Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

By reading and using the thesis, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.

2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.

3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

ROLES OF DEPTH OF VOCABULARY KNOWLEDGE IN

EFL LEARNERS' WRITING PROFICIENCY

LIN HUA FANG

Ph.D

The Hong Kong Polytechnic University

2015

THE HONG KONG POLYTECHNIC UNIVERSITY

Department of English

ROLES OF DEPTH OF VOCABULARY KNOWLEDGE IN

EFL LEARNERS' WRITING PROFICIENCY

LIN HUA FANG

A thesis submitted in partial fulfilment of the requirements

for the degree of Doctor of Philosophy

March 2015

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____ (Signed)

_____LIN HUA FANG_____ (Name of student)

ABSTRACT

The principal objective of the study is to examine the association between EFL (English as a foreign language) learners' depth of lexical knowledge and their writing proficiency, and the role that contextual factors play in determining that association. A total of 150 students, 67 from Hong Kong and 83 from mainland China, participated in the study. Nation's Vocabulary Levels Tests (2001), Read's Word Associates Test (1998) and a writing test were the main instruments applied within the study. Measures for assessing lexical richness in learners writing, such as lexical diversity, lexical sophistication and lexical frequency profile, were also deployed within the overall analyses.

The cardinal element of the study examines whether EFL learners' depth or extensiveness of vocabulary knowledge is closely related to their writing ability; and whether the depth dimension of word knowledge can provide a distinctive prediction of their writing scores, over and above the contribution made by the breadth of lexical knowledge. In addition to this, the study also explores whether contextual factors affect the association between learners' lexical knowledge and their writing proficiency. Through in-depth investigation, the completed overall study collects both quantitative and qualitative data, originating from two vocabulary tests, a writing test, five case studies, three focus-group interviews, a questionnaire survey and a learner corpus analysis.

Amongst other conclusions, the study found 1) that the depth of EFL learners' lexical knowledge provides a unique prediction of their writing scores, over and above the contribution made by the breadth of lexical knowledge; and 2) that

different language learning contexts affect the association between these learners' depth of vocabulary and their writing proficiency.

The present study both traverses and fills a gap in the existing literature appertaining to the relationship between EFL learners' depth of word knowledge and their language proficiency. Importantly, the study also draws researchers' attention to the imperative influence of language learning contexts in determining the predictive power of lexical knowledge (in particular the depth of vocabulary knowledge) on learners' writing proficiency. Pedagogically, the study both confirms observations made in previous studies which suggest that the National Matriculation English Test in China cannot objectively measure candidates' writing ability, and reasserts the pressing need for fundamental examination reform in China.

ACKNOWLEDGEMENTS

I am deeply grateful to the many people who assisted me during this project. Above all, I thank the student participants who volunteered to be studied for the project. Without them, the study would not have been possible. My gratitude goes particularly to the mainland Chinese participants who expressed great enthusiasm and strong belief in the value of the study. I was very touched when a mainland student, after meeting me only two or three times, provided me with a NMET (National Matriculation English Test) preparation book, which his parents posted to Hong Kong from their home town. We discussed how heavily many students in mainland China relied on such books to train their writing skills. He surprised me by asking his parents to post the one he used in the secondary school so that I could better understand the impact of such books on the writing proficiency of mainland Chinese students.

I feel most deeply indebted to my chief supervisor, Prof. David Qian. His research and publications inspired me and shed light on the direction of this study. I thank him for the intellectual freedom he gave me, the trust he had in my independence in conducting research, the encouragement he offered for me to explore new horizons in the field, and the demands he placed on me to use my best endeavours to produce high-quality academic work. I am also grateful to my co-supervisor, Prof. Martin Warren, for his strong support and guidance.

I owe special thanks to Prof. Stephen Evans who allowed me to attend his classes and workshops as a mature student even though we have been colleagues for many years. His encouraging comments on my academic work, and, in particular, his inspiring publications on language use in Hong Kong, have greatly helped me

in my investigation into the disparities between Hong Kong and mainland China in their English language learning contexts.

I owe my eternal gratitude to Dr. Bruce Morrison, the director of my English Language Centre. He kindled my initial interest in vocabulary studies by asking me to conduct a joint project with him. When I decided to develop this interest into a PhD study, he understood the difficulties of my dual roles of being a full-time member of teaching staff and a part-time student, and therefore generously granted me a three-month study leave. These three months allowed me to focus fully on the study and were instrumental in the completion of this dissertation.

My thanks also go to Prof. Tom Cobb. He generously allows the public to use his website *Compleat Lexical Tutor* for teaching and research. Without his generosity, part of the data processing for this study would have been impossible.

I am very grateful to Miss Anna Cheung, from whom I obtained much statistical assistance. Her curiosity about the project was also very important to me. Many of her questions made me think more carefully about my interpretation of some data.

I would like to thank many of my friends, such as Lucy, Jan and Alice, who always demonstrated their full understanding of my frequent "no" answer to their social invitations. My thanks go particularly to Kate and Julia, to whom I forever owe my gratitude for their always being there for me. Whenever I encountered difficulties, they never failed to provide a listening ear and offered as much help as they could. I am immensely thankful to my family, in particular my parents and my youngest sister, Peony. Their love and unwavering support have made the difficulties during this study easier to manage.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF MAJOR ABBREVIATIONS

CET: College English Test

CET-4: College English Test, Band 4

CLT: Communicative language teaching

DVK: Depth of vocabulary knowledge

EFL: English as a foreign language

ESL: English as a second language

EW: Essay writing

GSEEE: Graduate School Entrance English Examination

HKALE: Hong Kong Advanced Level Examinations

HKDSE: Hong Kong Diploma of Secondary Education

IELTS: International English Language Testing System

L1: First language

L2: Second language

MC: Multiple choice

MOI: Medium of instruction

MPS: Maximum possible score

NMET: National Matriculation English Test

RF: Ratio of frequency

SBA: School based assessment

TOEFL: Test of English as a Foreign Language

TBLT: Task based language teaching

UE: Use of English (test)

VLT: Vocabulary levels test

VS: Vocabulary size

VS 3,000: VS at the 3,000 words level

VS 5,000: VS at the 5,000 words level

WAT: Word associates test

CHAPTER ONE

INTRODUCTION

1.1. Overview

This chapter provides a brief overview of the concepts pertaining to the lexical development of learners of English as a second and foreign language. The chapter begins by discussing the role of vocabulary knowledge in L2 (second language) learning and then examines the differences between L1 (first language) and L2 lexical development processes. The section concludes by providing a detailed review of the breadth and depth dimensions of lexical knowledge.

1.2. Vocabulary knowledge in L2 learning

Vocabulary knowledge has always been regarded as a crucial part of language learning. Research has shown that achieving certain levels and qualities of lexical knowledge is one of the important prerequisites for successful language learning (Zareva, Schwanenflugel & Nikolova, 2005; Schoonen & Verhallen, 2008). Gass and Selinker (2008) even assert that "language learning is largely lexical learning" (p. 173). This assertion is sufficiently supported by empirical data and explicitly confirmed by language acquisition researchers, in particular those who study learners of English as a second language (ESL) and English as a foreign language (EFL). Examples are as follows.

1. An adequate knowledge of words is a prerequisite for effective language use (Read, 2000, p. 83).

2. All other things being equal, learners with bigger vocabularies are more proficient in a wide range of language skills than learners with smaller vocabularies (Meara, 1996, p. 37).

One of the reasons for this widely recognised importance of vocabulary knowledge in language learning is the inextricably intertwined relationship between lexis and grammar. The strong association between the two language forms has led to the appearance of lexicogrammar (Halliday, 2013), a linguistic view that focuses on the continuity rather than the disparity between grammar and vocabulary. Nevertheless, when the two language forms are compared, in particular in the process of L2 acquisition, lexical competence is often considered more important than syntactical knowledge in achieving effective communication. This view, again, is well-documented in the literature as can be seen from the two examples below.

1. Knowing words is the key to understanding and being understood. The bulk of learning a new language consists of learning new words. Grammatical knowledge does not make for great proficiency in a language (Vermeer, 1992, p. 147).

2. No matter how well the student learns grammar, no matter how successfully the sounds of L2 are mastered, without words to express a wider range of meanings communication in an L2 just cannot happen in any meaningful way (McCarthy, 1999, p. viii).

Considering the critical role of lexical knowledge in language learning, it is hardly surprising to see that lexical errors are the most numerous in L2 learners' language production; and that inappropriate lexical use often causes the most

serious hindrance to successful communication of L2 learners (Llach, 2011). These observations are largely derived from studies in the domain of ESL and EFL (e.g. Ellis, 2008a; Laufer, 2013). One of the notable examples of such studies is Santos (1998) who elicited responses of tertiary teaching staff members to the mistakes made by their EFL students in academic writing. Results of the study reveal that lexical errors are the most severe obstruction to the understanding of the students' texts. This finding accords with Djokic's (1999) study which reports that vocabulary errors "cause momentary confusion" and "bring about misunderstanding" (p. 128). This confusion and misunderstanding not only cause an interruption to communication but may also eventually lead to a communication breakdown (Llach, 2011). Difficulties in lexical use have also been deeply felt by many L2 students who believe that of all the error types in their language production, the ones pertaining to vocabulary use are most damaging (Gass & Selinker, 2008).

1.3. Lexical development in L1 and L2 learning

Many researchers (e.g. Ellis, 2008a; Jiang, 2000) believe that lexical errors cause more communication problems in the language production by L2 learners than that by L1 learners. This belief is largely derived from the different lexical development processes of the two types of learners. The lexical development of an L1 child is relatively quicker due to his/her extensive and highly contextual exposure to the language. This exposure makes it possible for the child to develop the semantic, syntactic and morphological knowledge of a word while becoming familiar with its form. These three kinds of information about the word form an integrated part of the lexical entry in the child's lexicon, which is automatically

activated when the word is used. This automatic and simultaneous activation of all three types of information is crucial for the child's appropriate and efficient contextual use of the word.

The lexical development of L2 learners, on the other hand, is much more onerous. When L2 learners, in particular those who learn the target language in classroom settings, learn a word, they face two practical constraints. The first constraint is a lack of input opportunities both quantitatively and qualitatively, i.e. a shortage of opportunities for learners to have sufficient and highly contextual exposure to the target language. This restriction causes significant difficulties in the learners' extraction of the semantic, syntactic and morphological knowledge about a word. As a corollary, many learners turn to their L1 representations, and in particular the L1 translations of the word, in the learning process. This could make the specifications of the L2 word incomplete in the learners' lexicon.

The second constraint, which is likely to have a more serious impact on L2 learning, is the existence of an established conceptual/semantic system in the learners' lexicon. The presence of the established L1 lexical system can greatly facilitate the acquisition of L2 words by providing a source for learners to draw on (Ellis & Shintani, 2013). This source often allows learners to understand and use L2 words almost immediately. The immediate use of L2 words provides a sense of achievement for learners, especially beginning learners, and therefore motivates them to move forward with the cumulative development of their L2 learning. This is what is known as positive L1 lexical transfer (Ellis & Shintani, 2013). However, the established L1 lexical system could also cause L2 learners, in particular adult learners, to over-depend on the L1 semantic and syntactic

knowledge about a word while learning new words. This problem manifests itself particularly in learners' over-reliance on the L1 translations of L2 words. Since "not all L2 words have corresponding L1 translations and not all L1 translations have the same degree of semantic overlap with L2 words" (Jiang, 2000, p. 67), this over-dependence could trigger lexical errors in the learners' language production. When there is not a complete overlap between an L2 word and its L1 translation, either semantically or grammatically, lexical errors in word choices and morphological forms could occur. This is called negative L1 lexical transfer (Ellis & Shintani, 2013).

The evidence of L1 negative lexical transfer has been documented in many studies. One of such studies was undertaken by Zughoul (1991). This study revealed that over 73% of the English lexical errors in the learners' written texts were triggered by interference via L1 transfer. Another study by Wang (2003) examined the writing performances of different proficiency-levels of Chinese learners of English in two writing tasks. Through think-aloud protocols, retrospective interviews and questionnaires, the study concluded that learners at all proficiency levels switch to L1 for lexical searching. In the retrospective interview, a high proficiency participant openly admitted using L1 lexical sources in the completion of his L2 writing tasks and explicitly explained the reasons for such use (p. 363):

> Using Chinese is easy for me to retrieve English words and formulate conceptual information quickly. Otherwise, it takes longer time for me to search for a word in English. And, most of the time I am not sure the word I found in English could express my intended meaning precisely.

Of the two types of lexical errors, inappropriate word choices and morphological forms, the former causes more communication problems (Ellis, 2008a). This can be explained from two perspectives. The main factor is that errors in morphological forms do not have a severe impact on the understanding of the message. For example, the morphological form of *teach* in Sentence 2 below is incorrect, but the error does not affect the conveyance of the meaning of the sentence.

1. She *teaches* every day.

2. She *teach* every day.

However, errors in word choices can cause confusion, misunderstanding or even a communication breakdown. This can be seen in Sentences 3 and 4 below. In both sentences there is a lexical distinction between the employed words (*works* in 3 and *finding* in 4) and the intended words (in parentheses, *jobs* in 3 and *result* in 4) in L2, but not in L1. Sentence 3, for example, was written by a native Arabic speaker. The Arabic translations for both of the English words, *work* and *job*, are the same (Zughoul, 1991). This situation has contributed to, if is not the sole cause of, the inappropriate use of the selected word "*work*" instead of the intended word "*job*" by the learner.

3. There are many works (*jobs*) in the city (Zughoul, 1991)

4. Could you tell me the finding (*result*) of my exam? (Llach, 2011)

Errors in word choices not only frequently appear in the language production of L2 beginners but also in that of high-proficiency learners (Gass & Selinker, 2008). This is partially caused by lexical fossilisation, a cessation of lexical development in the language acquisition of some advanced learners (Jiang, 2000), but more

likely is due to the fact that words in two different languages rarely share exactly the same semantic and syntactic properties (Ellis, 2008a). Taking these two points into consideration, it is hardly surprising to see high-proficiency L2 learners make inappropriate lexical choices in their language production. This problem could be more persistent with learners who largely depend on L1 semantic and syntactic specifications, especially L1 translations, in their L2 lexical learning.

