not consistent with the two sections of my outlined VLS questionnaire

which are strategies for learning the meanings of new words and strategies for learning the use of new words. Such two factors categorization was therefore dismissed.

Because of the failure of the EFA of the VLS questionnaire with two factors, principal component analysis (PCA) was used to identify the possible number of underlying factors or dimensions of the VLS questionnaire. There are three techniques which are frequently adopted to decide the retention of the number of factors: Kaiser's criterion; scree plot; and parallel analysis (PA) (Pallant, 2013). According to Kaiser's criterion, only factors with an eigenvalue of 1.0 or above are retained for further investigation. Thus, following this principle, fourteen factors with an eigenvalue of 1.0 or above in total were retained in the present study as presented in Table 3.10.

**Table 3.10: Total variance explained**

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	11.11	21.36	21.36	11.11	21.36	21.36
2	3.14	6.05	27.41	3.14	6.05	27.41
3	2.49	4.80	32.20	2.49	4.80	32.20
4	2.06	3.97	36.17	2.06	3.97	36.17
5	1.81	3.48	39.65	1.81	3.48	39.65
6	1.74	3.35	43.01	1.74	3.35	43.01
7	1.65	3.18	46.18	1.65	3.18	46.18
8	1.46	2.81	48.99	1.46	2.81	48.99
9	1.30	2.49	51.48	1.30	2.49	51.48
10	1.29	2.47	53.95	1.29	2.47	53.95
11	1.21	2.32	56.27	1.21	2.32	56.27
12	1.12	2.15	58.42	1.12	2.15	58.42
13	1.11	2.13	60.55	1.11	2.13	60.55
14	1.03	1.99	62.54	1.03	1.99	62.54
Extraction Method: Principal Component Analysis.						

Kaiser's criterion has been criticized by other scholars for resulting in too many factors in some cases (Pallant, 2013). Thus, it was thought that the parallel analysis (Pallant, 2013) will help here. By running the program called Monte Carlo PCA, three pieces of information were needed: the number of variables (which is 52 in the present study); the number of participants in the sample (419 in the present study); and the number of replications (usually specified and set as 100). This program generated 100 sets of random data of the same size as the real data file (52 variables  $\times$  419 cases) and calculated the average eigenvalues for these 100 random samples as presented in Table 3.11.

**Table 3.11: Results of Monte Carlo Parallel Analysis (PA)**

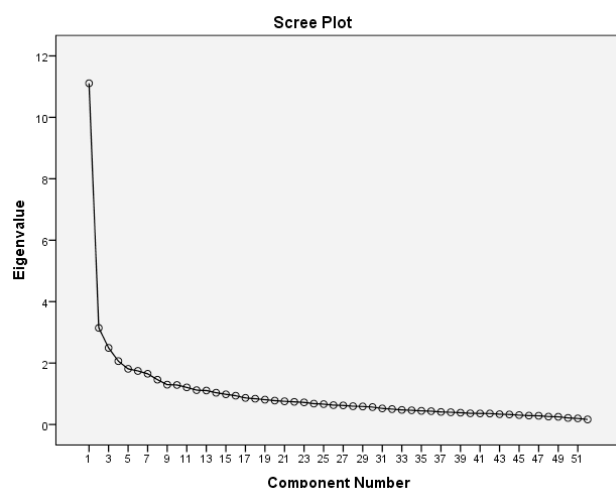
Component	Random Eigenvalue	<i>SD</i>
1	1.76	0.04
2	1.69	0.04
3	1.63	0.03
4	1.58	0.03
5	1.54	0.02
6	1.50	0.02
7	1.46	0.02
8	1.43	0.02
9	1.39	0.02
10	1.36	0.02
11	1.34	0.02
12	1.31	0.02

The first eigenvalue obtained in PCA with the corresponding random eigenvalue generated by parallel analysis were compared in the following table. According to Pallant (2013), if the initial eigenvalue is larger than the PA eigenvalue, the factor should be retained, otherwise it should be deleted. Based on this principle, 8 factors were retained as plausible underlying factors for VLS questionnaire (see Table 3.12).

**Table 3.12: The comparison of initial eigenvalues and PA eigenvalues**

Component	Initial Eigenvalues	PA Eigenvalues	Accepted/Rejected
1	11.11	1.76	Accepted
2	3.14	1.69	Accepted
3	2.49	1.63	Accepted
4	2.06	1.58	Accepted
5	1.81	1.54	Accepted
6	1.74	1.50	Accepted
7	1.65	1.46	Accepted
8	1.46	1.43	Accepted
9	1.30	1.39	Rejected
10	1.29	1.36	Rejected
11	1.21	1.34	Rejected
12	1.12	1.31	Rejected

The third technique for determining the retention of the number of factors is Cattell's scree test (Cattell, 1966). This involves looking for a change (or elbow) in the shape of the plot and only those components above the point should be retained. Figure 3.3 shows that after component 8 or 9, the line becomes almost flat. Thus based on this scree plot, it is probably safe to retain 7 or 8 factors.



**Figure 3.3: Scree plot**

After checking the above three techniques, factor analyses were run again with the identified 8 factors and 7 factors. Based on the principles suggested by Pallant (2013) and Tabachnick and Fidell (2013), it is up to researchers to finally determine the number of factor that can best demonstrate the underlying dimensions of the variables. Thus, according to the findings of the previous studies on vocabulary learning strategies, 7 factors were retained as the satisfactory solution for the present study, which accounted for 46.18% of the total variance of vocabulary learning strategy use. The results of the exploratory factor analysis are presented in Table 3.13. Of the other factors, repetition and social strategies were identified as a single factor in the present study. The result indicates that in my study repetition and social strategies strongly correlated with each other in that, rather than standing as two separate factors, they were combined to form a single factor impacting on learners' vocabulary learning.

**Table 3.13: Factors underlying VLS questionnaire**

Categories	Strategy Items
Self-initiated	2.12, 2.13, 2.14, 2.10, 2.11, 1.10
Repetition/Social	2.1, 2.2, 1.11, 2.21
Using dictionary	1.8, 2.4, 1.9, 2.24, 2.3, 1.6, 2.5, 2.20, 2.25
Production-based	2.22, 2.17, 2.23, 2.28, 2.15, 2.16, 2.18, 2.26, 2.27, 2.19
Note-taking	2.7, 2.6, 1.23, 2.9, 2.8, 1.24, 1.12
Association and imagination	1.16, 1.17, 1.15, 1.18, 1.20, 1.19, 1.14, 1.13
Inferencing	1.2, 1.3, 1.21, 1.4, 1.5, 1.1

The descriptive statistics for the seven categories of VLSs are presented in Table 3.14. The descriptive results reveal that inferencing strategies were reportedly most commonly used by the participants ( $M = 3.60$ ,  $SD = .59$ ), which was followed by using dictionary ( $M = 3.36$ ,  $SD = .63$ ). Note-taking strategies ( $M = 2.59$ ,  $SD = .74$ ) were reportedly the least

frequently used by the participants along with production-based strategies ( $M = 2.67$ ,  $SD = .61$ ) and association and imagination strategies ( $M = 2.75$ ,  $SD = .63$ ). Normality of these variables was checked through histograms. Although they were not perfectly normal, no abnormality was observed.

**Table 3.14: Descriptive statistics for the seven categories of vocabulary learning strategies reportedly used by the participants ( $N = 419$ )**

Categories	$M$	$SD$	Skewness	Kurtosis
Inferencing	3.60	0.59	0.13	-0.20
Using dictionary	3.36	0.63	-0.29	0.21
Repetition/Social	3.34	0.69	0.04	-0.25
Self-initiated	2.76	0.76	0.20	-0.37
Association and imagination	2.75	0.63	0.12	-0.50
Production-based	2.67	0.61	0.31	0.21
Note-taking	2.59	0.74	0.06	-0.40

**Research question 5:** Are there any differences in the participants' reported vocabulary learning strategies in terms of their gender, discipline, and proficiency levels?

A set of one-way between-groups MANOVA was conducted to investigate any probable differences in the participants' reported vocabulary learning strategies in terms of their gender, discipline, and proficiency levels. The seven dependent variables were: association and imagination strategies, dictionary strategies, self-initiated strategies, inference strategies, note-taking strategies, production-based strategies, and repetition/social strategies. The independent variables were gender, discipline, and proficiency levels. Before conducting the main MANOVA analyses, preliminary assumptions of the test were checked for normality, univariate and multivariate outliers, linearity, homogeneity of variance-covariance matrices, equality of variance and multicollinearity. There was no serious violation of the assumptions, except for four multivariate outliers for the dependent variables (ID = 274, ID = 21, ID = 262, ID = 144) which were excluded from the analysis. Table 3.15 shows the differences in the participants' overall reported vocabulary learning strategies in terms of their gender, discipline, and proficiency levels. The results indicate that there were statistically significant differences in the overall reported vocabulary learning strategy use (combined dependent variables), between the male and female students ( $F(7, 408) = 9.01$ ,  $p = .000^{**}$ ; Wilks' Lambda = .87; partial eta squared = .134); between the humanities and social sciences students and the science and engineering students ( $F(7, 408) = 6.01$ ,  $p = .000^{**}$ ; Wilks' Lambda = .91; partial eta squared = .093); and also among the students in different proficiency groups ( $F(14, 814) = 1.92$ ,  $p = .021^{*}$ ; Wilks' Lambda = .94; partial eta squared = .032).

**Table 3.15: Summary of MANOVA test for the differences in the participants' overall reported vocabulary learning strategies in terms of their gender, discipline, and proficiency levels**

Effect	Wilks' Lambda	<i>F</i>	Hypothesis <i>df</i>	Error <i>df</i>	<i>p</i>	Partial Eta Squared
Gender	0.87	9.01	7	408	0.000**	0.134
Discipline	0.91	6.01	7	408	0.000**	0.093
Proficiency	0.94	1.92	14	814	0.021*	0.032

\* $p < .05$ , \*\* $p < .001$

Table 3.16 presents the results when the seven dependent variables were considered separately. By using a Bonferroni adjusted alpha level of .007, the significant differences between the male and female students were found in the note-taking strategies ( $F(1, 414) = 54.57, p = .000^{**}$ , partial eta squared = .116). An inspection of the mean scores indicates that the female students reported using more note-taking strategies ( $M = 2.81, SD = .70$ ) than the male students ( $M = 2.29, SD = .70$ ). In addition, the significant difference in note-taking strategies was also found between the students in the two broad disciplines ( $F(1, 414) = 19.54, p = .000^{**}$ , partial eta squared = .045). The humanities and social sciences students reported using more note-taking strategies ( $M = 2.74, SD = .73$ ) than the science and engineering students ( $M = 2.43, SD = .72$ ). Finally, the only significant difference found among the proficiency groups was the inferencing strategies ( $F(2, 413) = 5.39, p = .005^{*}$ , partial eta squared = .025).

**Table 3.16: Summary of MANOVA test for the differences in each group of participants' reported vocabulary learning strategies in terms of their gender, discipline, and proficiency levels**

Source	Dependent Variable	Type III Sum of Squares	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>	Partial Eta Squared
Gender	Self-initiated	1.00	1	1.00	1.78	0.183	0.004
	Using dictionary	2.78	1	2.78	7.29	0.007	0.017
	Production-based	0.11	1	0.11	0.30	0.584	0.001
	Note-taking	26.64	1	26.64	54.57	0.000**	0.116
	Association and imagination	0.55	1	0.55	1.41	0.237	0.003
	Inferencing	0.02	1	0.02	0.05	0.832	0.000
	Repetition/Social	2.50	1	2.50	5.30	0.022	0.013
Discipline	Self-initiated	0.22	1	0.22	0.39	0.533	0.001
	Using dictionary	0.44	1	0.44	1.13	0.289	0.003
	Production-based	0.37	1	0.37	1.02	0.312	0.002
	Note-taking	10.31	1	10.31	19.54	0.000**	0.045
	Association and imagination	0.51	1	0.51	1.32	0.251	0.003
	Inferencing	0.00	1	0.00	0.00	0.967	0.000

	Repetition/Social	1.39	1	1.39	2.94	0.087	0.007
Proficiency	Self-initiated	3.91	2	1.96	3.51	0.031	0.017
	Using dictionary	1.22	2	0.61	1.57	0.208	0.008
	Production-based	3.51	2	1.75	4.92	0.008	0.023
	Note-taking	2.32	2	1.16	2.11	0.122	0.01
	Association and imagination	2.58	2	1.29	3.35	0.036	0.016
	Inferencing	3.59	2	1.79	5.39	0.005*	0.025
	Repetition/Social	0.07	2	0.04	0.08	0.928	0.000

\* $p < .007$ , \*\* $p < .001$

In order to find out where exactly the significant differences lie among the proficiency groups, the post-hoc comparisons using the Tukey HSD test was conducted. Results, as presented in Table 3.17, indicate that the significant difference in using inferencing strategy only lies between the high proficiency group ( $M = 3.82$ ,  $SD = .60$ ) and the low proficiency group ( $M = 3.43$ ,  $SD = .52$ ). The intermediate proficiency group ( $M = 3.60$ ,  $SD = .58$ ) did not differ significantly from either the high proficiency group or the low proficiency group.

**Table 3.17: Results of the Tukey HSD test**

Dependent Variable	(I) Proficiency	(J) Proficiency	Mean Difference (I-J)	SD	$p$	95% Confidence Interval	
						Lower Bound	Upper Bound
Inferencing strategies	Low	Intermediate	-0.17	0.09	0.134	-0.39	0.04
		High	-.39	0.09	0.003*	-0.67	-0.11
	Intermediate	Low	0.17	0.09	0.134	-0.04	0.39
		High	-0.21	0.09	0.045	-0.42	0.00
	High	Low	0.39	0.09	0.003*	0.11	0.67
		Intermediate	0.21	0.09	0.045	0.00	0.42

\* $p < .007$

**Research question 6:** Is there any relationship between participants' reported vocabulary learning strategies and their receptive vocabulary knowledge (VST and WAT)?

The relationship between participants' reported vocabulary learning strategies and their receptive vocabulary knowledge was investigated using the Pearson product-moment correlation coefficient. Preliminary analyses were performed to ensure no violation of the

assumptions of normality, linearity and homoscedasticity. Table 3.18 indicates that there was a significant, positive correlation between learners' vocabulary size (VST) and their reported use of inferencing strategies ( $r = .16, p < .01$ ) and there was a significant, negative correlation between learners' VST scores and their reported use of repetition and social strategies ( $r = -.11, p < .05$ ). In addition, there was a significant, positive correlation between learners' depth of vocabulary knowledge (WAT) and their reported use of inferencing strategies ( $r = .15, p < .01$ ). No significant correlations were found between participants' receptive VK and other reported vocabulary learning strategies.

**Table 3.18: Pearson product-moment correlations between participants' reported vocabulary learning strategies and their receptive vocabulary knowledge (VST and WAT) ( $N = 419$ )**

		Inferencing	Association & imagination	Dictionary	Production-based	Note-taking	Repetition /social	Self-initiated
VST	Pearson Correlation	0.16**	0.01	-0.01	0.07	-0.03	-0.11*	0.06
	Sig.(2-tailed)	0.001	0.800	0.846	0.143	0.588	0.024	0.196
WAT	Pearson Correlation	0.15**	0.09	0.04	0.08	0.05	-0.06	0.08
	Sig.(2-tailed)	0.003	0.078	0.374	0.094	0.312	0.199	0.126

\* $p < .05$ , \*\* $p < .01$

**Research question 7:** Is there any relationship between participants' reported vocabulary learning strategies and their productive vocabulary knowledge as represented by the lexical indices of their essays?

The relationship between participants' reported vocabulary learning strategies and their productive vocabulary knowledge was investigated using Pearson product-moment correlation coefficient. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity and homoscedasticity. Table 3.19 indicates that there was a significant, positive correlation between learners' lexical diversity and their reported use of self-initiated strategies ( $r = .11, p < .05$ ) and inferencing strategies ( $r = .17, p < .01$ ). There was a significant, but negative correlation between learners' lexical frequency and their reported use of self-initiated strategies ( $r = -.10, p < .05$ ). In addition, a significant, but negative correlation existed between semantic co-referentiality and learners' reported use of association and imagination strategies ( $r = -.12, p < .05$ ). The word meaningfulness significantly, but negatively correlated with both learners' reported use of self-initiated strategies ( $r = -.13, p < .01$ ) and association and imagination strategies ( $r = -.11, p < .05$ ).

Word imaginability also significantly, but negatively correlated with learners' reported use of self-initiated strategies ( $r = -.11, p < .05$ ). Finally, word familiarity was significantly, but negatively correlated with their reported use of self-initiated strategies ( $r = -.14, p < .01$ ) and inferencing strategies ( $r = -.14, p < .01$ ). No significant correlations were found between participants' productive VK and other reported vocabulary learning strategies.

**Table 3.19: Pearson product-moment correlations between participants' reported vocabulary learning strategies and their productive vocabulary knowledge ( $N = 419$ )**

		Self-initiated	Dictionary	Production-based	Note-taking	Association and imagination	Inferencing	Social/repetition
Lexical diversity	Pearson Correlation	0.11*	0.04	0.07	0.08	0.05	0.17*	0.00
	Sig. (2-tailed)	0.024	0.458	0.139	0.095	0.313	0.000	0.955
Lexical frequency	Pearson Correlation	-0.10*	-0.06	-0.09	0.00	-0.05	-0.09	0.03
	Sig. (2-tailed)	0.037	0.265	0.060	0.962	0.272	0.058	0.483
Hypernymy	Pearson Correlation	-0.03	0.06	0.00	0.00	-0.04	0.03	0.02
	Sig. (2-tailed)	0.591	0.212	0.929	0.998	0.428	0.546	0.635
Polysemy	Pearson Correlation	-0.02	-0.04	-0.06	0.02	-0.09	-0.01	0.05
	Sig. (2-tailed)	0.674	0.468	0.248	0.749	0.074	0.861	0.317
Semantic co-referentiality	Pearson Correlation	-0.04	-0.07	-0.08	-0.03	-0.12*	-0.08	0.03
	Sig. (2-tailed)	0.429	0.133	0.096	0.537	0.017	0.094	0.497
Word meaningfulness	Pearson Correlation	-0.13*	-0.01	-0.06	-0.08	-0.11*	-0.07	-0.07
	Sig. (2-tailed)	0.009	0.805	0.239	0.088	0.029	0.160	0.136
Word concreteness	Pearson Correlation	-0.10	0.01	-0.05	-0.06	-0.05	-0.03	-0.02
	Sig. (2-tailed)	0.052	0.918	0.266	0.254	0.313	0.551	0.761
Word imaginability	Pearson Correlation	-0.11*	-0.01	-0.06	-0.07	-0.05	-0.03	-0.04
	Sig. (2-tailed)	0.026	0.881	0.236	0.172	0.321	0.616	0.371

Word familiarity	Pearson Correlation	-0.14*	-0.09	-0.09	-0.03	-0.02	-0.14*	-0.02
	Sig. (2-tailed)	0.006	0.064	0.055	0.555	0.725	0.005	0.727

\* $p < .05$ , \*\* $p < .01$

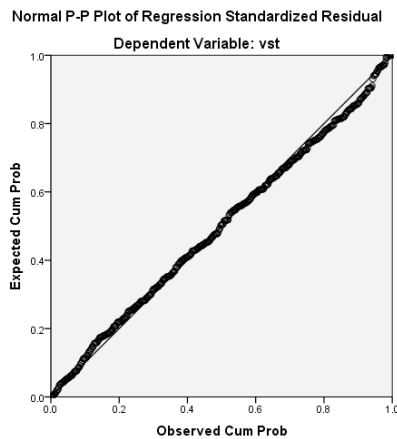
**Research question 8:** What is the contribution of individual and overall variables (gender, discipline, proficiency levels and vocabulary learning strategies) to the participants' receptive vocabulary knowledge?

In order to discover which variables could best predict the participants' receptive vocabulary knowledge (VST and WAT), multiple linear regression was performed in SPSS 21. The variables of gender, discipline, proficiency levels and vocabulary learning strategies, were considered as predictor variables and the participants' receptive vocabulary knowledge (VST and WAT) were taken as the criterion variable respectively.

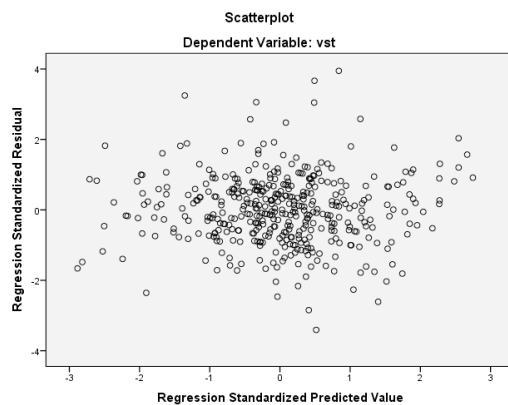
#### Standard multiple regression for Vocabulary Size Test (VST)

Before running the multiple regression, underlying assumptions were checked. Firstly, multicollinearity was checked. Multicollinearity exists when the predictor variables are highly correlated with a correlation coefficient of .70 or more. In order to examine this assumption, the table labeled "Correlations" in SPSS was examined. The table shows that while there were some relations between the criterion variable and the predictor variables, the correlations between predictor variables were all below .70 meaning that there was no multicollinearity. As such, all the variables were retained and included in the analysis. In addition, the Tolerance and VIF values were checked for confirming the above result. If the Tolerance value is very small, less than .10, it suggests the possibility of multicollinearity. While if VIF value is above 10, it would indicate multicollinearity as well (Pallant, 2013). The Tolerance and VIF values of the data in the present study show that the assumption of multicollinearity was not violated.

Then, the "Normal P-P Plot" was examined for normality. Figure 3.4 shows the points lie in a reasonably straight diagonal line from the bottom left to top right, which indicates there is no major deviation from normality. The scatterplots of the standard residuals could provide information of normality, linearity, and homoscedasticity. In the present study (see Figure 3.5), the residuals are roughly rectangularly distributed, with most of the scores concentrated in the center. This distribution of residuals suggests that the assumptions of normality, linearity, and homoscedasticity have not been violated.



**Figure 3.4: Normal P-P plot of regression standardized residuals**



**Figure 3.5: Scatterplots of the standard residuals**

Furthermore, outliers were checked for both criterion and predictor variables by inspecting the Boxplots and a few of the outliers were found in these variables. In order to decide whether these outliers should be removed, I compared 5% trimmed mean values with mean values of these variables and found that the two mean values were quite similar for all these variables. This means that these values are not too different from the remaining distribution, and thus I retained these cases in the data file (Pallant, 2013). The other way of detecting outliers is by comparing the Mahalanobis Distances values with the critical value (29.59). The maximum value of Mahal. Distance was 38.108 in my study, exceeding the critical value. It suggests that there should be some outliers in the variables. I sorted the MAH-1 variable in a descending order to find which values of cases were above the critical value (29.59). There were two cases with Mahal. Distance values above the critical value (ID = 21, ID = 274). Therefore, these two outliers were excluded from the analysis.

The other information in the output concerning unusual cases is obtained with the assistance of “Casewise Diagnostics” in SPSS. This provides information about cases with standardized residual values above 3.0 or below -3.0. In this sample, I have found six cases (ID = 139, ID = 175, ID = 274, ID = 316, ID = 331, ID = 390) with a residual above 3.0 or below -3.0. To check whether the strange cases are having any undue influence on the results for my model as a whole, “Cook’s Distance” values were examined. According to Tabachnick and Fidell (2013), cases with values larger than 1 are a potential problem. In my example, the maximum value for Cook’s Distance was .156, suggesting no major problems. Therefore, I retained these cases in my analysis.

After removing the outliers, the multiple regression analysis was performed again. R Square explained how much of the variance in the criterion variable (VST scores) is explained by the model (including gender, discipline, proficiency, self-initiated strategies, repetition/social strategies, dictionary strategies, production-based strategies, note-taking strategies, association and imagination strategies, and inferencing strategies). In the present study, the value of R Square was .132. This means that my model explained 13.2% of the variance in participants’ vocabulary size test scores. In other words, the predictor variables were able to account for only 13.2% of the variance in VST (see Table 3.20).

**Table 3.20: Model summary of multiple regression of all predictor variables on VST (N = 419)**

Model	R	R Square	Adjusted R Square	<i>p</i>
	0.363	0.132	0.111	0.000**

\*\* $p < .001$

In the present study, I was concerned with the contribution of each predictor variable. To do so, I used the standardized coefficient beta values. Based on George and Mallery (2003), the larger beta value means the stronger effect of the predictor variable on the criterion variable. In the present study, proficiency level showed the largest beta value (.21), which means this variable could make the strongest unique contribution to explain the criterion variable VST. Gender (Beta = .14), repetition/social strategies (Beta = - .14), inferencing strategies (Beta = .13), and discipline (Beta = .12) also made unique contributions to the prediction of the criterion variable VST (see Table 3.21).

**Table 3.21: Summary of correlation coefficients of all predictor variables on VST**

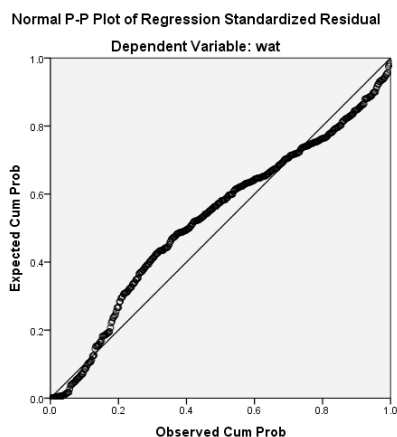
Predictor variable	Beta	<i>t</i>	<i>p</i>
Proficiency	0.21	4.28	0.000**
Gender	0.14	2.72	0.007*
Repetition/ Social strategies	-0.14	-2.78	0.006*
Inferencing strategies	0.13	2.45	0.015*
Discipline	0.12	2.36	0.019*
Production-based strategies	0.06	0.96	0.338
Note-taking strategies	-0.02	-0.31	0.758
Self-initiated strategies	-0.03	-0.53	0.598
Dictionary strategies	-0.01	-0.19	0.853
Association and imagination strategies	-0.01	-0.21	0.831

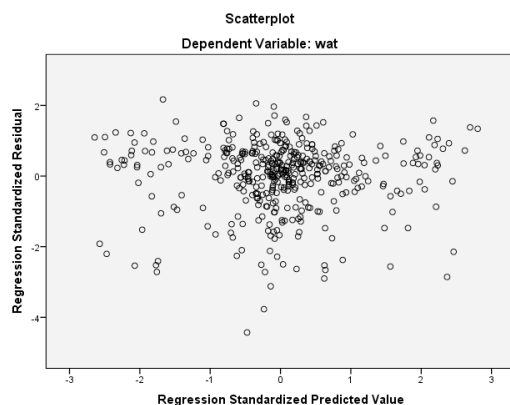
\* $p < .05$ , \*\* $p < .001$

### Standard multiple regression for Word Associates Test (WAT)

The underlying assumptions were checked again before conducting the multiple regression for WAT. The results of the correlation analysis show that there were some relations between the criterion variable and the predictor variables. Additionally, the correlations between the predictor variables which are the same as those used in the above section, were all below .70. In addition, the Tolerance and VIF values also confirmed that the assumption of multicollinearity was not violated.

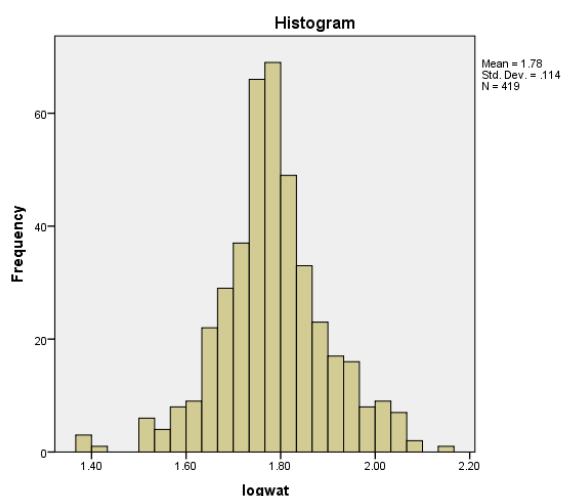
Then, I checked the “Normal P-P Plot” for normality, as presented in Figure 3.6. The points lie in a relatively curved diagonal line from the bottom left to top right, which indicates that there are some deviations from normality. The shape of “Scatterplot of the standardized residuals” also suggests that some of the assumptions of normality, linearity, and homoscedasticity might have been violated (see Figure 3.7). Therefore it is necessary to further observe the distributions of the variables.