Sentences 5 and 6 below are manifestations of the problem, one in spoken and the other in written form. The two sentences were produced by two advanced Chinese learners of English studying at a university where English is used as the medium of instruction. Both learners were participants in the current study. Sentence 5 appeared at the end of an academic oral presentation and Sentence 6 occurred in an academic essay.

5. This is the end of my presentation. If you have any problems (*questions*), I will be happy to answer them.

6. Parents should concern (*take care of* or *pay attention to*) the physical and also the mental health of their children.

A careful examination of the above two sentences indicates that the most likely cause of the lexical errors is the same as that in Sentences 3 and 4 above: there is a lexical distinction between the employed words and the intended words/expressions in English, but not in the learners' native language, Chinese. In Sentence 5, for example, the Chinese translation for both the used word, *problem,* and the intended word *question* (italicized and in parentheses) is *wenti.* This may have led to the inappropriate use of *problem* in the sentence. The cause of the error in Sentence 6 is the same: the translation for both the word *concern* and the

expression *taking caring of* is *guanxin,* which has very likely given rise to the inappropriate choice of *concern (*instead of *taking care of* or *looking after)* in the sentence.

The above examples provide evidence for the negative influence of L1 lexical transfer. However, as Ellis and Shintani (2013) rightly point out, a learner's first language does not always impede L2 learning; instead it is "a source" that can be drawn on to facilitate the learning of an L2 (p. 240). Built on this understanding, some researchers have started using the term "cross-linguistic influence" instead of "L1 transfer" to acknowledge the positive function of L1 influence. Another important consideration along this line is that an established L1 lexical system is present in the lexicon of all L2 learners, which means that L1 lexical influence is by and large inevitable in L2 lexical learning. This unavoidable factor suggests that L2 learners should focus more on eliminating the first aforementioned restraint faced by L2 learners in their lexical acquisition, i.e. the lack of sufficient and highly contextual exposure to the target language. This means L2 learners should try to enhance their input opportunities, both quantitatively and qualitatively. The qualitative aspect refers to the type and intensity of the exposure and the quantitative part largely pertains to the length of the exposure, which is a determinant element in L2 learners' lexical competence (Llach, 2011).

1.4. Lexical breadth and lexical depth

It is worth noting that lexical competence is different from lexical knowledge. The latter comprises learners' knowledge of the form and meaning of a word which can be simply *remembered*. The former, on the other hand, refers to learners' ability to use the word appropriately and efficiently. This ability has to be

acquired (Meara, 1996). The acquisition process requires extensive and high contextual exposure to the target language. This points to the prominence of contextual factors in the process of ESL and EFL learning.

Working along the same theoretical lines, some researchers (e.g. Qian, 1999; Nation, 2001; Read, 2007) divide the word knowledge in a learner's lexicon into two facets: receptive and productive vocabulary. Receptive vocabulary refers to the words a learner needs for reading and listening. Use of receptive vocabulary only involves recognising the form of a word and retrieving its meaning. Productive vocabulary, on the other hand, refers to the words used in speaking and writing. Productive use of vocabulary requires activating words for language production and thus demands more knowledge about each individual word. The distinction of receptive-productive dimension suggests that assessing a learner's lexical knowledge requires the identification of not only the total number of words one recognises but also how well s/he can use each individual word in language production, of which writing is an essential part.

The development from receptive to productive vocabulary, according to Færch *et al* (1984), can be seen on a continuum, starting from superficial familiarity with the word and ending with an ability to use the word correctly in free production. The process of progressing on this continuum is the development of qualities or depth of one's lexical knowledge. *Depth* of vocabulary knowledge refers to the "extensiveness" of knowledge of each individual word (Schoonen & Verhallen, 2008). This extensiveness defines the quality of a learner's lexical knowledge, or how well s/he knows a word, in contrast to the *breadth* of one's lexical knowledge (also regarded as vocabulary size), which describes the quantity or number of

words one knows in a language. Recognising *breadth/size* and *depth* as two prime dimensions of vocabulary knowledge is crucial in understanding the relationship between vocabulary knowledge and other language skills (Qian, 1999; Schoonen & Verhallen, 2008).

The *size* dimension of a learner's lexical knowledge is single-faceted and thus quantifiable. This allows teachers or/and researchers to set targets for learners to achieve. Nation (2006), for example, suggests that L2 learners should gain a vocabulary size of 6,000 to 7,000 word-families to manage spoken texts and 8,000 to 9,000 to cope with written texts. Nation (2014) makes the target even more specific by suggesting that "a vocabulary size of 9,000 words or more is a sensible long-term goal for unassisted reading of un-simplified texts". Nation's suggestions were based on empirical findings which demonstrated that 98% lexical coverage of a text was needed for learners to gain unassisted comprehension.

The *depth* dimension of a learner's lexical knowledge, on the other hand, is more complex. This construct is multi-faceted. It is not only unquantifiable, but also difficult to conceptualise. Some researchers regard it as "the wooliest, least definable, and least operationalisable construct in the entirety of cognitive science" (Schmitt, 2014, p. 920). Because of this, an increasing number of studies, for example, Milton (2009) and Read, (2004), have called into question whether *depth* should be treated as a single construct.

The primary cause of the difficulties in conceptualising this construct is that learners' size and depth of lexical knowledge are closely linked, and the depth dimension of their knowledge appears only after a certain vocabulary size has

been attained (Milton, 2009). This can create confusions to researchers/teachers who may wonder whether the measured lexical knowledge is the size or depth dimension.

Another important issue is that, in accordance with Nation's (2001) specification, learners' vocabulary knowledge is said to consist to eighteen separate components (nine elements with each containing receptive and productive levels of mastery). These eighteen knowledge aspects do not exist dichotomously in a learner's mental lexicon, i.e. in a known or unknown manner. Instead, they are developmental in nature, with each aspect progressing at a very different rate along the developmental continuum of the learner's lexicon. This can be seen from the illustration of Schmitt's hypothetical graph of a learner's developing knowledge of a word. This situation makes it extremely difficult, if possible at all, to measure the overall depth of a learner's lexical knowledge.
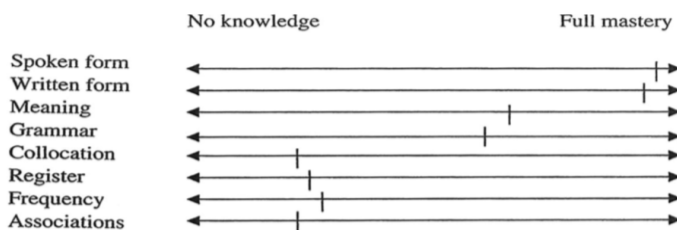
Figure 1-1. Developing knowledge of a word (Schmitt, 2010, p. 38)

The last and perhaps the major cause is the organisation of the eighteen aspects of vocabulary knowledge in a learner's mental lexicon. The depth dimension refers to the quality aspect of a learner's vocabulary knowledge, which depends greatly upon the associations between words in the learner's lexicon (Milton, 2009; Schmitt, 2010; Schmitt; 2014). Factors such as the number of links a word has

with other words, the way these links are organised and the level of strength between the links, could all have an impact on the learner's language production.

1. *Number of links a word has with other words*

Many of the above word knowledge aspects (e.g. collocation, connotation and register constraints) are implicit in character. Their acquisition thus requires extensive and highly contextual exposure to the target language. This could mean some aspects of a learner's word knowledge, for example, polysemous meaning senses, have been well developed because these parts of the knowledge are amenable to intentional learning. Other aspects, such as word associations and collocations which requires much exposure to diverse contexts, however, have not been acquired by the learner and thus links in these aspects have not been fully developed.

2. *Ways that the links are organised*

According to a study by Meara (1982), words are linked very differently in the lexicon of an EFL learner and a native speaker. A native speaker's lexis is primarily organised in paradigmatic associations whilst that of an EFL learner is largely in syntagmatic or "clang" associations (how similar words sound). This disparity in organising word knowledge components could affect the appropriateness and effectiveness of EFL learners in word use. Another factor is that most EFL learners develop their vocabulary links according to their existing L1 lexical links, and perhaps also their individual habits in building networks in the lexicon (Milton, 2009). These L1-related networks could lead to further stigmatisation of EFL learners' language performance.

3. *Levels of strength between the links*

Many types of depth lexical knowledge, for example, word associations and connotations, are intuitive and difficult to explain and thus demand extensive exposure to diverse contexts. Insufficiency in either frequency or duration in the exposure could affect the strength of the links. This means links between many words may have been developed but some of them are not very strong. This lack of strength could exert a substantial impact on the learner's ability to effectively use many lexical items in the lexicon. As a consequence, Daller, Milton and Treffers-Daller (2007) subdivide the quality aspect of lexical knowledge into *depth* and *fluency*. This division creates a three-dimensional space to view a learner's vocabulary proficiency i.e. *breadth, depth* and *fluency*. Here fluency is defined as the ease (speed and accuracy) with which words are used in language production. This division is important because it "moves the conceptualisation of lexical proficiency onward from simple knowledge to the ability to use that knowledge" (Schmitt, 2014, p. 920). To improve learners' vocabulary fluency is to increase the automaticity of their lexical use and should be the ultimate goal of most language learners.

Notwithstanding the complexity in conceptualising the depth construct of vocabulary knowledge, researchers' views on one point are consistent: the depth dimension is more associated with the two productive language skills, i.e. speaking and writing. Of these two skills, writing, and in particular formal writing, exerts a higher demand on learners' depth of word knowledge due to various conventions required in different genres of writing. This could be why some researchers (e.g. Leki & Carson, 1994; Walters & Wolf, 1996) assert that lexical use is one of the most important features that determine EFL and ESL learners'

writing quality. This observation possibly also means that the depth dimension of lexical knowledge can better predict the writing ability of EFL (including ESL) learners. Considering the prominent role of contextual exposure to the target language in the acquisition of depth construct, contextual factors may affect the predicting power of depth of lexical knowledge.

Based on the above discussion, the current study proposes to investigate the role of depth of vocabulary knowledge in predicting the writing proficiency of EFL learners. Through an in-depth investigation, the study aims to examine the following research questions: (1) whether EFL learners' depth of word knowledge correlates with their writing proficiency and, if so, to what extent they relate to each other; (2) whether depth of lexical knowledge is a more reliable predictor for the writing proficiency of EFL learners compared with vocabulary size; and (3) whether different language learning contexts affect the association between EFL learners' depth of vocabulary knowledge and their writing proficiency and, if so, how learners from unconducive learning contexts are affected.

1.5. Summary

This chapter has briefly overviewed concepts pertaining to L2 lexical learning and its close relationship with L2 writing. It has, in particular, examined the roles of *size* and *depth* of vocabulary knowledge in a learner's mental lexicon. A more comprehensive and detailed review of the literature pertaining to each of the above research questions is presented in Chapter 2.

CHAPTER TWO

LITERATURE REVIEW

2.1. Overview

This chapter reviews the literature pertaining to the three research questions of the current project. The first part focuses on measures that have been hitherto used to assess EFL learners' vocabulary knowledge. By critically examining the important functions and limitations of each measure, it identifies the need for multiple measures in assessing learners' productive knowledge. The second part centres on the role of contextual factors in determining the lexical knowledge in learners' mental lexicon. By comparing the language learning contexts in Hong Kong and mainland China, it suggests a possible disparity in the productive language skills of the two groups of participants even though they have a similar vocabulary size. The first part bridges the way to Research Questions 1 and 2 and the second part leads to Research Question 3.

2.2. Vocabulary tests

2.2.1. Tests measuring vocabulary size

Despite the importance of deep word knowledge in language learning, investigations into second language (L2) acquisition have mostly been into breadth rather than depth of vocabulary (Vermeer, 2001). This is mainly because of the practicality of vocabulary size tests. Size tests only measure one dimension of learners' lexical knowledge, i.e. the amount of one's vocabulary or the number of words one knows. This construct makes size tests more focused and therefore relatively less complicated in test design. It is also this one-dimensional construct

that makes most standardized size tests take the form of levels tests. For example, the two widely accepted tests measuring vocabulary size, Nation's (1990) *Vocabulary Levels Test* (VLT) and Meara and Milton's (2003) *X-Lex*, are both levels tests. These levels tests can function as an efficient measure for placement and admission in language teaching programmes. They can also act as a relatively simple research instrument providing researchers and language teachers with evidence showing the growth in vocabulary size before and after an experimental intervention.

There are a number of established size tests including *Eurocentre's Vocabulary Size Test* (Meara & Johnes, 1990), *EFL Vocabulary Test* (Meara, 1992), *Vocabulary Levels Test* (Nation, 1990, 2001), *X-Lex* (Meara & Milton, 2003) and *Vocabulary Size Test* (Nation & Beglar, 2007; Beglar, 2010). Up till now, no vocabulary test has achieved "the same wide acceptance" as the VLT (Read, 2007, p. 114). See Section 2.4 for a more detailed discussion about the VLT.

The VLT, together with other size tests, has played an important role in predicting success in learners' proficiency, and even academic achievement (Saville-Troike, 1984; Laufer, 1997; Roche & Harrington, 2013). Vocabulary size measures are found to correlate well with scores on achievement tests such as IELTS (International English Language Testing System) and TOEFL (Test of English as a Foreign Language). Stæhr (2008), for example, compares examination grades on listening, reading and writing papers with the vocabulary size of the testees and identifies a connection between vocabulary size and examination performance, particularly in reading.

Size tests have also been used in measuring other areas of learning and teaching. Fan (2001), for example, identifies the vocabulary needs for academic achievement of tertiary students in Hong Kong by employing the VLT. Lin and Morrison (2010) even use the test to gauge the impact of change of government language policy on tertiary students' academic studies.