**Figure 3.6: Normal P-P plot of regression standardized residuals**



**Figure 3.7: Scatterplots of the standard residuals**

SPSS EXPLORE was used to examine the distributions of the variables. The results show that the criterion variable WAT has violated the assumption of normal distribution (Skewness =  $-.995$ , Kurtosis =  $1.763$ ), which explains, at least in part, the problems in the residuals scatterplot. “Reflect and logarithm” was applied as an appropriate method to transform WAT, and the transformed distributions were checked once again for Skewness and Kurtosis (Skewness =  $.077$ , Kurtosis =  $.960$ ). Figure 3.8 shows that the transformation of WAT reduced both Skewness and Kurtosis values.

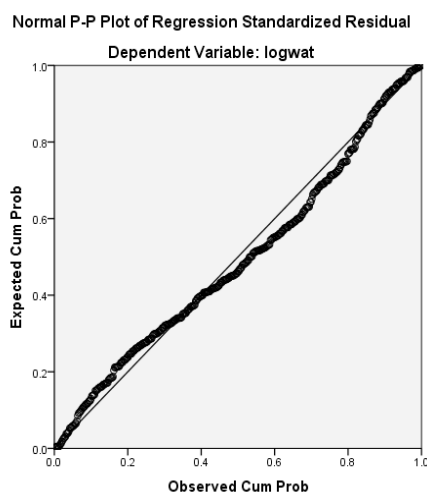


**Figure 3.8: The distribution of the transformed WAT**

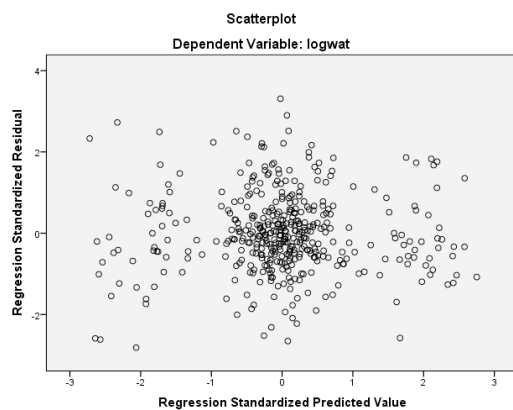
By examining the Boxplot, several outliers of WAT were detected. However, the 5% trimmed mean value of WAT was quite similar to its mean value implying that these cases were not too different from the remaining distribution. For this reason, they were retained in the data file (Pallant, 2013). Then by comparing the Mahal. Distance values with the critical value (29.59), two outliers were detected (ID = 21, ID = 274) and were thus excluded from the analysis. The outliers of each predictor variable have already been checked in the above section. The other information in the output, as mentioned in the

above section, concerning unusual cases is obtained with the assistance of “Casewise Diagnostics” in SPSS. In this sample, I found one case (ID = 117) with a residual above 3.0 or below -3.0. Then “Cook’s Distance” was checked to identify whether the strange case had any undue influence on the result. It shows that the maximum value for “Cook’s Distance” was .053, suggesting no major problems.

After removing the outliers, the standard multiple regression was performed again. The result reveals that although the distributions were still not perfect, both shapes of “Normal P-P Plot” (see Figure 3.9) and “Scatterplot” (see Figure 3.10) have been improved with the transformed criterion variable WAT.



**Figure 3.9: Normal P-P plot of regression standardized residuals**



**Figure 3.10: Scatterplots of the standard residuals**

As Table 3.22 shows, the R Square value was .137. This means that my model explained only 13.7% of the variance in participants' Word Associates Test (WAT) scores based on the predictor variables included in the study.

**Table 3.22: Model summary of multiple regression of all predictor variables on WAT ( $N = 419$ )**

Model	R	R Square	Adjusted R Square	$p$
	0.370	0.137	0.116	0.000**

\*\* $p < .001$

Then, I checked which variables included in the model could best predict the criterion variable (WAT). The results show again that the predictor variable, proficiency, made the strongest unique contribution to predict the dependent variable WAT with the largest beta value (Beta = - .30), followed by inferencing strategies (Beta = - .13) (see Table 3.23).

**Table 3.23: Summary of correlation coefficients of all predictor variables on WAT**

Predictor variable	Beta	$t$	$p$
Proficiency	-0.30	-5.94	0.000**
Inferencing strategies	-0.13	-2.37	0.019*
Gender	-0.04	-0.83	0.410
Discipline	-0.06	-1.16	0.248
Self-initiated strategies	0.03	0.47	0.638
Repetition/Social strategies	0.04	0.76	0.448
Dictionary strategies	0.01	0.16	0.875
Production-based strategies	0.00	0.04	0.971
Note-taking strategies	-0.04	-0.74	0.461
Association and imagination strategies	-0.03	-0.48	0.629

\* $p < .05$ , \*\* $p < .001$

**Research question 9:** What is the contribution of individual and overall variables (gender, discipline, proficiency levels and vocabulary learning strategies) to the participants' productive vocabulary knowledge as represented by the lexical indices of their essays?

The underlying assumptions were checked again before conducting the multiple regression for each lexical index of their essays. The results of the correlation analysis indicate that there were some relations between each criterion variable and the predictor variables. Moreover, the correlations between the predictor variables which are the same as those used in Research question 8 were all below .70. In addition, the Tolerance and VIF values also confirmed that the assumption of multicollinearity was not violated. Then, the "Normal P-P Plot" of each lexical index indicates that the normality was not violated. In

addition, the “Scatterplots of the standard residuals” of each lexical index suggests that the assumptions of normality, linearity, and homoscedasticity were not violated.

Moreover, the criterion variables were checked for outliers and several outliers were detected in each variable by inspecting their Boxplots. However, as illustrated in the Research question 8, the 5% trimmed mean value of each variable was quite similar to its mean value, thus these cases were retained in the data file. The outliers for each predictor variable have been checked in the Research question 8. Then by comparing the Mahal. Distance values with the critical value (29.59), two outliers (ID = 21, ID = 274) were detected and removed from the analysis. Moreover, Cook’s Distance value was checked for each lexical index. The results show that the maximum values were all below 1, suggesting the strange cases did not have major influence on the results.

After removing the outliers, the standard multiple regression was performed again for each lexical index. Table 3.24 indicates that my model explained 11.8% of the variance in lexical diversity, 7% of the variance in word familiarity, 6.7% of the variance in lexical frequency, and 4.9% of the variance in word meaningfulness.

**Table 3.24: Model summary of multiple regressions of all predictor variables on the lexical indices ( $N = 419$ )**

Criterion variable	R	R Square	Adjusted R Square	$p$
Lexical diversity	0.344	0.118	0.096	0.000**
Lexical frequency	0.259	0.067	0.044	0.001**
Hypernymy	0.127	0.016	-0.008	0.759
Polysemy	0.145	0.021	-0.003	0.555
Semantic co-referentiality	0.212	0.045	0.022	0.041
Word meaningfulness	0.221	0.049	0.026	0.024*
Word concreteness	0.168	0.028	0.004	0.305
Word imaginability	0.165	0.027	0.003	0.338
Word familiarity	0.264	0.070	0.047	0.001**

\* $p < .05$  \*\* $p < .001$

Then, I checked which variables in the model could best predict these four criterion variables. Table 3.25 shows that proficiency levels made the strongest unique contribution to predict the criterion variable lexical diversity with the largest beta value (Beta = .27), followed by inferencing strategies (Beta = .14).

**Table 3.25: Summary of correlation coefficients of all predictor variables on lexical diversity**

Predictor variable	Beta	<i>t</i>	<i>p</i>
Gender	0.07	1.30	0.195
Discipline	-0.01	-0.09	0.928
Proficiency	0.27	5.45	0.000**
Self-initiated strategies	0.04	0.70	0.483
Repetition/Social strategies	-0.02	-0.46	0.646
Dictionary strategies	-0.05	-0.79	0.433
Production-based strategies	-0.03	-0.39	0.698
Note-taking strategies	0.07	1.15	0.251
Association and imagination strategies	-0.03	-0.49	0.625
Inferencing strategies	0.14	2.63	0.009*

\* $p < .01$ , \*\* $p < .001$

The following table presents the result for lexical frequency. It shows that only proficiency levels made a unique contribution to the prediction of the criterion variable lexical frequency (Beta = - .20) (see Table 3.26).

**Table 3.26: Summary of correlation coefficients of all predictor variables on lexical frequency**

Predictor variable	Beta	<i>t</i>	<i>p</i>
Gender	-0.08	-1.45	0.147
Discipline	0.00	-0.04	0.967
Proficiency	-0.20	-3.85	0.000**
Self-initiated strategies	-0.07	-1.07	0.286
Repetition/Social strategies	0.05	0.95	0.345
Dictionary strategies	-0.01	-0.16	0.875
Production-based strategies	-0.05	-0.76	0.447
Note-taking strategies	0.08	1.25	0.212
Association and imagination strategies	-0.02	-0.26	0.799
Inferencing strategies	-0.02	-0.37	0.713

\*\* $p < .001$

Table 3.27 shows that only proficiency levels made a unique contribution to predict the criterion variable word meaningfulness, with the largest beta value (Beta = - .12).

**Table 3.27: Summary of correlation coefficients of all predictor variables on word meaningfulness**

Predictor variable	Beta	<i>t</i>	<i>p</i>
Gender	0.03	0.47	0.642
Discipline	-0.03	-0.63	0.530
Proficiency	-0.12	-2.37	0.018*
Self-initiated strategies	-0.11	-1.74	0.083
Repetition/Social strategies	-0.05	-1.02	0.310
Dictionary strategies	0.10	1.52	0.130

Production-based strategies	0.05	0.72	0.475
Note-taking strategies	-0.08	-1.24	0.215
Association and imagination strategies	-0.08	-1.31	0.192
Inferencing strategies	-0.02	-0.38	0.707

\* $p < .05$

For the criterion variable word familiarity, Table 3.28 indicates that only proficiency levels made a unique contribution to explain the variances in word familiarity, with the largest beta value (Beta = - .17).

**Table 3.28: Summary of correlation coefficients of all predictor variables on word familiarity**

Predictor variable	Beta	<i>t</i>	<i>p</i>
Gender	-0.07	-1.23	0.219
Discipline	-0.02	-0.39	0.694
Proficiency	-0.17	-3.36	0.001**
Self-initiated strategies	-0.11	-1.75	0.081
Repetition/social strategies	0.00	0.01	0.989
Dictionary strategies	-0.02	-0.36	0.722
Production-based strategies	-0.01	-0.22	0.829
Note-taking strategies	0.04	0.60	0.551
Association and imagination strategies	0.07	1.22	0.222
Inferencing strategies	-0.09	-1.57	0.118

\*\* $p < .01$

### 3.4 Summary of the Quantitative Study

This chapter presented the results of the quantitative analyses in order to (a) explore the participants' vocabulary knowledge and vocabulary learning strategies, (b) to identify the potential influence of the variables such as gender, discipline, and proficiency level upon their vocabulary knowledge and vocabulary learning strategies, and (c) to examine the contribution of the variables (gender, discipline, proficiency levels and vocabulary learning strategies) to the participants' vocabulary knowledge. My analyses have produced the following findings:

1) The average vocabulary size of the participants measured by VST was 6,494 English word families, and the average score of depth of vocabulary knowledge measured by WAT was 98.07, amounting to 61.29 per cent of the maximum possible score of WAT.

2) Significant gender difference was found in the participant's scores of VST. The results show that the vocabulary size of the female students ( $M = 65.71$ ) was larger than that of the male students ( $M = 63.13$ ). It is also found that the science and engineering

students ( $VST = 66.07$ ,  $WAT = 100.86$ ) performed significantly better than their counterparts in humanities and social sciences ( $VST = 63.37$ ,  $WAT = 96.25$ ). Moreover, the significant differences were also found among the three proficiency groups both in their scores of VST and WAT showing that the VST of the high proficiency students ( $M = 71.83$ ), was larger than that of the intermediate proficiency group students ( $M = 64.29$ ) and the low proficiency group students ( $M = 59.44$ ). Moreover, the high proficiency students ( $M = 109.44$ ) also performed better in their scores of WAT than the intermediate proficiency students ( $M = 98.15$ ), and the low proficiency group students ( $M = 88.56$ ).

3) The high proficiency group was significantly different from both the intermediate and the low proficiency groups in terms of lexical diversity, lexical frequency and lexical familiarity. However, the high proficiency group was only significantly different from the low proficiency group in lexical meaningfulness. Moreover, no statistically significant difference was found between the students in the two broad disciplines and between the male and female students in terms of their productive vocabulary knowledge as represented by Coh-Metrix indices.

4) The exploratory factor analysis identified seven factors underlying the VLS questionnaire. These factors were: inferencing strategies, association and imagination strategies, note-taking strategies, production-based strategies, dictionary strategies, repetition and social strategies, and self-initiated strategies. The descriptive results reveal that inferencing strategies were reportedly most commonly used by the participants, which was followed by using dictionary and repetition and social strategies. Note-taking strategies were reportedly the least frequently used by the participants along with production-based strategies and association and imagination strategies.

5) Statistically significant differences were found in the overall reported vocabulary learning strategy use, between the male and female students, between the students in the two broad disciplines, and also among different proficiency groups. The significant differences between the male and female students were found in the note-taking strategies. The female students reported using note-taking strategies more frequently than the male students. In addition, the significant difference in the reported use of note-taking strategies was found between the students in the two broad disciplines. The humanities and social sciences students reported using note-taking strategies more frequently than the science and engineering students. Finally, there was a significant difference between the high proficiency students and the low proficiency students in terms of their reported use of

inferencing strategies. That is to say, students in the high proficiency group reported adopting inferencing strategies more frequently than did those in the low proficiency group.

6) There was a significant and positive correlation between learners' vocabulary size (VST) and their reported use of inferencing strategies. However, a strong but negative correlation was found between learners' vocabulary size and their reported use of repetition and social strategies. In addition, there was a strong and positive correlation between learners' depth of vocabulary knowledge (WAT) and their reported use of inferencing strategies.

7) There was a significant and positive correlation between learners' lexical diversity and their reported use of self-initiated strategies and inferencing strategies. There was a significant but negative correlation between learners' lexical frequency and their reported use of self-initiated strategies. In addition, a significant but negative correlation was found between the semantic co-referentiality of learners' essays and their reported use of association and imagination strategies. The word meaningfulness significantly but negatively correlated with both learners' reported use of self-initiated strategies and association and imagination strategies. Word imaginability significantly but negatively correlated with learners' reported use of self-initiated strategies. Finally, word familiarity significantly but negatively correlated with their reported use of self-initiated strategies and inferencing strategies.

8) The regression model (with gender, discipline, proficiency levels and vocabulary learning strategies as predictor variables) only explained 13.2% of the variance in the participants' vocabulary size (VST), to which the proficiency levels made the strongest unique contribution, followed by gender, repetition and social strategies, inferencing strategies and discipline. The regression model explained 13.7% of the variance in the participants' depth of vocabulary knowledge (WAT) and proficiency levels still made the strongest contribution, only followed by inferencing strategies.

9) Finally, for the participants' productive vocabulary knowledge, my model could explain 11.8% of the variance in lexical diversity, 7% of the variance in word familiarity, 6.7% of the variance in lexical frequency, and 4.9% of the variance in word meaningfulness. Proficiency levels made the strongest unique contribution to the lexical diversity, followed by inferencing strategies. While for word familiarity, word meaningfulness, and lexical frequency, only proficiency levels made unique contributions to predict these variables.

In Chapter 5, I will discuss the quantitative findings and synthesize them with qualitative findings. Before that, in Chapter 4, I will present the qualitative findings.



## **Chapter 4 Qualitative Study**

I have discussed the method and presented the results of the quantitative study in the previous chapter. In this chapter, I will focus on the qualitative interview study. The main purpose of interviewing 15 participating Chinese EFL students is to achieve a better and deeper understanding of their perspectives on vocabulary knowledge (VK), vocabulary learning strategies (VLSs), and the interplay between VK and VLSs. Accordingly, the analysis of this set of interview data aims to seek any possible convergence and corroboration with the results drawn on the analysis of the quantitative data, or to shed more light on the quantitative findings where such convergence is not found. To do so, this chapter first presents the qualitative research question, then describes the methods employed to collect and analyze qualitative interview data, and finally presents the results drawn from the analysis of interview data.

### **4.1 Research Question**

In accord with the aim of the interview study in this research project, the following research question has been raised:

What are the Chinese EFL students' perspectives on vocabulary knowledge and vocabulary learning strategies, and how do they think vocabulary learning strategies affect their vocabulary knowledge?

### **4.2 Methods**

This section presents research methods that I have followed in conducting the qualitative research, including the selection of participants, data collection procedures and methods for analyzing data.

#### **4.2.1 Participants**

The interviewees were selected from the participant pool of the quantitative study on a voluntary basis. In the selection process, I first explained the purpose of the interview and what participants were supposed to do during the interview to all the participants of the quantitative study. I also emphasized that the interview data would be kept confidentially in line with the ethics requirements of my research and that their participation was on a

voluntary basis. Consequently, 36 students among the four universities expressed their willingness to participate in the follow-up interviews. Of these 36 volunteers from the four universities, 15 participants were selected mainly based on their available time for the interviews, their proficiency levels and their universities. For those volunteers who were not selected in the follow-up interviews, I provided them my contact details and invited them to discuss with me their concerns over English vocabulary learning. Each university was represented by three or four interviewees and each proficiency level was well represented. Consequently, five of these 15 participants were from the high proficiency group, five from the intermediate proficiency group, and another five from the low proficiency group. Seven participants were female while eight were male. Eight of them majored in humanities and social sciences, while seven majored in science and engineering. After the selection of participants was finalized, each of them was given a code to replace their real name in an effort to preserve their anonymity. Accordingly, they were referred to in this study as H1 (F), H2 (M), H3 (M), H4 (F), H5 (F) for those in the high proficiency group, M1 (M), M2 (M), M3 (F), M4 (F), M5 (M) for intermediate proficiency students and L1 (M), L2 (F), L3 (M), L4 (M), L5 (F) for low proficiency students.

#### **4.2.2 Data Collection Procedures**

Semi-structured interviews were used to interview these 15 voluntary interviewees at the four universities. Semi-structured interviews were adopted mainly because it, on the one hand, allows the researcher to have some control over the collection of data, while on the other, participants can provide some additional information extending beyond the interview questions (Creswell, 2009). In other words, the questions were raised to elicit but not to restrict the interviewees' responses. All interviews were guided by an interview protocol (see Appendix G). The questions on the interview protocol were primarily related to the significance and difficulty of English vocabulary learning, participants' use of vocabulary learning strategies and their perspectives about vocabulary knowledge. Open-ended questions as presented in the interview protocol were raised during the interviews so as to enable the interviewees to present and explain their ideas in more detail. The interviews were conducted in available teachers' lounges, and each interview lasted for about 30 minutes. Before implementing interviews, the consent forms written in Chinese were distributed to the participants and they signed the form on a voluntary basis. All the interviews were conducted in Chinese and tape-recorded with participants' permission and for subsequent transcription and analysis.

### **4.2.3 Data Analysis Procedures**

All the interviews were firstly transcribed in Chinese before they were carefully translated into English. The translation of the interviews was later reviewed and checked by another Chinese Ph.D. student in linguistics who is native in Chinese and fluent in English, so as to guarantee the accuracy and reliability of the translation. No serious inconsistencies regarding the translation of the interviews were found between, except for some suggestions on improving a few English expressions, such as tense coherence. The analysis of the interview data was conducted by using the NVivo 10 computer software, which is a useful software program in assisting researchers to organize and analyze various types of data, such as audio, video, and word documents etc. in a systematic and comprehensive manner. This program starts creating a new project to save all the documents that are related to the qualitative data. Then, the qualitative data needs to be uploaded into NVivo 10. In the present study, the audio files of interviews were transcribed and then translated into English as a word document. This word document of the interview text was then uploaded into NVivo 10. After uploading the interview file, the transcripts of interviews were examined firstly by the researcher to have an overall sense of the data and some of the initial insights and observations were noted in memos which is another facility in NVivo program. Then, the coding scheme was developed following the iterative process by which the transcripts of interviews were examined several times to identify the emerging categories and sub-categories (Bryman, 2008). Accordingly, by reading, classifying, and interpreting the data, the main categories and subcategories were established. These emerging categories and sub-categories were created as Nodes in NVivo 10.

The coding scheme displays a hierarchical system consisting of three tiers of codes. Tier 1 presents the general constructs addressed in research questions and interview questions. These constructs are organized into four categories: (1) Significance of learning English vocabulary; (2) Difficulties of learning English vocabulary; (3) Current vocabulary learning strategies (4) Perspectives about vocabulary knowledge (e.g., size and depth). The second and third tiers elaborate further on the first tier (see Appendix H for the coding scheme).

## 4.3 Results

### 4.3.1 Significance of Learning English Vocabulary

Fourteen participants (93.33%) reported that English vocabulary learning was quite important for them; however, only one participant (L1, low proficiency student) said that it was not quite important. The responses of these fourteen students could be categorized into four aspects: English vocabulary is quite important for reading comprehension, for writing, for speaking and for listening comprehension (see Table 4.1).

**Table 4.1: The significance of learning English vocabulary for the four English communication skills**

	High prof. group					Intermediate prof. group					Low prof. group				
	H1	H2	H3	H4	H5	M1	M2	M3	M4	M5	L1	L2	L3	L4	L5
Reading comprehension	√	√	√		√	√		√	√	√		√	√	√	√
Listening comprehension	√		√		√	√	√		√	√		√	√	√	√
Writing	√		√	√	√		√			√		√	√	√	
Speaking			√	√	√	√	√	√				√		√	√

Table 4.1 indicates that English vocabulary was considered quite important for English reading comprehension by most of the participants, while a large portion of the participants also mentioned its significance for English listening comprehension, writing and speaking. Of these participants, some students (H3, H5, L2) claimed that learning English vocabulary was the foundation for all of the four English communication skills. For instance, H3 stated:

I think English vocabulary learning is quite important for me. It is the foundation for English reading comprehension, listening comprehension, speaking, and writing. Without English vocabulary, you can't understand, speak or write anything in English.

The rest of the students considered English vocabulary learning very important only for certain English communication skills and their reasons differed from each other. For instance, some participants (H1, M1, M4, M5, L3, L5) claimed that English vocabulary learning was quite important for reading comprehension, listening comprehension or writing in relation to the English test, as M4 said,

S: Learning English words is very important for me. It is quite important for English reading comprehension and listening comprehension.

R: Why it is very important only for reading and listening comprehension? Do you think it is also important for other English skills?

S: Maybe, but in my case, I think it is quite important for reading and listening because these two parts take a large proportion in the English test. I seldom write anything in English or communicate with others in English.

Likewise, H1 said:

English vocabulary is quite important for me because there are lots of new English words in the English test, such as CET-4. If you don't know these new words in the English test, you could not achieve high scores in the listening comprehension, reading comprehension or the writing task of the test.

In the meantime, two participants (H4 and M2) stated their reasons relevant to their weakness in English communication skills. For instance, H4 considered English vocabulary learning quite important only for speaking and writing, as she said,

I think English vocabulary is very important for English speaking, such as daily communication and English writing. Both of my oral English and writing are poor probably because of my limited English vocabulary size. I think if my vocabulary size is large enough, I can write and speak freely in English.

Similarly, some of them (M1, M2, M3, L5) also considered English vocabulary learning very important for English communication (speaking) because it was necessary for them to know a variety of English words to express their meaning. In addition, two participants (H2 and M3) specifically pointed out that English vocabulary learning was very important for reading comprehension as they would meet many new English words in the reading, thus it would be quite helpful for them if they had a large vocabulary size.

By contrast, only one student (L1, low proficiency student) reported that English vocabulary learning was not quite important. He did not consider learning English vocabulary very important as he thought even for passing CET-4 there was no need to learn many words.

I think English vocabulary learning is not quite important because in my opinion, a certain vocabulary size would be enough for passing the CET-4 test, so it does not make much sense to have a very large English vocabulary size.

Overall, my research participants from the three proficiency levels shared the similar views on the significance of learning English vocabulary. Almost all of them highlighted the important role of English vocabulary for the development of their English communication skills and some participants also made a connection between their vocabulary learning and English communication skills, English test, and their weaknesses. However, only one low proficiency student questioned the importance of expanding one's vocabulary size on the view that a limited vocabulary size would be enough to pass tests such as CET-4.

### 4.3.2 Difficulties of Learning English Vocabulary

Table 4.2 presents the difficulties that the participants encountered when learning English vocabulary.

**Table 4.2: Difficulties of learning English vocabulary**

	High prof. group					Intermediate prof. group					Low prof. group				
	H1	H2	H3	H4	H5	M1	M2	M3	M4	M5	L1	L2	L3	L4	L5
Hard to remember new English words for a long time.			√	√		√	√				√		√	√	√
Having no efficient methods to remember new English words.			√					√			√			√	√
Difficult to remember the long words, the words with irregular spellings and the infrequent words	√				√				√						
Hard to learn the depth of vocabulary knowledge (various meanings of a word and how to use it)										√		√			
Too many English words to remember										√					
Some English words you can recognize, but you cannot produce them											√				
No problem of memorizing words		√													

When talking about the difficulties of learning English vocabulary, most participants found learning English vocabulary difficult. Among these students, some of them (H3, H4, M1, M2, L1, L3, L4, L5) emphasized that it was hard to remember new English words for

a long time. It seems that they have suffered a common problem of maintaining the long-term memory of the learnt words, as L3 mentioned:

It is hard for me to remember new English words for a long time. When I try to remember a new word, I always forget it very quickly. I think I should keep reviewing the learnt words frequently to memorize them, but I cannot persist in reviewing because sometimes I am a little bit lazy.

It should be noted that, akin to L3, some of them recognized that the reason why they could not maintain the learnt words in their long-term memory was that they did not review them regularly. While another two students considered their use of the rote-memory method as the primary reason. For instance, M1 said:

I used to remember new English words by using rote-memory, but it was hard for me to maintain the long-term memory of the learnt words by adopting this method. I always forget the learnt words very quickly. Now I am trying to remember new English words by reading some interesting English learning materials. This method seems better for me.

In addition, five interviewees (H3, M3, L1, L4, L5) mentioned that they had no efficient methods to remember new English words. Three of them found the strategy of rote memory boring and inefficient, but they had no other better methods at their disposal to remember new English words. For instance, H3 said:

I think I have no good methods to remember new English words. In my case, it is hard to remember new words by using rote memory, especially to remember the learnt words for a long time. I find remembering new words in English extensive reading quite helpful for me, however, it is quite time-consuming in the meanwhile.

In the meanwhile, three students (H1, H5, M4) specifically mentioned that it was difficult to remember the long words, the words with irregular spellings and the infrequent words. This view was clearly reflected in the response like:

I always remember English words by segmenting their pronunciations into syllables. I find that it is hard for me to remember the English words with irregular pronunciations. In this case, I cannot use this method any more. In addition, infrequent words are hard for me to remember too, because I don't have any chances to use them. (M4)

Moreover, two students (L2 and M5) commented that it was hard to learn the depth of vocabulary knowledge. M5 mentioned that it was difficult to retain different meanings of a word. L2 blamed her university English teachers for their poor teaching methods. It seems that after graduating from high school, L2 did not have any opportunity to learn how to use English words.

I think I have difficulties in learning how to use English words. Sometimes, I only know the meaning of a word, but I don't know how to use it properly. It seems that I don't have any chances to learn how to use new words after high school. In my high school, my English teacher always gave us an example sentence of a new English word along

with its collocations and some grammatical information. However, my current university English teacher seldom illustrates English words in a very detailed way. (L2)

In addition, one student complained that there are too many English words to remember and sometimes he got confused with some words with similar spelling:

I think English words are too many for me to remember. It is quite hard for me to remember such a big amount of English words and it is still hard to distinguish some English words from those with similar spellings. Sometimes, I confuse some English words with those spelt or pronounced in similar ways. (M5)

Finally, one student described another difficulty for him as some English words you can recognize, but you cannot produce them:

I find that for some English words, I can only know their meanings in the reading, but when I want to use them in my writing or speaking, I just can't recall their spellings. (L1)

However, it was quite interesting to find that only one interviewee (H2) reported that he had no problem of memorizing new English words. This student believed that English vocabulary should be learned through extensive reading and new English words could be remembered when encountering them many times.

I think English vocabulary should be learned through the English extensive reading. For example, I like reading some English novels and some English academic articles. When I encounter the English words several times in the reading, I could remember them well. (H2)

As mentioned above, almost all of the participants have encountered various kinds of difficulties in learning English vocabulary. Most of these difficulties were similar among the three proficiency level students. However, with a close scrutiny, some differences were detected. Firstly, in general the low proficiency students mentioned more difficulties in learning English vocabulary than did both high and intermediate proficiency students. Secondly, compared to the intermediate and high proficiency students, more low proficiency students claimed that it was hard to remember new English words for a long time and they had no efficient methods to remember new English words. Thirdly, the difficulty indicated only by the high and intermediate proficiency students was remembering long words, the words with irregular spellings and the infrequent words, which was not mentioned by any low proficiency student. Finally, the low and intermediate proficiency students seemed to suffer more problems on learning the depth of vocabulary knowledge, that is, various meanings of a word and how to use a word, than did the high proficiency students.