2.2.2. Tests measuring depth of word knowledge

Despite these robust initiatives in using size tests for various language learning and research purposes, the limitations of size tests have been recognised by researchers. Size tests, after all, only measure the quantity but not the quality of a learner's lexical knowledge. Knowing a word, as pointed out in Richards (1976) and Nation (2001), involves much more than associating the form of a word with a simple statement of its meaning. Various types of knowledge of a word, ranging from spelling, pronunciation and morphological forms to syntactical and semantic features, constraints in use, collocations and frequency (Nation, 1990, p. 31), are required of a learner for successful use of the word. These multidimensional components of deep word knowledge makes it more complicated to assess the depth of lexical knowledge. To counter this difficulty, most researchers choose to focus on the key aspects of word knowledge. This can be reflected in the WAT, *Word Associates Test* (Read, 1993, 1998). Another test measuring deep word knowledge that has also received attention is Paribakht and Wesche's (1993) *Vocabulary Knowledge Scale*. This test, however, due to its various problems, including insensitivity to many aspects of deep word knowledge and unreliability in its function as a scale test (Milton, 2009, pp. 160-161), has not been as widely used as the WAT.

The WAT was initially developed for the tertiary students in a New Zealand university. After its development, the test has undergone several revisions and validation tests. The format of the revised version has been adapted by a number of Dutch scholars (e.g. Greidanus, Beks & Wakely, 2005) to design their versions of associates tests and the tests have been found to be an effective measure of lexical knowledge for advanced EFL learners. Schoonen and Verhallen (2008) even produced a revised version of the associates test for primary school pupils in the Netherlands, where English is a second language (L2). Their results indicate the test is "a reliable and efficient method" in understanding the pupils' vocabulary knowledge (p. 211). This is possibly why the WAT, together with variations of its formats, has been the "most-used measure of depth" (Schmitt, 2014, p. 938). See Section 2.6 for a more detailed discussion of the test.

The WAT has also been used to predict success in language learning. Nassaji (2004), for example, identifies a significant relationship between L2 learners' deep word knowledge and their lexical referencing strategy by comparing the learners' scores from the WAT with their degree and type of reading strategy use. Using a modified version of the test with a sample size of 217, Qian (2002) investigates the association between vocabulary knowledge and L2 reading ability. A new version of the WAT was also designed by Qian and Schedl (2004) to explore the possibility of using this measure in the reading comprehension part of a new TOEFL test. Given that the WAT measures more dimensions of a learner's lexis, it is possible that this test can serve as a more precise indicator of learners' language skills than size tests. This hypothesis was, again, tested by using the WAT in Qian (1999), in which he confirms the observation that vocabulary size

has a high and positive correlation with L2 reading comprehensive ability. Meanwhile, he provides empirical evidence showing that depth of vocabulary knowledge makes a unique contribution in predicting L2 learners' ability in reading comprehension, over and above the contribution made by size of vocabulary. Based on his findings, Qian also makes a theoretical assumption that depth of word knowledge should also play an important role in predicting proficiency of other language skills, such as listening and writing.

This hypothesis, however, is not fully supported by Stæhr (2009) empirically. Using a version of the WAT especially designed for the project (with headwords being either adjectives or verbs and all selected from Coxhead's (2000) Academic Word List), Stæhr documents a strong correlation between learners' vocabulary knowledge (in the dimension of both breadth and depth) and their listening comprehension; but meanwhile he reports that depth of vocabulary knowledge adds "very little" (only 2%) to the variance already explained by vocabulary size in predicting listening comprehension (p. 592).

Although the disparity in the above two studies can be explained by the different frequency levels of lexis required in listening and reading materials (Qian, personal communication, 2014), more research is necessary to explore the association between depth of lexical knowledge and L2 learners' language proficiency. This is not only because of the differences in the research results of these studies but also because of the insufficient "attention" paid to this area (Zareva, Schwanenflugel & Nikolova, 2005, p. 568). The very few studies in the literature are mostly related to connections between deep lexical knowledge and skills in reading and listening, for which only receptive skills are required. More

studies, therefore, are needed to explore the relationship between deep word knowledge and the productive language modalities, i.e. speaking and writing. Extending the studies by Qian (1999) and Stæhr (2009), the present study aims to explore the relationship between depth of word knowledge and an important language modality: writing. Based on Qian's theoretical assumption, the study proposes the first hypothesis:

Hypothesis 1: EFL learners' depth of word knowledge positively correlates with their writing proficiency.

To test the above hypothesis, measures to test vocabulary knowledge in learners' mental lexicon are important. Sections 2.3 - 2.6 below discuss the details of these assessment measures.

2.3. Vocabulary Levels Test

The Vocabulary Levels Test (VLT) is "the *de facto* standard vocabulary size test" (Meara (2010). Designed by Nation in the early 1980's, this test gauges learners' receptive vocabulary by assessing their performance in word-definition matching, as in the example below:

1. adequate

2. internal   ____1____ enough

3. mature   ____3____ fully grown

4. profound   ____5____ alone away from other things

5. solitary

6. tragic

This test was developed to estimate a learner's lexical knowledge according to the frequency level of a word, i.e. how often the word appears in text and discourse. There are six levels in the test: the 1,000, 2,000, 3,000, UWL (University Word List), 5,000 and 10, 000 levels. The test was used to measure learners' vocabulary size for a number of years, and was regarded as the "nearest thing we have to a standard test in vocabulary" (Meara, 1996, p. 38) until it was revised in 2001. This revision led to three new and sufficiently validated versions of the test (see Schmitt, Schmitt & Clapham, 2001). The most important changes in the revised versions are the extended test items (from 18 in the old version to 30 in the new versions) and the change of source for testing academic lexis. The original source is Xue and Nation's UWL (1984), whereas the source for the updated versions of the test is Coxhead's Academic Word List (AWL) (Coxhead, 1998).

The VLT was originally designed as a "diagnostic tool" for teaching, i.e. helping classroom teachers to prepare appropriate vocabulary programmes for students, but later on its uses were extended to evaluation and research purposes. Apart from these functions, the test, together with other size tests, has also played an important role in predicting success in learners' proficiency, and even academic achievement (Laufer, 1997; Saville-Troike, 1984). Results from the test are found to correlate well with scores on achievement tests such as IELTS (International English Language Testing System) and TOEFL (Test of English as a Foreign Language). Stæhr (2008), for example, compares examination grades on listening, reading and writing papers with the vocabulary size of the testees and identifies a connection between vocabulary size and examination performance, particularly in reading.

This test has also been deployed in measuring other areas of learning and teaching. Fan, for example, identifies the vocabulary needs for academic achievement of tertiary students in Hong Kong (Fan, 2001) and the vocabulary learning strategies used by these students (Fan, 2003). With the same test, Lo and Murphy (2010) study the difference in vocabulary knowledge and vocabulary growth between students in Hong Kong's two major language teaching programmes, immersion and regular classroom teaching.

Although being "the *de facto* standard vocabulary size test", the VLT can only measure four levels (the first 2,000, 3,000, 5,000, and 10,000 frequency levels). The UWL contains words from different frequency levels and therefore cannot be considered as one level of one's lexical knowledge. Words at other frequency levels cannot be tested. This limitation of the VLT has led to the design of the VST (Vocabulary Size Test, see Nation & Beglar, 2007; Beglar, 2010). This test has made it viable to measure word families at any of the first 14,000 frequency levels and has been increasingly accepted as a research tool (Nation, 2012).

Despite these initiatives in using size tests for various language learning and research purposes, the limitations of size tests have been recognised by researchers. Size tests, after all, only measure the quantity but not the quality of learners' lexical knowledge. The quality dimension of learners' vocabulary knowledge involves not only associating the form of a word with a simple statement of its meaning but also understanding other types of knowledge of the word, for example, spelling, pronunciation, morphological forms, collocations, frequency, syntactical and semantic features, and constraints in use (Nation, 1990,

Comment [LL[1]:

p. 31). To counter these limitations, Laufer and Nation (1995) designed another levels test, the *Productive Levels Test*.

2.4. Productive Levels Test

This test was developed by Laufer and Nation (1995; 1999). Its purpose is to measure learners' productive or active vocabulary by asking them to fill in the missing letters of a word in a sentence (see example below).

    In a hom_____ class all students are of a similar proficiency.

The test adopts the overall structure of the original VLT, i.e. six levels and 18 test items at each level, and assesses the same target words as the VLT. It was sufficiently validated (see Laufer & Nation, 1999) but has not been revised since the time it was developed.

One of the key reasons for not updating the test could be the uncertainty of what the test actually measures. Although called Productive Levels Test, the test, according to Read (2000, p. 155), can only estimate learners' *recall* ability, the ability of eliciting the target word from memory when stimulus is provided. On the continuum of a learner's vocabulary development, there is a considerable distance between his/her recall ability and productive ability. Productive ability demands more knowledge about each individual word, such as its grammatical functions, collocations, register and frequency (Nation, 2001, p. 27). Laufer and Nation acknowledge the limitation of the productive test by referring to it as a *controlled* productive test (Laufer & Nation, 1999; Laufer, 1998).

Despite the limitation, this test has been found to be a reliable and practical measure and can distinguish learners at different proficiency levels (Laufer &

Nation, 1999). Together with the VLT and other measures for vocabulary knowledge, this test has also been used for a variety of research purposes, for example, to gauge the impact of change of government language policy on tertiary students' academic studies (Lin & Morrison, 2010), to identify the relationship between incidental vocabulary learning and the retention of productive vocabulary knowledge (Yamamoto, 2011), and to explore the association between learners' different facets of vocabulary knowledge and their quality of writing (Shi & Qian, 2012).

Given that the Productive Levels Test is not ideal in measuring learners' ability in productive use of vocabulary, researchers have designed other tests to assess depth of word knowledge. In designing these tests, researchers have to consider the multidimensional components of deep word knowledge, which makes the assessment more complicated. One way to address this issue is to focus on the key aspects of word knowledge. This can be seen in the *Word Associates Test* (Read, 1993, 1998). Another test measuring deep word knowledge that has also received attention is Paribakht and Wesche's (1993) *Vocabulary Knowledge Scale*. This test, however, due to its various problems, including insensitivity to many aspects of deep word knowledge and unreliability in its function as a scale test (Milton, 2009, pp. 160-161), has not been as widely used as the Word Associates Test.

2.5. Word Associates Test

Initially known as the *Word Associates Format*, the WAT was developed for tertiary students in a New Zealand university by John Read in 1993. In the following few years, the test underwent a series of revisions and validations, and a high level of reliability, validity and utility of the test format has now been

established (Read, 1994, 1995; Qian, 1999; Lee, 2003). The latest version covers 40 test items (Read, 1998). Each item contains a target word, which is an adjective, and eight other words. Of the eight words, four are either semantically or collocationally associated with the tested word, and the other four are not related to the stimulus word in any sense. See one example below:

dense

crowded hot noisy thick	forest handle smoke weather

The WAT has the three essential features of a vocabulary test for capturing "lexical dimensions with respect to the language proficiency of the L2 learners" (Zareva, Schwanenflugel & Nikolova, 2005, p. 568), i.e. practical, theory-driven and empirically-tested. Contributing to these three features, the WAT has been "the most-used measure of depth" in studies pertaining to the qualities of learners' vocabulary knowledge (Schmitt, 2014, p. 938).

2.5.1. Practical

Compared with other methods of deep word knowledge testing, for example definition and description, the format of the WAT is simple and efficient in use. There are only 40 tested items, and testees do not need to write down much information except for ticking the words that are related to the stimulus words. This is time-efficient, and allows testees to demonstrate their understanding of "some key elements of the core meaning of the target word" in a reasonably short time (Read, 2007, p. 113).

2.5.2. Theory-driven

Due to the multifaceted dimensions of deep word knowledge, researchers have reached an agreement that, instead of testing the full range of depth of word knowledge, tests on depth of lexical knowledge should focus on some key aspects of word knowledge. Following this line, the WAT is set to test word associations, i.e. various semantic and collocational relationships that a word has with other words in a language. By doing this, the test identifies several key components that constitute the knowledge involved in knowing a word, such as form-meaning relationship, concept and referents, associations and collocations. These aspects are important components listed in the analysis about word knowledge in Nation (2001, 2008), one of the most influential researchers in the area of L2 vocabulary studies. See the summary of Nation's analysis below:

Form: pronunciation, spelling, word parts

Meaning: form-meaning relationship, concept and referents, associations

Use: grammatical functions, collocations, constraints on use (register, frequency...)

2.5.3. Empirically tested

The WAT adopts effective arrangements to reduce chances of guessing in the test. As stated before, the relationship between the target word and the other eight words are primarily semantic and collocational. The meaning-related or paradigmatic words are posited in the left-hand box and the collocation related or syntagmatic ones in the right-hand box. The paradigmatic words in the example are *crowded* and *thick* (both are synonymous with the target word *dense* in one

respect or another) and the syntagmatic words are *forest* and *smoke* (both collocate with the target word *dense*). The example has two correct answers on each side by chance, but this is not always the case. Since there are always four correct choices, there could be two alternatives in the choices: three correct answers on the right-hand side and one on the left-hand side, or vice versa. This arrangement largely minimises the possibilities of guessing, which partially contributes to the high reliability and validity rates of the test. In one of the validation tests for the WAT (N=94), Read (1995) reported a high correlation of .82 between the WAT and a vocabulary matching test. In this test, the reliability rating of the WAT was .93 and that of the matching test was .90.

Although practical, theory-driven and empirically tested, the WAT has its limitations. The first one is that it only measures two aspects, i.e. meaning (including synonymy and polysemy) and collocation (Qian, 2002; Qian & Schedl, 2004), of the various facets of lexical knowledge listed in Nation (2001). Other aspects, such as grammatical functions, register and frequency, are not tested. The second limitation is that words involved in each item, both the target word and the other eight words, are either adjectives or nouns. Verbs and adverbs are excluded. Nevertheless, as argued in Qian (2002), this measure is an "efficient instrument" for research (p. 525) and the tested components in the measure are among the most important in discussions about lexical knowledge in related literature. It is therefore a well-accepted research instrument.

2.6. Other measures of lexical knowledge

There is little doubt that the above vocabulary tests can help identify learners' lexical knowledge and even their language proficiency. However, it has also been

widely recognised by researchers that in a learner's mental lexicon, there is a distance between one's vocabulary knowledge and vocabulary fluency (Milton, 2013; Daller, Milton & Treffers-Dallar, 2007). This distance is between "what one knows and the extent to which one is consciously aware and in control of vocabulary" (Zareva, Schwanenflugel & Nikolova, 2005, p. 572). Words that one is in control of are those that one can call to mind freely and use appropriately. This dimension of lexical knowledge cannot be fully measured in any vocabulary tests, in which tested words are elicited. This is because a learner could choose correct answers for the stimulus words in a test but may not be able to freely use them in free language production. Since "language is meant to be communicative" (Milton, 2009, p. 125), teachers and researchers are more interested in identifying the lexis that a learner is in control of. Efforts made in this area are meant to measure the lexical richness in language production, such as a piece of writing or a spontaneous speech. These measures include lexical originality, lexical density, lexical diversity, lexical sophistication and lexical frequency profile. Of these measures, lexical diversity, lexical sophistication and lexical frequency profile are often deployed to predict learners' language competence.

2.6.1. Lexical diversity

Lexical diversity, also referred to as lexical variation, measures the "variety of active vocabulary deployed by a speaker or writer" (Malvern & Richards, 2002, p. 78). It calculates the type-token ratio (TTR) in a piece of language production. This measure draws on an assumption that the higher a learner's proficiency is in a language, the better s/he is able to call on a variety of words in his mental lexicon when speaking or writing. This means an essay by a learner of high

proficiency should contain more diversified lexis and therefore the TTR derived from the essay should be higher than that from an essay by a low-proficiency learner. The measure, although still being much used, has been criticised for its limitations. The one that causes most concern is its over sensitiveness to length of texts (TTR curve falls with the increase of text size). This causes unreliable results when sample size varies, which explains why TTR is now only used "with texts of equal length" (Milton, 2009, p. 126).

To address this issue, several other measures, such as Guiraud's (1954) *Index,* Malvern and Richard's (1997, 2004) *D*, and Daller *et al.*'s (2003) *Advance Guiraud Index*, have been developed to measure lexical diversity. Data drawn from these measures demonstrate a significant correlation between learners' lexical diversity and their language proficiency (see Daller & Phelan, 2007). However, the very same study also shows there is no relationship between a learner's essay grade and the lexical variation derived from his/her essay, either in D or TTR. Similar findings can also be found in Yu (2009), who reports that D has no significant correlation with the essay grades for his Philippine and Chinese participants, the two biggest groups in his study. These findings are not surprising because, after all, lexical diversity only captures the feature of how varied words are in a piece of writing but not how well these words are used in the writing. Apart from this, lexical variation could also be affected by activity types (e.g. a genuinely free conversation versus a controlled conversation) and by the topics in the same activity (e.g. a familiar topic versus an unfamiliar topic in essay writing). This is why many researchers, (e.g. Yu, 2009; Laufer, 1994) suggest that other

measures for lexical richness, such as lexical frequency profile, should also be deployed to better predict a learner's language productive ability.

2.6.2. Lexical sophistication

Lexical sophistication refers to the proportion of infrequent words in a text. Infrequent words are also known as "non-basic" (Laufer, 1995, p. 23), "advanced" (Laufer & Nation, 1995, p. 309) and "rare" (Read & Nation, 2002) words. There is no rule as to what consists of infrequent words. In Meara and Bell's *P_Lex* (2001), a computer program designed to capture the degree of word complexity of a written text, any words outside the most frequent 1,000 words belong to this category. In other programs, such as Nation and Laufer's (1995) *Lexical Frequency Profile* (LFP) and Heatley *et al.*'s (2002) *Range*, words outside the most frequent 2,000 words in West's (1953) General Service Word List are placed in this category. This is why Laufer (1994) also uses "beyond 2,000" words to refer to advanced lexis in a learner's language production.

According to Nation and Laufer (1995), the higher a learner's proficiency is, the more likely s/he is able to deploy infrequency words in language production (p. 316). This finding has been empirically supported in Daller and Phelan (2007) and Lerenzo-Dus (2007), in which they find lexical sophistication is an effective indicator of a learner's writing, and in Lu (2011) who identifies a strong connection between a learner's speaking ability and the lexical sophistication in his/her oral production.

This measure, however, is also affected by length of texts. Sample sizes either too long (over 300 words) or too short (below 200 words) can cause unreliable results (Meara, 2005; Smith, 2005). The sensitivity level of the tool may be affected if

sample sizes fall beyond its scope. To address this concern, Meara and Bell (2001) proposes another measure, *P_Lex*. This measure calculates the value *lambda*, the number of infrequency words (words beyond the first 1,000 word range) in every 10-word chunk in a text. In *P_Lex*, the more infrequent words a text contains, the higher the lambda score is. *P_Lex* has helped a number of researchers (e.g. Read & Nation, 2006) successfully assess the language proficiency of EFL and ESL learners.

2.6.3. Lexical Frequency Profile

Drawn from different word lists, such as General Service Word List (West, 1953), Academic Word List (Coxhead, 2000) and BNC (British National Corpus) Lists (Nation, 2006), various word profiles of a learner can be produced. Of these profiles, the LFP (Lexical Frequency Profile) is the most widely used one.

The LFP was developed by Laufer and Nation (see Laufer, 1994, 1995; Laufer & Nation, 1995) to evaluate essays written by EFL learners. This measure calculates the free productive knowledge or what Read (2000) calls the *use* ability of vocabulary. It does so by identifying the lexical richness at four frequency levels: the first 1,000 words, the second 1,000 words, words from the AWL, and words "not-in-the-list", which are normally low frequency words. Take a piece of writing consisting of 600 words, for example. Of these 600 words, 400 belong to the first 1,000 most frequent words, 100 to the second 1,000, 60 to the AWL, and 40 to the "not-in-the-list" category. When these numbers are converted into percentages, the LFP of this piece of writing is 67%-16%-10%-7%. In the LFP, a word is defined as a base form plus its inflected and derived forms, i.e. a word family.

This measure was validated by Laufer and Nation (1995) and has proved to be an effective research and assessment measurement. Many researchers have not only used this assessment tool to identify the lexical richness of learners' written work but also to predict learners' language performance. Morris (2001) and Morris and Cobb (2004), for instance, use this tool to predict the academic and pedagogic performance of TESL (Teaching English as a Second Language) trainees, while Coniam (1999) employs it to provide an indicator of the written language proficiency of Grade 13 students in Hong Kong. Laufer also deploys this measure to produce "condensed lexical profiles" (1994, p. 27) to compare the changes in the basic words (the first 2,000 most frequent words) and the advanced words (words in AWL and not-in-the list) in a learner's writing.

A free on-line program called Vocabprofiler (Cobb, 2011) has made the LEP more widely used in research and teaching. This program not only analyses the frequency profile of a piece of writing but also shows the word variation level (type-token rate) in the writing.

Nevertheless, concerns have been expressed regarding Laufer and Nation's claims for this instrument (Meara, 2005; Smith, 2005). These concerns include the sensitivity level of the tool and length of texts. Although Laufer and Nation (1995) contend that the LFP is a sensitive research tool and that sample sizes of over 200 words produce stable results, Smith (2005) finds that this tool is only sensitive if the study focus is on "group differences rather than individual differences". He also demonstrates that only consistent sample sizes of around 300 words can generate reliable results (p. 448). Abbasian and Parizad (2011) conclude that "the LFP indexes are reliable and valid to some extent but not so

strong as to be used as a stand-alone measure for the assessment of lexical richness" (p. 226).

It seems there is no single perfect tool that can definitely measure productive vocabulary of learners. As a result, Nation (2007) suggests that multiple measures be used to gain a full picture of a learner's productive knowledge.

2.6.4. Corpus-informed measures

The lexical-richness measures discussed above focus on the overall vocabulary knowledge in learners' language production. However, to identify specific problems in word use of a group of learners, corpus-informed measures seem more effective. This is particularly true after the emergence of a computer-software called Wordsmith (Scott, 2012) and the development of collections of learner language, i.e. learner corpora.

A learner corpus is a collection of learner language. Learner corpora vary according to their size and type. There could be a corpus of 20 learners in a secondary school class, a corpus of a particular group of learners, for example, Corpus of Chinese Learners of English (Granger, Dagneaux, Meunier & Paquot, 2009) and a corpus of all learners of a language, i.e. International Corpus of Learner English. This flexibility in size and type of learner corpora allows researchers and teachers to identify patterns of lexical use in a certain genre, for example, reporting verbs in academic papers (Bloch, 2010); or by a particular group of learners, for example, dialogic and hortatory writing in IELTS Test by Chinese candidates (Mayor, 2006); overuse, underuse and misuse of the pronoun *it* by Chinese learners of English (Lin, 2002); inappropriate use of conjunctions/sentence connectors by students in Hong Kong (Bolton, Nelson &

Hung, 2002); and overuse of first person pronouns (*I*) and its reflexive forms (*my, mine*) in the academic texts written by Hong Kong and mainland learners (Li, 2014).

The majority of the above studies were conducted by using contrastive interlanguage analysis (Granger, 1996), i.e. comparing a learner corpus to "a control corpus of expert production" which is usually a native speaker corpus (Granger, 2012, p. 3). This method allows researchers to uncover features that are typical of the language use in the learner language and therefore helps to shed light to the lexical items and grammatical structures overused or underused by the learners. Since underuse and overuse of a language item often lead to misuse of it, contrastive analysis could help researchers and language specialists better understand the causes of problems in learners' language use. This method is particularly effective in providing insight into the "foreign-soundingness of perhaps otherwise error-free advanced interlanguage" (p. 4). Investigation into learner corpora has indicated that inappropriate lexical use including word choice, collocation (word co-occurrence on the lexical level, including semantic prosody), colligation (word co-occurrence on the lexico-syntactic and/or syntactic level) and phraseology in a broad sense, are the most serious errors in learners' interlanguage (p. 4).

These robust corpus-driven studies on word use of EFL learners have informed teaching and thus indirectly helped students improve their language proficiency.

The above measures that have been hitherto used to assess learners' vocabulary knowledge have indicated the importance of lexical knowledge in learners' writing. To gain a deeper insight into the relationship between these two factors, it

is necessary to examine how learners' lexical knowledge contributes to their writing proficiency.

2.7. Vocabulary knowledge and writing proficiency

One's lexical knowledge and writing proficiency are closely interrelated. Sufficient lexical knowledge contributes to the effectiveness of writing. Meanwhile writing practice is most conducive to imprinting newly learnt words into memory (Laufer, 2013). This interrelated relationship applies to writing in one's first language (L1) as well as in a second language (L2). However, compared with L1 writing, writing in L2 is more effortful because some learners' linguistic knowledge required for writing may not be fully developed (Schoonen *et al*, 2003; Kormos, 2013). To understand this, it is important to discuss the roles of various facets of knowledge (linguistic, cognitive and metacognitive) in the writing process.

2.7.1. Linguistic and cognitive knowledge in the writing process

Writing is a demanding task. Various cognitive and motivational factors exert influence on the completion of a writing task. Cognitive variables mainly consist of one's language learning aptitude and capacity of working memory. Motivational variables include one's writing needs, attitudes towards writing and perceived value of the given task (Kormos, 2012).

Writing is also a complex task, which is fully manifested in the writing process. One of the most influential models in this regard was proposed by Kellogg (1996). This model divides writing into three, albeit recursive, processes: formulation, execution and monitoring. The first process involves writers' formulating and

organising ideas, which demands their cognitive and metacognitive resources. The second process is to execute the formulated plans, i.e. to translate ideas into linguistic forms. This stage requires writers' cognitive and linguistic knowledge, including retrieving related lexical items, encoding clauses and sentences syntactically, and establishing cohesive relationships in the written text. The last process is to monitor the quality of the created text. It is to ascertain whether the composed text effectively expresses the writer's intention and, if not, revision needs to be undertaken. This stage requires the use of writers' linguistic, cognitive and metacognitive resources.

Kellogg's model indicates clearly that a writing task requires different facets of the writer's knowledge, some at lower-order (linguistic) and some at higher-order (cognitive and metacognitive). Studies have provided empirical evidence showing that cognitive and metacognitive knowledge is a major discriminating factor of the performance of L1 writers (Victori, 1999) whereas linguistic knowledge is more instrumental in predicting the writing proficiency of L2 learners (Sasaki & Hirose, 1996). This difference can be explained by the theory of working memory capacity.

2.7.2. Working memory in writing processes

Working memory refers to a person's mental capacity to store and orchestrate resources connected to a task (Baddeley, 2003). The role of working memory is described as "a bottleneck for learning" (Gathercole & Alloway, 2008, p. 12). During the process of composing a text, for example, a writer's linguistic, cognitive and metacognitive knowledge stored in long-term memory has to be processed by working memory at the same time. If a writer has obtained automatised linguistic

knowledge, as is often the case in L1 writing, the writer can attend more closely to the cognitive aspects of writing such as organising ideas and enhancing the persuasiveness of the text. However, many L2 and EFL writers are found to be "tied up with word- or sentence-level process" (Sasaki & Hirose, 1996, p. 158). When this lower-order process load increases, the working memory capacity available for the higher-order (cognitive) aspects of writing, which are required in all the three processes in Kellogg's model, is reduced. The corollary of this insufficient activation of cognitive knowledge is very likely reduced quality of the written product.