### 4.3.3 Current Vocabulary Learning Strategies

This theme consists of two sub-categories: *use of vocabulary learning strategies and perceived usefulness, and sources of vocabulary learning strategies*. For the sake of clarity, comparisons among the high-, mid- and low- proficiency students were presented based on each sub-category.

#### 4.3.3.1 Use of Vocabulary Learning Strategies and Perceived Usefulness

Table 4.3 summarizes the vocabulary learning strategies reported by the three proficiency groups in the interviews.

**Table 4.3: The reported vocabulary learning strategies**

Category	Reported strategies	High prof.	Inter prof.	Low prof.
Self-initiated	Watching English movies or TV series	√	√	√
	Watching English online video	√		
	Reading English newspapers combined with checking the dictionary and using inferencing strategies	√		
	Reading English novels with the Chinese translation	√		√
	Reading English academic articles combined with checking the dictionary and using inferencing strategies	√	√	
	Listening to English radio program on the English website combined with checking some key words in the dictionary	√		
	Listening to English songs combined with checking some new words in the lyrics in the dictionary	√	√	√
	Using English vocabulary learning software on the cell phone (Youdao)	√		
Inferencing	Reading English newspapers combined with checking the dictionary and using inferencing strategies	√		
	Reading English academic articles combined with checking the dictionary and using inferencing strategies	√	√	
	Doing the reading comprehension of CET-4/6 combined with using inferencing strategies and checking the dictionary	√	√	
Using dictionary	Reading English newspapers combined with checking the dictionary and using inferencing strategies	√		
	Reading English academic articles combined with checking the dictionary and using inferencing strategies	√	√	
	Listening to English radio program on the English website combined with checking some key words in the dictionary	√		
	Listening to the English songs combined with checking some new words in the lyrics in the dictionary	√	√	√
	Doing the reading comprehension of CET-4/6 combined with using inferencing strategies, and checking the dictionary	√	√	
	Doing the reading comprehension of CET-4 combined with checking the dictionary			√
Using context to remember new words	Remembering new words on vocabulary book of CET-4/6 (using associative methods, example sentences, collocations and rote memory)	√		

	Remembering new words in the context of English text	√	√	√
Note-taking	Learning some new words in English classes by using note-taking method	√		√
Production-based	Trying to find opportunities to communicate with others in English	√	√	
Association	Connecting the pronunciations of some English words with the pronunciations of some Chinese words	√		
	Remembering new words on the vocabulary book of CET-4/6 (using associative methods, example sentences, collocations and rote memory)	√		
	Using affixes to remember new words			√
Repetition	Remembering new words on vocabulary book of CET-4/6 by using rote memory	√	√	√
	Remembering the word list of the English textbook out of the context			√
	Remembering some new words on the English textbook by using rote memory before the English class		√	
	Making a list of new words to remember them		√	
	Remembering some English passages by using rote memory		√	
	Learning some new words on the English textbook by doing dictations in English classes		√	

### High proficiency students

The strategies reported by the high proficiency students for learning English vocabulary were quite diversified, ranging from self-initiated strategies, such as watching English movies, listening to English radios, to inferencing strategies, using dictionary, note-taking, social, association strategies, using context and repetition strategies.

The interview data indicate that the high proficiency students reported using a wider range of self-initiated strategies than did both intermediate and low proficiency students. Watching English movies or TV series was the most popular strategy among the high proficiency students. It was noteworthy that all of the five high proficiency students tended to focus on vocabulary learning when they watching English movies or TV series, for example, they all reported they would pay attention to the meanings of new words when watching English movies or TV series. Three of them (H1, H2, H3) also emphasized that they would pay their attention to how words were used when watching English movies or TV series. For instance, H2 stated:

I like watching my favorite English TV series repeatedly. One of my favorite is Friends and I have watched it many times. I sometimes pay my attention to the subtitles so as to learn some new English words including their meanings and usages. Now, I even could repeat some lines of this TV series.

In addition to watching English movies and TV series, watching English online videos, such as TED was also mentioned by H5. She illustrated her special method in a very detailed way:

When I watch TED videos, I always watch them twice. For the first time, I focus on the Chinese subtitle, while for the second time, I pay my attention to the English subtitle. I think the first time of watching could help me understand the talk, while the second time could help me learn some new English words.

Concerning the strategies related to reading, two interviewees (H1 and H4) talked about reading English newspapers in combination with checking the dictionary and inferring the meanings of new words, such as using contexts, affixes or parts of speech to infer a word's meaning. For instance, H4 stated that:

I always read English newspapers on my smart phone, such as CNN and BBC. For some key words, I always look them up in the dictionary, whereas for some unimportant words which I don't think they would affect my understanding of news, I always guess their meanings from the context where they occur or employ some other techniques, such as analyzing their affixes, roots or the parts of speech of words. In the process of reading, sometimes I also pay my attention to the usage of some English words.

Two interviewees (H2 and H3) stated that they read English novels with the Chinese translation and always pay attention to new words. In addition to reading English novels, H2 also favored reading English academic articles combined with checking the dictionary and inferring the meanings of new words.

I like reading some English academic articles in my field which are recommended by my teachers. These articles are quite difficult for me in the beginning, because there are lots of new technical terms which I don't know. I always look up these words in the dictionary. Besides, I sometimes guess the meanings of some words by using their context, or their affixes. (H2)

Regarding the vocabulary learning strategies associated with listening, H3 favored listening to English radio programs on the English website combined with checking some key words in the dictionary, while H5 reported using dictionary to check some new words in the lyrics when listening to English songs. What was the most evident in the data was that the high proficiency students tended to combine using dictionary and inferencing strategies with English extensive reading and to combine using dictionary with some English listening activities. In addition, it seems that most of them pay their attention to vocabulary learning in the processes of doing these English reading and listening activities.

Moreover, H4 mentioned using English vocabulary learning software on the cell phone to remember new words, and provided an example to illustrate this strategy:

You can create a vocabulary notebook in a software and the software will remind you to review these words regularly. I sometimes use Youdao<sup>®</sup> to remember some new English words.

In addition, two of them (H1 and H2) reported that they always did the reading comprehension of CET-4/6 in combination with using inferencing strategies, and checking dictionary.

Moreover, the high proficiency students preferred to use more context or associative methods to remember new English words, rather than only using repetition strategies. For instance, H5 mentioned that sometimes she connected the pronunciations of some English words with those of Chinese words to remember some English words. She thought that this method could help her remember new English words for a long time, but that this method may not be applied to learning those words that present no such pronunciation connections to her native language. Two interviewees (H3 and H4) also reported their techniques to cope with the vocabulary book of CET-4/6, as H3 stated that he always remembered new English words in the example sentences available in the book. H4 tended to use more techniques, as she stated:

When I remember the new English words on the vocabulary book of CET-4 or 6, I always use the associative methods provided by the book, such as semantic relations, words' affixes and roots. In addition, I always read example sentences and some collocations of a word on the book. However, for some simple English words, I only use rote memory to remember them.

Finally, two high proficiency students also reported using social strategy and note-taking strategy, respectively. H5 reported that she learned some English words in English classes by using the note-taking method and H2 stated that he always tried to find opportunities to communicate with others in English, such as some foreigners.

Among the reported VLSs adopted by the high proficiency students, most self-initiated strategies, using dictionary, inferencing strategies, using context, association and note-taking strategy were considered useful for enlarging vocabulary size, whereas repetition and social strategies were regarded as less helpful. In comparison with perceived effective strategies for enlarging vocabulary size, the high proficiency students mentioned a limited number of strategies which were perceived as useful for improving the depth of vocabulary knowledge including several self-initiated strategies, such as paying attention to how words are used when watching English movies or TV series, using context, note-taking strategy and checking dictionary for collocations and example sentences of a word.

### Intermediate proficiency students

In comparison to the high proficiency students, the intermediate proficiency students reported adopting fewer self-initiated strategies. The most common strategy reported by all of the intermediate proficiency students, resembling that reportedly used by the high proficiency students, was watching English movies and TV series. Among these five students, four of them (M1, M2, M3, M4) said they would pay attention to the meanings of new words when watching the movies, while three of them (M1, M2, M3) reported they also paid attention to how to use the words. M1 described his process of watching movies in a detailed way:

At the very beginning, I only focus on the Chinese subtitles, then after twenty to thirty minutes, I shift to focus on the English subtitles. The reason why I use this method is that if at first I focus on the English subtitles, it will be hard for me to understand the movie. Consequently, I will lose my interest. Thus, I always focus on the meaning of the movie first, and then transfer to focus on both the meanings of new words and how to use them.

In contrast with the responses of the other four interviewees in this group, M5 said he did not pay much attention to new English words when watching the movies.

Similar to the high proficiency students, the intermediate proficiency students tended to combine using dictionary and inferencing strategies with English reading and listening activities. Regarding the strategies related to English extensive reading, M1 reported reading English academic articles combined with checking the dictionary and inferring the meanings of new words, which was similar with the response of H2, as M1 said,

Sometimes, I read some academic papers in my field. When I encounter some unknown words, sometimes I look them up in the dictionary directly and sometimes I infer the meanings of the words before checking them in the dictionary. I always use the context to infer a word's meaning and sometimes I also analyze a word's affixes or use my common sense to infer its meaning.

Concerning the strategies related to listening, M2 reported using the same strategy with that reported by H5 which was listening to English songs combined with checking some new words in the lyrics in the dictionary. In addition, two of them also reported that they sometimes did the reading comprehension of CET-4/6 combined with using dictionary and inferencing strategies, such as using context or affixes to infer, which was similar to two of the high proficiency students.

However, in contrast to the high proficiency students, the intermediate proficiency students tended to adopt more repetition strategies. For instance, some intermediate proficiency students (M2, M4, M5) tended to use rote memory to remember new English words on the vocabulary book which was quite different from those strategies reported by

the high proficiency students such as using associative methods or exemplary sentences, as M2 stated:

S: I try to remember the vocabulary book of CET-4/6 by using rote memory. It is boring and inefficient for me. It is hard for me to remember English words by using this method, especially hard to have a long-term memory of the learnt words. I couldn't persist in doing that.

R: So why do you use this method if you don't think it is useful for you.

S: This method is simple to me, so I just keep using repetition though it is boring and not quite helpful for me.

M5 seemed to favor rote memory and held a negative attitude towards the associative methods:

I find using associative methods quite troublesome and time-consuming, while rote memorization is simple and efficient for me. I always use both oral repetition and written repetition to remember new words.

In addition, M2 stated that he always prepared the English classes and remembered some new English words by using rote memory before the class, but most of time, he treated it only as a compulsory task. Thus, it seemed not helpful for him. M3 favored remembering some English passages by rote memory and M2 reported that he used to make a list of new English words to remember them, but it was hard for him to persist in doing that.

Finally, only one student M3 reported that she always remembered new English words in the context of English text and found it a good way to remember new words. For the social strategy, M1 said he tried to find opportunities to communicate with others in English.

In comparison with high proficiency students, intermediate proficiency students mentioned fewer VLSs they considered effective for expanding vocabulary knowledge. Similar to the high proficiency students, they considered some self-initiated strategies, inferencing strategies, using dictionary, and using context useful for enlarging vocabulary size. However, dissimilar from the high proficiency students, intermediate proficiency students considered repetition strategies useful. Concerning the facilitative VLSs for the depth of vocabulary knowledge, the intermediate proficiency students only suggested one self-initiated strategy (learning how words are used when watching movies) and two repetition strategies such as remembering some English passages by using rote memory.

### Low proficiency students

The low proficiency students reported using fewer self-initiated strategies and dictionary strategies compared to the intermediate and high proficiency students and they did not mention any inferencing strategy used in their English reading. The strategy of watching English movies or TV series was mentioned by two low proficiency students. Their purpose of watching movies was however primarily for entertainment and they seldom focused on learning new English words, which was quite different from both high and intermediate proficiency students, as L1 said:

I sometimes watch English movies. It is not quite often. I only watch them for entertainment and I seldom focus on the new English words.

In the meanwhile, similar to M2 and M5, L4 stated that sometimes he learned new words from the lyrics of some English songs by checking the dictionary. In addition, L5 reported that she read English novels with Chinese translation and sometimes she paid attention to new English words when reading. Moreover, some low proficiency students did the reading comprehension of CET-4 combined with checking the dictionary, which was reported by three of them. However, unlike the high proficiency students, two of them (L1 and L3) emphasized that they just sometimes checked the dictionary after reading, while another one (L4) said he seldom looked up the new English words in dictionary. Moreover, L1 emphasized that he just sometimes did the reading comprehension of CET-4 and actually he did not read much. It is also noteworthy that they did not mention using any inferencing strategy in the process of reading, which was quite different from the high and intermediate proficiency students.

Similar to the intermediate proficiency students but different from the high proficiency students, all low proficiency students reported adopting rote memory to remember new words on the vocabulary book of CET-4. Despite its popularity, this method was considered as boring and inefficient by three low proficiency students (L1, L2, L3):

I use both oral and written repetitions to remember new English words on the vocabulary book of CET-4, but it seems not quite useful for me. It is quite hard for me to remember new English words by using this method. (L3)

In addition to using repetition strategies, L3 and L5 stated that they sometimes remembered new words in the context of English text and L2 said she learned some new English words in the English classes by using note-taking strategy. Finally, L5 reported sometimes she used affixes to remember new English words, however, she also stated that the affixes themselves were hard for her to remember.

The low proficiency students stated only a few VLSs they perceived useful for enlarging vocabulary size, encompassing self-initiated strategies, using dictionary and using context. However, different from the high and intermediate proficiency students, they did not suggest any inferencing strategy. As for the depth of vocabulary knowledge, the low proficiency students considered one self-initiated strategy and two using context strategies useful.

### **Summary**

In general, the 15 participants reported using a wide range of VLSs encompassing self-initiated strategies, using dictionary, inferencing, note-taking, social, association strategies, repetition and using context to remember new English words. Among these strategies, self-initiated strategies, using dictionary, inferencing, and repetition were more frequently mentioned by these participants, whereas, note-taking, social, and association strategies were less frequently mentioned. The high proficiency students reported adopting a wider range of VLSs than did both intermediate and low proficiency students, especially they reported more self-initiated strategies, dictionary strategies and inferencing strategies. The intermediate and low proficiency students seemed to use more repetition strategies than did the high proficiency students. Finally, the low proficiency students did not mention using any inferencing strategy.

As for the useful VLSs these participants suggested, the number descends from the high proficiency group to the intermediate and low proficiency groups. The high proficiency students suggested a wide range of strategies which were facilitative for enlarging vocabulary size, except for repetition and social strategies. In opposite to the high proficiency students, some intermediate proficiency students considered repetition strategies effective for expanding vocabulary size. Different from both high and intermediate proficiency students, the low proficiency students did not consider inferencing strategies useful for enlarging vocabulary size. Concerning the perceived useful VLSs for improving the depth of vocabulary knowledge, the participants suggested fewer than those for enlarging vocabulary size, including several self-initiated strategies, using context, note-taking, using dictionary and repetition strategies. It was unexpected to find that inferencing strategies were not considered useful for improving the depth of vocabulary knowledge by all the participants.

### 4.3.3.2 Sources of the Vocabulary Learning Strategies

Table 4.4 summarizes how participants discovered the vocabulary learning strategies.

**Table 4.4: The participants' knowledge sources about VLS**

	High prof. group					Intermediate prof. group					Low prof. group				
	H1	H2	H3	H4	H5	M1	M2	M3	M4	M5	L1	L2	L3	L4	L5
Explored by themselves	√	√	√	√		√				√	√	√			
Introduced by their English teachers	√	√	√	√	√	√	√	√	√	√			√	√	√
Introduced by others (e.g., sister, classmates)							√		√						

The interview data reveal that there were three main knowledge sources of the participants' strategy use: strategies explored by themselves, those introduced by their English teachers and those introduced by others, such as their classmates and relatives. For the high proficiency students, most of them stated that some of their VLSs were advised by their English teachers, in the meanwhile, some were explored by themselves. It is apparent that these high proficiency students tended to combine their self-initiated strategies with teachers' instruction in terms of VLSs use. As for the intermediate proficiency students, all of them mentioned that they followed their English teachers' advice on strategy use, while two students stated that some VLSs were learned from their classmates or sisters. However, only two of them mentioned that they sometimes also found VLSs by themselves. For the low proficiency students, two of them said they only explored VLSs by themselves while another three students preferred to follow their English teachers' instructions.

### 4.3.4 Perspectives about Vocabulary Knowledge

#### 4.3.4.1 Perspectives about Vocabulary Size

When talking about their perspectives on vocabulary size, there were mainly four concepts that emerged from the interview data of the 15 participants: knowing the basic meaning, quantity, ease of learning, and foundation for English learning. Table 4.5 presents the interviewees' perspectives about these four concepts.

**Table 4.5: Perspectives about vocabulary size**

	High prof. group					Intermediate prof. group					Low prof. group				
	H1	H2	H3	H4	H5	M1	M2	M3	M4	M5	L1	L2	L3	L4	L5
Knowing the basic meaning	√	√	√	√	√	√	√	√	√	√	√	√	√		√
Quantity		√		√	√		√	√	√	√		√	√	√	
Ease of learning			√				√				√				√
Foundation for English learning				√			√			√					

For most interviewees, vocabulary size means knowing the basic meaning of a word, as H2 stated:

Vocabulary size requires you to know the meaning of a word. Usually, you only need to know its basic meaning so that you can understand this word when doing reading or listening comprehension in English.

In addition, the idea of quantity appeared in the interview data, which indicates how many words someone knows. The third common idea about vocabulary size shared among some students was that it was easy to learn. For instance, as mentioned by L5:

I think increasing vocabulary size is relatively easy for me compared with broadening the depth of vocabulary knowledge. Depth of vocabulary knowledge involves many aspects of vocabulary knowledge and thus it is difficult for me to master.

Finally, three of them considered vocabulary size as the foundation for English learning, as stated by M3:

I think vocabulary size is the basis for English learning. For example, I always encounter lots of new English words in the reading comprehension. If I don't know the meanings of these words, it will hinder my understanding of the passage.

Overall, the interview data did not indicate many differences among the three proficiency groups in terms of their perspectives about vocabulary size. The only difference detected was that vocabulary size was considered as the foundation for English learning by some high and intermediate proficiency students, but not by low proficiency students.

#### 4.3.4.2 Perspectives about Vocabulary Depth

As for the perspectives about the depth of vocabulary knowledge, three common ideas suggested by the interviewees were: quality, difficult to learn, and knowing how to use a word. Table 4.6 presents the interviewees' perspectives about the depth of vocabulary knowledge.

**Table 4.6: Perspectives about the depth of vocabulary knowledge**

	High prof. group					Intermediate prof. group					Low prof. group				
	H1	H2	H3	H4	H5	M1	M2	M3	M4	M5	L1	L2	L3	L4	L5
Quality	√	√		√	√	√	√	√	√	√	√		√		√
Difficult to learn			√	√	√			√		√	√	√	√		√
Knowing how to use a word			√									√		√	

A large portion of the interviewees considered that vocabulary depth means how well you know a word, such as its collocations, various meanings, and how to use the word, therefore, we used the word quality to summarize this concept. This idea was reflected in many talks of the students, as stated by H5:

Vocabulary depth refers to how well you know a word. It is more complicated than vocabulary size. You have to know the various meanings of a word, grammatical information and even its usage.

Moreover, nine interviewees considered depth of vocabulary knowledge difficult to learn, as illustrated in one of the low proficiency students:

Vocabulary depth involves so many aspects of word knowledge, so it is quite hard for me to master. For example, for some English words, I only know their meanings, but I don't know how to use them. (L3)

Finally, three participants (H3, L2, L4) specifically emphasized vocabulary depth requires knowing how to use a word in real life. This concept was supposed to be included in the category of quality, but here, I separated them because I intended to distinguish these three interviewees from those interviewees of category one, as they only emphasized on one aspect (how to use a word) rather than mentioning various aspects of the depth of vocabulary knowledge.

In general, similar perspectives about the depth of vocabulary knowledge were shared by the participants from the three proficiency levels. The interview data indicate that most participants had some understanding of the depth of vocabulary knowledge, but that despite their proficiency levels, most of them found depth of vocabulary knowledge difficult for them to master.

#### **4.3.4.3 Perspectives about the Importance of Vocabulary Size/Depth**

When talking about the importance of vocabulary knowledge, most intermediate and low proficiency students considered vocabulary size more important for them, while most high proficiency students reported the significant role of vocabulary depth in English learning.

In addition, two students (H4 and M5) said both were equally important for them. Table 4.7 indicates the interviewees who share each idea.

**Table 4.7: Perspectives about the importance of vocabulary size and depth**

	High prof. group					Intermediate prof. group					Low prof. group				
	H1	H2	H3	H4	H5	M1	M2	M3	M4	M5	L1	L2	L3	L4	L5
Vocabulary size					√	√	√	√			√	√	√	√	√
Vocabulary depth	√	√	√						√						
Both				√						√					

Four high proficiency students explained the reasons for the importance of vocabulary depth. For instance, H1 stated that:

Depth is more important, because you need to know various meanings of a word and its usages. Only knowing the basic meaning of a word is not enough, especially for English speaking and writing.

Similarly, H2 mentioned about the effect of vocabulary depth on the English reading comprehension, which is illustrated below:

Sometimes, various meanings of a word might appear in English reading passages, so if you only know the basic meaning of a word, it would affect your reading comprehension.

Meanwhile, H3 and H4 emphasized that the aim of learning English was mainly for communication, so they considered that it would not make much sense if you could not use it.

As for the reasons for the importance of vocabulary size, H5 stated that:

Vocabulary size is important for me because I don't have many opportunities to use English, such as English writing and speaking. While for English reading and listening, I think vocabulary size would be more important than the depth of vocabulary knowledge.

Likewise, M3 mentioned that vocabulary size was important for English test because reading and listening comprehension took a pretty large proportion of the English test. She considered that vocabulary size was essential for getting a high mark of listening and reading parts. In addition, M2 explained that vocabulary size was more important because it was the foundation for English learning. Finally, M1 and all of the low proficiency students suggested that they considered vocabulary size important because their vocabulary size was quite limited, so enlarging their vocabulary size was their primary aim at the current stage.

H4 and M5 perceived vocabulary size and depth as equally important, as H4 explained:

I think both of them are important because on the one hand, vocabulary size is the foundation for English learning, while on the other, learning English is mainly for communication. Thus if we don't have the depth of vocabulary knowledge, it doesn't make much sense to learn English.

Likewise, M5 commented that:

Sometimes, you have to use different meanings of a word, so vocabulary depth is important. Meanwhile, vocabulary size plays the fundamental role in English learning. Thus both of them are important for me.

Overall, most high proficiency students reported the significant role of the depth of vocabulary knowledge for various reasons, such as its importance for speaking, writing, and reading in English. By contrast, most intermediate and low proficiency students considered vocabulary size more important for them and provided the reasons, such as it was more important for the reading and listening comprehension in the English test and their vocabulary size was limited in this regard. In addition, only one high and one intermediate proficiency student considered both vocabulary size and depth equally important.

#### **4.4 Summary of the Interview Analyses**

The interview data clearly demonstrated that, first of all, most interviewees considered English vocabulary quite important for reading, listening, speaking and writing in English, whereas only one low proficiency interviewee questioned the importance of learning English vocabulary since he thought that a certain amount of vocabulary would be enough for passing the English test.

Secondly, regardless of their proficiency levels, a predominant number of interviewees encountered various difficulties in learning English vocabulary, for example, it was hard to remember new English words for a long time; they had no efficient methods to remember new English words; it was hard to remember the long words, the words with irregular spellings and the infrequent words. The low proficiency students had more difficulties in learning English vocabulary than did both high and intermediate proficiency students, especially in maintaining the long-term memory of learnt words, and lacking of efficient methods to remember new words. In addition, the low and intermediate proficiency students confronted more problems relating to the learning of vocabulary depth than did the high proficiency students.

Thirdly, the participants reported using self-initiated strategies, using dictionary, inferencing, note-taking, social, association strategies, repetition and using context to

remember new English words. Among these strategies, self-initiated strategies, using dictionary, inferencing, and repetition were more frequently mentioned by these participants, whereas, note-taking, social, and association strategies were less frequently reported by them. In addition, the high proficiency students reported using a wider range of VLSs than did both intermediate and low proficiency students, especially they mentioned more self-initiated strategies, dictionary strategies, and inferencing strategies. By contrast, the intermediate and low proficiency students reported more repetition strategies than did the high proficiency students. The low proficiency group did not mention using any inferencing strategy.

Fourthly, when discussing the useful strategies for enlarging vocabulary size or improving the depth of vocabulary knowledge, the high proficiency students mentioned more strategies than did both intermediate and low proficiency students. However, despite their proficiency levels, they all reported fewer useful strategies for improving the depth of vocabulary knowledge than those suggested for enlarging vocabulary size.

Fifthly, as for the participants' knowledge sources of strategy use, they reported three aspects: strategies explored by themselves, those introduced by their English teachers and those introduced by others (e.g., sister, classmates). The high proficiency students tended to combine their self-initiated strategies with their English teachers' advice in terms of the use of VLSs. By contrast, most intermediate and low proficiency students mainly relied on advice from their English teachers or others.

Sixthly, the participants from the three proficiency levels shared relatively similar perspectives about vocabulary knowledge. Vocabulary size was regarded as quantity, easy to learn, knowing the basic meaning and foundation for English learning, while vocabulary depth was considered as quality, difficult to learn and knowing how to use a word.

Finally, when discussing the importance of vocabulary size/depth, most intermediate and low proficiency students regarded vocabulary size as more important than depth, while most high proficiency students emphasized the significant role of vocabulary depth. Another two students (one high- and one mid- proficiency students) considered both were equally important for them (See Table 4.8 for a summary of the interview analyses).