The coordination function of working memory on different facets of a writer' resources in L1 and L2 writing points to the imperative importance of automatisation of EFL learners' lexical and syntactical knowledge. For advanced L2 learners, whom the present project is focused on, lexical knowledge is more important. At this proficiency level, most of these learners have gained a sufficient understanding of syntactical rules (You, 2010). Their level of automaticity in lexical use has thus become more instrumental in composing a text. This possibly explains why Gass and Selinker (2008) believe that "language learning is largely lexical learning" (p. 173). Support for Gass and Selinker's assertion can be found in a study of tertiary students in China by Gan, Humphreys and Hamp-Lyon (2004, p. 234), in which one of the students commented that

> *I think I've already mastered the basic English grammar at secondary school; now the main task for me is to expand my vocabulary.*

A learner's lexical automaticity in vocabulary use, also known as lexical fluency, refers to the ease (i.e. speed and accuracy) with which words are retrieved for

language use. Although this topic is beyond the scope of the current study, understanding its basic parameters is crucial. Achieving lexical automaticity amounts to reaching the strongest level of the depth development of one's lexical knowledge, and should therefore be the ultimate goal of most language learners. Attaining this goal, however, could be challenging for many learners. This is due to the fact that only lexical items with right connections with other words in a learner's mental lexicon can be retrieved quickly and used appropriately in language production (Meara, 1996). Factors, such as how many links a word has with other words, how strong these links are, and how the words are organised (e.g. primarily in paradigmatic or "clang" associations), exert a significant impact on the automaticity of the learner's lexical use (see Section 1.4 for more detailed discussions). Following this deduction, the current study proposes that, compared with vocabulary size, the depth dimension of lexical knowledge could better predict the writing proficiency of EFL learners, hence, the second hypothesis of the present study:

Hypothesis 2: Compared with vocabulary size, depth of lexical knowledge is a more reliable predictor of the writing proficiency of EFL learners.

Apart from the above linguistic and cognitive factors, motivational factors are also important determinants in writing processes. Motivational variables are not only related to learners' interest in a given task but also connected to the social, cultural and educational values of writing activities. This applies to both L1 and L2 writing. To L2 writers, instructional practices also exert a force on their attitude towards writing (Kormos, 2012). These variables are language learning contextual factors and will be discussed in Sections 2.8-2.10.

2.8. Contextual factors in EFL learning

Learning context refers to the cultural and social environment where learning takes place (Ellis, 2008b; Gu, 2003a). Many contextual factors affect second and foreign language learning and they play a crucial role in determining learning strategies, classroom ethos and learning outcomes (Ellis, 2008b; Schmitt, 2008). These factors are evident largely at two levels, albeit intertwined with each other (Ellis, 2008b; Butler, 2011). The first one is at the conceptual level, and is related mainly to learner beliefs about language learning and teacher beliefs about language teaching. The other one is at the institutional-societal level, which could include choices of teaching and assessment approaches; and the availability of input/output opportunities, i.e. the opportunities to use the target language (or lack thereof).

2.9. Contextual factors at conceptual level

2.9.1. Learner beliefs

Learners adopt a range of learning strategies, and particularly in language learning. These strategies steer their approaches to complete learning tasks. For example, in learning vocabulary, one learner might choose to memorise a variety of word lists because, compared with others, this strategy is more effective in helping him/her accumulate words. Another learner might believe that words in his/her mental lexicon should be accrued via language use, such as reading, listening and communicating with people, because words acquired in such language contexts could be more easily activated in language production. Language learning

strategies are therefore closely related to a learner's philosophy of learning, which Horwitz refers to as learner beliefs (1987).

Learner beliefs could be "consciously or unconsciously" held by learners (Borg, 2001, p. 186). Despite being unconscious to many learners, learner beliefs could to some extent determine the strategies learners adopt, which could then affect learners' language proficiency (Abraham & Vann, 1987).

Learner beliefs about language learning stem from two major conceptions learners hold (Ellis, 2008b), the conceptions of *what language is* and *how to learn*. Based on the understanding of these two conceptions, Benson and Lor (1999) divide learners into two categories, quantitative/analytic learners and qualitative/experiential learners. Quantitative/analytic learners believe that what is most important in second language learning is to accumulate vocabulary and understand grammar. Memorisation of word lists, texts and grammar rules is crucial for these learners. Quantitative/analytic learners, on the other hand, have a strong faith in EFL being learned via language use. To accomplish this, these learners guess words in reading and take every possible opportunity to use the language. Learner beliefs are past-experience-based and also possibly "cultural determined" (Ellis, 2008b, p. 10). These beliefs could be either changed or reinforced by an important factor in language learning, that is: learning context.

Learner beliefs also cover learners' self- efficacy which, according to Butler (2014), refers to learner beliefs about his/her own competency in mastering the target language. Beliefs in this connection are important in social-cognitive psychology because they are found to be an "important predictor in one's academic success" (p. 25).

43

2.9.2. Teacher beliefs

Teacher beliefs, also known as teachers' pedagogical philosophy, refer to teachers' beliefs on their roles in the process of teaching and assessing students (Carless, 2013). The literature in this connection has divided teacher beliefs into two main categories: (1) teacher as transmitter of knowledge; (2) teacher as mediator of learning (Özmen, 2012; Kane, Sandretto & Heath, 2002). Teachers taking on the transmissive views tend to see themselves as the source of knowledge and thus as having the requisite authority over their students. As a consequence, their students play a relatively passive role in receiving knowledge in the process of teaching. Teachers in the other category take the constructivist view of teaching. They see themselves more as mediators or even facilitators in the classroom and their students thus play a more active role in negotiating meaning and constructing knowledge (Chai, Teo & Lee, 2009).

Similar to learner beliefs, teacher beliefs are often culturally determined and therefore could be consciously or unconsciously held. Teacher beliefs stem largely from teachers' previous educational experiences (Özmen, 2012). In-service teacher training programmes could, to some extent, re-centre a teacher from being a transmitter of knowledge to a mediator of learning (Lee, 2010).

Teacher beliefs are one of the key factors contributing to learner beliefs, which means they could be passed on from generation to generation. Understanding teacher beliefs is thus crucial, particularly in second/foreign language teaching. These beliefs, when "mediated through societal values", could become the most significant factor in determining the success of implementing language educational innovations (Carless, 2013, p. 185).

2.9.3. Contextual factors at institutional-societal level

Contextual factors at institutional-societal level shape the socio-culturo-political environment in which one studies the target language. This environment has also been referred to as language learning context (Gu, 2003a). This context covers:

- in-classroom environment such as adopted teaching approaches, teachers' linguistic, sociocultural and strategic competence, classroom ethos; and

- out-of-class environment such as language education policies, high-stakes examinations as well as the richness of input and output opportunities (the availability of language learning materials and chances to use the target language).

These factors could "impinge on the dynamics of language teaching and learning in various ways, for example, by creating or withholding opportunities to use and experience the target language and by shaping learner perceptions, learning strategies and classroom behaviours" (Hu, 2003, p. 303).

2.10. Contextual factors in mainland China

2.10.1. Conceptual level

Many researchers have found that most learners in China are quantitative/analytic learners (e.g. Gu, 2003b; Jin & Cortazzi, 2006). They regard learning English as a simple process of accruing vocabulary and understanding grammar rules. Some learners even believe that "learning English is largely a matter of learning new words" (Jin & Cortazzi, 2006, p. 11). One learner metaphorically describes this learning belief: "Words are the bricks a building is made up of. Without bricks, where will the building be?" (Gan, Humphreys & Hamp-Lyons, 2004, p. 234). So,

to learners like this one, learning English is a matter of collecting bricks (i.e. new words). These beliefs, to a very large extent, shape the strategies these learners adopt in learning English.

Research evidence indicates that most Chinese learners employ mainly "intentional repetition and practice strategies" (Gu, 2003b, p. 192). These strategies include reciting selected texts and memorising word lists. Text memorisation is regarded as one of the most effective learning strategies by many Chinese learners (Yu, 2012). These learners believe that text memorisation helps to "cultivate language sense" and "build self-confidence" (p. 180). Some conscientious learners are reported to recite all the texts in secondary school English textbooks (Gu, 2003b). These reciting exercises are sometimes requested by teachers but more often are self-imposed. Many learners regard these exercises as "independent learning" or "self-study" (Jin & Cortazzi, 2006, p. 12). To accumulate vocabulary, most learners in China memorise long word lists by rote, which means much attention is paid to spelling and form-meaning association. To link the form of a word and its meaning, these learners use Chinese as the medium (Gan, Humphreys & Hamp-Lyons, 2004). This is to say, memorising a word means remembering a particular meaning of the word in Chinese to these learners. Word-list memorisation can be a ruthless exercise for many learners. It has been documented that some learners memorise as many as 200 words per day. This is why some learners jokingly refer to their vocabulary books as holy books (Zhang, 2012).

Ellis (2008b) contends that learner beliefs are "culturally determined" (p. 10) and past-experience-based. These two factors are evident in China. Learner beliefs in

the country are largely shaped by the traditional Confucian cultural heritage, which values "effort, perseverance and willpower" (Gu, 2003b, p. 98). One of the most salient reflections of this tradition in education is text memorisation. As a popular saying goes, "Master 300 Tang poems and you will become a poet yourself". Beliefs like this have triggered a common practice by many parents, who demand that their children recite classical Chinese poems at the age of three or four. When coming to school age, these children are often required to memorise long Chinese texts (Yu, 2012; Jin & Cortazzi, 2006). The habit of reciting Chinese texts, and to many learners the sense of achievement from their recitation of Chinese texts, has then motivated them to recite English texts (Yu, 2012).

Another factor that works to shape the learner beliefs in China is closely related to the language learning context in the country, including adopted teaching approaches, richness of input and output opportunities as well as high-stakes national English tests.

2.11. Institutional and societal level

2.11.1. Adopted teaching approaches

The traditional grammar-translation approach dominated English teaching in China for many years. Many learners trained with this approach believe that discrete-point knowledge of grammar and vocabulary is crucial in English learning, and this can cause some learners a severe lack of communication skills in English. Notwithstanding a vigorous promotion of the Communicative Language Teaching (CLT) by the education authorities in the early 1990s (Hu, 2005a) and their introduction of Task Based Language Teaching (TBLT) in 2001

(Zhang & Hu, 2010), English teaching in China remains largely teacher-centred and book-based (Jin & Cortazzi, 2006; Qi, 2007; Zhang & Hu, 2010), particularly in the economically under-developed regions (Hu, 2008). Studies have found that the discrepancy between the official and the *de facto* curricula is largely attributed to the examination system in the country (Deng & Carless, 2010; Hu, 2005b; Littlewood, 2007). Since many teachers believe that CLT/TBLT only produces students who "speak loud in class but scratch their heads in tests and exams" (Quyang, 2000, p. 410), they continue centring their teaching on grammar and vocabulary.

2.11.2. High-stakes national English tests

China "has the largest English-learning population in the world" (Cheng, 2008, p. 17). English tests are required in many areas such as entering prestigious schools, gaining tertiary education, studying overseas, obtaining some high-income jobs and seeking promotions within a workplace. These tests, in particular the high-stakes ones, have to some extent shaped the language learning context in China.

The two most influential high-stakes national English tests in China are the National Matriculation English Test (NMET) and the College English Test, Band 4 (CET-4). The NMET is the test on candidates' English proficiency in the university entrance test battery in China. Together with the tests on Chinese and mathematics, it is one of the three compulsory tests for all candidates and is thus crucial in university admission decisions. Approximately 10 million candidates sit the test each year (Cheng, 2008). The CET-4 is a nation-wide test administered by the National English Testing Committee. It aims to assess the English proficiency

of the tertiary graduates in the country. Most colleges and universities make passing CET-4 as one of the graduation requirements (Paltridge, 2007).

The NMET and CET-4 dictate the English learning and teaching in the country (Qi, 2007; You, 2010). The NMET, for example, has been playing the role of "a traffic wand" in China ever since it was introduced in 1985 (Cheng & Qi, 2006, p. 64). This is to say that the regimes of the NMET define what is being taught and the approaches to teaching that teachers adopt in the classroom (Gu, 2014). Until very recently, both tests focused mainly on candidates' ability in reading and their knowledge of English grammar and vocabulary (Hu, 2003; You, 2010). Little attention is paid to productive skills. For example, speaking is absent from the NMET and is only an optional sub-test on the CET-4, which most students opt not to take. Writing is tested in both tests. However, due to the task requirements and the adopted assessment criteria, the writing component in the tests cannot objectively measure the writing ability of candidates (Qi, 2007; You, 2010).

Although aimed at different levels of candidates, the NMET and the CET-4 share a number of similarities in their writing sections, in particular task requirements, and adopted assessment criteria:

    a. They both require that candidates compose a short text (around 100 words for NMET, not less than 100 words for CET-4) in the form of guided writing. The prompts for the tasks, which are often in Chinese, list all the main points that candidates should cover (see Appendix C). These task-descriptions are sometimes in such detail that the initial sentence of each paragraph (normally serving as the topic sentence) is provided. This practice virtually turns writing into translation (Wu, 2008; You, 2010).

b. Candidates' discrete-point knowledge of English grammar and vocabulary is the core assessment criterion (Hu, 2003). This requirement may not be explicitly written into the official assessment criteria, but in practice, language accuracy plays a crucial role in deciding on scores for the candidates (Qi, 2007). You (2004) points out that this practice makes the "correct form rather than well-developed thought" most valued in the writing tasks of these tests (p. 104).

c. The judgment on language-use has also extended to candidates' ability to use complex sentence structures and set-sayings, including proverbs and clichés in the writing (Paltridge, 2007; Qi, 2007). Because of this, many candidates produce unnecessarily long and overly complex sentences and use many inappropriate clichés in their writing, ignoring the requirements in tone and style for academic English (You, 2010).

d. Candidates' ability of writing to a word limit and maintaining a desirable appearance of their writing scripts, which includes clear and neat handwriting, is assessed (Paltridge, 2007; Qi, 2007). These two requirements, unwritten rules though they may be, play critical roles in deciding on candidates' scores. They are so important that some researchers (e.g. Wu, 2008) contend that they, together with the criteria on language accuracy and inclusion of main points prescribed in the task description, are the four key assessment criteria for assessing candidates' writing.

These task requirements and, to a larger extent, the assessment criteria, either explicitly or implicitly stated, have caused negative washback-effects to the teaching and learning of writing in China (Cheng & Qi, 2006; Qi, 2007; Cheng,

2008), which, in turn, have impeded learners' development in language proficiency and communication abilities in the country (Jin & Cotazzi, 2006).

The most salient washback effect is possibly the widespread teacher practice of "teaching to the test" (Cheng & Qi, 2006, p. 63). This practice is evident in what is taught in the test year - the last year of senior secondary schools and the secondary year of tertiary institutes. According to research (e.g. Qi, 2004; Deng & Carless, 2010; You, 2010), a large majority of schools and tertiary institutes take the test year as their revision year, which means teachers spend most, if not all, of the year reviewing what has been taught and preparing for the intimidating test. Because of this, teaching content required in the official syllabi for the year is virtually ignored.

The washback effects are possibly more revealing in what is going on in the English classes in the test year. To help students obtain high scores in the tests, teachers are encouraged by schools and tertiary institutions to guess the topic for the writing task far before the test date. After collecting the most likely topics for the year, some elite schools organise teachers to write exemplary essays on the topics and then publish them in commercial books. These books are often popular with students and teachers because of the guidance and test strategies they provide for the NMET or CET-4 of that year. Apart from sample essays, these books highlight what is commonly known as "beautiful structures" in China. These structures fall into two main categories:

- frequently used phrases or sentence structures in the academic writing of native speakers of English; and

- set-phrases, idioms, proverbs and clichés.

Little concern has been expressed in the books about a clash of tone and style if both types of phrases are used in the same genre of writing. One of the most popular CET-4 preparation book writers even regards these structures as a "panacea" and therefore advises his readers to use them. He asserts that "it is these beautiful and native-sounding sentences that will make your writing stand out" (You, 2010, p. 154). His books are well-received and methods recommended by him are taken by many students and teachers as the "default approach" to prepare for the CET-4 writing (p. 154).

Such strategy training deepens many students' misconception of what counts as good writing. They thereby memorise exemplary texts and "beautiful structures". If a candidate happens to have memorised an essay similar to the topic in the writing task, s/he can slightly modify the essay or simply copy the essay onto the answer sheet. This behaviour is even approved by some CET-4 testing centres (Paltridge, 2007). Those who unfortunately have not memorised the right essay are encouraged to integrate "relevant chunks from the samples" into their own writing (You, 2010, p. 154).

The strategy training is not only prevalent amongst people involved in the NMET and the CET but also the GSEEE (Graduate School Entrance English Examination), the English test for non-English major candidates who wish to undertake postgraduate studies in China. He (2010) reports "all-purpose" exemplary essays for the GSEEE. These "all-purpose" essays provide the beautiful sentences/structures that candidate may need for the test. More importantly, they feed candidates with chunks of texts that can be memorised and then regurgitated during the test. The two examples reported in the study explain

how regurgitation is possible in such a high-stakes English test. The writing task on the test intends to promote communicative language use. To achieve this purpose, candidates are required to discuss a social problem, such as the widening wealth gap between the developed and underdeveloped areas in China. Their discussions should normally be based on a picture provided as a part of the test prompt. To prepare candidates for the test, some tutorial schools/programmes prepared a number of "one-for-all" sample essays, i.e. essays that can be used for all given social problems to be discussed for the test. One of the examples cited in He's study is as follows:

> *It goes without saying that the symbolic meaning conveyed should be given deep consideration. What on earth can be derived from these interesting and instructive drawings? There is no doubt that what the painter virtually aims to convey is deep and profound. Primarily, we can learn that such case is far from rare and upsetting parallels can be found anywhere from our neighbourhood and around the world. What's more, there has been a growing concern nowadays over the worsening phenomenon. It is hard to imagine what our society would be like years after such pervasive trend goes unchallenged* (p. 154*).*

It is evident from the above sample that the discussed social problem is not specified and this deliberate vagueness of the text makes it possible for candidates to apply it to discussions of any social problem that may appear in the writing task. According to the study, 20 percent of the essays written for the GSEEE in 2007 in Zhejiang Province contained chunks of texts from exemplary samples similar to the above. This has caused severe problems of fairness in the marking process.

Some raters gave a score as high as 19 (out of 20) for essays with such chunks of texts because they were impressed with the "richness of vocabulary", "well-organised paragraphs" and "use of discourse markers" (p. 155). However, when some raters read more essays similar to the above (with minor changes in some candidates' texts) and gained a better understanding of where the texts were from, they decided to punish this plagiaristic behaviour by giving a score as low as 2 (out of 20) for such essays (p. 155). These extremely high and low scores for similar essays could have a serious impact on the reliability and status of the GSEEE.

The tutorial schools/programmes for high-stakes English tests such as the NMET, the CET and the GSEEE have created highly profitable business opportunities in China. According to He (2010), the annual market value of such schools/programmes for the GSEEE alone (including test-preparation materials such as books) is approximately three billion Chinese Yuan. Meanwhile, many students have also successfully achieved their goal, i.e. to obtain high scores by imitating sample essays at the tests. This is possibly one of the reasons why many Chinese learners are very positive about text memorisation (Yu, 2012). This seemingly win-win situation appears to satisfy the needs of many examination candidates as well as the related business sector at present, but the negative impact of such practices on society and, in particular, the education system of the country, could be detrimental in the long run.

The washback effects of this situation have also been manifested in teachers' training of students' writing skills in general. The training has mainly focused on strategies designed to obtain high scores in the writing tasks of high-stakes

examinations but not learner's written communication skills. The strategies include use of writing templates and intensive exercises in grammar and vocabulary. Writing templates are highly recommended by teachers and authors of examination-preparation books for the writing tests in various levels of high-stakes English tests in the country, such as the NMET, the CET and the GSEEE. Below is a template recommended in an NMET preparation book for writing expository essays (Wang, 2013, p. 20). The italicised parts (in brackets) were in Chinese in the original template and were translated into English by the researcher for the purpose of the current study (more examples and detailed discussions of writing templates such as below can be found in Chapter 7 of this dissertation):

There is a widespread concern over the issue that _____ (*essay topic*).   But it is well known that the opinion concerning this hot topic varies from person to person. A majority of people think that _____ (*View 1*). In their views there are two factors contributing to this attitude as follows. In the first place, _____ (*Reason 1*). Furthermore, in the second place, _____ (*Reason 2*). So it goes without saying that _____ (*View 1*).

People, however, differ in their opinions on this matter. Some people hold the idea that _____ (*View 2*). In their points of view, on the one hand, _____ (*Reason 1*). On the other hand, _____ (*Reason 2*). Therefore, there is no doubt that _____ (*View 2*).

As far as I am concerned, I firmly support the view that _____ (*your own view*). It is not only because _____ (*Reason 1*), but also because_____ (*Reason 2*).

Another main strategy is teachers' intensive training relating to students' grammar and vocabulary knowledge. According to Qi (2007), instead of focusing on the communicative functions, which is the main objective of writing, teachers' attention is largely paid to linguistic accuracy in student writing. That is to say, discrete items in grammar and vocabulary usage are always highlighted in the teaching, while other writing skills, for example, consistency in tone and style and appropriateness in communication context, are virtually ignored (Qi, 2004). This misplaced focus in teaching, together with many teachers' own limited knowledge and experience in English writing (Jin & Cortazzi, 2006), might cause students to misunderstand what good writing should be like. As a consequence, many candidates concentrate on skills to cope with the writing requirements of the tests, but their writing skills are not much improved thereby (Wu, 2008; You, 2010).

Since productive skills are not demanded overmuch in these tests, teaching and learning in China are largely focused on the discrete-point knowledge of English grammar and vocabulary. One teacher vividly describes how she prepares her students for the NMET (Qi, 2004).*

The most important training is on vocabulary. Every year I print the NMET vocabulary list for my students and ask them to memorise every single word on the list. I facilitate their memorisation by conducting frequent quizzes in class. At each quiz, I give students 100 English words and ask them to write the corresponding Chinese meaning of these words and then vice versa, i.e. give students 100 Chinese words and ask them to write down the corresponding English meaning. The quizzes take place every week.

*. Translation of the researcher. Original text in Chinese.

After undergoing such training, many students might have gained strategies for getting high scores at the tests. However, these strategies "backfire" when the learners encounter the demands of academic writing in their overseas studies (Edward, Ran, & Li, 2007). Another impact of the training is that it reinforces many learners' beliefs that learning English is a process of accumulating vocabulary and understanding grammar rules. The end result is that even learners who are aware of the importance of contextual vocabulary use, i.e. automaticity of lexical production in context, also focus mainly on the meaning retention of words (Wei, 2007).

2.11.2.1. Richness of input and output opportunities

Another crucial factor attributed to the language learning context in China is a shortage of input and output opportunities (the availability of language learning materials and chances to use the target language). English is a foreign language in China. There are thus very limited sources (input opportunities) for English learners, for example, English-medium television, radio and newspapers. Most learners rely on textbooks, dictionaries, vocabulary lists, grammar books and even examination-preparation books for their English study (Parris-Kidd & Barnett, 2011). Their output opportunities are hardly better. Most learners seldom have any chances to be engaged in a meaningful English conversation (Liu & Jackson, 2011).The reason for this is that English is little used in the community. Butler (2014), for example, interviewed 527 students from different socio-economic background in an eastern costal city and found that "hardly any parents" of these participants need English in their professional life (p. 22). With the learning context in China as such, learning English for most learners there becomes largely

instrumentally motivated (Cheng, 2008), i.e. to pass various English examinations and, in particular, high-stakes tests (Cheng, 2008; Fox & Curtis, 2010).

Another impact of this lack of input and output opportunities is on the learners' understanding of the rhetorical conventions in English writing. Rhetorical conventions in Chinese and English differ markedly (Chu, Swaffar & Charney, 2002). This is manifested mainly in three aspects. The first is related to idea sequencing. The Chinese discourse structure, particularly in expository writing, prefers inductive approaches (providing elaboration, analysis and examples before stating the thesis statement of a text) whilst the preference in English is deductive (stating the main thesis of a text before setting out supporting ideas and evidence). Kachru (1998) comments that the inductive approach often leads to a "delayed topic statement" (p. 55). Another difference lies in the sequencing of main and subordinating information. Discourse cues, such as signposts for sentential and intersentential links, are crucial in English writing. Such linear linkages, however, are not as important in Chinese writing (Normant, 1986). The third disparity lies in the level of directness in expressing authorial intent. Native English-speaking authors, like those in most other Western cultures, prefer directly presenting their arguments while traditional Chinese rhetorical heritage, on the other hand, privileges subtlety. Because of this, analogies and metaphors are often used in Chinese prose, including expository writing (Jensen, 1998).

The Chinese rhetorical tradition, as a part of the cultural heritage, has become an artefact and has helped to create many masterpieces of world-class literature. However, the differences in rhetorical practices of the two languages have also pointed to the need for Chinese learners of English to be fully aware of rhetorical

conventions in English writing. A lack of such awareness could create obstructions for the learners' ability in reading (Chu, Swaffar & Charney, 2002) and more importantly writing, particularly in academic writing in English (Edward, Ran, & Li, 2007). Unfortunately, the shortage of input and output opportunities in mainland China has deprived many mainland learners of their chance to develop awareness of this, which could exert a negative impact on their learning effectiveness.

2.11.3. Impact on learners' writing proficiency in mainland China

The above analysis suggests that the motivation to learn English is misplaced by most learners in China. This misplaced motivation is possibly manifested most in the teaching and learning of English writing in the country. Since writing is not heavily weighted in the NMET, writing skills are not emphasised in the curricula for secondary schools. Writing skills are normally not taught in secondary schools until Senior III, the year students sit the NMET. Before this year, writing practice is fundamentally at the sentence level, i.e. combining simple sentences to make complex sentences and translating sentences from Chinese into English (Qi, 2007). This practice has extended into the curricula for tertiary students. According to You (2010), until very recently, writing instructions in the university English textbooks for freshmen were still largely centred on sentence-combining. Paragraph writing was not introduced until their sophomore year. Even in the recently published textbooks, writing instructions still start with paragraph writing. When the formal training on essay writing starts, the centrality of the training, both in secondary schools and tertiary institutes, is on test-strategies rather than writing skills (Qi, 2007; Paltridge, 2007; You, 2010).

You (2004) illustrates the commonly adopted training practice in China. The teacher provides students with the outline and keywords for an assigned topic and then asks students to write an essay. After the writing is completed, only very few student essays are selected for the teacher to give feedback on. This is mainly attributed to the large student numbers (normally more than 50 students per group) in the class and teachers' workload (You, 2010; Cheng, 2008). The teacher's comments on student essays are fundamentally centred on lexical and syntactical mistakes. For the large majority of students whose essays are not given feedback on, the teacher provides them with an exemplary essay and requires that all students in the group memorise it. This is why several studies (e.g. Yu, 2012; You, 2010; Singh & Fu, 2008; Zhao, 2009; Paltridge, 2007) find that most students in China learn to write by memorising sample essays. Because of the availability of exemplary essays, some teachers even never give feedback on students' compositions (Zhao, 2009). This is partially because these sample essays contain language components, such as "beautiful structures" and "shining phrases", complex lexis and sentences, and sentence connectors, that many teachers expect their students to use in examinations, and are therefore very popular with students. After such training, most students know little about writing skills and have to rely heavily on their resources in L1 to complete a writing task. One student in You's (2004) study illustrates how most students in the mainland write (p. 102):

I write according to my instinct. I think of Chinese sentences first, and then translate them one by one into English. I feel the biggest problem I face is my small vocabulary. In terms of grammar, we have learnt all of it in high school.

Apart from the above translation strategy, many students also ensure that they insert chunks of phrases and sentences they have memorised from exemplary essays into their compositions, a strategy strongly recommended by many teachers and authors of examination preparation books (You, 2010). As a consequence, many learners from the mainland are found to write "flowery prose" (Singh & Fu, 2008, p. 121). Flowery prose contains the aforementioned "beautiful structures", "shining phrases" as well as hyperboles (over-statements), which are a manifestation of the influence from their L1 writing (Singh & Fu, 2008). The hortatory function of these overstatements may generate a polemical tone in these learners' writing. Another feature related to the polemical tone in the writing by these learners is a dialogic tone, as manifested in the use of personal pronouns, interrogatives and imperatives in their expository essays (Mayor, 2006). This feature indicates their low awareness of genre differences, a result of a lack of training in English medium writing. The third characteristic, as previously mentioned, is their adoption of different rhetorical moves in their expository writing. Instead of presenting their arguments deductively, which is common in the Western argumentative writing approaches, learners from Chinese mainland often use inductive approaches (beginning with examples or/and explanations which lead to an argument) to present their views (Singh & Fu, 2008). This is another manifestation of the influence of their L1 writing.