**Table 4.8: A summary of the interview analyses**

	Proficiency group		
	high	intermediate	low
Significance of learning English Voc.	Reading comprehension	Reading comprehension	Reading comprehension
	Listening comprehension	Listening comprehension	Listening comprehension
	Writing	Writing	Writing
	Speaking	Speaking	Speaking
Difficulties of learning English Voc.	Hard to remember new English words for a long time	Hard to remember new English words for a long time	Hard to remember new English words for a long time
		Having no efficient methods to remember new English words	
	Having no efficient methods to remember new English words	Difficult to remember the long words, the words with irregular spellings and the infrequent words	Having no efficient methods to remember new English words
	Difficult to remember the long words, the words with irregular spellings and the infrequent words	Hard to learn the depth of vocabulary knowledge	Hard to learn the depth of vocabulary knowledge
	No problem of memorizing words	Too many English words to remember	Some English words you can recognize, but you cannot produce them.
Strategies for learning English Voc.	Self-initiated strategies	Self-initiated strategies	Self-initiated strategies
	Inferencing strategies	Inferencing strategies	Using dictionary
	Using dictionary	Using dictionary	Using context to remember new words
	Using context to remember new words	Using context to remember new words	Repetition strategies
	Note-taking strategies	Social strategies	Association strategies
	Social strategies	Repetition strategies	
	Association strategies		
Repetition strategies			
VLS perceived useful for enlarging Voc. size	Self-initiated strategies	Self-initiated strategies	Self-initiated strategies
	Inferencing strategies	Inferencing strategies	
	Using dictionary	Using dictionary	Using dictionary
	Note-taking strategies	Repetition strategies	
	Using context	Using context	Using context
	Association strategies		
VLS perceived useful for improving Voc. depth	Self-initiated strategies	Self-initiated strategies	Self-initiated strategies
	Using context		
	Note-taking strategies	Repetition strategies	Using context
	Using dictionary		
Knowledge sources of strategy use	Explored by themselves	Explored by themselves	Explored by themselves
	Introduced by their English teachers	Introduced by their English teachers	Introduced by their English teachers
		Introduced by others (e.g., sister, classmates)	
Perspectives	Quantity	Quantity	Quantity

on Voc. size	Easy to learn	Easy to learn	
	Knowing the basic meaning	Knowing the basic meaning	Easy to learn
	Foundation for English learning	Foundation for English learning	Knowing the basic meaning
Perspectives on Voc. depth	Quality	Quality	Quality
	Difficult to learn		Difficult to learn
	Knowing how to use a word	Difficult to learn	Knowing how to use a word
Perspectives on the importance of Voc. size/depth	Vocabulary size	Vocabulary size	Vocabulary size
	Vocabulary depth	Vocabulary depth	
	Both	Both	

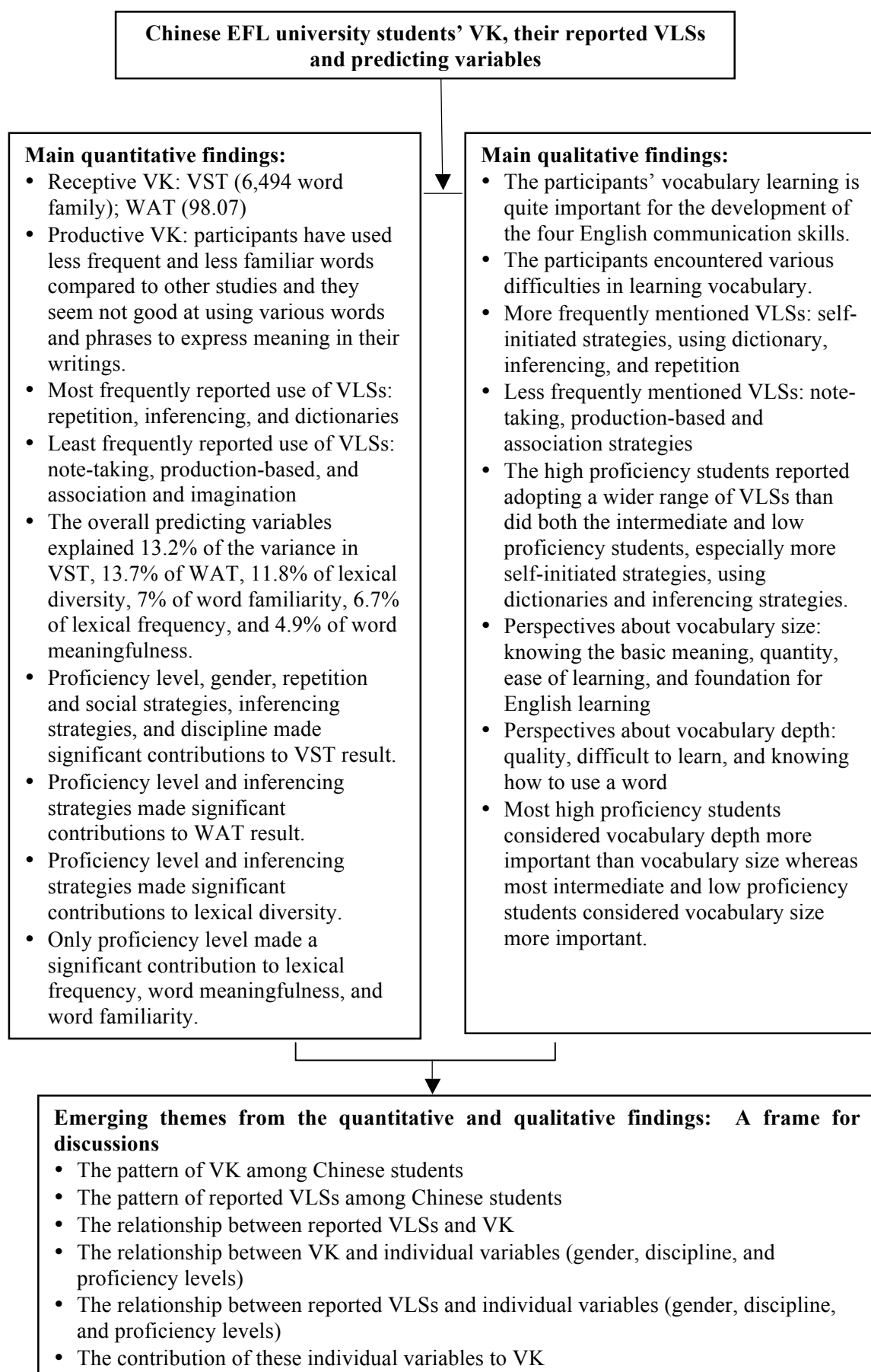
In the next chapter, I will interpret and discuss both quantitative and qualitative findings in the context of relevant literature and integrate the findings from these two parts.

## **Chapter 5 Discussion of the Quantitative and Qualitative Findings**

Chapters three and four have presented the methods and findings of the quantitative and qualitative investigation of vocabulary knowledge (VK) and vocabulary learning strategies (VLSs) of Chinese non-English major university students. In this chapter, I focus on the discussion of the quantitative and qualitative findings in accordance with the six themes that emerged from these two sets of findings.

### **5.1 Discussion of Quantitative and Qualitative Findings**

Figure 5.1 presents a synthesis of the quantitative and qualitative findings. As can be seen from this figure, six themes relevant to the research aims of the current study emerged from the two phases of the study. These themes include (1) the pattern of VK among Chinese students; (2) the pattern of reported VLSs among Chinese students; (3) the relationship between reported VLSs and VK; (4) the relationship between VK and individual variables (gender, discipline, and proficiency levels); (5) the relationship between reported VLSs and individual variables (gender, discipline, and proficiency levels); and (6) the contribution of these individual variables to VK.



**Figure 5.1: A summary of the quantitative and qualitative findings as a frame for discussions**

### 5.1.1 The Pattern of VK among Chinese Students

This section firstly reiterates the main quantitative findings in relation to the participants' receptive and productive vocabulary knowledge; it then compares these results with those of other relevant studies in the literature; and finally it presents the corresponding interview results so as to explain the quantitative findings in light of the qualitative interview data.

The descriptive results presented previously (see Table 3.2) showed that the average vocabulary size of the participants measured by VST was 6,494 English word families. However, as mentioned in Section 2.1.1, the family-based VST may have overestimated the vocabulary size of these participants since Schmitt (1998) found that L2 learners may have only a partial knowledge of a word family. Therefore, the results should be treated with caution. Nation (2006) indicated that in order to get 98% coverage of a text, the L2 reader needs to know 9,000 word families when reading novels, 8,000 word families when reading newspapers, 7,000 word families for understanding most words in spoken English (proper nouns are included) and 6,000 word families for comprehending most words in children's movies. Furthermore, as Nation and Beglar (2007) contended, 8,000 word families constitute a threshold for L2 learners to achieve a comprehensive understanding of non-simplified spoken and written texts. Based on these criteria, the average vocabulary size of the participants of the present study falls short of the suggested criteria for unassisted reading of authentic learning materials, and as such they may face some challenges in comprehending non-simplified written and spoken texts.

This result corroborates some findings found in previous studies. For instance, Nation and Beglar (2007) reported that the vocabulary size of the undergraduate non-native speakers at an English speaking university was around 5,000-6,000 word families and that the competent non-native speaking Ph.D. students had round 9,000 word families. In addition, by adopting a bilingual version of VST, Nguyen and Nation (2011) found that Vietnamese third-year English major students' average vocabulary size was 6,661 English word families. In comparison, it seems that the average vocabulary size of the participants of the present study was larger than that of those participants of Nation and Beglar (2007) and was close to that of Vietnamese third-year English major students in Nguyen and Nation's (2011) study.

As for the participants' depth of vocabulary knowledge, the average score of WAT was 98.07, amounting to 61.29 per cent of the maximum possible score of WAT. This result was lower than that reported in Qian's (1999) study which amounts to 118.24 for the Chinese ESL university learners, 120.54 for the Korean ESL university learners and 148.62 for the native English speakers. One reason for such a difference may be that most of the participants in Qian's (1999) study were university graduates or senior university students and they were ESL learners, whereas the participants of the present study were the second and third year EFL university students. In comparison, Chinese ESL students in Qian's study seem to have a broader depth of vocabulary knowledge than that of the Chinese EFL university students in the present study, which seems natural given the rate of exposure ESL students have to English language. In addition, there was a big gap between my participants' (Chinese EFL students) depth of vocabulary knowledge and that of the native speakers.

Concerning the participants' productive vocabulary knowledge, the mean score for their Measure of Textual Lexical Diversity (MTLD) was 67.68, a score that was lower than that drawn on the non-native speakers in Gonzalez's (2013) study which was 69.12. The participants in Gonzalez's (2013) study were those non-native speakers enrolled in an intensive English program which prepared them to enter university or college. They were at the second semester or last semester of this learning program. In comparison, the words used by the participants in their English essays in the present study were less diversified than those used by the non-native speakers in Gonzalez's (2013) study.

In addition, Coh-Metrix provides three CELEX indices: the average word frequency for content words (WRDFRQc), the average word frequency for all words (WRDFRQa), and the average minimum word frequency in sentences (WRDFRQmc). The present study examined the first two frequency indices. CELEX word frequency scores range from 0 to 6, with 0 suggesting the most uncommon words used in the English language, and 6 indicating the most common words (Baayen et al., 1995). However, it is difficult to interpret the Coh-Metrix scores on their own. The mean score of the participants in the present study for word frequency was 2.48 (CELEX word frequency for content words) and 3.03 (CELEX word frequency for all words). González (2013) reported that the mean score of the non-native speakers' word frequency (CELEX word frequency for all words) was 3.09. In comparison, my research participants have used slightly less frequent words than the non-native speakers in González (2013).

Other indices of productive vocabulary knowledge include word imaginability, word meaningfulness, word familiarity and word concreteness all ranging from 100 to 700. The higher scores show that the words used by language learners have a higher level of imagery and have more associations with other words, and that they have used more familiar and more concrete than abstract words. Again, while the scores of these lexical indices are hard to interpret on their own, the results of the present study showed that the average score for word meaningfulness was 443.23, for word concreteness was 376.95, for word imaginability was 402.77, and for word familiarity was 577.23. Crossley et al. (2011) examined a corpus of the written texts produced by three different proficiency levels of L2 learners and reported that the mean values of word imaginability, word concreteness, word familiarity and word meaningfulness decreased from low proficiency students to high proficiency students. Simply put, the words used by low proficiency students had a higher level of imagery and had more associations with other words, and these students used more familiar and more concrete words than the high proficiency students. It is worth mentioning that Crossley et al. (2011) examined word imaginability, word concreteness and word meaningfulness for all words and only word familiarity for content words, while in the present study only content words were examined for these indices. Crossley et al. (2011) found that mean values of word familiarity for content words ranged from 592.13 to 585.13 in line with an ascending order of language proficiency levels, while in the present study, the average mean value was 577.23, suggesting that the participants in the present study used less familiar words than did those participants in Crossley et al.'s (2011) study. This finding supports the word frequency findings for this cohort of participants as discussed above. That is, on average, participants of the current study used less familiar and less frequent words than those used by participants in other studies.

The value of "LSA sentence adjacency" which measures how conceptually and semantically similar each sentence is to the next sentence, ranges from 0 (low cohesion) to 1 (high cohesion). In the present study, the mean value was 0.196. Crossley, Salsbury, and McNamara (2010b) discovered that the LSA sentence adjacent value increased with time spent learning English and reported that the mean value of LSA increased from 0.16 to 0.32 across around 50-week English training, suggesting that the students began to use more words that had closer semantic similarity. Crossley et al. (2010b), however, used the spoken data whereas the present study examined the written data, and the expectation for

written texts is to have higher levels of cohesion. In comparison, therefore, the essays produced by the participants of this study were less cohesive.

In addition, hypernymy index shows the rate of use of cover terms like “color” which is a cover term for all individual colors. Hypernymy values are calculated with “1” being the highest possible hypernymy (the most abstract cover term) and scores higher than one indicate the more frequent use of concrete words (Crossley et al., 2009). In the present study, the mean hypernymy value was 1.776. By examining spoken language of the L2 learners, Crossley et al. (2009) reported that the mean hypernymy value decreased from 1.44 to 1.05 across the 52-week English training, indicating that as learners’ proficiency increased, they developed their lexical proficiency toward more abstract words. This finding was supported by Crossley et al. (2010) who found that the more concrete words a writer used in their writings, the more likely they had low lexical proficiency. Comparing the hypernymy value in my study (i.e. 1.776) with that in Crossley et al.’s (2009) study (i.e. 1.05-1.44), my research participants used more concrete words.

Coh-Metrix also reports the mean WordNet polysemy values for all content words in a text. The higher the value of polysemy, the more senses a word contains. The results of the present study showed that the mean value of polysemy for all content words in participants’ essays was 4.65, indicating that the content words produced by the participants had an average of 4.65 different meanings. Crossley et al. (2010a) found that as language proficiency increased, L2 learners produced words that had more potential senses. In their study, the mean polysemy values for underachieved L2 learners’ spoken texts ranged from 3.39 to 4.03 after around 50-week English training. The participants of the present study used the content words which have more potential meanings than the words used by the participants in Crossley et al.’s (2010a) study. The quantitative findings of the participants’ vocabulary knowledge may thus be summarized in the light of the above discussion in Table 5.1.

**Table 5.1: A summary of discussion of the participants' vocabulary knowledge**

		Present study	Nation and Beglar (2007)	Nguyen and Nation (2011)	Comparison of results
Receptive Voc.	Voc. size	6,494 word families (Chinese non-English majors)	1) 5,000-6,000 (ESL undergraduates); 2) 9,000 (non-native speaking PhD students); 3) 8,000 (a threshold for dealing with non-simplified spoken and written texts)	6,661 (Vietnamese third-year English majors)	1) Larger than that of the ESL undergraduates in Nation and Beglar (2007); 2) Similar to that of Vietnamese third-year English majors; 3) Having difficulties in comprehending non-simplified spoken and written texts
	Voc. depth	98.07 (60 per cent of the maximum score)	Qian (1999) 1) 118.24 (Chinese ESL university students) 2) 120.54 (Korean ESL university students) 3) 148.62 (native speaker)		Having narrower vocabulary depth than that of Chinese ESL university students in Qian's (1999) study
Productive Voc.	Lexical diversity (MTLD)	67.68	González (2013) 69.12 (non-native speaker)		Less diversified than the words used by the non-native speakers in González (2013)
	Lexical frequency (WRDFRQa)	3.03	0-6	González (2013)	Slightly less frequent than the words used by the non-native speakers in González (2013)
				3.09 for non-native speakers	
	Word imaginability	402.77	100 to 700		N/A
	Word meaningfulness	443.23	100 to 700		N/A
	Word familiarity	577.23	100 to 700	Crossley et al. (2011)	Less familiar than those words used by the participants in Crossley et al. (2011)
				592.13 to 585.13 (ascending order of proficiency levels)	
	Word concreteness	376.95	100 to 700		N/A

	LSA sentence adjacent	0.196	0 (low cohesion) to 1 (high cohesion)	Crossley, Salsbury, et al. (2010b)	Relatively low in terms of cohesive writing skills
				0.16 to 0.32 across 50-week English training	
	Hypernymy	1.776	scores higher than “1” indicate the more use of concrete words	Crossley et al. (2009)	Using more concrete words than abstract words
				1.44 to 1.05 across 52-week English training	
	Polysemy	4.65	Crossley, Salsbury, et al. (2010a)		Having an average of 4.65 different potential meanings
			3.39 to 4.03 after around 50-week English training		

The interview data provided richer information about the participants' perspectives on English vocabulary learning, which can be used to explain the quantitative findings. The interview study showed that almost all participants regarded learning English words quite important for the development of the four English communication skills. Some participants made connections between certain language skills and CET-4 and how this might relate to their weaknesses. It seemed that passing the CET-4 was a major objective for some interviewees, which motivated them to learn English words. This is while some other interviewees believed that learning and improving their vocabulary knowledge would be helpful for improving their English writing and speaking skills, which they thought they could perform poorly. Therefore, the importance of vocabulary learning as perceived by these participants might motivate them to devote more efforts to learn English words. This may explain the quantitative finding that the participants' vocabulary size was satisfactory in comparison to other studies.

However, despite the importance of vocabulary learning for these interviewees, almost all of them faced difficulties in learning and using English vocabulary. For instance, M5 stated that it was difficult to retain various meanings of a word and L2 (the second low proficiency level interviewee) mentioned that she had difficulties in learning how to use a word. In addition, many interviewees considered depth of vocabulary more difficult to deal with than the breadth of vocabulary or vocabulary size given the former involves many aspects of vocabulary knowledge. These interview results may explain why their depth of vocabulary knowledge was less satisfactory compared to other studies as discussed above.

Moreover, in the Chinese EFL context, students did not have many chances to use English in real life situations, which results in their poor performance in English speaking or writing. For instance, some interviewees were aware of their weaknesses of speaking and writing in English. H4 stated she was poor in these two communication skills. In addition, L1 (the first low proficiency level interviewee) mentioned that one difficulty for him was that he can recognize some English words, but cannot spell them correctly. These interview findings may explain the reason why some values of lexical indices in the participants' essays seemed less satisfactory compared to other relevant studies. For example, these participants used more concrete words in their essays, and their essays turned out to be less cohesive. In addition, the words used by the participants of this study were less diversified than those words used by the non-native speakers in Gonzalez's (2013) study. However, it appears that the participants of this study used slightly less frequent words than the non-native speakers in Gonzalez's (2013) study. The participants of this study also adopted less familiar words than those adopted by the participants of Crossley et al. (2011). These findings suggest that the participants of the present study were more likely to use less frequent and less familiar words in their English essays, however, they seem not good at using a variety of words and phrases to express meaning (lexical diversity) in their writings. There are two possible explanations for these findings. Firstly, CET-4 and CET-6 require test takers to master 4000 and 6000 English words respectively. Based on the interviews, most participants remembered English words by using vocabulary books for the CET test, therefore, they may remember some less frequent words in the vocabulary book. Secondly, these participants have few opportunities in their learning processes to use the words, and they may be less proficient in choosing different words of similar meanings to express themselves in writing.

### **5.1.2 The Pattern of Reported VLSs among Chinese Students**

This section firstly synthesizes both quantitative and qualitative findings concerning the participants' reported use of VLSs, and then it provides some explanations for these findings. Finally, the findings of the present study are compared with those of other relevant empirical studies.

The quantitative results revealed that inferencing strategies were reportedly most frequently used by the participants of this study which included inferring the meaning of a new word from its affixes and roots, from its textual context, from the real situations in daily life and from one's common sense and general information. The participants also

reported that they used dictionaries very frequently for learning both the meaning and the use of new words. Repetition and social strategies ranked the third most frequently used strategies reported by the participants. However, when separating repetition strategies from social strategies, repetition strategies emerged as the most frequently used strategies, exceeding inferencing strategies, whereas social strategies were reportedly less frequently used by the participants. Note-taking strategies were the least favored ones reported by the participants along with production-based strategies and association and imagination strategies. The quantitative findings thus revealed the participants of this study reported using repetition, inferencing, and dictionaries as the main strategies for learning new English words. The qualitative study showed consistent results with those of the quantitative study, except for the reported use of self-initiated strategies. In the interviews, the most popular VLSs reported by the participants encompassed self-initiated strategies, using dictionaries, inferencing strategies and repetition strategies, however, note-taking, association and production-based strategies were less frequently mentioned by the participants.

It was unsurprising to find that the participants reported using both oral and written repetition strategies most frequently, as Chinese learners are generally believed to be good at using rote memorization. This result may be influenced by the Chinese culture and Chinese students' L1 learning background. There is a popular old saying in China which goes as "If you have recited thoroughly 300 Tang poems, even if you cannot write a poem, you will be able to chant some". This old saying emphasizes the importance of rote memorization in general which may also extend to learning a foreign language. Moreover, Chinese students are more likely to use rote memory to learn their L1 as they are encouraged by some Chinese teachers to recite some texts in Chinese textbooks and to memorize some exemplar Chinese essays. Therefore, they may transfer those learning strategies of their L1 learning when learning an L2. In addition, using dictionaries and inferencing strategies were also frequently reported by the participants and these two kinds of strategies were generally considered as two popular VLSs among Chinese EFL students (Gu & Johnson, 1996). On the contrary, this study found that the participants seldom reported using production-based strategies. This finding was quite understandable as in the Chinese EFL context, the students have few opportunities to use English in their daily life. In addition, they did not favor association and imagination strategies which may be due to language distance between Chinese and English and also their L1 learning background.

Nevertheless, it was unexpected to find that note-taking strategies were not frequently reported by these participants.

In comparison with previous studies, my findings were partially consistent with those reported by Gu and Johnson (1996). Both my study and Gu and Johnson's study found that Chinese EFL learners reported using inferencing strategies and dictionaries most frequently, whereas they reported adopting production-based strategies (activation strategies) and association and imagination strategies (encoding strategies) least frequently. Nevertheless, there were also some differences between the findings of these two studies. In the present study, both oral and written repetitions were the most popular strategies reported by the participants, whereas in Gu and Johnson's (1996) study, the participants were less likely to use written repetition strategies compared to other strategies. Moreover, note-taking strategies were reportedly the least frequently used strategies by the participants in the present study, however, in Gu and Johnson's (1996) study, note-taking strategies ranked the third most frequently used strategies. In addition, the results of the present study endorsed the findings presented in other relevant studies on EFL learning. For instance, Catalán (2003), Kojic-Sabo and Lightbown (1999), Riazi and Alvari (2004), and Schmitt (1997) found that EFL learners tended to make use of dictionaries, inferencing strategies, note-taking and repetition strategies, and used less mnemonic and association strategies. Nevertheless, one difference between the findings of the present study and those of the above studies was that the participants in the present study infrequently reported using note-taking strategies when learning English vocabulary. In the ESL learning context, Fan (2003) found that inferencing strategies were most frequently used by the Chinese ESL university students in Hong Kong, indicating that inferencing strategies were favored by both Chinese EFL and ESL university students. Fan (2003) also found that Hong Kong students did not favor association strategies, which resembled the finding of the present study. However, the difference between the findings of these two studies was that Chinese ESL university students in Fan's (2003) study tended to use more strategies for consolidating known words, such as revising newly learnt words, while they used less repetition strategies in comparison to the EFL university students of the present study.

The above discussion suggests that, similar to the results of the other relevant EFL studies, the participants of the present study reported using repetition strategies, inferencing strategies, and dictionaries most frequently, while disfavored production-based strategies and association and imagination strategies. Nevertheless, the only difference

noted between this and other studies was in the use of note-taking strategies, which was less favorable in the present study compared to other relevant EFL vocabulary studies. In comparison to Fan's (2003) study, both Chinese ESL and EFL students favored inferencing strategies whereas disfavored association strategies. However, the Chinese ESL university students adopted more strategies for consolidating known words, while they used fewer repetition strategies than did the Chinese EFL students of the present study.

### **5.1.3 The Relationship between Reported VLSs and VK**

This section provides an explanation of the quantitative findings concerning the relationship between VLSs and VK, followed by a comparison of these findings with other relevant research studies. A synthesis the quantitative and qualitative findings will be presented as well.

The quantitative findings suggested that inferencing strategies made a significant and positive contribution to the participants' vocabulary size, depth of vocabulary knowledge and productive lexical diversity. One explanation may be that inferencing strategies involve learners' hypothesis formation and testing about word meaning (Ellis, 1994; Haastrup, 1991) and such cognitive processing may result in better comprehension of a text and learning of some lexical items (Read, 2000). In addition, Schouten-van Parreren (1989) argued that the presence of psychological and linguistic context of a text could facilitate learners memorizing new words. Moreover, De Bot, Paribakht, and Wesche (1997) suggested that the process of inferring the meaning of an unknown word was one of gathering multiple sources of information, such as morphology, syntax, word association and derivation, so as to construct the meanings of unknown words. In this process of meaning construction, the learner may internalize some features of a new word, such as its word association and derivation, which are parts of vocabulary depth. This may provide some explanations for the finding that inferencing strategies were effective for expanding the depth of vocabulary knowledge in addition to vocabulary size. These results were consistent with a number of studies (e.g., Fraser, 1999; Haastrup, 1991; Huckin & Bloch, 1993; Hulstijn, 1992; Nation & Coady, 1988; Schouten-van Parreren, 1989) in which inferencing strategies were found to be effective for both English reading comprehension and vocabulary learning. In addition, the present results were partially consistent with those advanced by Gu and Johnson (1996) who found that inferencing strategies were significantly and positively related to learners' vocabulary size. Moreover, the results of the present study, to some extent, corroborated Nassaji's (2006) findings which showed a

significant relationship between learners' depth of vocabulary knowledge and their use of inferencing strategies.

In addition, repetition and social strategies made a significant, but negative contribution to the participants' vocabulary size in the present study. The possible explanation for this finding could be that repetition strategies may be more effective for short-term word retention (Hummel, 2010). As the participants stated in the interviews, they felt it was hard to maintain the long-term retention of words by using repetition strategies because they seldom reviewed the learnt words. The other possible explanation could be attributed to the participants' English proficiency levels. Some previous studies (e.g., Hummel, 2010; Van Hell & Mahn, 1997) showed that repetition strategies were useful for low or intermediate proficiency L2 learners in word retention. However, the participants of the present study were advanced EFL learners who have studied English for more than 8 years. These findings partially corroborated those of Gu and Johnson (1996) who found that visual repetition strategy had the significant, but negative effect on learners' vocabulary size. Moreover, Fan (2003) found that the less proficient L2 learners adopted repetition strategies significantly more often than those high proficient learners.

It is also noteworthy that self-initiated strategies showed significant relations with the participants' productive vocabulary knowledge, including lexical diversity, lexical frequency, word meaningfulness, and word imaginability. The explanation for this finding may be attributed to the link between learners' input and output, as using self-initiated strategies involves performing a great amount of English input activities, such as reading English newspapers, reading English novels, and watching online English talks etc. These findings were consistent with some theories concerning L2 acquisition. For instance, Ellis (1994) proposed three different theoretical approaches concerning L2 acquisition. Behaviorist theories emphasized a direct relationship between input and output, mentalist theories mainly focused on learners' internal mechanisms, and interactionist theories emphasized on both of input and internal mental processing. Although these theories differ in terms of the importance of input on L2 acquisition, they recognize the link between input and output. Therefore, it was not surprising to find that self-initiated strategies were significantly and positively related to many aspects of the learners' productive vocabulary knowledge such as lexical diversity, lexical frequency, word meaningfulness, and word imaginability.

Meanwhile, the interview findings showed that the useful VLSs considered by the high proficiency interviewees for enlarging their vocabulary size were similar to the quantitative results. High proficiency students believed that combining self-initiated strategies with using dictionaries and inferencing strategies would be helpful for enlarging their vocabulary size while they prioritized using context over repetition strategies in their vocabulary learning approach. Similar to the high proficiency students, the intermediate proficiency students considered some self-initiated strategies, inferencing strategies, using dictionary, and using context useful for enlarging vocabulary size while dissimilar from the high proficiency students, they considered repetition strategies useful. In addition, the low proficiency students believed that self-initiated strategies, using dictionary and using context were useful, but they did not consider using inferencing strategies as helpful for enlarging their vocabulary size. Therefore, these two sets of findings indicated a gap between the interview-based elicited VLSs that were perceived as useful by both intermediate and low proficiency students for enlarging vocabulary size, and their usefulness drawn from the quantitative findings. Moreover, as for the effective VLSs for vocabulary depth, the interview-based elicited useful VLSs were not consistent with those drawn from the quantitative study. The interview data revealed that inferencing strategies were not suggested as useful for expanding the depth of vocabulary knowledge by the interviewees. Instead, some students considered that self-initiated strategies, using context, note-taking, using dictionary and repetition strategies were helpful for improving their depth of vocabulary knowledge. Therefore, there was a discrepancy between the interview and quantitative data in terms of the reported useful VLSs for expanding vocabulary depth. One possible explanation for this finding may be that the participants only considered inferencing strategies useful for enlarging their vocabulary size because these strategies were primarily used by them for discovering new words' meanings. However, they did not mention the usefulness of inferencing strategies for improving their depth of VK.

Overall, the quantitative findings suggested that using repetition and social strategies made a significant but negative contribution to the participants' vocabulary size. Instead, inferencing strategies were found helpful for improving participants' vocabulary size, depth and productive lexical diversity, which were consistent with a large number of previous studies. In addition, self-initiated strategies were significantly related with several aspects of the participants' productive vocabulary knowledge probably due to the link between input and output. Finally, the synthesized results revealed that there were some

gaps between the interview-based elicited VLSs that were perceived useful by the participants and those found in quantitative study, especially the useful VLSs for expanding vocabulary depth.

#### **5.1.4 The Relationship between VK and Individual Variables (gender, discipline, and proficiency levels)**

The results of the quantitative study showed that the average vocabulary size of the female students was significantly larger than that of the male students (see Section 3.3). The previous studies indicated three patterns on the relationship between gender and VK. Firstly, some studies (e.g., Gu, 2002; Halpern, 2010; Kimura, 1999; Nyikos, 1990) showed that female students outperformed male students in foreign language vocabulary learning. For instance, Gu (2002) found that female students significantly outperformed the male counterparts in terms of their vocabulary size. In addition, Nyikos (1990) reported that female students of German as a foreign language excelled over their male counterparts in a memorization test of German vocabulary. Secondly, some studies (Catalán, 2010; Grace, 2000) found no significant gender difference in terms of vocabulary learning. For instance, in the Spanish EFL primary school learning context, Catalán (2010) found that there was no significant difference between the 12-year-old male and female students in terms of two vocabulary tests – the 1000 Word Test (Nation, 1983) and the 2000 word frequency band from Vocabulary Level Test (Schmitt et al., 2001). Thirdly, Boyle (1987) reported that male university students performed better than female students in terms of the listening vocabulary (comprehension of heard vocabulary) in the Hong Kong ESL learning context.

In comparison with the previous research, the findings of the present study corroborated those of Gu (2002), Halpern (2010), Kimura (1999), and Nyikos (1990), suggesting that female students excelled over male students in word learning. One possible explanation for this finding may be that, as suggested by Gu (2002), female students were expected to be more successful in language learning in China and as such they might have a higher motivation in learning English. Moreover, some studies (e.g., Nyikos, 1990) indicated that female students seemed to have stronger memorization power than male students. In addition to these two explanations, the present study found that female students used each group of VLSs more frequently than their male counterparts, in particular, they significantly used more note-taking strategies. The previous research (e.g., Fan, 2003; Gu & Johnson, 1996; Lawson & Hogben, 1996) showed that the frequent use of VLSs was positively related to vocabulary learning or language proficiency. Therefore, the

higher the frequency of the VLSs use by the female students in this study may partly explain the reasons why they performed better than the male students in terms of vocabulary size. However, the results of the present study were inconsistent with those of Catalán (2010) and Grace (2000). One possible explanation for the contradictory results may be due to the different participants involved especially in terms of their age and proficiency levels. The participants of the above two studies were beginning L2 learners, whereas the participants of the present study were the university students who have studied English for more than 8 years. Therefore, it seems that the gender difference in vocabulary size is more likely to emerge at a later stage of L2 learning, rather than at an early stage. In addition, these findings were also contradictory to those of Boyle's (1987) study, which suggests that gender difference may be affected by various vocabulary learning tasks adopted in different studies.

As for the depth of vocabulary knowledge, no significant difference was found between male and female students in the present study. This result coincides with that reported by Espinosa (2010) despite substantial differences in research methodologies. Espinosa (2010) found gender similarity in terms of the late primary school boys' and girls' word association responses, which include syntagmatic and paradigmatic word associations. Nevertheless, Scarcella and Zimmerman (1998) found that male Year-1 ESL university students including Chinese, Vietnamese, Spanish and others, performed better than the female counterparts in terms of their scores on the Test of Academic Lexicon which adopted the test format of Vocabulary Knowledge Scale (Paribakht & Wesche, 1993). These contradictory results might be partially attributed to the participants' different ethnic backgrounds, different tests of vocabulary depth adopted, and different learning contexts.

In terms of participants' difference of VK across disciplines, this study found that the science and engineering students significantly outperformed the humanities and social sciences students in terms of their vocabulary size (see Section 3.3). Significant differences were also found between these two groups of students in terms of their depth of vocabulary knowledge. These findings showed that the science and engineering students had much deeper and wider vocabulary knowledge than did the humanities and social sciences students. The possible reasons for these findings may be partially attributed to the differences in the students' use of VLSs. This study found that the science and engineering students were more inclined to learn new words in context than the humanities and social sciences students, such as using self-initiated strategies, inferencing strategies and

production-based strategies, although the difference was not significant statistically. These results were partially in accord with Gu's (2002) findings. Both studies found the average vocabulary size of the science and engineering students was larger than that of the humanities and social sciences students although the difference was not significant in Gu's (2002) study.

Moreover, the high proficiency students performed significantly better both in their vocabulary size and depth of vocabulary knowledge than did the intermediate and low proficiency students (see Section 3.3). These results therefore showed that there was a strong link between the size and depth of vocabulary knowledge and students' overall English proficiency levels. These findings partially corroborated some previous research which indicated that there were significant correlations between EFL learners' vocabulary size/depth and their English competence, such as their reading comprehension ability (e.g., Henriksen, Albrechtsen, & Haastrup, 2004; Laufer 1992; Qian, 1999, 2002; Stæhr, 2008), listening comprehension ability (e.g., Stæhr, 2008) and also their overall English proficiency (Gu & Johnson, 1996).

Concerning the participants' productive vocabulary knowledge, the quantitative results showed that there was no significant difference in the participants' productive vocabulary knowledge in terms of their gender and disciplines. These results were partially consistent with those reported by Prados (2010) who found that there was no significant difference between the male and female university students in terms of the lexical density index (the ratio of content words to function words and the number of content words per sentence) of the participants' essays. In addition, Morris (1998) reported that there was no significant lexical difference between the male and female college students in terms of their written essays which were assessed for linguistic accuracy, readability and conformity to assignment guidelines. In contrast, in the Spanish EFL secondary education learning context, Fontecha (2010) found that girls significantly outperformed boys in terms of their English learning motivation and productive English vocabulary achievement. In addition, Catalán (2010) reported that 12-year-old Spanish female EFL learners outperformed their male counterparts in terms of two productive tests (composition and cue word tests) assessing their produced type and token number. These inconsistent results might possibly suggest that at an early stage of L2 learning, female learners perform significantly better than might male learners do in terms of their productive vocabulary learning, however, this difference becomes smaller or vanished at a later stage in the process of L2 learning.

The significant differences were found in the participants' productive lexical diversity, lexical frequency, lexical familiarity and word meaningfulness in terms of English proficiency levels. These findings indicated that the participants' language proficiency was also closely related to their productive vocabulary knowledge in addition to receptive vocabulary knowledge. The high proficiency students performed significantly better than did both intermediate and low proficiency students in terms of lexical diversity, lexical frequency and lexical familiarity. However, the high proficiency students performed only significantly better than did the low proficiency students in terms of lexical meaningfulness (see Section 3.3). Different from receptive vocabulary knowledge, no significant difference was found between the intermediate and low proficiency students in terms of their productive vocabulary knowledge. These findings suggested that the high proficiency students used significantly more diverse, less frequent and less familiar words in their English essays than did both intermediate and low proficiency students. In addition, the high proficiency students tended to use significantly more English words with fewer associations with other words than did the low proficiency students.

These results were consistent with those reported by some previous studies. Crossley, Salsbury, McNamara, et al. (2010) maintained that the higher value of lexical diversity index implied the more proficient and larger lexicons. That is to say, advanced learners were likely to have more diverse and larger vocabulary than underachieved learners. In addition, some research (e.g., Laufer & Nation, 1995; Meara & Bell, 2001) found that the underachieved L2 learners tended to use higher frequency words than did the advanced L2 learners. Moreover, Crossley, Salsbury, McNamara, et al. (2010) claimed that lexical familiarity was closely related to lexical frequency as familiar words were likely to occur more frequently in a text. Therefore, the higher frequency English words used by the low proficiency students were likely to be more familiar than those used by the high proficiency students. In addition, Ellis and Beaton (1993) argued that learners were likely to acquire English words with more associations first. Likewise, Salsbury, Crossley, and McNamara (2011) found that as learners' English proficiency increased, they tended to use more difficult words that had fewer associations with other words. Finally, the present findings completely corroborated Crossley et al.'s (2011) findings which showed that as L2 learners' English proficiency increased, they used the English words which were more diverse, less familiar, less frequent and having fewer associations with other words in their English free writings.

Overall, female university students of the present study performed significantly better than their male counterparts in vocabulary size, which was consistent with a large number of previous studies. The science and engineering students excelled over the humanities and social sciences students in both their vocabulary size and depth of vocabulary knowledge, which was partially supported by the findings of Gu (2002). Moreover, language proficiency level had strong relationships with both receptive and productive vocabulary knowledge, which corroborated many other relevant studies.

#### **5.1.5 The Relationship between Reported VLSs and Individual Variables (gender, discipline, and proficiency levels)**

The quantitative results indicated that female students reported using significantly more VLSs than did their male counterparts. To be specific, they reported using significantly more note-taking strategies than did the male students (see Section 3.3). There might be two possible explanations for these findings. Firstly, the higher frequency of VLSs use reported by the female students might be attributed to their higher degree of motivation compared to the male students towards English learning in general and vocabulary learning in particular. Several studies (Bacon & Finnemann, 1992; Fontecha, 2010; Kaylani, 1996) revealed that female students had more positive attitudes towards language learning and were more motivated than their male counterparts. Specifically in the Chinese context, as Gu (2002) suggested, female students were expected to be more successful in language learning in China and as such they might have a higher motivation in learning English and become more active in using strategies. Secondly, note-taking strategies are normally regarded as the traditional learning strategies in China, and it might be that Chinese female students are more willing than male students to follow the conventions. This is also an observation reported by Oxford and Nyikos (1989). These findings partially coincide with those reported by previous research on sex difference in terms of the adoption of language learning strategies and vocabulary learning strategies. Some studies in this strand (e.g., Catalán, 2003; Gu, 2002; Oxford & Nyikos, 1989; Oxford, Nyikos, & Ehrman, 1988; Stoffer, 1995; Wen & Johnson, 1997) showed that female students used more strategies than did their male counterparts. The most comparable study was conducted by Gu (2002) who found that the female students used significantly more VLSs that were found relevant to the success of language/vocabulary learning by Gu and Johnson (1996) than did the male students, such as metacognitive strategies, contextual guessing, dictionary, note-taking, activation strategies, contextual encoding and oral repetition. In addition, Catalán

(2003) reported that the female students used significantly more vocabulary learning strategies than did their male counterparts. In particular, female students used significantly more formal rule strategies (e.g., analyzing parts of speech), input elicitation strategies (e.g., asking a teacher), and rehearsal strategies, whereas male students used more strategies involving forming an image of a word's meaning. Nevertheless, in the present study, I found that the significant difference between the male and female students on reported use of VLSs only existed in note-taking strategies.

The results of the present study also showed the significant differences in the overall reported use of VLSs between the humanities and social sciences and the science and engineering students. The humanities and social sciences students reported using significantly more note-taking strategies than did the science and engineering students (see Section 3.3). This result was consistent with Gu (2002) who found that arts students adopted significantly more note-taking strategies than did sciences students.

In addition, the only significant difference found among the proficiency groups was reported use of inferencing strategies. The significant difference in reported using inferencing strategies lies only between the high proficiency group and low proficiency group (see Section 3.3). The interview results were partially consistent with the quantitative results which found that the high proficiency students reported more VLSs than did both the intermediate and low proficiency students, especially more self-initiated strategies, dictionary strategies and inferencing strategies. In addition, the high proficiency students tended to combine these three strategies in their vocabulary learning, for instance, some participants reported reading English newspapers in combination with using inferencing strategies and dictionaries. By contrast, the low proficiency students did not mention using any inferencing strategy. The interview data revealed that low proficiency students rarely focused on vocabulary learning when using the self-initiated strategies, such as watching English movies or reading English novels, which was quite different from the high proficiency students. This point could be explained by the "Noticing Hypothesis" proposed by Schmidt (1990) which argued that noticing was the essential starting point for language acquisition. Therefore, insufficient attention to vocabulary learning when using the self-initiated strategies might result in poor vocabulary learning. Moreover, as for the knowledge sources of VLSs, the interview data indicated that the high proficiency students combined their self-initiated strategies with their English teachers' advice. By contrast, most intermediate and low proficiency students mainly relied on

advice from their English teachers or others. The use of self-initiated strategies by high proficiency students indicated that these students were more likely to search for various VLSs beyond the limited strategies given by their English teachers. From this view point, this finding may explain in part the reason why the high proficiency students adopted more VLSs than both intermediate and low proficiency students. These quantitative and qualitative findings were consistent with Gu and Johnson (1996) who found that the most successful group of Chinese EFL students, whom they called Readers, learned English vocabulary mainly through reading, inferencing and contextual encoding. In addition, Gu (2003a) and Fan (2003) found that inferencing strategies were frequently adopted by the successful EFL and ESL Chinese learners in their English vocabulary learning.

In general, there is a consensus among researchers that female students reported using significantly more VLSs than male students. In addition, the humanities and social sciences students reported using significantly more note-taking strategies than the science and engineering students, which was consistent with the finding of Gu (2002). Finally, high proficiency students reportedly used significantly more inferencing strategies than low proficiency students, which corroborated previous studies.

#### **5.1.6 The Contribution of Overall Variables (gender, discipline, proficiency levels and VLSs) to VK**

The overall predictor variables including gender, discipline, proficiency levels, and VLSs accounted for 13.2% of the variance in the participants' vocabulary size. Among these predictor variables, proficiency level was the strongest predictor, followed by gender, repetition and social strategies, inferencing strategies, and discipline. The explanation rate of these variables on vocabulary size was lower than that reported by Gu and Johnson (1996) who found that the participants' vocabulary-related beliefs and vocabulary learning strategies explained around 20% of the variance of their vocabulary size. Beliefs about vocabulary learning included 17 items involving whether vocabulary should be memorized, should be picked up naturally or should be studied and used. For instance, repetition was found to be the best way to remember words.

Concerning the participants' depth of vocabulary knowledge, the overall predicting factors explained 13.7% of the variance, which was similar to that of vocabulary size. Proficiency levels still made the strongest unique contribution, followed by inferencing strategies. As for the participants' productive vocabulary knowledge, the overall variables explained less of the variance, which was 11.8% of the variance in the participants' lexical

diversity, 7% in word familiarity, 6.7% in lexical frequency, and 4.9% in word meaningfulness. As for lexical diversity, proficiency level was the strongest predictor, followed by inferencing strategies, whereas for the other three indices, only proficiency level made a significant contribution.

These findings indicate that the overall variables could better predict the participants' receptive vocabulary knowledge than did the productive vocabulary knowledge. Both explanation rates were relatively low. The low explanation rates may be due to other related variables that were excluded from the present study, such as learners' language learning beliefs, motivation, and attitude, which may also have predicting powers on vocabulary knowledge. Among the variables included, the proficiency level was the most relevant variable to the participants' vocabulary size and depth, productive lexical diversity, word familiarity, lexical frequency and word meaningfulness (see Section 5.1.4 for a discussion). In addition, inferencing strategies was a strong and positive predictor of the participants' vocabulary size, depth and productive lexical diversity, while repetition and social strategies showed a strong, but negative contribution to their vocabulary size (see Section 5.1.3 for a discussion). Comparing these two groups of strategies, repetition and social strategies had slightly better predicting power for the vocabulary size than inferencing strategies. Moreover, gender and discipline could only significantly predict the participants' vocabulary size (see Section 5.1.4 for a discussion), and the former had stronger predicting power than the latter.

## **5.2 Conclusion**

This chapter summarized the main findings of both quantitative and qualitative studies and discussed and synthesized both sets of findings according to six themes which emerged from the findings including 1) the pattern of VK among Chinese students; 2) the pattern of VLSs among Chinese students; 3) the relationship between reported VLSs and VK; 4) the relationship between VK and individual variables (gender, discipline, and proficiency levels); 5) the relationship between reported VLSs and individual variables (gender, discipline, and proficiency levels); and 6) the contribution of the overall variables to VK (see Table 5.2 for the integration of both sets of conclusions). In the next chapter, I will present both empirical and pedagogical implications of the present study, pinpoint the limitations and provide some recommendations for future research.

**Table 5.2: The integration of quantitative and qualitative conclusions**

The patterns of VK	QUAN results	<ol style="list-style-type: none"> <li>1) Voc. size on VST (6,494 word families), a size that is satisfactory compared to other studies.</li> <li>2) Voc. depth on WAT (98.07): indicating that my participants have narrower vocabulary depth compared to Qian's (1999) study.</li> <li>3) My participants used less frequent and less familiar words compared to other studies. Nevertheless, they seem not good at using various words and phrases to express meaning in their writings.</li> <li>4) My participants used more concrete words and their essays were less cohesive.</li> </ol>
	QUAL results	<ol style="list-style-type: none"> <li>1) Vocabulary learning is quite important for the development of the four English communication skills.</li> <li>2) The participants encountered various difficulties in learning vocabulary, with some related to the depth of VK.</li> <li>3) The participants considered that depth of VK was more difficult to learn than voc. size.</li> <li>4) The participants did not have many chances to use English in real life situations, such as writing and speaking.</li> </ol>
	Synthesized results and meta-inferences	<ol style="list-style-type: none"> <li>1) The importance of vocabulary learning as perceived by these participants might motivate them to devote more efforts to learn English words. This may explain why the participants' vocabulary size was satisfactory.</li> <li>2) The participants had difficulties in learning depth of VK and they considered depth of VK more difficult to learn than voc. size. In addition, they did not have many chances to use English. These findings might explain why their depth of VK on WAT and productive VK were less satisfactory.</li> </ol>
The patterns of VLSs	QUAN results	<ol style="list-style-type: none"> <li>1) Most frequently reported VLSs were repetition, inferencing strategies, and using dictionaries.</li> <li>2) Least frequently reported VLSs were note-taking, production-based strategies, and association and imagination.</li> </ol>
	QUAL results	<ol style="list-style-type: none"> <li>1) More frequently mentioned VLSs were self-initiated strategies, using dictionaries, inferencing strategies, and repetition.</li> <li>2) Less frequently mentioned VLSs were note-taking, production-based strategies and association strategies.</li> </ol>
	Synthesized results and meta-inferences	The quantitative and qualitative results confirmed each other.
The relationship between reported VLSs and VK	QUAN results	<ol style="list-style-type: none"> <li>1) Inferencing strategies made a significant and positive contribution to the participants' voc. size, depth of VK and productive lexical diversity.</li> <li>2) Repetition and social strategies made a significant but negative contribution to the participants' voc. size.</li> <li>3) Self-initiated strategies significantly and positively correlated with some aspects of the participants' productive VK.</li> </ol>

	QUAL results	<ol style="list-style-type: none"> <li>1) High proficiency students believed that combining self-initiated strategies with using dictionaries and inferencing strategies would be helpful for enlarging voc. size.</li> <li>2) The intermediate proficiency students considered repetition strategies as useful for enlarging voc. size in addition to the perceived useful VLSs mentioned by high proficiency students.</li> <li>3) The low proficiency students did not consider using inferencing strategies as helpful for enlarging voc. size.</li> <li>4) Inferencing strategies were not suggested as useful for expanding the depth of VK by interviewees.</li> </ol>
	Synthesized results and meta-inferences	<ol style="list-style-type: none"> <li>1) The useful VLSs considered by the high proficiency students for enlarging voc. size elicited from the interviews were similar to the quantitative results.</li> <li>2) There was a gap between the interview-based elicited VLSs that perceived useful by both intermediate and low proficiency students for enlarging voc. size and their usefulness drawn from the quantitative findings.</li> <li>3) There was a discrepancy between the interview and quantitative data in terms of the useful VLSs for expanding depth of VK.</li> </ol>
	The relationship between VK and individual variables (gender, discipline, and proficiency levels)	<ol style="list-style-type: none"> <li>1) The voc. size of female students was significantly larger than that of male students.</li> <li>2) Science and engineering students significantly outperformed humanities and social sciences students in terms of their voc. size and depth of VK.</li> <li>3) High proficiency students performed significantly better in their voc. size, depth of VK, lexical diversity, lexical frequency and word familiarity than did intermediate and low proficiency students.</li> <li>4) High proficiency students performed significantly better than did low proficiency students in terms of word meaningfulness.</li> </ol>
	QUAL results	N/A
	Synthesized results and meta-inferences	N/A
	The relationship between reported VLSs and individual variables (gender, discipline, and proficiency levels)	<ol style="list-style-type: none"> <li>1) Female students reported using significantly more note-taking strategies than did male students.</li> <li>2) Humanities and social sciences students reported using significantly more note-taking strategies than did science and engineering students.</li> <li>3) High proficiency students reported using significantly more inferencing strategies than did low proficiency students.</li> </ol>
	QUAL results	<ol style="list-style-type: none"> <li>1) High proficiency students reported in particular more self-initiated strategies, dictionary strategies and inferencing strategies and they tended to combine these VLSs. By contrast, low proficiency students did not mention using any inferencing strategy.</li> <li>2) High proficiency students reported paying attention to vocabulary learning when using self-initiated strategies, such as reading novels, whereas low proficiency students did not</li> </ol>

<hr/>		
		report paying much attention to vocabulary learning.
		3) As for the knowledge sources of VLSs, high proficiency students combined their self-initiated strategies with their English teachers' advice. By contrast, most intermediate and low proficiency students mainly relied on advice from their English teachers or others.
	Synthesized results and meta-inferences	<p>1) High proficiency students reported using significantly more inferencing strategies than did low proficiency students. In addition, high proficiency students tended to combine self-initiated strategies with dictionary strategies and inferencing strategies.</p> <p>2) High proficiency students reported that they paid attention to word forms when using self-initiated strategies.</p> <p>3) High proficiency students reported that they combined their self-initiated strategies with their English teachers' advice to discover new VLSs, whereas most intermediate and low proficiency students mainly relied on advice from their English teachers or others. This finding may explain in part the reason why high proficiency students reported adopting more VLSs than both intermediate and low proficiency students.</p>
The contribution of the overall variables to VK	QUAN results	The overall predicting variables explained 13.2% of the variance in VST, 13.7% of WAT, 11.8% of lexical diversity, 7% of word familiarity, 6.7% of lexical frequency, and 4.9% of word meaningfulness.
	QUAL results	N/A
	Synthesized results and meta-inferences	N/A
<hr/>		



## **Chapter 6 Conclusions, Implications, Limitations, and Recommendations for Further Research**

This chapter concludes the thesis by firstly summarizing the quantitative and qualitative studies; secondly outlining the overall findings of the studies; thirdly discussing the empirical and pedagogical implications of these findings; fourthly explicating the limitations of the present study; and finally recommending future research in this strand of research.

### **6.1 Summary of the Quantitative and Qualitative Studies**

The main quantitative findings of the study can be summarized in dot-point formats as presented below:

- The receptive VK of the participants was 6,494 word family on VST and 98.07 on WAT.
- As for the participants' productive VK, they used less frequent and less familiar words compared to other studies and they seemed not good at using various words and phrases to express meaning in their writings.
- The most frequently reported use of VLSs by the participants were repetition, inferencing, and using dictionaries.
- The least frequently reported use of VLSs by the participants were note-taking, production-based strategies, and association and imagination.
- The overall predicting variables explained 13.2% of the variance in the participants' vocabulary size on VST, 13.7% of their depth of VK on WAT, 11.8% of lexical diversity, 7% of word familiarity, 6.7% of lexical frequency, and 4.9% of word meaningfulness of their English essays measured by Coh-Metrix.
- Proficiency level, gender, repetition and social strategies, inferencing strategies, and discipline made significant contributions to the participants' vocabulary size.
- Proficiency level and inferencing strategies made significant contributions to the participants' depth of VK.
- Proficiency level and inferencing strategies made significant contributions to lexical diversity of the participants' English essays. Only proficiency level made a significant

contribution to lexical frequency, word meaningfulness, and word familiarity of their essays.

The main qualitative findings of the study can be summarized as presented below:

- Participants indicated in their interviews that vocabulary learning was quite important for them to develop their four English communication skills.
- Participants encountered various difficulties in learning vocabulary. For instance, it was hard for them to remember English words for a long time, they had no efficient methods to remember new English words, and it was difficult to remember the long words, the infrequent words and those with irregular spellings.
- Participants indicated in the interviews that the VLSs they used more frequently were self-initiated strategies, using dictionaries, inferencing, and repetition strategies.
- Participants indicated in the interviews that the VLSs they used less frequently were note-taking, production-based strategies and association strategies.
- The interviews indicated that the high proficiency students reported adopting a wider range of VLSs than did both the intermediate and low proficiency students, especially more self-initiated strategies, using dictionaries and inferencing strategies.
- Participants' perspectives about vocabulary size included knowing the basic meaning, quantity, ease of learning, and foundation for English learning.
- Participants' perspectives about vocabulary depth included quality, difficult to learn, and knowing how to use a word.
- The interviews indicated that most high proficiency students considered vocabulary depth as more important than vocabulary size whereas most intermediate and low proficiency students considered vocabulary size as more important than vocabulary depth.

## **6.2 Conclusions: Main Findings**

### **6.2.1 Main Findings Related to VK**

The present study firstly showed that the participants' depth of VK was less satisfactory in comparison to their vocabulary size. Concerning the participants' productive VK, they used less frequent and less familiar words compared to other studies, however, they seemed not good at using various words and phrases to express meaning in their English writings. In addition, they used more concrete words and their cohesive writing skills were relatively low.

In the interviews, the participants reported that vocabulary depth was more difficult for them to learn than vocabulary size and that they did not have many chances to use English in real life contexts. The high proficiency students performed significantly better than both the intermediate and low proficiency students in terms of the vocabulary size and depth of VK and some aspects of productive VK. Most high proficiency students emphasized the significant role of vocabulary depth while most intermediate and low proficiency students regarded vocabulary size as more important than depth.

### **6.2.2 Main Findings Related to Reported VLSs**

This study explored a 7-factors structure for the VLS questionnaire by using exploratory factor analysis (EFA), encompassing inferencing strategies, association and imagination strategies, note-taking strategies, production-based strategies, dictionary strategies, repetition and social strategies and self-initiated strategies. It showed that the participants reported favoring repetition strategies, inferencing strategies and using dictionaries, but reported disfavoring note-taking strategies, production-based strategies and association and imagination strategies. The high proficiency students reported using more VLSs than did both intermediate and low proficiency students, in particular, they reported more inferencing strategies. In addition, the high proficiency students tended to combine self-initiated strategies, dictionary strategies, and inferencing strategies in vocabulary learning. Moreover, as for the participants' knowledge sources of strategy use, the high proficiency students combined their self-initiated strategies with their English teachers' advice. By contrast, most intermediate and low proficiency students mainly relied on advice from their English teachers or others.

### **6.2.3 Main Findings Related to the Relationship between VK and Reported VLSs**

Concerning the relationship between reported VLSs and VK, inferencing strategies made a significant and positive contribution to the participants' vocabulary size and depth of VK and productive lexical diversity, whereas repetition and social strategies made a significant but negative contribution to their vocabulary size. In addition, self-initiated strategies were found significantly and positively related to some aspects of their productive VK including lexical diversity, lexical frequency, word meaningfulness, and word imaginability.

The overall variables (i.e. gender, discipline, proficiency levels, and VLSs) explained 13.2% of the variance in the participants' vocabulary size. Among these predicting variables,

proficiency level was the strongest predictor, followed by gender, repetition and social strategies, inferencing strategies, and discipline. Moreover, the overall predicting variables explained 13.7% of the variance in the participants' depth of VK. Proficiency level still made the strongest unique contribution, followed by inferencing strategies. Concerning the participants' productive VK, the overall variables explained less of the variance, which was 11.8% in lexical diversity, 7% in word familiarity, 6.7% in lexical frequency, and 4.9% in word meaningfulness. As for lexical diversity, proficiency level was the strongest predictor, followed by inferencing strategies, whereas for the other three lexical indices, only proficiency level made a significant contribution.

### **6.3 Implications of the Study**

Firstly, this study provides valuable information concerning the status quo of the Chinese non-English major EFL students' VK. The participants' vocabulary size was satisfactory compared to other studies, but they still faced challenges in dealing with authentic English texts. As such, EFL teachers may select more simplified English reading and listening materials that match their students' current proficiency levels, such as graded readers or special Voice of America (VOA) listening materials, if their students find it difficult to understand authentic English texts. In addition, the participants' depth of VK was less satisfactory in comparison with other studies. Therefore, it would be desirable for teaching and learning material designers and EFL teachers to incorporate vocabulary depth into learning materials, teaching syllabuses and teaching activities in the classroom. In addition, EFL teachers are invited to inform EFL learners of the importance of vocabulary depth. Finally, the present study indicates that the participants may use less frequent and less familiar words in their English writings whereas they seem not good at using various words to express meaning. As such, EFL teachers are invited to inform their EFL learners of the unique contribution that lexical diversity makes to the students' writing performance and to provide necessary trainings to their students about how to increase their lexical diversity rather than simply encouraging learners to use less frequent and less familiar words in their writings. In addition, EFL learners are expected to improve their use of cohesive devices in their writing.

Secondly, the present study provides valuable information concerning the strategy use of Chinese non-English major EFL students when learning English vocabulary. Based on the findings of this study, it would be desirable for EFL teachers to help students see the relevance of strategy use to vocabulary learning, to introduce more strategies, such as some

production-based strategies and to give students more opportunities to try these strategies. In addition, it would be highly desirable for EFL teachers to provide more trainings on using inferencing strategies, and to encourage EFL learners to combine self-initiated strategies with inferencing strategies and using dictionaries. Nevertheless, it is worth noticing that VLSs training should be tailored to suit students' proficiency levels. Therefore, EFL teachers are invited to help EFL learners figure out their own suitable VLSs instead of blindly encouraging them to use the strategies adopted by high proficiency students. Moreover, EFL learners are expected to know more VLSs, to know the strategies which are useful for VK and to develop their own tailored VLSs.