The above three features in the writing by learners from the mainland are largely a product of the language learning context in China. Some factors in the learning context cause their misplaced motivation to writing, i.e. learning to write for passing exams but not for communicative purposes. This misplaced motivation

has shaped their views of what counts as good writing. As a consequence, the development of their writing skills is impeded. This is evidenced by the study of Fox & Curtis (2010) who find that learners from the mainland are considerably weak in writing. Although some of them, through effort and perseverance, can achieve "stunning levels of success" in some international English proficiency tests such as IELTS and TOEFL (Gu, 2005, p. 84), these learners find it challenging to communicate with people in an English-speaking country (Parris-Kidd & Barnett, 2011), and even more of a challenge to meet the demands of academic writing in their overseas studies (Edward, Ran, & Li, 2007; Plicher, Cortazzi & Jin, 2011).

Having presented the characteristics of the writing by learners from the Chinese mainland, it is worth noting that not all of these features are problematic or need to be changed. The deductive rhetorical moves, for example, are different from the Western norms, but may need to be learnt by Western scholars. Even the flowery rhetorical style of writing might be effective in certain genres of writing in English. After all, understanding and even adopting different cultures, including writing styles, should be part of globalisation. Nevertheless, before Western scholars and university teachers have gained a deeper insight into the traditional Chinese writing conventions, it is worth raising the awareness of these features. This should benefit both Western scholars/university teachers as well as Chinese learners themselves. After all, English language is "a core component of the networks and systems of globalisation" (Evans, 2013, p. 318), at least for the time being.

2.12. Contextual factors in Hong Kong

2.12.1. Conceptual level

Learners in mainland China and Hong Kong share a similar cultural background (Gan, 2009, 2013). Notwithstanding Hong Kong's 150 years of colonial influence, both places are much influenced by Confucian ideologies (Jin & Crotazzi, 2006). Contextual factors at the conceptual level (learner/teacher beliefs) are therefore similar in many ways. For example, due to the Chinese competitive examination tradition that can be traced back to the Han Dynasty over 2,000 years ago, learners/teachers in both places regard examinations as a way to provide "a level playing-field and a means for social mobility" (Carless, 2013, p. 175). This means the discriminatory power of examinations is crucial. The corollary is that until very recently the centrality of examinations (in particular, high-stakes tests in both places) is premised on a supposed reliability which is often "at the expense of validity" (p. 175). This practice also limits the range of types of assessment which might be used.

This is, however, not to say that learners/teachers in these two language contexts share the same level of traditional beliefs. Learners in Hong Kong, for example, have been reported to rely much less on rote-learning (Gan, 2009) when compared with their mainland counterparts; and teachers in Hong Kong have also been documented to be more ready to adopt innovative teaching approaches than their counterparts from mainland China (Carless, 2012; Luk, 2012; Lee, 2011). What is even more important is the divergent "institutional teaching contexts and social environments" in these two places. This divergence has resulted in a

marked difference in learners' engagement in their English learning (Gan, 2009, p. 49).

2.12.2. Institutional-societal level

The 150 years' colonial legacy is clearly evident at the institutional-societal level. This legacy, together with Hong Kong's status as an international centre in finance, business and tourism, has contributed to a favourable English-learning environment in the territory. This is reflected at both institutional and societal levels.

2.12.2.1. Institutional context

One of the most significant British legacies to Hong Kong is the medium of instruction (MOI) in Hong Kong classrooms. This is the area where "mainland China cannot hold a candle to Hong Kong" (Hu, 2004, p. 17). As of now, the MOI of all the universities in the territory (except for Chinese University of Hong Kong which implements a bilingual policy) is English (Lee, 2010). This indicates that a command of English is a prerequisite for the tertiary education system here. Insofar as the secondary school education is concerned, over 90% of Hong Kong's secondary schools adopted English as their official medium of instruction (Evans, 2009) before the handover in 1997. Notwithstanding the hitherto common practices of "code mixing" (teaching conducted mainly in Cantonese but mixed with English technical terminology) and "code switching" (teaching delivered through teachers' constant switching between Cantonese and English) in many schools, their teaching materials and assessments were in English (Pennington, 1997; Falvey, 1998; Evans, 2011a, 2013; Poon, 2013). The availability of these

materials increased the richness of students' language input and thus allowed them to be more engaged in incidental language learning, particularly in acquiring vocabulary.

The code mixing/switching practices resulted in the SAR government's implementation of a controversial language policy for secondary schools in 1998 which permitted only about one third (35%) of secondary schools (the 'elite' schools) to retain English as their instructional language (Luk, 2010; Evans, 2011a). This policy incurred strong opposition of many stakeholders, particularly students and parents (Poon, 2013), and is regarded as one of the most unpopular policies to be introduced by the new SAR administration (Tsui, 2007). This opposition evidenced the status and continuing presence of English within the community.

English was the "high" language in Hong Kong, related to "success, stylishness and academic achievement" in the colonial era (Pennington, 1998, p. 13). This prestigious status, and other factors, led to a situation in which Chinese-medium school graduates suffered from a significant disadvantage, when compared with their counterparts from English-medium schools, in attaining admission to university (Tsang, 2008), and also in advancing their university studies (Lin & Morrison, 2010); and this, in turn, compelled the government to "fine-tune" the controversial language policy. The new policy issued in 2009 allows schools which had hitherto been obliged to operate a Chinese medium of instruction more flexibility in their choice of MOI, and this has resulted in a situation in which more students and more courses in these schools are now taught in English (Evans, 2011a).

Apart from the benefit in MOI, Hong Kong also has an advantage over the Chinese mainland in human resources in English-language teaching. Hong Kong has more comprehensive and effective pre-service and in-service training programmes (Gan, 2009; Lee, 2013) and therefore more competent language teachers (Gan, 2009, 2013) who are more ready to adopt innovative teaching approaches (Tang & Nesi, 2003; Luk, 2012; Carless, 2013). This competence not only includes their linguistic ability but also their socio-cultural knowledge and strategic skills. Skills in this regard allow them to focus more on the meaning rather than form of the language in teaching. The focus of their teaching could thus be on communication skills rather than knowledge transmission. This possibility, together with the government's effort in implementing innovative teaching approaches such as CLT and TBLT, has contributed to Hong Kong students' development of both receptive skills (reading and listening) and productive skills (speaking and writing).

Speaking and writing are important skills to be developed in Hong Kong schools, according to the official English curricula. These two skills are taught, practised and tested through students' schooling, starting from kindergarten (Lee, 2010; Qian, 2008). In many schools, oral activities such as oral presentations, group/peer discussions and debates are parts of normal lessons (Luk, 2012), and compositions are written on a regular basis (Lee, 2013). In the implementation of the recent three-year senior secondary curriculum, students' writing and speaking abilities are further emphasised. In speaking, for example, students are required to develop their ability to "initiate, respond, and negotiate meaning in situated performances" (Luk, 2010, p. 25). This reform has been positively received

because the new scheme makes learning "more enjoyable" and permits students "a more active role in engaging with stimulating English materials" (Carless, 2013, p. 181).

This is, however, not to claim that all Hong Kong English teachers warmly embrace these new educational initiatives. On the contrary, many educational reforms, in particular, the "inquiry-oriented or student-centred approaches" have not been well-received by many Hong Kong teachers because elements in the approaches are incongruent with local norms and values (Carless & Harfitt, 2013, p. 174). Some teachers, therefore, re-interpret the initiatives in line with their own experiences. For example, Adamson and Tong (2008) report that some schools implemented versions of TBLT which were less strong than that stipulated in the official guidelines. This s the existence of a gap between the intended and enacted curricula.

Nevertheless, compared with the mainland, this detachment is possibly not so serious a matter because, in China, instead of just a gap, there is a major disjunction between the official curriculum and the instructional practices taking place. Zhang (2007), for example, observes that in the supposed TBLT classrooms she visited no features of what one would normally consider as TBLT could be found. One of the major causes of this problem, according to related studies (see Zhang & Hu, 2010; Zhang, 2007; Qi, 2007), is the prevalence of high-stakes examinations in the country. The official or *de jure* curriculum aims at developing learners' communicative abilities, while the enacted or *de facto* curriculum focuses on strategies which enable students to obtain high scores in high-stakes tests (Zhang & Hu, 2010).

2.12.2.2. Societal level

A number of factors affect English learning at the societal level of Hong Kong. The two most influential ones are the high-stakes English tests and the input and output opportunities in the community.

Hong Kong's examination system, in particular that related to high-stakes tests, has been much influenced by both the traditional Confucian ideology and the British colonial assessment culture (Qian, 2008). The Confucian tradition regards examinations as "gatekeepers for making selection decisions" (p. 87), and as crucial determinants of social mobility (Poon, 2013). In this connection the examination system in Hong Kong and that in mainland China are similar. On the other hand, due to the 150 years colonial legacy, Hong Kong has "traditionally looked to major Anglophone countries" for education reforms (Carless & Harfitt, 2013, p. 174). Although many of these innovations have met resistance from local teachers, the centrality of the high-stakes examinations has gradually "shifted from grammatical accuracy to communicative competence" (Luk, 2010, p. 25).

The percentage of speaking and writing components in the university entrance English examinations, in both the Hong Kong Advanced Level Examinations (HKALE) in the previous four-year secondary curriculum, and the Hong Kong Diploma of Secondary Education (HKDSE) in the current three-year secondary curriculum, is substantial. In the HKALE, both speaking and writing carry 18% of the weightage. In the HKDSE, speaking weights 40% and writing 30% (Qian, 2008). Compared with the NMET in China, in which speaking is not tested while the weighting for the writing component is only 16.5% (Butler, 2013; Qi, 2007),

the productive skills in Hong Kong university entrance examinations are attached much more attention to.

The two productive skills are not only tested in Hong Kong high-stakes examinations but also tested with a communicative purpose. The writing paper on the UE (*Use of English*) of the HKALE, for example, requires that candidates write extended English discourse (minimum 500 words) based on a very brief prompt (e.g. *discuss the popularity of comic books*, see more examples in Crossley and McNamara, 2012). This test not only examines candidates' language accuracy but also the organisation and coherence of the argument presented in the writing (HKEAA, 2013). In the recent three-year senior secondary curriculum, emphasis on writing has also been placed on "genre features, style and register" (Lee, 2012, p. 5). This is very different from the writing paper in the NMET, which is a short guided-writing test (around 100 words), and in which no genre requirement is explicitly stipulated. Prompts for this test are often highly detailed. The detailed prompts, together with the explicit and implicit marking criteria (see Section 2.9.2.2 for detailed discussions), have reduced the reliability as well as validity of the test. As a result, Qi (2007) believes that the writing test cannot discriminate between different candidates' writing proficiency.

The speaking test in the UE requires more communicative skills and is thus more demanding. Candidates have to undergo two oral tasks and are examined in groups of four. In Part I, each candidate is required to deliver a presentation on a given text and, in Part II, the four candidates are asked to participate in a 10-minute discussion on a topic connected to the texts they presented. This test

measures candidates' speaking fluency as well as their communicative strategies such as interaction skills.

With the replacement of the HKALE by the HKDSE in 2012, there was a high-profile introduction of SBA (School Based Assessment). This new system demands even higher communicative strategies. In this scheme, students are required to participate in oral tasks within the school which are graded by their own teachers (Carless, 2012). SBA weights 15% in the HKDSE and, to complete it, candidates need to deliver an oral presentation on or participate in a group discussion about a topic related to the texts they have read or viewed. One of the text types is movies and many candidates have been reported to use this medium to prepare for SBA (Luk, 2010). This scheme intends to cultivate "a stronger alignment between curriculum, pedagogy and assessment" (Carless, 2012, p. 349). Although a number of issues derived from SBA, such as increased teacher workload which is directly related to teacher motivation, a lack of teaching training, which could negatively impact upon teachers' willingness and ability to execute the programme, and the traditional Chinese culture, which may dampen some students' enthusiasm for participation in the programme, still remain unsettled (Qian, 2014; Carless, 2013), SBA has been by and large accepted by the school community (HKEAA, 2010) and is expected to profoundly affect the "behaviours and responses of students and teachers" (Carless, 2013, p. 183).

A greater difference between the language context in Hong Kong and China lies perhaps in the input and output opportunities (availability of language learning materials and chances to use the target language). This should be largely attributed to Britain's 150 year colonial legacy. In the colonial era, English was

the official language of the government. A high proficiency in English was a prerequisite to more elite education, a passport for prestigious employment and also a vehicle for the grassroots to reach the upper level of the social ladder (Poon, 2013). After the handover, English remains one of the official languages of the SAR government. In addition, the new administration developed a language policy of transforming Hong Kong into a bi-literate and trilingual community at ease with Cantonese, English and Putonghua (Lin, 2009, Evans, 3013). To further improve the English learning context, the government also launched a workplace-English campaign (Gan, 2009) and initiated the NET (Native-speaking English Teacher) scheme in schools. Under the NET scheme all English-medium government-subsidised schools are allowed to hire at least one NET, and all Chinese-medium government-subsidised schools up to two (Luk, 2012). These measures taken by the new administration to maintain the use of English and, more importantly, the longstanding belief of many Hong Kong Chinese in providing English medium education for their offspring (Poon, 2013), have made it necessary for part of the mass media in Hong Kong (for example, radio, television programmes, newspapers and the newly developed social media such as Facebook, YouTube and WhatsApp) to remain in English. This part of the mass media enhances the input opportunities for English learners, in particular young learners, in the community. For example, some of them watch English films and television programmes, and some listen to English songs (Evans, 2011c; Evans, 2013). Activities such as these greatly increase these learners' exposure to English and thus provide them with opportunities to gain vocabulary in context.