#### **6.4 Limitations of the Study**

This study has several limitations. Firstly, although a relatively large sample was selected from four universities located in three different areas of mainland China, only one university is considered an ordinary university, while the other three universities are top universities among Chinese universities. Therefore, the sample may not truly represent Chinese EFL learners in Mainland China. In addition, the participants were chosen from among second-year and third-year university students, thus the findings may not be generalizable to students of other educational levels.

Secondly, the present study has adopted the CET-4 score as the criterion for identifying the participants' English language proficiency levels. While CET-4 is the most reliable and influential one in mainland China at the present time and all Chinese university students are obligated to take this test, it only consists of listening comprehension, reading comprehension and writing tests, excluding a speaking test (only a few students are eligible to sit for the CET spoken English test) (CET-SET). Accordingly, the scores of CET-4 may not have been able to precisely represent participants' overall English language proficiency.

Thirdly, the VLS questionnaire used in this study was only composed of strategies for learning the meaning and use of new words, whereas the strategies for learning word form were excluded. Therefore, the content of this questionnaire is not comprehensive in this regard. In addition, I borrowed and revised some items from the questionnaires proposed by Gu and Johnson (1996) and Schmitt (1997), which may have some impact on the construct of current VLS questionnaire.

Fourthly, the present study has adopted the bottom-up approach to the analysis of VLS questionnaire, that is, using factors resulting from exploratory factor analysis. The resulting

factors were meaningful in this study, by which I mean they could be largely interpreted in relation to established categories of VLS (e.g. dictionary, inferencing strategies). However, it is noteworthy that the loadings of some factors could be caused by chance or subjective interpretation of the researcher (Alderson & Banerjee, 1996, cited in Petric & Czarl, 2003; Gardner & Glikzman, 1982). Hence, those factors should be treated with caution.

Fifthly, this study has used the bilingual Mandarin version of VST. There has been no hard evidence in literature regarding the differences between results of monolingual and bilingual versions of VST. Therefore, the VST results of the present study cannot easily be compared with those of other studies which use a monolingual version of VST.

Finally, both the questionnaire and interview data were based on self-reports of the students. Despite the researcher's efforts to ensure the reliability of responses, such as informing participants that their answers would not affect their academic scores or making their responses anonymous, it is still questionable whether the self-report data were completely consistent with reality.

## **6.5 Recommendations for Future Research**

In the light of the limitations as indicated in the above section, this study provides several recommendations for future research.

Firstly, the participants of this study were mainly recruited from key universities in mainland China. Therefore, the findings based on the sample size are hardly generalizable. Future research with participants from various levels of universities in mainland China are definitely desirable. In addition, the participants of this study were chosen from among second-year and third-year university students. Future studies are invited to recruit participants from other educational levels.

Secondly, the VLS questionnaire of this study was only composed of strategies for learning the meaning and use of new words. It is worthwhile for future studies to add and include strategies for learning words' form and some other strategies through doing focus groups or interviews to develop a more comprehensive and tailored VLS questionnaire to a particular research context.

Thirdly, the present study has deployed the participants' CET-4 scores as the criterion to classify participants into three English language proficiency levels, however, CET-4 only consists of reading, listening and writing tests, excluding speaking test. Future research may

adopt other English language proficiency tests which could better precisely represent participants' overall English language proficiency.

Finally, the present results were based on the students' survey reports and their oral recall of vocabulary learning experience. These self-report data may not be consistent with reality. In order to achieve more reliable findings, future research may be beneficial in using ethnographic methods to observe the use of VLSs by L2 learners.



## References

- Ahmed, M. O. (1989). Vocabulary learning strategies. In P. Meara (Ed.), *Beyond words* (pp. 3-14). London: Centre for Information on Language Teaching and Research.
- Alderson, J. C., & Banerjee, J. (1996). *How might impact study instruments be validated?* Unpublished manuscript, IELTS Impact Study.
- Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77-117). Newark, DE: International Reading Association.
- Atkinson, R. C. (1975). Mnemotechnics in second-language learning. *American Psychologist*, 30(8), 821-828.
- Avila, E., & Sadoski, M. (1996). Exploring new applications of the keyword method to acquire English vocabulary. *Language Learning*, 46(3), 379-395.
- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (CD-ROM). Philadelphia, PA: University of Pennsylvania, Linguistic Data Consortium.
- Bacon, S., & Finnemann, M. D. (1992). Sex differences in self-reported beliefs about foreign-language learning and authentic oral and written input. *Language Learning*, 42(4), 471-495.
- Bartlett, J. E., Kotrlik, J. W., & Higgins, C. C. (2001). Organizational research: Determining appropriate sample size in survey research. *Information Technology, Learning, and Performance Journal*, 19(1), 43-50.
- Bartlett, M. S. (1954). A note on the multiplying factors for various chi square approximations. *Journal of the Royal Statistical Society*, 16 (2), 296-298.
- Bauer, L., & Nation, I. S. P. (1993). Word families. *International Journal of Lexicography*, 6(4), 253-279.
- Baxter, J. (1980). The dictionary and vocabulary behavior: A single word or a handful? *TESOL Quarterly*, 14(3), 325-336.
- Beglar, D. (2010). A Rasch-based validation of the Vocabulary Size Test. *Language Testing*, 27(1), 101-118.
- Bengeleil, N. F., & Paribakht, T. S. (2004). L2 reading proficiency and lexical inferencing by university EFL learners. *Canadian Modern Language Review/La Revue canadienne des langues vivantes*, 61(2), 225-250.
- Bialystok, E. (1990). *Communication strategies: A psychological analysis of second-language use*. Oxford: Basil Blackwell.
- Boyle, J. P. (1987). Sex differences in listening vocabulary. *Language Learning*, 37(2), 273-284.
- Brown, T. S., & Perry, F. L. (1991). A comparison of three learning strategies for ESL vocabulary acquisition. *TESOL Quarterly*, 25(4), 655-670.
- Bryman, A. (2008). *Social research methods* (3 ed.). Oxford: Oxford University Press.
- Catalán, R. M. J. (2003). Sex differences in L2 vocabulary learning strategies. *International Journal of Applied Linguistics*, 13(1), 54-77.
- Catalán, R. M. J. (2010). *Gender perspectives on vocabulary in foreign and second languages*. London: Palgrave Macmillan.

- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1(2), 245-276.
- Chan, A. Y. W. (2011). Bilingualised or monolingual dictionaries? Preferences and practices of advanced ESL learners in Hong Kong. *Language, Culture, and Curriculum*, 24(1), 1-21.
- Chan, A. Y. W. (2012). Cantonese ESL learners' use of grammatical information in a monolingual dictionary for determining the correct use of a target word. *International Journal of Lexicography*, 25(1), 68-94.
- Chen, H. (2001). The vocabulary learning strategies of Chinese non-English major students (非英语专业的中国学生学习英语词汇的策略). *Foreign Language Education (外语教学)*, 22(96), 46-50.
- Chen, Y. (2011). Studies on bilingualized dictionaries: The user perspective. *International Journal of Lexicography*, 24(2), 161-197.
- Cheng, L. (2008). The key to success: English language testing in China. *Language Testing*, 25(1), 15-37.
- Cheng, L., & Wang, H. (2004). Understanding professional challenges faced by Chinese teachers of English. *TESL-EJ*, 7(4), 1-14.
- Clark, H. H. (1970). Word associations and linguistic theory. In J. Lyons (Ed.), *New horizons in linguistics* (pp. 271-286). Harmondsworth: Penguin.
- Cochran, W. G. (1977). *Sampling techniques*. New York: John Wiley and Sons.
- Cohen, A. D. (1987). The use of verbal and imagery mnemonics in second-language vocabulary learning. *Studies in Second Language Acquisition*, 9(01), 43-61.
- Cohen, A. D. (1998). *Strategies in learning and using a second language*. London New York: Longman.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238.
- Cozby, P. C. (2009). *Methods of behavioral research* (10 ed.). New York, NY: McGraw-Hill.
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3 ed.). Los Angeles: Sage.
- Cronbach, L. J. (1942). An analysis of techniques for diagnostic vocabulary testing. *The Journal of Educational Research*, 36(3), 206-217.
- Crossley, S. A., Salsbury, T., & McNamara, D. S. (2009). Measuring L2 lexical growth using hypernymic relationships. *Language Learning*, 59(2), 307-334.
- Crossley, S. A., Salsbury, T., & McNamara, D. S. (2010a). The development of polysemy and frequency use in English second language speakers. *Language Learning*, 60(3), 573-605.
- Crossley, S. A., Salsbury, T., & McNamara, D. S. (2010b). The development of semantic relations in second language speakers: A case for Latent Semantic Analysis. *Vigo International Journal of Applied Linguistics*, 7, 55-74.
- Crossley, S. A., Salsbury, T., & McNamara, D. S. (2011). Predicting the proficiency level of language learners using lexical indices. *Language Testing*, 29(2), 243-263.
- Crossley, S. A., Salsbury, T., McNamara, D. S., & Jarvis, S. (2010). Predicting lexical proficiency in language learner texts using computational indices. *Language Testing*, 28(4), 1-20.
- D'Anna, C. A., Zechmeister, E. B., & Hall, J. W. (1991). Toward a meaningful definition of vocabulary size. *Journal of Literacy Research*, 23(1), 109-122.
- Dahlen, K., & Caldwell-Harris, C. (2013). Rehearsal and aptitude in foreign vocabulary learning. *The Modern Language Journal*, 97(4), 902-916.
- Dale, E. (1965). Vocabulary measurement: Techniques and major findings. *Elementary English*, 42(8), 895-901.

- De Bot, K., Paribakht, T. S., & Wesche, M. B. (1997). Toward a lexical processing model for the study of second language vocabulary acquisition. *Studies in Second Language Acquisition*, 19(03), 309-329.
- Deese, J. (1966). *The structure of associations in language and thought*. Baltimore: Johns Hopkins University Press.
- Deng, Z. C., & Zeng, Z. P. (1998). An investigation of English vocabulary size of undergraduates and teaching reflection (本科生英语词汇量调查与教学思考). *Foreign Language Teaching and Research (外语教学与研究)*, 10, 19-20.
- Dörnyei, Z. (2005). *The psychology of the language learner*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Dörnyei, Z. (2003). *Questionnaires in second language research: Construction, administration and processing*. London: Lawrence Erlbaum Associates.
- Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.
- Ellis, N. C. (1994). Consciousness in second language learning: Psychological perspectives on the role of conscious processes in vocabulary acquisition. *AILA Review*, 11, 37-57.
- Ellis, N. C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*, 43(4), 559-617.
- Espinosa, L. (2010). Assessment of young English-language learners. In E. E. Garcia & E. Frede (Eds.), *Young English-language learners: Current research and emerging directions for practice and policy*. New York: Teacher's College Press.
- Fan, M. Y. (2003). Frequency of use, perceived usefulness, and actual usefulness of second language vocabulary strategies: A study of Hong Kong learners. *The Modern Language Journal*, 87(2), 222-241.
- Fellbaum, C. (1998). *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.
- Folse, K. (2001). *A survey of ESL/EFL teacher preferences regarding student dictionary use*. Unpublished work.
- Fontecha, A. F. (2010). Gender and motivation in EFL vocabulary production. In R. M. J. Catalán (Ed.), *Gender perspectives on vocabulary in foreign and second languages* (pp. 93-116). London: Palgrave Macmillan.
- Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.
- Frantzen, D. (2003). Factors affecting how second language Spanish students derive meaning from context. *The Modern Language Journal*, 87(2), 168-199.
- Fraser, C. A. (1999). Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition*, 21(2), 225-241.
- Gardner, R. C., & Glikson, L. (1982). On "Gardner on affect": A discussion of validity as it relates to the attitude/motivation test battery: a response from Gardner. *Language Learning*, 32(1), 191-200.
- George, D., & Mallery, P. (2003). *SPSS for windows step by step: A simple guide and reference* (4 ed.). Boston, MA: Allyn and Bacon.
- Goh, C. C. M. (1998). How ESL learners with different listening abilities use comprehension strategies and tactics. *Language Teaching Research*, 2(2), 124-147.
- González, M. C. (2013). *The intricate relationship between measures of vocabulary size and lexical diversity as evidenced in non-native and native speaker academic compositions*. (PhD), University of Central Florida Orlando, Florida.
- Goulden, R., Nation, I. S. P., & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics*, 11(4), 341-363.

- Grace, C. A. (2000). Gender differences: Vocabulary retention and access to translations for beginning language learners in CALL. *The Modern Language Journal*, 84(2), 214-224.
- Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, and Computers*, 36(2), 193-202.
- Greidanus, T., & Nienhuis, L. (2001). Testing the quality of word knowledge in a second language by means of word associations: Types of distractors and types of associations. *The Modern Language Journal*, 85(4), 567-577.
- Gu, Y. (2002). Gender, academic major, and vocabulary learning strategies of Chinese EFL learners. *RELJ Journal*, 33(1), 35-54.
- Gu, Y. (2003a). Fine brush and freehand: The vocabulary-learning art of two successful Chinese EFL learners. *TESOL Quarterly*, 37(1), 73-104.
- Gu, Y. (2003b). Vocabulary learning in a second language: Person, task, context and strategies. *TESL-EJ*, 7(2), 1-25.
- Gu, Y. (2005). *Vocabulary learning strategies in the Chinese EFL context*. Singapore: Marshall Cavendish Academic.
- Gu, Y. (2010). Learning strategies for vocabulary development. *Reflections on English Language Teaching*, 9(2), 105-118.
- Gu, Y. (2012). Learning strategies: Prototypical core and dimensions of variation. *Studies in Self-Access Learning Journal*, 3(4), 330-356.
- Gu, Y., & Johnson, R. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning*, 46(4), 643-679.
- Haastrup, K. (1991). *Lexical inferencing procedures, or, talking about words: Receptive procedures in foreign language learning with special reference to English*. Tübingen: Gunter Narr Verlag.
- Halpern, D. F. (2010). *Sex differences in cognitive abilities* (3 ed.). Mahwah, NJ: Erlbaum.
- Hammadou, J. (1991). Interrelationships among prior knowledge, inference, and language proficiency in foreign language reading. *The Modern Language Journal*, 75(1), 27-38.
- Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition*, 21(02), 303-317.
- Henriksen, B., Albrechtsen, D., & Haastrup, K. (2004). The relationship between vocabulary size and reading comprehension in the L2. In D. Albrechtsen, K. Haastrup, & B. Henriksen (Eds.), *Angles on the English speaking world: Writing and vocabulary in foreign language acquisition* (pp. 129-147). Copenhagen: Museum Tusculanum Press.
- Hermann, F. (2003). Differential effects of reading and memorization of paired associates on vocabulary acquisition in adult learners of English as a second language. *TESL-EJ*, 7(1), 1-16.
- Hirsh, D., & Nation, I. S. P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, 8, 689-689.
- Hu, M., & Nassaji, H. (2012). Ease of inferencing, learner inferential strategies, and their relationship with the retention of word meanings inferred from context. *Canadian Modern Language Review*, 68(1), 54-77.
- Hu, M., & Nassaji, H. (2014). Lexical inferencing strategies: The case of successful versus less successful inferencers. *System*, 45, 27-38.
- Hu, M., & Nation, I. S. P. (2000). Vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403-430.
- Huckin, T., & Bloch, J. (1993). Strategies for inferring word-meanings in context: A cognitive model. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 153-178). Norwood, NJ: Ablex.

- Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 113-125). Basingstoke: Macmillan.
- Hulstijn, J. H. (1993). When do foreign-language readers look up the meaning of unfamiliar words? The influence of task and learner variables. *The Modern Language Journal*, 77(2), 139-147.
- Hulstijn, J. H. (1997). Mnemonic methods in foreign language vocabulary learning: Theoretical considerations and pedagogical implications. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 203-224). Cambridge: Cambridge University Press.
- Hulstijn, J. H. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258-286). Cambridge: Cambridge University Press.
- Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal*, 80(3), 327-339.
- Hummel, K. M. (2010). Translation and short-term L2 vocabulary retention: Hindrance or help? *Language Teaching Research*, 14(1), 61-74.
- Ishii, T., & Schmitt, N. (2009). Developing an integrated diagnostic test of vocabulary size and depth. *RELC Journal*, 40(1), 5-22.
- Joe, A. (1995). Text-based tasks and incidental vocabulary learning. *Second Language Research*, 11(2), 149-158.
- Joe, A. (1998). What effects do text-based tasks promoting generation have on incidental vocabulary acquisition? *Applied Linguistics*, 19(3), 357-377.
- Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika*, 35(4), 401-415.
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36.
- Kaivanpanah, S., & Moghaddam, M. S. (2012). Knowledge sources in EFL learners' lexical inferencing across reading proficiency levels. *RELC Journal*, 43(3), 373-391.
- Kaylani, C. (1996). The influence of gender and motivation on EFL learning strategy use in Jordan. In R. L. Oxford (Ed.), *Language learning strategies around the world: Cross-cultural perspectives* (pp. 75-88). Honolulu, HI: University of Hawaii Press.
- Kimura, D. (1999). *Sex and cognition*. Cambridge, MA: MIT press.
- Knight, S. (1994). Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. *The Modern Language Journal*, 78(3), 285-299.
- Kojic-Sabo, I., & Lightbown, P. M. (1999). Students' approaches to vocabulary learning and their relationship to success. *The Modern Language Journal*, 83(2), 176-192.
- Kucera, H., & Francis, W. N. (1967). *A computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Landauer, T., McNamara, D. S., Dennis, S., & Kintsch, W. (2007). *LSA: A road to meaning*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. J. L. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 126 -132). London: Macmillan.
- Laufer, B. (2001). Quantitative evaluation of vocabulary: How it can be done and what it is good for. In C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, T. McNamara, & K. O'Loughlin (Eds.), *Experimenting with uncertainty* (pp. 241-250). Cambridge: Cambridge University Press.

- Laufer, B., & Hadar, L. (1997). Assessing the effectiveness of monolingual, bilingual, and "bilingualised" dictionaries in the comprehension and production of new words. *The Modern Language Journal*, 81(2), 189-196.
- Laufer, B., & Kimmel, M. (1997). Bilingualised dictionaries: How learners really use them. *System*, 25(3), 361-369.
- Laufer, B., & Nation, I. S. P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307-322.
- Laufer, B., & Nation, I. S. P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, 16(1), 33-51.
- Lawson, M. J., & Hogben, D. (1996). The vocabulary-learning strategies of foreign-language students. *Language Learning*, 46(1), 101-135.
- Lawson, M. J., & Hogben, D. (1998). Learning and recall of foreign-language vocabulary: Effects of a keyword strategy for immediate and delayed recall. *Learning and Instruction*, 8(2), 179-194.
- Lessard-Clouston, M. (1996). ESL vocabulary learning in a TOEFL preparation class: A case study. *Canadian Modern Language Review*, 53(1), 97-119.
- Lessard-Clouston, M. (2008). Strategies and success in technical vocabulary learning: Students' approaches in one academic context. *Indian Journal of Applied Linguistics*, 34, 31-63.
- Lessard-Clouston, M. (2012). Technical vocabulary use in English-medium disciplinary writing: An L1/L2 case study. *The Linguistics Journal*, 6, 127-150.
- Liu, N., & Nation, I. S. P. (1985). Factors affecting guessing vocabulary in context. *REL C Journal*, 16(1), 33-42.
- Lotto, L., & De Groot, A. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning*, 48(1), 31-69.
- Lupescu, S., & Day, R. R. (1993). Reading, dictionaries, and vocabulary Learning. *Language Learning*, 43(2), 263-287.
- Ma, G. H. (2001). A study of English vocabulary proficiency of science and engineering majors (理工科大学生英语词汇水平研究). *Foreign Language Education (外语教学)*, 22(2), 48-52.
- Macaro, E. (2006). Strategies for language learning and for language use: Revising the theoretical framework. *The Modern Language Journal*, 90(3), 320-337.
- McCarthy, P. M., & Jarvis, S. (2010). MTL D, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42(2), 381-392.
- Meara, P. (1980). Vocabulary acquisition: A neglected aspect of language learning. *Language Teaching*, 13(3-4), 221-246.
- Meara, P. (1982). Word associations in a foreign language. *Nottingham Linguistic Circular Nottingham*, 11(2), 29-38.
- Meara, P. (1990). Some notes on the Eurocentres vocabulary tests. In J. Tömmola (Ed.), *Foreign language comprehension and production* (pp. 103-113). Turku: AFinLa Yearbook.
- Meara, P. (1992). *EFL vocabulary tests*. Swansea: Centre for Applied Language Studies, University of Wales, Swansea.
- Meara, P. (1996a). The dimensions of lexical competence. In G. Brown, K. Malmkjaer, & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 35-53). Cambridge: Cambridge University Press.
- Meara, P. (1996b). The vocabulary knowledge framework. Retrieved 14/Oct., 2012, from <http://www.lognostics.co.uk/vlibrary/index.htm>

- Meara, P., & Bell, H. (2001). P-Lex: A simple and effective way of describing the lexical characteristics of short L2 tests. *Prospect*, 16(3), 5-19.
- Meara, P., & Buxton, B. (1987). An alternative to multiple choice vocabulary tests. *Language Testing*, 4(2), 142-154.
- Meara, P., & Jones, G. (1988). Vocabulary size as a placement indicator. In P. Grunwell (Ed.), *Applied linguistics in society* (pp. 80-87). London: CILT.
- Meara, P., & Milton, J. L. (2003). *X\_Lex: the Swansea Vocabulary Levels Test*. Newbury: Express Publishing.
- Meara, P., & Wolter, B. (2004). V\_Links: Beyond vocabulary depth. In D. Albrechtsen, K. Haastup, & B. Henriksen (Eds.), *Angles on the English speaking world* (pp. 85-96). Copenhagen: Museum Tusculanum Press.
- Mondria, J.-A., & Wit-de Boer, M. (1991). The effects of contextual richness on the guessability and the retention of words in a foreign language. *Applied Linguistics*, 12(3), 249-267.
- Moore, J. C., & Surber, J. R. (1992). Effects of context and keyword methods on second language vocabulary acquisition. *Contemporary Educational Psychology*, 17(3), 286-292.
- Morris, L. A. (1998). Differences in men's and women's ESL writing at the junior college level: Consequences for research on feedback. *Canadian Modern Language Review/La Revue canadienne des langues vivantes*, 55(2), 219-238.
- Nassaji, H. (2003). L2 vocabulary learning from context: Strategies, knowledge sources, and their relationship with success in L2 lexical inferencing. *TESOL Quarterly*, 37(4), 645-670.
- Nassaji, H. (2006). The relationship between depth of vocabulary knowledge and L2 learners' lexical inferencing strategy use and success. *The Modern Language Journal*, 90(3), 387-401.
- Nation, I. S. P. (1983). Testing and teaching vocabulary. *Guidelines*, 5(1), 12-25.
- Nation, I. S. P. (1986). *Vocabulary lists: Words, affixes, and stems*. English Language Institute, Victoria University of Wellington.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York: Newbury House Publishers.
- Nation, I. S. P. (1993). Using dictionaries to estimate vocabulary size: essential, but rarely followed, procedures. *Language Testing*, 10(1), 27-40.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review/La Revue canadienne des langues vivantes*, 63(1), 59-82.
- Nation, I. S. P. (2012). The Vocabulary Size Test. Retrieved 15 December 2014, from <http://www.victoria.ac.nz/lals/about/staff/paul-nation>
- Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9-13.
- Nation, I. S. P., & Coady, J. (1988). Vocabulary and reading. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 97-110). London: Longman.
- Nation, I. S. P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6-19). Cambridge: Cambridge University Press.
- Nguyen, L. T. C., & Nation, P. (2011). A bilingual vocabulary size test of English for Vietnamese learners. *RELC Journal*, 42(1), 86-99.
- Nyikos, M. (1990). Sex-related differences in adult language learning: Socialization and memory factors. *The Modern Language Journal*, 74(3), 273-287.

- Oller, J. W. (1982). Gardner on affect: A reply to Gardner. *Language Learning*, 32(1), 183-189.
- O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.
- Onwuegbuzie, A. J., & Johnson, R. B. (2004). Mixed method and mixed model research. In B. Johnson & L. Christensen (Eds.), *Educational research: Quantitative, qualitative, and mixed approaches* (pp. 408-431).
- Oxford, R. L. (1989). Use of language learning strategies: A synthesis of studies with implications for strategy training. *System*, 17(2), 235-247.
- Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. New York: Newbury House Publishers.
- Oxford, R. L. (2013). *Teaching and researching: Language learning strategies*. New York: Routledge.
- Oxford, R. L., & Nyikos, M. (1989). Variables affecting choice of language learning strategies by university students. *The Modern Language Journal*, 73(3), 291-300.
- Oxford, R. L., Nyikos, M., & Ehrman, M. (1988). Vive la difference? Reflections on sex differences in use of language learning strategies. *Foreign Language Annals*, 21(4), 321-329.
- Pallant, J. (2013). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS* (5 ed.). Sydney, Melbourne, Auckland, London: Allen & Unwin.
- Paribakht, T. S. (2005). The influence of first language lexicalization on second language lexical inferencing: A study of Farsi-speaking learners of English as a foreign language. *Language Learning*, 55(4), 701-748.
- Paribakht, T. S., & Wesche, M. B. (1993). Reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada Journal*, 11(1), 9-29.
- Paribakht, T. S., & Wesche, M. B. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 174-200). Cambridge: Cambridge University Press.
- Paribakht, T. S., & Wesche, M. B. (1999). Reading and "incidental" L2 vocabulary acquisition: An introspective study of lexical inferencing. *Studies in Second Language Acquisition*, 21(2), 195-224.
- Parry, K. (1993). Too many words: Learning the vocabulary of an academic subject. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 109-129). Norwood, NJ: Ablex.
- Petric, B., & Czarl, B. (2003). Validating a writing strategy questionnaire. *System*, 31, 187-215.
- Postman, L., & Keppel, G. (1970). *Norms of word association*. New York: Academic Press.
- Prados, M. D. (2010). Gender and L1 influence on EFL learners' lexicon. In R. M. J. Catalán (Ed.), *Gender perspectives on vocabulary in foreign and second languages*. London: Palgrave Macmillan.
- Prince, P. (1996). Second language vocabulary learning: The role of context versus translations as a function of proficiency. *The Modern Language Journal*, 80(4), 478-493.
- Pulido, D. (2007). The effects of topic familiarity and passage sight vocabulary on L2 lexical inferencing and retention through reading. *Applied Linguistics*, 28(1), 66-86.
- Pulido, D. (2009). How involved are American L2 learners of Spanish in lexical input processing tasks during reading? *Studies in Second Language Acquisition*, 31(01), 31-58.

- Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review/La Revue canadienne des langues vivantes*, 56(2), 282-308.
- Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, 52(3), 513-536.
- Qian, D. D. (2005). Demystifying lexical inferencing: The role of aspects of vocabulary knowledge. *TESL Canada Journal*, 22(2), 34-54.
- Read, J. (1988). Measuring the vocabulary knowledge of second language learners. *REL C Journal*, 19(2), 12-25.
- Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing*, 10(3), 355-371.
- Read, J. (1998). Validating a test to measure depth of vocabulary knowledge. In A. J. Kunnan (Ed.), *Validation in language assessment: Selected papers from the 17th Language Testing Research Colloquium, Long Beach* (pp. 41-60). Mahwah, New Jersey: Lawrence Erlbaum.
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge university press
- Riazi, A. M., & Alviri, A. (2004). Strategy activation in learning English words. *Academic Exchange Quarterly*, 8(2), 199.
- Riazi, A. M., & Candlin, C. N. (2014). Mixed-methods research in language teaching and learning: Opportunities, issues and challenges. *Language Teaching*, 47(02), 135-173.
- Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quarterly*, 10(1), 77-89.
- Rodriguez, M., & Sadoski, M. (2000). Effects of rote, context, keyword, and context/keyword methods on retention of vocabulary in EFL classrooms. *Language Learning*, 50(2), 385-412.
- Sagarra, N., & Alba, M. (2006). The key is in the keyword: L2 vocabulary learning methods with beginning learners of Spanish. *The Modern Language Journal*, 90(2), 228-243.
- Salsbury, T., Crossley, S. A., & McNamara, D. S. (2011). Psycholinguistic word information in second language oral discourse. *Second Language Research*, 27(3), 343-360.
- Sanaoui, R. (1992). *Vocabulary learning and teaching in French as a second language classrooms*. (Unpublished doctoral dissertation), University of Toronto, Toronto, Ontario, Canada.
- Sanaoui, R. (1995). Adult learners' approaches to learning vocabulary in second languages. *The Modern Language Journal*, 79(1), 15-28.
- Scarcella, R., & Zimmerman, C. (1998). Academic words and gender. *Studies in Second Language Acquisition*, 20(1), 27-49.
- Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129-158.
- Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 199-227). Cambridge: Cambridge University Press.
- Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language Learning*, 48(2), 281-317.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Basingstoke, England: Palgrave Macmillan.
- Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework. *Studies in Second Language Acquisition*, 19(01), 17-36.
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, 18(1), 55-88.

- Schmitt, N., & Zimmerman, C. B. (2002). Derivative word forms: What do learners know? *TESOL Quarterly*, 36(2), 145-171.
- Schoonen, R., & Verhallen, M. (2008). The assessment of deep word knowledge in young first and second language learners. *Language Testing*, 25(2), 211-236.
- Schouten-van Parreren, C. (1989). Vocabulary learning through reading: Which conditions should be met when presenting words in texts. *AILA Review*, 6(1), 75-85.
- Shen, M.-Y. (2010). Effects of perceptual learning style preferences on L2 lexical inferencing. *System*, 38(4), 539-547.
- Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, 36(2), 139-152.
- Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3 ed.). Boston: Houghton Mifflin.
- Stoffer, I. (1995). *University foreign language students' choice of vocabulary learning strategies as related to individual difference variables*. (Unpublished doctoral dissertation), University of Alabama.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6 ed.). Boston: Pearson Education.
- Takac, V. P. (2008). *Vocabulary learning strategies and foreign language acquisition*. Clevedon: Multilingual Matters.
- Tashakkori, A., & Teddlie, C. (2003). *Handbook on mixed methods in the behavioral and social sciences*. Thousand Oaks, CA: Sage.
- Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York: Teacher's College, Columbia University.
- Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago: University of Chicago Press.
- Tseng, W. T., & Schmitt, N. (2008). Toward a model of motivated vocabulary learning: A structural equation modeling approach. *Language Learning*, 58(2), 357-400.
- Van Hell, J. G., & Mahn, A. C. (1997). Keyword mnemonics versus rote rehearsal: Learning concrete and abstract foreign words by experienced and inexperienced learners. *Language Learning*, 47(3), 507-546.
- Vidal, K. (2003). Academic listening: A source of vocabulary acquisition? *Applied Linguistics*, 24(1), 56-89.
- Wang, A. Y., & Thomas, M. H. (1992). The effect of imagery-based mnemonics on the long-term retention of Chinese characters. *Language Learning*, 42(3), 359-376.
- Wei, Z. (2015). Does teaching mnemonics for vocabulary learning make a difference? Putting the keyword method and the word part technique to the test. *Language Teaching Research*, 19(1), 43-69.
- Wen, Q., & Johnson, R. K. (1997). L2 learner variables and English achievement: A study of tertiary-level English majors in China. *Applied Linguistics*, 18(1), 27-48.
- Wesche, M. B., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. *Canadian Modern Language Review*, 53(1), 13-40.
- West, M. (1953). *A general service list of English words*. London: Longman, Green.
- Wilkins, D. A. (1972). *Linguistics in language teaching*. London: Edward Arnold.
- Wilson, M. (1988). The MRC psycholinguistic database: Machine readable dictionary. *Behavioral Research Methods, Instruments and Computers*, 20(1), 6-11.
- Xue, G., & Nation, I. S. P. (1984). A university word list. *Language Learning and Communication*, 3(2), 215-229.
- Zhang, B., & Li, C. (2011). Classification of L2 vocabulary learning strategies: Evidence from exploratory and confirmatory factor analyses. *RELC Journal*, 42(2), 141-154.

## **Appendix A**

### **The Vocabulary Learning Strategy Questionnaire**

Dear students,

We would greatly appreciate your participation in this survey which is concerned with how you learn English words. The study is being conducted to meet the requirements of the Doctor of Philosophy Degree under the supervision of A/P Mehdi Riazi and Ms Jean Brick of the Department of Linguistics of Macquarie University. The purpose of the study is to better understand the Chinese non-English majors' vocabulary knowledge in relation to the variables, such as the academic discipline, gender, English proficiency level and vocabulary learning behaviors. This questionnaire will take approximately 30 minutes of your time.

This is not a test so there are no “right” or “wrong” answers. Please choose the answer that exactly describes your vocabulary learning habit. Your answers will be kept strictly confidential and will not affect your academic scores, so please try to choose answers as accurately as possible. The success of the research highly depends on your true answers. If you wish to be informed about the results of our research, please leave your email address at the end of the questionnaire.

Thank you very much for your cooperation.

Na Fan, PhD student

A/P Mehdi Riazi and Ms. Jean Brick (supervisors)

## The participants' demographic information

Name \_\_\_\_\_ Gender \_\_\_\_\_ Major \_\_\_\_\_

CET-4 score \_\_\_\_\_ Grade \_\_\_\_\_

### Instruction:

The following is a list of the things students usually do in order to learn English words. We would like to know what you actually do, NOT what you should do or want to do. We would like you to indicate how often you have used a certain action by putting an “X” in the box. If you do not use the action at all, please mark the word *never*. If you use it, please mark one of the options, *seldom*, *sometimes*, *often* or *always*, depending on how often you use it. Thank you very much for your help.

### Levels of Your Own Vocabulary Learning Actions

“**Always**” means that you use the action 100% of the time.

“**Often**” means that you use the action 75% of the time.

“**Sometimes**” means that you use the action 50% of the time.

“**Seldom**” means that you use the action 25% of the time.

“**Never**” means that you use the action 0% of the time.

**Example: If you always use the following action, please mark as**

Strategy	Never	Seldom	Sometimes	Often	Always
I guess the meaning of the word from its part of speech (e.g. whether it is a noun, a verb, an adjective, or an adverb)					X

## Part 1: Strategies for learning the meanings of new words

	Levels of Your Own Vocabulary Learning Actions				
	Never	Seldom	Sometimes	Often	Always
I guess the meaning of a new word from its part of speech (e.g., whether it is a noun, a verb, an adjective, or an adverb).					
I guess the meaning of a new word from its affixes and roots (e.g., replay – “re” means “do it again”).					
I guess the meaning of a new word from its textual context (e.g., surrounding words).					
I guess the meaning of a new word from real situations in daily life (e.g., guessing from some traffic signs).					
I guess the meaning of a new word from my common sense and general information.					
I use an English to Chinese dictionary to learn the meaning of a new word.					
I use an English to English dictionary to learn the meaning of a new word.					
I pay attention to various meanings of a new word when I look it up in the dictionary.					
I take down the meanings of a new word when I look it up in the dictionary.					
I use the translation of the sentence (e.g., whether it is in a bilingual advertising, or English film with Chinese scripts etc.) in which the word embedded to learn the meaning of a new word.					
When I don't know the meaning of a new word, I ask others (e.g., my English teachers or friends) for its meaning.					
I make a list of new words to remember their meanings.					
I try to group new words together within a storyline to remember the meanings of new words (e.g., I make up a story by using the new words).					
I try to use semantic maps to remember the meanings of new words (e.g., whether they are synonyms or antonyms etc.).					
I try to attach physical sensations to certain words (e.g., relating them to senses) to remember their meanings.					
I try to create an imaginary context to remember the meaning of a new word (When remembering the word “scissor”, I try to imagine the context where I want to cut sth into pieces).					
I try to connect a new word to a personal experience to remember its meaning (e.g., when I remember the word “happy”, I connect the word to the happy things I've					

experienced).					
I try to remember the meaning of a new word together with the sentence in which the word is used.					
I try to remember the meaning of a new word in a short phrase or an expression.					
I try to link new words to the physical objects or physical actions to remember their meanings (e.g., when I try to remember the word “camera”, I link the word “camera” to my camera.).					
I try to use affixes and roots to remember the meaning of a new word.					
I try to paraphrase or translate a new word to remember its meaning.					
I try to write new words and their meanings on small pieces of paper, and put them around my room to remember the meanings of new words.					
I try to test myself with word tests to remember the meanings of new words.					

## Part 2: Strategies for learning how to use new words

	Levels of Your Own Vocabulary Learning Actions				
	Never	Seldom	Sometimes	Often	Always
I try to orally repeat the new word, the word with its collocations, or the whole sentence in which the new word is used.					
I try to write several times of the new word, the word with its collocations, or the whole sentence in which the new word is used.					
I pay attention to the stylistic features of a new word (e.g., whether it is formal, informal; whether it is negative or positive) when I look it up in the dictionary.					
I pay attention to the grammatical information of a new word (e.g., whether it is a countable noun or an uncountable noun) when I look it up in the dictionary.					
I pay attention to the collocations of a new word (What words it typically occurs with? e.g., fast food; united states; commit murder) when I look it up in the dictionary.					
I take down the grammatical information of a new word when I look it up in the dictionary.					
I take down the collocations of a new word when I look it up in the dictionary.					
I take down example sentences of a new word when I look it up in the dictionary.					
I make a note when I see a useful expression or phrase.					
I pay attention to how words are used when I attend English lectures or presentations.					
I pay attention to how words are used when I listen to English radio program.					
I pay attention to how words are used when I watch English movies or English TV program.					
I pay attention to how words are used when I read English novels.					
I pay attention to how words are used when I read English newspapers.					
I try to use new words as much as possible in my speaking.					
I try to use new words as much as possible in my writing (e.g., in my diaries).					
I try to use different meanings of a new word in different contexts.					
I try to use some synonyms of a new word in my speaking or writing.					
I try to make up my own sentences using the new words I just learned.					
When I don't know how to use a word, I look it up in the dictionary.					
I ask others (e.g., my English teachers or friends) for how to use a word.					

I try to use the collocations of the new words in my speaking.					
I try to use the collocations of the new words in my writing.					
When I want to use a word, I pay attention to its grammatical features.					
When I want to use a word, I pay attention to whether it is frequently used in this type of context.					
When I want to use a word, I will choose formal or informal words according to the contexts.					
I try to use a new word in an imaginary context.					
I guess the words I think frequently co-occur with the new word in my speaking or writing.					

**If you used other vocabulary learning strategies which are not included in the above questionnaire, please indicate them in the following blank box.**

**The end.**

**Thanks so much again for your participation!**

## Appendix B

### 英语词汇学习策略调查问卷

亲爱的各位同学，

非常感谢你们参加本次英语词汇学习方法问卷研究。本研究是在澳大利亚麦考瑞大学语言学系 A/P Mehdi Riazi 和 Ms. Jean Brick 指导下的博士学位论文研究，旨在探讨中国非英语专业学生词汇知识及其与专业、性别、英语水平和词汇学习方法之间的关系。完成本问卷大概需要 40 分钟。

本问卷不同于测试，因此没有正确或者错误的答案。请选择出能够真实反映你们词汇学习方法的答案。你们的选择将会被严格的保密，并且不会影响到你们考试成绩，所以请尽量选出最准确的答案。你们答案的真实性将会直接影响到本次研究结果的有效性。如果你们对本次研究结果感兴趣，请把你们的电子邮箱地址写在问卷的最后空白处。

非常感谢你们的合作。

范娜，在读博士研究生

导师：A/P Mehdi Riazi 和 Ms. Jean Brick

### 基本信息:

姓名\_\_\_\_\_ 性别\_\_\_\_\_ 专业\_\_\_\_\_

四级成绩\_\_\_\_\_ 年级\_\_\_\_\_

以下列表是学习英语词汇时经常用到的一些方法。我们想要了解的是你们的真实做法，而不是你们认为应该如何做或者想要如何做。请在相应的空格中打“X”来表示你是否会经常使用表格中描述的方法。如果你从来没有用过这种方法，请选择“从不”。如果你平时用到所描述的方法，请根据其使用的频率相应地选择“很少”“有时”“经常”“总是”。感谢你们的合作。

#### 词汇学习方法频率等级

“总是”代表使用这个方法频率在 100% 左右。

“经常”代表使用这个方法频率在 75%左右。

“有时”代表使用这个方法频率在 50%左右。

“很少”代表使用这个方法频率在 25%左右。

“从不”代表使用这个方法频率在 0%左右。

例如：如果你总是使用以下方法，请在空格内打 X

词汇学习方法	从不	很少	有时	经常	总是
我会根据单词的词性猜测单词的意思（例如：根据这个单词是名词，动词，形容词或者副词来猜测它的意思）					X

## 第一部分：学习新单词意思的方法

	词汇学习方法使用频率等级				
	从不	很少	有时	经常	总是
我会根据新单词的词性猜测新单词的意思（例如：根据这个单词是名词，动词，形容词或者副词来猜测它的意思）。					
我会根据新单词的前缀或者后缀猜测新单词的意思（例如：replay- 前缀“re”的意思是重新）。					
我会根据文章上下文猜测新单词的意思。					
我会从新单词出现的真实场景中猜测新单词的意思（例如：从一些交通标志图片来猜测单词的意思）。					
我会根据自己的一些基本常识来猜测新单词的意思。					
我会用英汉词典学习新单词的意思。					
我会用英英词典学习新单词的意思。					
当我查词典时，我会注意词典中单词的各种不同意思。					
当我查词典时，我会记下词典中单词的各种不同意思。					
我会从中英文对照的句子中学习新单词的意思（例如：具有中文翻译的英文广告，具有中文字幕的英文电影等）。					
当我不知道单词意思的时候，我会向其他人询问单词的意思（例如：我的英语老师或者朋友）。					
我把新单词制作成单词表来记忆新单词意思。					
我会把新学到的单词编在一个故事里来记忆新单词意思。					
我会根据单词的语义关系来记忆新单词意思（例如：同义词，反义词等）。					
我会用感官动作记忆新单词意思（例如：把新单词和嗅觉，听觉等联系起来）。					
我会想象出一个可以使用新单词的场景来记忆新单词意思（例如：当我记忆“剪刀”这个单词时，我会想象出一个我想要剪东西的场景）。					
我会把新单词和自身经历结合起来记忆单词意思（例如：当学习“happy”这个词时，我会联想到以前一些快乐的事情）。					
我会把新单词连同它出现的句子一同记忆从而记忆新单词的意思。					
我会把新单词放在词组或者短语中来记忆新单词的意思。					
我会把新单词和现实事物或者肢体动作联系起来记忆新单词的意思（例如：当我学习“camera”这个词，我会想到照相机）。					
我会通过新单词的词根和词缀来记忆新单词意思。					
我会通过翻译或者用英文解释新单词的意思来记忆新单词意思。					
我会把写有新单词和新单词意思的纸条贴在房间里来记忆新单词意思。					
我会进行自我词汇测试来记忆新单词意思。					

## 第二部分：学习新单词如何使用的方法

	词汇学习方法使用频率等级				
	从不	很少	有时	经常	总是
我会口头重复新单词，以及包含新单词的短语或者句子。					
我会重复书写新单词，以及包含新单词的短语或者句子。					
当我查词典时，我会注意新单词的文体特征（例如：这个单词是否正式，是褒义词还是贬义词）。					
当我查词典时，我会注意新单词的语法信息（例如：这个单词是可数名词还是不可数名词）。					
当我查词典时，我会注意新单词的搭配（这个单词经常和哪些单词连用？例如：fast food; united states; commit murder）。					
当我查词典时，我会把新单词的语法信息记在笔记本上。					
当我查词典时，我会把新单词的搭配记在笔记本上。					
当我查词典时，我会把新单词的例句记在笔记本上。					
当我看到一个有用的搭配或者词组时，我会把它记在笔记本上。					
当我参加一些英文课程，英文报告时，我会注意单词的用法。					
当我听英文广播时，我会注意单词的用法。					
当我看英文电影或者英文电视节目时，我会注意单词的用法。					
当我读英文小说时，我会注意单词的用法。					
当我读英文报纸时，我会注意单词的用法。					
我会尽量在口语中使用新单词。					
我会尽量在写作中使用新单词（例如：我的日记）。					
我会在不同语境中使用新单词的不同意思。					
我会尽量在口语或者写作中使用新单词的同义词。					
我会用新单词造个句子。					
当我不知道单词应该如何使用时，我会查词典。					
当我不知道单词应该如何使用时，我会询问他人（例如：我的英语老师或者朋友）。					
我会在口语中尽量用到新单词的搭配。					
我会在写作中尽量用到新单词的搭配。					
当我使用一个单词时，我会注意它的语法信息。					
当我使用一个单词时，我会注意它在这种语境中是否常用。					

当我使用一个单词时，我会根据不同语境选择正式或者非正式的单词。					
我会想象出一个语境使用新单词。					
我会在口语或者写作中猜测我认为经常和新单词搭配的一些单词。					

如果您用到其它不包括在以上问卷中的英语词汇学习方法，请写在以下空白方格中。

问卷到此结束，再次感谢您的参与！

## Appendix C

### Word Associates Test

Name: \_\_\_\_\_

**Example:**

**Sudden**

beautiful   quick   surprising   thirsty	change   doctor   noise   school
--	----------------------------------

The words here on the left side may help to explain the meaning of “sudden”. (左边方框里的词是有可能解释“sudden”的词)	The words here on the right side are nouns that may come after “sudden” in a phrase or a sentence. (右边方框里的词是有可能和“sudden”搭配的词)
--	---

There are eight words in the box, but only **four** of them are correct. You have to choose which are the four correct words.

“Sudden” means “happening quickly and unexpectedly”, so the correct answers on the left side are “quick” and “surprising”.	We do not normally say “a sudden doctor” or “a sudden school”, but we often say “a sudden change” and “a sudden noise”, so “change” and “noise” are the correct answers on this side.
--	---

说明：方框中的八个词中一共有四个词是正确选项。在四个正确选项中，有可能左边方框出现两个，右边方框出现两个；或者左边方框出现一个，右边方框出现三个；或者左边方框出现三个，右边方框出现一个。请在正确选项上打√。请选出尽量多的你认为正确的选项。

**1 beautiful**

enjoyable expensive free loud	education face music weather
-------------------------------	------------------------------

**2 bright**

clever famous happy shining	colour hand poem taste
-----------------------------	------------------------

**3 calm**

open quiet smooth tired	cloth day light person
-------------------------	------------------------

**4 natural**

expected helpful real short	foods neighbours parents songs
-----------------------------	--------------------------------

**5 fresh**

another cool easy raw	cotton heat language water
-----------------------	----------------------------

**6 general**

closed different usual whole	country idea reader street
------------------------------	----------------------------

**7 bare**

empty heavy uncovered useful	cupboard feet school tool
------------------------------	---------------------------

**8 acute**

hidden often rich sharp	angle hearing illness stones
-------------------------	------------------------------

**9 common**

complete light ordinary shared	boundary circle name party
--------------------------------	----------------------------

**10 complex**

angry difficult necessary sudden	argument passengers patterns problem
----------------------------------	--------------------------------------

**11 broad**

full moving quiet wide	night river shoulders smile
------------------------	-----------------------------

**12 conscious**

awake healthy knowing laughing	face decision effort student
--------------------------------	------------------------------

**13 convenient**

easy fresh near suitable	experience sound time vegetable
--------------------------	---------------------------------

**14 dense**

crowded hot noisy thick	forest handle smoke weather
-------------------------	-----------------------------

**15 curious**

helpful interested missing strange	accident child computer steel
------------------------------------	-------------------------------

**16 distinct**

clear famous separate true	advantage meanings news parents
----------------------------	---------------------------------

**17 dull**

cloudy loud nice secret	colour knife place rock
-------------------------	-------------------------

**18 direct**

honest main straight wide	fence flight heat river
---------------------------	-------------------------

**19 favourable**

helpful legal possible positive	habit response teacher weather
---------------------------------	--------------------------------

**20 secure**

confident enjoyable fixed safe	game job meal visitor
--------------------------------	-----------------------

**21 tight**

close rough uncomfortable wet	bend pants surface wood
-------------------------------	-------------------------

**22 violent**

expected smelly strong unlucky	anger death rubbish storm
--------------------------------	---------------------------

**23 chronic**

continuing local serious unplanned	accident examination illness shortage
------------------------------------	---------------------------------------

**24 compact**

effective small solid useful	group kitchen medicine string
------------------------------	-------------------------------

**25 crude**

clever fair rough valuable	behaviour drawing oil trade
----------------------------	-----------------------------

**26 domestic**

home national regular smooth	animal movement policy speed
------------------------------	------------------------------

**27 profound**

bright deep exact great	effect machine taste thought
-------------------------	------------------------------

**28 fertile**

dark growing private special	business egg mind soil
------------------------------	------------------------

**29 formal**

fast loud organised serious	bomb education growth statement
-----------------------------	---------------------------------

**30 independent**

changed equal important separate	child country ideas prices
----------------------------------	----------------------------

**31 original**

careful closed first proud	condition mind plan sister
----------------------------	----------------------------

**32 sensitive**

feeling interesting sharp thick	clothes instrument skin topic
---------------------------------	-------------------------------

**33 professional**

paid public regular religious	advice manner musician transport
-------------------------------	----------------------------------

**34 critical**

clear dangerous important rough	festival illness time water
---------------------------------	-----------------------------

**35 synthetic**

artificial electronic expensive simple	drug meal radio sound
--	-----------------------

**36 liberal**

free moderate plenty valuable	crops furniture parents transport
-------------------------------	-----------------------------------

**37 dramatic**

exciting official surprising worried	adventure change patient salary
--------------------------------------	---------------------------------

**38 conservative**

hopeful safe together traditional	clothes estimate meeting signal
-----------------------------------	---------------------------------

**39 coherent**

clear normal recent together	crime health speech theory
------------------------------	----------------------------

**40 ample**

heavy large plentiful windy	amount climate feelings time
-----------------------------	------------------------------

## Appendix D

### (Vocabulary Size Test) 词汇量测试

#### First 1000

1. see: They **saw** it.
  - a. 切
  - b. 等待
  - c. 看
  - d. 开始
2. time: They have a lot of **time**.
  - a. 钱
  - b. 食物
  - c. 时间
  - d. 朋友
3. period: It was a difficult **period**.
  - a. 问题
  - b. 时间
  - c. 要做的事情
  - d. 书
4. figure: Is this the right **figure**?
  - a. 答案
  - b. 地方
  - c. 时间
  - d. 号码
5. poor: We **are poor**.
  - a. 贫穷的
  - b. 感到幸福的
  - c. 很感兴趣的
  - d. 不喜欢努力工作的
6. drive: He **drives** fast.
  - a. 游泳
  - b. 学习
  - c. 扔球
  - d. 开车

7. jump: She tried to **jump**.
  - a. 漂浮
  - b. 跳
  - c. 停车
  - d. 跑
8. shoe: Where is your **shoe**?
  - a. 父或母
  - b. 钱包
  - c. 钢笔
  - d. 鞋子
9. standard: Her **standards** are very high.
  - a. 后跟
  - b. 分数
  - c. 要价
  - d. 标准
10. basis: I don't understand the **basis**.
  - a. 原因
  - b. 话
  - c. 路标
  - d. 中心议题

#### Second 1000

1. maintain: Can they **maintain** it?
  - a. 维持
  - b. 扩大
  - c. 改善
  - d. 得到
2. stone: He sat on a **stone**.
  - a. 石头
  - b. 凳子
  - c. 垫子
  - d. 树枝

3. upset: I am **upset**.
  - a. 疲倦的
  - b. 著名的
  - c. 富足的
  - d. 不高兴的
  
4. drawer: The **drawer** was empty.
  - a. 抽屉
  - b. 车库
  - c. 冰箱
  - d. 鸟笼
  
5. patience: He **has no patience**.
  - a. 没有耐心
  - b. 很忙
  - c. 没有信心
  - d. 不公正
  
6. nil: His mark for that question was **nil**.
  - a. 很差的
  - b. 什么也没有的
  - c. 很好的
  - d. 中等的
  
7. pub: They went to the **pub**.
  - a. 酒吧
  - b. 银行
  - c. 商场
  - d. 游泳池
  
8. circle: Make a **circle**.
  - a. 素描
  - b. 空白
  - c. 圆圈
  - d. 大洞
  
9. microphone: Please use the **microphone**.
  - a. 微波炉
  - b. 麦克风
  - c. 显微镜
  - d. 手机

10. pro: He's a **pro**.
  - a. 间谍
  - b. 傻瓜
  - c. 记者
  - d. 职业运动员

### Third 1000

1. soldier: He is a **soldier**.
  - a. 商人
  - b. 学生
  - c. 金属工艺制造者
  - d. 士兵
  
2. restore: It has been **restored**
  - a. 重复
  - b. 重新分配
  - c. 降价
  - d. 复原
  
3. jug: He was holding a **jug**.
  - a. 罐子
  - b. 聊天
  - c. 贝雷帽
  - d. 枪
  
4. scrub: He is **scrubbing** it.
  - a. 抓
  - b. 修理
  - c. 刷洗
  - d. 作素描
  
5. dinosaur: The children were pretending to be **dinosaurs**.
  - a. 海盗
  - b. 仙女
  - c. 龙
  - d. 恐龙
  
6. strap: He broke the **strap**.
  - a. 诺言
  - b. 盖子
  - c. 盘子
  - d. 带子

7. pave: It was **paved**.

- a. 堵塞
- b. 分开
- c. 镶金边
- d. 铺路

8. dash: They **dashed** over it.

- a. 猛冲
- b. 磨蹭
- c. 争吵
- d. 瞥见

9. rove: He couldn't stop **roving**.

- a. 喝醉
- b. 漂泊
- c. 哼曲子
- d. 努力工作

10. lonesome: He felt **lonesome**.

- a. 不领情的
- b. 疲倦的
- c. 孤独的
- d. 精力充沛的

#### Fourth 1000

1. compound: They made a new **compound**.

- a. 协议
- b. 复合物
- c. 公司
- d. 预言

2. latter: I agree with the **latter**.

- a. 牧师
- b. 理由
- c. 后者
- d. 答案

3. candid: Please be **candid**.

- a. 小心的
- b. 表示同情的
- c. 公平的
- d. 直率的

4. tummy: Look at my **tummy**.

- a. 围巾
- b. 肚子
- c. 松鼠
- d. 拇指

5. quiz: We made a **quiz**.

- a. 箭筒
- b. 错误
- c. 竞赛
- d. 鸟巢

6. input: We need more **input**.

- a. 输入
- b. 工人
- c. 填料
- d. 钱

7. crab: Do you like **crabs**?

- a. 蟹
- b. 薄脆饼干
- c. 又紧又硬的领子
- d. 蟋蟀

8. vocabulary: You will need more **vocabulary**.

- a. 词汇
- b. 技巧
- c. 钱
- d. 枪

9. remedy: We found a good **remedy**.

- a. 矫正问题的方法
- b. 餐馆
- c. 食谱
- d. 等式

10. allege: They **alleged** it.

- a. 辩解
- b. 剽窃
- c. 证明
- d. 反抗

## Fifth 1000

1. deficit: The company **had a large**
  - a. 出现赤字
  - b. 贬值
  - c. 有这笔大开销的计划
  - d. 在银行里有很多存款
2. weep: He **wept**.
  - a. 毕业
  - b. 哭
  - c. 死
  - d. 担心
3. nun: We saw a **nun**.
  - a. 蠕虫
  - b. 事故
  - c. 修女
  - d. 天空中无法解释的亮光
4. haunt: The house is **haunted**.
  - a. 充满了装饰物
  - b. 已被出租
  - c. 空的
  - d. 闹鬼
5. compost: We need some **compost**.
  - a. 大力支持
  - b. 扶持
  - c. 混凝土
  - d. 堆肥
6. cube: I need one more **cube**.
  - a. 大头针
  - b. 立方体
  - c. 缸子
  - d. 卡片
7. miniature: It is a **miniature**.
  - a. 微型画
  - b. 显微镜
  - c. 微生物
  - d. 在书法中把字母连在一起细小的连
8. peel: Shall I **peel** it?
  - a. 浸泡
  - b. 削皮
  - c. 烫洗
  - d. 切成薄片

9. fracture: They found a **fracture**.
  - a. 裂口
  - b. 碎片
  - c. 夹克衫
  - d. 稀有的宝石
10. bacterium: They didn't find a single **bacterium**.
  - a. 细菌
  - b. 开有红色或橘黄色花的植物
  - c. 骆驼
  - d. 赃物

## Sixth 1000

1. devious: Your plans are **devious**.
  - a. 诡计多端的
  - b. 成熟的
  - c. 考虑不周详的
  - d. 过于昂贵的
2. premier: The **premier** spoke for an hour.
  - a. 律师
  - b. 讲师
  - c. 冒险家
  - d. 总理
3. butler: They have a **butler**.
  - a. 男管家
  - b. 锯
  - c. 家庭教师
  - d. 地窖
4. accessory: They gave us some **accessories**.
  - a. 签证
  - b. 官方命令
  - c. 选择
  - d. 额外的零件
5. threshold: They raised the **threshold**.
  - a. 旗子
  - b. 门槛
  - c. 天花板
  - d. 利息

6. thesis: She has completed her **thesis**.
  - a. 论文
  - b. 归纳证词
  - c. 试用期
  - d. 延期治疗
7. strangle: He **strangled** her.
  - a. 掐死
  - b. 宠坏
  - c. 绑架
  - d. 赞美
8. cavalier: He treated her in a **cavalier** manner.
  - a. 慢待的
  - b. 礼貌的
  - c. 尴尬的
  - d. 兄长的
9. malign: His **malign** influence is still felt.
  - a. 邪恶的
  - b. 好的
  - c. 非常重要的
  - d. 秘密的
10. veer: The car **veered**.
  - a. 改变方向或路线
  - b. 剧烈晃动
  - c. 发生逆火引起爆鸣
  - d. 打滑
3. stealth: They did it by **stealth**.
  - a. 花费大量的钱
  - b. 逼迫
  - c. 悄悄的或秘密的行动
  - d. 没有注意到所遇到的问题
4. shudder: The boy **shuddered**.
  - a. 低语
  - b. 差点摔倒
  - c. 发抖
  - d. 大声叫喊
5. bristle: The **bristles** are too hard.
  - a. 问题
  - b. 短而硬的毛发
  - c. 折叠床
  - d. 鞋底
6. bloc: They have joined this **bloc**.
  - a. 乐队
  - b. 小偷帮
  - c. 侦察员
  - d. 集团
7. demography: This book is about **demography**.
  - a. 土地使用模式研究
  - b. 用图片表示数字事实的研究
  - c. 水文学
  - d. 人口学

## Seventh 1000

1. olive: We bought **olives**.
  - a. 橄榄
  - b. 康乃馨
  - c. 男人的游泳衣
  - d. 清除杂草的工具
2. quilt: They made a **quilt**.
  - a. 遗嘱
  - b. 合同
  - c. 被子
  - d. 羽毛笔
8. gimmick: That's a good **gimmick**.
  - a. 高空作业时所站的东西
  - b. 钱包
  - c. 引人注意的行为或事物
  - d. 花招
9. azalea: This **azalea** is very pretty.
  - a. 杜鹃花
  - b. 由天然棉所制成的很轻的材料
  - c. 莎丽
  - d. 扇贝

10. yoghurt: This **yoghurt** is disgusting.

- a. 淤泥
- b. 伤口
- c. 酸奶
- d. 榲桲

7. locust: There were hundreds of **locusts**.

- a. 飞蝗
- b. 志愿者
- c. 素食者
- d. 颜色鲜艳的野花

## Eighth 1000

1. erratic: He was **erratic**.

- a. 完美的
- b. 很坏的
- c. 很有礼貌的
- d. 不可靠的

8. authentic: It is **authentic**.

- a. 真的
- b. 非常吵闹的
- c. 老的
- d. 干旱的

2. palette: He lost his **palette**.

- a. 装鱼的篮子
- b. 胃口
- c. 年轻的女伴
- d. 调色板

9. cabaret: We saw the **cabaret**.

- a. 壁画
- b. 卡巴莱
- c. 蟑螂
- d. 美人鱼

3. null: His influence was **null**.

- a. 具有好的结果
- b. 毫无帮助的
- c. 没有效果的
- d. 持久的

10. mumble: He started to **mumble**.

- a. 集中精力
- b. 颤抖
- c. 远远地落后于其他人
- d. 咕哝

4. kindergarten: This is a good **kindergarten**.

- a. 消遣
- b. 幼儿园
- c. 背包
- d. 图书馆

## Ninth 1000

1. hallmark: Does it have a **hallmark**?

- a. 表明什么时候之前应该被使用的印
- b. 纯度印记
- c. 表明经过皇室同意的标记
- d. 禁止复制的标记

5. eclipse: There was an **eclipse**.

- a. 飓风
- b. 泼溅
- c. 大屠杀
- d. 日食

2. puritan: He is a **puritan**.

- a. 喜欢被人注意的人
- b. 具有严格道德标准的人
- c. 吉普赛人
- d. 守财奴

6. marrow: This is the **marrow**.

- a. 吉祥物
- b. 骨髓
- c. 操纵杆
- d. 增加工资

3. monologue: Now he has a **monologue**.

- a. 单眼镜
- b. 独白
- c. 专制
- d. 把字母有趣地连在一起的图画

## Tenth 1000

4. weir: We looked at the **weir**.
  - a. 行为古怪的人
  - b. 红树属植物
  - c. 通过吹来演奏的古老金属乐器
  - d. 拦河坝
5. whim: He had lots of **whims**.
  - a. 古老的金币
  - b. 母马
  - c. 没有任何动机的奇异想法
  - d. 疼痛的红肿块
6. perturb: I was **perturbed**.
  - a. 被迫同意的
  - b. 烦恼的
  - c. 困惑的
  - d. 湿透的
7. regent: They chose a **regent**.
  - a. 不负责任人的人
  - b. 暂时主持会议的人
  - c. 摄政者
  - d. 代表
8. octopus: They saw an **octopus**.
  - a. 猫头鹰
  - b. 潜水艇
  - c. 直升飞机
  - d. 章鱼
9. fen: The story is set in the **fens**.
  - a. 沼泽
  - b. 山地
  - c. 贫民窟
  - d. 很久以前
10. lintel: He painted the **lintel**.
  - a. 过梁
  - b. 渡船
  - c. 长有伸展树枝和绿色果实的美
  - d. 戏院中显示场景的板子
1. awe: They looked at the mountain with **awe**.
  - a. 担心
  - b. 兴趣
  - c. 惊奇
  - d. 尊重
2. peasantry: He did a lot for the **peasantry**.
  - a. 当地人
  - b. 寺庙
  - c. 商人俱乐部
  - d. 农民
3. egalitarian: This organization is very **egalitarian**.
  - a. 保密的
  - b. 保守的
  - c. 诉讼的
  - d. 平等主义的
4. mystique: He has lost his **mystique**.
  - a. 体格
  - b. 神秘性
  - c. 情人
  - d. 胡子
5. upbeat: I'm feeling really **upbeat** about it.
  - a. 苦恼的
  - b. 乐观的
  - c. 受伤害的
  - d. 迷惑的
6. cranny: We found it in the **cranny**!
  - a. 旧杂物义卖
  - b. 小洞
  - c. 阁楼
  - d. 大箱子
7. pigtail: Does she have a **pigtail**?
  - a. 辫子
  - b. 长的衣、袍、裙等拖在地上的部
  - c. 开有一串下垂浅粉色花的植物
  - d. 情人

8. crowbar: He used a **crowbar**.  
 a. 撬棍  
 b. 化名  
 c. 锥子  
 d. 轻的金属拐杖
9. ruck: He got hurt in the **ruck**.  
 a. 骨盆  
 b. 打架  
 c. (运动员或竞赛者等的) 散乱一  
 d. 在雪地上奔跑
10. lectern: He stood at the **lectern**.  
 a. 讲台  
 b. 圣餐桌  
 c. 酒吧  
 d. 边缘
5. puma: They saw a **puma**.  
 a. 由坯建成的小房子  
 b. 来自炎热而干旱国家的树  
 c. 飓风  
 d. 美洲狮
6. pallor: His **pallor** caused them concern.  
 a. 非正常的高温  
 b. 对一切都毫无兴趣  
 c. 一帮朋友  
 d. 苍白的皮肤
7. aperitif: She had an **aperitif**.  
 a. 躺椅  
 b. 家庭教师  
 c. 带有很高羽毛的大帽子  
 d. 开胃酒

### Eleventh 1000

1. excrete: This was **excreted** recently.  
 a. 排泄  
 b. 澄清  
 c. 调查  
 d. 被列入不合法的事情之中
2. mussel: They bought **mussels**.  
 a. 玻璃弹球  
 b. 贻贝  
 c. 榅桲  
 d. 纸巾
3. yoga: She has started **yoga**.  
 a. 粗绳结的饰物  
 b. 瑜伽  
 c. 羽毛球  
 d. 东方国家的一种舞蹈
4. counterclaim: They made a **counterclaim**.  
 a. 在法律案件中一方所提出的诉求和另一方的诉求相同  
 b. 瑜伽  
 c. 羽毛球  
 d. 东方国家的一种舞蹈
8. hutch: Please clean the **hutch**.  
 a. 隔板  
 b. 行李箱  
 c. 轮毂  
 d. 兔笼
9. emir: We saw the **emir**.  
 a. 尾巴上长有两个长长卷曲羽毛的鸟  
 b. 奶妈  
 c. 埃米尔 (对穆斯林统治者的尊称)  
 d. (爱斯基摩人用坚硬雪快砌成的临时栖身用的) 拱型圆顶小屋
10. hessian: She bought some **hessian**.  
 a. 光泽油亮略呈粉红色的鱼  
 b. 大麻  
 c. 一种结实的粗麻布  
 d. 用来给食物调味的味道浓烈的根状

### Twelfth 1000

1. haze: We looked through the **haze**.  
 a. 舷窗  
 b. 薄雾  
 c. 窗帘  
 d. 花名册

2. spleen: His **spleen** was damaged.
  - a. 膝盖骨
  - b. 脾脏
  - c. 下水管
  - d. 自尊
3. soliloquy: That was an excellent **soliloquy**!
  - a. 六个人唱的歌曲
  - b. 修饰词或描述性词语
  - c. 带有灯光和音乐的娱乐
  - d. 独白
4. reptile: She looked at the **reptile**.
  - a. 手稿
  - b. 爬行动物
  - c. 挨家兜售货物的人
  - d. 水粉画
5. alum: This contains **alum**.
  - a. 取自一种普通植物的有毒物质
  - b. 一种由人造棉所制成的柔软材料
  - c. 鼻烟
  - d. 白矾
6. refectory: We met in the **refectory**.
  - a. 食堂
  - b. 签署法律文件的办公室
  - c. 宿舍
  - d. 温室
7. caffeine: This contains a lot of **caffeine**.
  - a. 一种让人瞌睡的物质
  - b. 由坚韧树叶织成的线状物
  - c. 错误的观点
  - d. 咖啡因
8. impale: He nearly got **impaled**.
  - a. 控告
  - b. 监禁
  - c. 被尖物刺中
  - d. 陷入争论之中
9. coven: She is the leader of a **coven**.
  - a. 合唱队
  - b. 集体企业
  - c. 秘密组织
  - d. 过着严格宗教生活的一群生活在教堂的女人

10. trill: He practised the **trill**.
  - a. 颤音
  - b. 小提琴
  - c. 投球
  - d. (跳芭蕾舞者的)单足旋转

### Thirteenth 1000

1. ubiquitous: Many weeds are **ubiquitous**.
  - a. 很难铲除的
  - b. 长有又长又壮根的
  - c. 在大部分国家都能够被发现的
  - d. 在冬天枯竭的
2. talon: Just look at those **talons**!
  - a. 山顶
  - b. 爪
  - c. 盔甲
  - d. 傻瓜
3. rouble: He had a lot of **roubles**.
  - a. 红宝石
  - b. 亲戚
  - c. 卢布
  - d. 心中的道德或者其他难题
4. jovial: He was very **jovial**.
  - a. 社会地位低下的
  - b. 总爱批评人的
  - c. 很幽默的
  - d. 很友好的
5. communiqué: I saw their **communiqué**.
  - a. 对一个机构的批评报告
  - b. 属于社区成员的花园
  - c. 用于做广告的印刷材料
  - d. 官方通告
6. plankton: We saw a lot of **plankton**.
  - a. 毒性传播很快的杂草
  - b. 生活在水中的很小的植物或动物
  - c. 坚硬木质的树木
  - d. 容易导致土地滑坡的灰色泥土

7. skylark: We watched a **skylark**.
  - a. 飞行表演
  - b. 人造卫星
  - c. 表演魔术的人
  - d. 一边鸣叫一边飞得很高的
8. beagle: He owns two **beagles**.
  - a. 车顶能够收起来的跑得很快的
  - b. 能够快速射中很多人的枪
  - c. 长有长长耳朵的小狗
  - d. 建在度假胜地的房子
9. atoll: The **atoll** was beautiful.
  - a. 环状珊瑚岛
  - b. 用细线织成图画的艺术品
  - c. 女人晚上所佩带的镶嵌有宝石的小皇冠
  - d. 河水流过大石间的缝隙
10. didactic: The story is very **didactic**.
  - a. 说教的
  - b. 难以置信的
  - c. 令人兴奋的
  - d. 令读者琢磨不透的
4. augur: It **augured** well.
  - a. 预示未来的好事情
  - b. 和预料的很吻合
  - c. 具有一种和其它东西搭配在一起很
  - d. 发出一种清脆而动听的声音
5. bawdy: It was very **bawdy**.
  - a. 难以预料的
  - b. 令人愉快的
  - c. 急促的
  - d. 粗鲁的
6. gauche: He was **gauche**.
  - a. 善谈的
  - b. 灵活的
  - c. 尴尬的
  - d. 坚决的
7. thesaurus: She used a **thesaurus**.
  - a. 一种字典
  - b. 一种化合物
  - c. 一种特殊的说话方式
  - d. 皮下注射

#### Fourteenth 1000

1. canonical: These are **canonical** examples.
  - a. 打破常规的
  - b. 取自一本宗教书的
  - c. 合乎原则并被广泛接受的
  - d. 最近发现的
2. atop: He was **atop** the hill.
  - a. 在...脚下
  - b. 在...顶部
  - c. 在这边
  - d. 在那边
3. marsupial: It is a **marsupial**.
  - a. 长有坚硬脚的动物
  - b. 生长几年的一种植物
  - c. 开出的花总是面向太阳的植物
  - d. 有袋类动物
8. erythrocyte: It is an **erythrocyte**.
  - a. 止痛的药
  - b. 血液中的红色的成分
  - c. 略显红色的白金属
  - d. 鲸家族中的一员
9. cordillera: They were stopped by the **cordillera**.
  - a. 特殊的法律
  - b. 装备有武器的船只
  - c. 山脉
  - d. 国王的长子
10. limpid: He looked into her **limpid**.
  - a. 清澈的
  - b. 含泪的
  - c. 深棕色的
  - d. 美丽的

## Appendix E

# An English Essay

“Developing countries should open their doors to foreign companies to open offices and build factories.” Do you agree or disagree with the above opinion? Use specific reasons to support your answer. (300 words)

[illegible]

## Appendix F

### The Frequency of the Use of VLSs

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
q1.1	419	1.00	5.00	3.2387	.99296
q1.2	418	1.00	5.00	3.7943	.89269
q1.3	417	1.00	5.00	3.9257	.85583
q1.4	417	1.00	5.00	3.6403	.90676
q1.5	418	2.00	5.00	3.6196	.91158
q1.6	419	1.00	5.00	3.7088	1.09640
q1.7	418	1.00	5.00	2.0885	.97416
q1.8	418	1.00	5.00	3.7368	1.02142
q1.9	417	1.00	5.00	2.9113	1.00685
q1.10	419	1.00	5.00	3.4893	.94173
q1.11	417	1.00	5.00	3.0600	1.02317
q1.12	419	1.00	5.00	2.1432	.99689
q1.13	415	1.00	5.00	1.6096	.85517
q1.14	418	1.00	5.00	2.9187	.89662
q1.15	418	1.00	5.00	2.4139	1.00287
q1.16	419	1.00	5.00	2.6802	1.08176
q1.17	415	1.00	5.00	2.6458	1.04365
q1.18	418	1.00	5.00	3.0120	.97565
q1.19	414	1.00	5.00	3.2923	.93575
q1.20	418	1.00	5.00	3.4019	1.05539
q1.21	418	1.00	5.00	3.3947	.98899
q1.22	417	1.00	5.00	3.0743	1.09708
q1.23	419	1.00	5.00	1.9618	.94260
q1.24	419	1.00	5.00	2.5107	1.06342
q2.1	419	1.00	5.00	3.6277	.90443
q2.2	418	1.00	5.00	3.6411	.97224
q2.3	419	1.00	5.00	2.8305	1.03638
q2.4	418	1.00	5.00	3.2823	1.05790
q2.5	418	1.00	5.00	3.4450	1.04254

q2.6	418	1.00	5.00	2.6842	1.13830
q2.7	419	1.00	5.00	2.9785	1.15000
q2.8	417	1.00	5.00	2.4844	1.12462
q2.9	419	1.00	5.00	3.3842	1.15863
q2.10	419	1.00	5.00	2.8234	1.08373
q2.11	419	1.00	5.00	2.2768	.96826
q2.12	419	1.00	5.00	2.8353	1.08009
q2.13	419	1.00	5.00	2.5513	1.05998
q2.14	418	1.00	5.00	2.5622	1.03983
q2.15	419	1.00	5.00	2.3819	.86520
q2.16	416	1.00	5.00	2.7933	.95466
q2.17	419	1.00	5.00	2.3986	.82491
q2.18	419	1.00	5.00	2.8305	.89251
q2.19	419	1.00	5.00	2.7184	.95443
q2.20	419	1.00	5.00	4.0024	.98917
q2.21	415	1.00	5.00	3.0265	1.05261
q2.22	417	1.00	5.00	2.4724	.80851
q2.23	419	1.00	5.00	2.9093	.91578
q2.24	416	1.00	5.00	3.2332	.97484
q2.25	418	1.00	5.00	3.0455	1.05728
q2.26	418	1.00	5.00	2.6675	1.02578
q2.27	419	1.00	5.00	2.6897	1.04179
q2.28	419	1.00	5.00	2.8401	.97858
Valid N (listwise)	376				

## **Appendix G**

### **Interview Questions**

#### **English vocabulary learning: Significance and difficulties**

1. How do you think about learning English vocabulary? Is it important or not? Why?
2. Do you have any difficulties in learning English vocabulary? Please explain.

#### **Current vocabulary learning strategies**

1. What specific techniques and methods do you use to learn English vocabulary?
2. How do you know whether they are useful or not? What are the effect of these techniques and methods on your vocabulary size and depth of vocabulary knowledge?
3. How did you find out these methods to learn English vocabulary?

#### **Other strategies**

1. Can you think of any other techniques you use to lean English vocabulary?
2. How did you find out these methods?
3. Have you found them useful? Why or why not?

#### **Vocabulary knowledge**

1. Share with me your perspectives about vocabulary size and depth? (how many words you know and how well you know them?)
2. Tell me which one is more important to you, vocabulary size or depth of vocabulary knowledge?

## Appendix H

### The Coding Scheme for Interviews

The coding scheme displays as a hierarchical system consisting of three tiers of codes. Tier 1 presents the general constructs addressed in research questions and interview questions. These constructs are organized into four categories: (1) *Significance of learning English words*; (2) *Difficulties of learning English words*; (3) *Current vocabulary learning strategies*; (4) *Perspectives about vocabulary knowledge (e.g., size and depth)*. Tier 2 codes break down each category identified in Tier 1 into more delicate sub-codes. By examining the interview data, I was able to categorize both *Significance of learning English words* and *Difficulties of learning English words* into polarities: *Important* and *Not important*; *Having difficulties* and *No difficulty*; to group *Current vocabulary learning strategies* into two sub-categories: *Use of vocabulary learning strategies and perceived usefulness*, and *Sources of vocabulary learning strategies* and to categorize *Perspectives about vocabulary knowledge* into three sub-categories: *Perspectives about vocabulary size*; *Perspectives about vocabulary depth*; *Perspectives about the importance of vocabulary size/depth*. The third tier of the coding scheme elaborates further on the second tier. It has been developed by drawing on a close examination of both interview data and relevant literature. The category – *Important*, is further divided into four features in line with different aspects of learning English as students emphasized: *Important for reading*; *listening*; *speaking*; and *writing*, while the category– *Not important* is further elaborated as: *Not important for the English test*. As for the category of *Having difficulties*, six features have been identified: *Having no efficient methods to remember new English words*; *Hard to remember new English words for a long time*; *Difficult to remember the long words, the words with irregular spellings and pronunciations and the infrequent words*; *Some words you can recognize, but can't produce*; *Too many English words to remember*; *Hard to learn the depth of vocabulary knowledge*. Regarding the *Use of vocabulary learning strategies and perceived usefulness*, some sub-groups have emerged, such as *self-initiated strategies*, *inferencing strategies*, and *using dictionary* etc.

Concerning *Sources of vocabulary learning strategies*, three main sub-categories have been identified: *Explored by themselves*; *Introduced by their English teachers*; *Introduced by others, such as their classmates*. Both *Perspectives about vocabulary size* and *Perspectives about vocabulary depth* are developed into some sub-categories, encompassing: *knowing the basic meaning, quantity, ease of learning, and foundation for English learning*; *quality, difficult to learn, and knowing how to use a word*, respectively. Finally, the category – *Perspectives about the importance of vocabulary size/depth* has been divided into a set of three features: *Vocabulary size is important*; *Depth of vocabulary knowledge is important*; *Both are important* (see the following table for the coding scheme and examples).

Categories (Tier 1)	Sub-categories (Tier 2)	Features (Tier 3)	Examples
<b>Significance of learning English words</b>	Important	Reading	e.g., English vocabulary is quite important. You can understand the English articles well only if you have a large vocabulary size. (M1)
		Writing	e.g., I think English vocabulary is very important for English speaking, such as daily communication and English writing. (H4)
		Speaking	e.g., English vocabulary is also quite important for oral English. You can express your ideas well in communication if you know lots of words. (M3)
		Listening	e.g., I think English vocabulary learning is quite important for me. It is the foundation for English reading comprehension, listening comprehension, speaking, and writing. Without English vocabulary, you can't understand, speak or write anything in English. (H3)
	Not important	English test	e.g., I think English vocabulary learning is not quite important because in my opinion, a certain vocabulary size would be enough for passing the CET-4 test, so it does not make much sense to have a very large English vocabulary size. (L1)
<b>Difficulties in learning English words</b>	Having difficulties	Having no efficient methods to remember new English words	e.g., I think I have no good methods to remember new English words. In my case, it is hard to remember new words by using rote memory, especially to remember the learnt words for a long time. I find remembering new words in English extensive reading quite helpful for me, however, it is quite time-consuming in the meanwhile. (H3)
		Hard to remember new English words for a long time	e.g., I always forget the learnt words quickly. I think it may be because I didn't review the learnt words frequently. (H4)
		Difficult to remember the long words, the words with irregular spellings and pronunciations and the infrequent words	e.g., Infrequent words are hard for me to remember because I seldom encounter and use them. (H5)
		Some words you can	e.g., I find that for some English

		recognize, but cannot produce	words, I can only know their meanings in the reading, but when I want to use them in my writing or speaking, I just can't recall their spellings. (L1)
		Too many English words to remember	e.g., I think English words are too many for me to remember. It is quite hard for me to remember such a big amount of English words and it is still hard to distinguish some English words from those with similar spellings. Sometimes, I confuse some English words with those spelt or pronounced in similar ways. (M5)
		Hard to learn the depth of vocabulary knowledge (various meanings of a word and how to use it)	e.g., It is difficult for me to remember various meanings of a word. Usually I can only remember one or two meanings of a word. (M5)
	No difficulty	Having no difficulty	e.g., I don't have difficulties in learning English vocabulary. I think English vocabulary should be learned through the English extensive reading. For example, I like reading some English novels and some English academic articles. When I encounter the English words several times in the reading, I could remember them well. (H2)
<b>Current vocabulary learning strategies</b>	Use of vocabulary learning strategies and perceived usefulness	Self-initiated	Watching English movies or TV series
			Watching English online video
			Reading English newspapers combined with checking the dictionary and using inferencing strategies
			Reading English novels with the Chinese translation
			Reading English academic articles combined with checking the dictionary and using inferencing strategies
			Listening to English radio program on the English website combined with checking some key words in the dictionary
			Listening to English songs combined with checking some new words in the lyrics in the dictionary
			Using English vocabulary learning software on the cell phone (Youdao)
		Inferencing	Reading English newspapers combined with checking the dictionary and using inferencing strategies

			Reading English academic articles combined with checking the dictionary and using inferencing strategies
			Doing the reading comprehension of CET-4/6 combined with using inferencing strategies and checking the dictionary
		Using dictionary	Reading English newspapers combined with checking the dictionary and using inferencing strategies
			Reading English academic articles combined with checking the dictionary and using inferencing strategies
			Listening to English radio program on the English website combined with checking some key words in the dictionary
			Listening to the English songs combined with checking some new words in the lyrics in the dictionary
			Doing the reading comprehension of CET-4/6 combined with using inferencing strategies, and checking the dictionary
			Doing the reading comprehension of CET-4 combined with checking the dictionary
		Using context to remember new words	Remembering new words on vocabulary book of CET-4/6 (using associative methods, example sentences, collocations and rote memory)
			Remembering new words in the context of English text
		Note-taking	Learning some new words in English classes by using note-taking method
		Production-based	Trying to find opportunities to communicate with others in English
		Association	Connecting the pronunciations of some English words with the pronunciations of some Chinese words
			Remembering new words on the vocabulary book of CET-4/6 (using associative methods, example sentences, collocations and rote memory)
			Using affixes to remember new words
		Repetition	Remembering new words on vocabulary book of CET-4/6 by using rote memory
			Remembering the word list of the

			English textbook out of the context
			Remembering some new words on the English textbook by using rote memory before the English class
			Making a list of new words to remember them
			Remembering some English passages by using rote memory
			Learning some new words on the English textbook by doing dictations in English classes
	Sources of vocabulary learning strategies	Explored by themselves	e.g., Some of my VLSs were introduced by my teachers and some were found by myself. For instance, sometimes I browsed some English online forums and learned some VLSs. (H1)
		Introduced by their English teachers	e.g., I do not have many methods for learning English words. These methods were mainly introduced by my English teacher. (L5)
		Introduced by others, such as their classmates	e.g., Sometimes, my elder sister told me some methods for learning English words. (M2)
<b>Perspectives about vocabulary knowledge</b>	Perspectives about vocabulary size	Knowing the basic meaning	e.g., Vocabulary size requires you to know the meaning of a word. Usually, you only need to know its basic meaning so you can understand this word when doing reading or listening comprehension in English. (H2)
		Quantity	e.g., Vocabulary size means how many English words you know. (M4)
		Ease of learning	e.g., Vocabulary size is relatively easy for me compared to learning the depth of vocabulary knowledge. (M2)
		Foundation for English learning	e.g., I think vocabulary size is the basis for English learning. For example, I always encounter lots of new English words in the reading comprehension. If I don't know the meanings of these words, it will hinder my understanding of the passage. (M3)
	Perspectives about vocabulary depth	Quality	e.g., Depth of vocabulary knowledge requires you to know the grammatical information or collocation of a word. (H4)
		Difficult to learn	e.g., It is difficult for me to know various meanings of a word and how to use the word. (H5)
		Knowing how to use a word	e.g., Depth of vocabulary knowledge means knowing how to use the word, such as how to use this word in conversation. (L2)

	Perspective about the importance of vocabulary size/depth	Vocabulary size is important	e.g., Vocabulary size is more important for me. My vocabulary size is very small, so I think I have to enlarge my vocabulary size first. (L5)
		Depth of VK is important	e.g., Sometimes, various meanings of a word might appear in English reading passages, so if you only know the basic meaning of a word, it would affect your reading comprehension. (H2)
		Both are important	e.g., I think both of them are important because on the one hand, vocabulary size is the foundation for English learning, while on the other, learning English is mainly for communication. Thus if we don't have the depth of vocabulary knowledge, it doesn't make much sense to learn English. (H4)



---

**RE: HS Ethics Application – Final Report Approved (5201300035)**

发件人 : Fhs Ethics <fhs.ethics@mq.edu.au>

时 间 : 2014年8月26日(星期二) 上午7:45

收件人 : Associate Professor Mehdi Riazi <mehdi.riazi@mq.edu.au>

抄 送 : Ms Jean Margaret Brick <jean.brick@students.mq.edu.au>; Miss Na Fan <na.fan@students.mq.edu.au>

---

Dear A/Prof Riazi,

Title of project: 'A study of Vocabulary Knowledge and Vocabulary Learning Strategies of Chinese EFL Learners' (Ref: 5201300035)

FINAL REPORT APPROVED

Your final report has been received and approved, effective 25th August 2014.

The Faculty of Human Sciences Human Research Ethics Sub-Committee is grateful for your cooperation and would like to wish you success in future research endeavours.

Yours sincerely,

Dr Simon Boag  
Acting Chair  
Faculty of Human Sciences  
Human Research Ethics Sub-Committee

-----  
Faculty of Human Sciences - Ethics  
Research Office  
Level 3, Research HUB, Building C5C  
Macquarie University  
NSW 2109

Ph: +61 2 9850 4197

Fax: +61 2 9850 4465

Email: fhs.ethics@mq.edu.au

<http://www.research.mq.edu.au/>