The output opportunities for many Hong Kong people, the chances for them to use English, are even more favourable. This is possibly the determinant factor in maintaining Hong Kong's English language context. Evans (2010), for example, finds that English still plays a critical role in both the public and private sectors. Written English is the "default medium of professional communication" and spoken English is very important in "presentations, seminars and conferences" (Evans, 2011b, p. 306). The study also identifies that "frequency of English use increases with rank" (p. 306), which indicates one needs more English in professional life as one advances on the career ladder. The practical need to use English in professional life intrinsically motivates many Hong Kong people to learn the language communicatively. This language context is very different from that in mainland China where English has "only limited use in the community" (Butler, 2013, p. 25). In Butler's study, "hardly any parents" of the 527 participants from an eastern costal city, regardless of their socio-economic disparity, use English in their professional life (p. 22). This difference indicates that Hong Kong's favourable language context is not only attributed to the fact that a strong command of English remains a prerequisite for tertiary education in the territory but also that Hong Kong, as an international city, is increasingly engaged in knowledge-intensive service industries such as finance and tourism, which are progressively integrated "into global economic networks" (Evans, 2010, p. 361). To maintain Hong Kong's position as an international city, the territory will have to retain and enhance the status of the English language in its business, legal and education system, which is "a core component of the networks and systems of globalisation" (Evans, 2013, p. 318).

The above discussion indicates that, compared with learners from the mainland, Hong Kong learners have stronger motivation to write for communicative purposes, for example, to obtain a prestigious job and/or to climb on a career ladder. Compared with learners who are not interested in and/or cannot perceive a need for writing in English, such as those from mainland China, these intrinsically motivated Hong Kong learners are more likely to engage themselves in meaningful writing practice.

This is, however, not to claim that the language learning context in Hong Kong has always been conducive. For example, Johnson and Ngor (1996) reported a "lexical processing" approach (p. 123), a survival strategy for most students in Hong Kong's EMI schools aimed at managing English texts that were often ''too advanced for their level of proficiency'' (p. 125). Because of this demand on dealing with a large number of such texts, a great deal of rote learning took place (Watkins, 1996). This rote-learning practice can still be found in many schools in the territory today (Lo & Lo, 2014). Nevertheless, owing to the 150 years' British colonial legacy and an increasing demand on English writing proficiency for prestigious jobs in the territory, many learners are motivated to write in English. Their engagement in writing, together with the instructional practices in schools may help to raise their awareness of genre differences in writing and provide them with output opportunities. These opportunities allow them to "notice and internalise new linguistic knowledge", and thus "promote automatisation" (Kormos, 2012, p. 392).

2.13. Lexical depth and writing proficiency

The discussion about the different learning contexts in Hong Kong and mainland China in this chapter suggests that a favourable learning context facilitates the automaticity of learners' linguistic knowledge. Insofar as lexical knowledge is concerned, automaticity is the final development stage in the 4-dimensional framework on vocabulary knowledge (lexical breadth, lexical depth, lexical organisation and lexical automaticity) proposed in Qian (2002). The level of automaticity is a manifestation of speed and accuracy with which words can be called upon in one's language production. The accuracy aspects here include the appropriate use of all components of learners' word knowledge, ranging from spelling, pronunciation and morphological forms, to syntactical and semantic features, constraints in use (e. g. register, frequency) and collocations (Nation, 2001). However, the WAT only measures three elements in the above list, i.e. synonymy, polysemy and collocation (Qian, 2002; Qian & Schedl, 2004). Other components, such as grammatical functions, register and frequency, are not covered in the test. This could mean that even if a learner can achieve a reasonably high score on the WAT, his/her ability of readily and automatically using lexis in his mental lexicon could still be limited. This possibility leads to another proposition of the current study: language learning contexts affect the predicting power of lexical depth in writing.

This proposition can also find its theoretical support from Ellis's (2008b) implicit and explicit framework. A learner, for example, who adopts explicit learning and learns most of his/her words in isolation, such as from word lists, could perform reasonably well on vocabulary tests, but may not be able to use the words s/he has

learnt appropriately to produce a coherent piece of writing. On the other hand, a learner who adopts implicit learning and acquires most of his/her words in natural meaningful contexts, such as in reading or oral interaction, could receive similar scores on the tests but may be able to employ the words s/he has acquired naturally to produce a piece of high-quality writing. This is so because natural language contexts can "disambiguate and delimit the meaning of a word" and can "expose learners to a word's range of meanings" (Zimmerman, 1997, p. 133). Given this situation, the current study proposes the third hypothesis:

Hypothesis 3: Language learning contexts affect the association between EFL learners' depth of vocabulary knowledge and their writing proficiency.

2.14. Summary

This chapter has briefly reviewed the literature appertaining to EFL learners' vocabulary learning, measures to test EFL learners' lexical knowledge and contextual factors in EFL learning. The literature review has led to the following three research questions for the current study:

1. Does EFL learners' depth of word knowledge positively correlate with their writing proficiency and, if so, to what extent do they relate to each other?

2. Compared to vocabulary size, is depth of lexical knowledge a better predictor for the writing proficiency of EFL learners?

3. Do different language learning contexts affect the association between EFL learners' depth of vocabulary knowledge and their writing proficiency and, if so, how are learners from unconducive learning contexts affected?

CHAPTER THREE

METHODOLOGY

3.1. Overview

This chapter explains the overall design of the current research project. Detailed information regarding the instruments, subjects, research procedures and data analysis for the present project is provided. The procedure section details the process of the vocabulary test delivery, questionnaire survey, focus-group interviews and case studies. The methods for building and analysing the two learner corpora are also discussed.

3.2. Instruments

The main instruments for the study involved two vocabulary tests and a writing test. One of the vocabulary tests was the revised Vocabulary Levels Tests (Test B, Nation, 2001, p. 416) and the other Read's (1998) Word Associates Test. See Appendices A and B for these two tests and see Sections 2.4 - 2.6 for detailed discussions about the two tests. The main reason for deploying the VLT instead of the VST, which allows for measurement of words at the first 14,000 levels, was to ensure the test results from the current study would be more comparable with those from Qian 1999 (2002) and Stæhr (2009).

It should be noted that the current study employed the VLT only at the 3,000 and 5,000 frequency levels. The VLT on the academic word list covers vocabulary from the 2nd to 5th frequency levels and is thus not targeted at words at a particular level. The exclusion of the 2,000 and 10,000 frequency levels is attributed to an earlier study conducted by the current project investigator in

which the VLT and the controlled VLT (see Section 2.4 for details of the test) were used to predict the writing proficiency of the same level of students. The results indicate the VLT at either the 2,000 or the 10,000 frequency level has little association with the students' writing ability. This is possibly because the 2,000 level is too basic for EFL students at the university level while the 10,000 level is far too difficult for most of them. Several other related studies also lend support to this choice of tests. Schmitt *et al* (2003), for example, deployed the VTL only at the 3,000 and 5,000 frequency levels to investigate formulaic language used by EFL students in the University of Nottingham, because "the 2,000 level was deemed too basic for the relatively advanced EAP (English for Academic Purposes) students, while the 10,000 level was still considered quite difficult" (p. 59). Laufer and Ravenhorst-Kalovski (2010) also excluded the 10,000 level when measuring the lexical threshold for reading comprehension of students in a university in Israel because this level of the test was "considered far too difficult in view of the background they had in English" (p. 21).

To examine the factors that may possibly affect the association level of depth of vocabulary knowledge and writing proficiency, Nation's Vocabulary Frequency Profiler (available on Tom Cobb's website) was also used to identify the lexical variation, lexical sophistication and lexical frequency profile in the essays of the two groups of learners (see Sections 2.7.1-2.7.3 for discussions about these measures of lexical knowledge). Two small learner corpora built from the essays by these two groups of participants were also used to investigate differences in lexical use in the essays by Hong Kong students and their mainland counterparts.

In order to identify the causes of any differences in these two groups, a questionnaire survey containing a total of 50 multiple-choice questions was administered. Focus-group interviews with selected participants were also conducted to obtain more detailed profiles of the learners regarding their language learning context, such as their beliefs, practices and experiences in English learning. To obtain deeper insight into the differences, five case studies were also undertaken.

IBM SPSS Statistics, computer software commonly used for quantitative statistical analysis for social sciences, was deployed to perform statistical analyses. The version used for the current study is IBM SPSS Statistics 21.0, the latest version available when the analyses took place.

Wordsmith 6.0 (Scott, 2012) was used to compare the differences in frequency and use of words in the two built learner corpora. Wordsmith is a suite of computer programs that analyses the behaviour of a word or a word-cluster in texts. It does so by generating occurrences of a word within a corpus in alphabetical and frequency orders. Its concordancing tool, Concord, allows displays of collocation and colligation of a word in the corpus. Data derived from Wordsmith allows researchers, teachers and even learners to compare the frequency and use (particularly in collocation, semantic prosody and colligation) of a word in a learner corpus with those in a native speaker corpus. This research tool has been much used in studies pertaining to lexical features in L2 and EFL learners' language production (e.g. Bolton, Nelson & Hung, 2002; Tang & Nesi, 2003).

3.3. Subjects

Two groups of first-year undergraduate students in the Hong Kong Polytechnic University were invited to participate in the study. One group consisted of local Hong Kong students and the other came from mainland China. The purpose of involving mainland students is to identify the impact of different contextual factors in Hong Kong and mainland China on test results. Such disparity has been documented in Tang and Nesi (2003), who compared differences between classroom vocabulary teaching and learning in Guangzhou and Hong Kong. The difference is also evident in the writing of a considerable number of mainland students who are currently studying at Hong Kong universities. These students employ many long and complex words in their writing; however, their writing often does not make complete - or sometimes very little - sense (see Appendix D for an example of such writing). This problem seems less apparent in the writing of local Hong Kong students.

A total of 190 students in the university participated in the project. After the data collection, profiles of these 190 participants were carefully checked. Data from those who failed to meet the requirements of the current project (40 in total) were excluded in the data analysis. These participants included:

- students whose information in the consent form was incomplete, for example, no signature or no university entrance examination result;

- students who were neither from Hong Kong nor the mainland;

- students who sat the university entrance examination in different years, i.e. not in the year the data were collected. This is to ensure that all participants

were first-year students in the university so as to minimise many of the variables in data analysis; and

- Students who did not achieve the satisfactory level, i.e. scoring 67% or higher, in the VLT at the 3,000 word families. This was to be in line with the data collection requirement in Qian (1999).

After the above-mentioned procedure was completed, the data from a total of 150 subjects (67 from Hong Kong and 83 from mainland China) was formally recorded for the main study. These participants were all fresh graduates from secondary schools (72 males and 78 females) with an age-range of 16 to 22. Detailed profiles of these respondents can be seen in Table 3-1 below.

Table 3-1. Profiles of participants in the main study

	Hong Kong	Mainland China
Number of participants	67	83
Gender ratio	37 (M)	35 (M)
	30 (F)	48 (F)
Age range	18-22	16-22
Education level	Tertiary (1st year)	Tertiary (1st year)

3.4. Procedure

3.4.1. Test administration

As stated above, the purpose of involving mainland students is to identify the impact of different language learning contexts on the test results. Due to this reason, all the tests were administered in the first three weeks of the new academic year in order to minimise the influence of a change of language learning context on the mainland participants. To further control external variables, the

current study added an item in the consent form asking the mainland students to declare whether or not they have received additional English support, such as private tutoring in their first three weeks in Hong Kong (see Appendix E). The information in the consent forms indicated no mainland students received such help.

A number of test sessions were arranged so that participants could attend the tests at their own convenience. During each session, all participants were asked to complete three vocabulary tests: VLT (Vocabulary Levels Test) at the 3,000 words level, VLT at the 5,000 words level and the WAT (Word Associates Test). They were then asked to write an essay. All these procedures were administered under controlled conditions, i.e. in a classroom environment monitored by an instructor. No dictionaries or electronic devices were allowed in the process. All participants were given the same amount of time for each of the tests.

Aligned with the *Prompts Design Guidelines* by Hamp-Lyons and Kroll (1996, p. 60), prompts for essay writing of the current study were "as brief as clarity allows". They normally contained a brief background statement and an instruction for the task. One of the examples is as follows:

> *Increasing concerns have been expressed in Hong Kong and mainland China about youngsters starting to experience sex at an early age, some even before ten years old. Discuss the causes of this issue and suggest ways to deal with it.*

The two participating groups were given different essay topics and each group were able to choose from a pool of eight or nine topics. This was to ensure the topics were "culturally accessible to intended interpretation" (Hamp-Lyons &

Kroll, 1996) and thus maximise the opportunity for each participant to write on a topic s/he was familiar with, and comfortable in writing about (see some of the topics in Appendix F).

3.4.2. Questionnaire survey

In addition to the vocabulary tests and essay-writing, a questionnaire survey was also administered. One hundred and fifty respondents completed the questionnaire (67 Hong Kong students and 83 mainland students, the same students who participated in the tests). The questionnaire was designed mainly to elicit subjects' beliefs and learning strategies in writing and vocabulary learning practised in the secondary school. It therefore drew on the rationale and construction of questionnaires employed in previous studies on learner beliefs and learning strategies (e.g. Oxford, 1990; Gu & Johnson, 1996). To understand the impact of contextual factors on the development of such beliefs and strategies, questions on teacher input and learner experiences in the two researched areas were also included (see Table 3-2).

Table 3-2. Items in questionnaire survey

Vocabulary learning (19 items)		Writing (31 items)		
learner beliefs &learning strategies (11 items)	teacher input (8 items)	learner beliefs &learning strategies (10 items)	learner writing experiences (9 items)	teacher input (12 items)

The questionnaire, which contained a total of 50 multiple-choice questions, was composed of two main sections, one on lexical learning and the other on writing. Both sections were sub-divided, with the first section divided into learner beliefs,

learning strategies and teacher input, and the second into learner beliefs, learner strategies, learner experiences and teacher input.

There were two versions of this questionnaire, one for mainland participants and the other for their Hong Kong counterparts. Both versions were in English and the items in the two versions were also the same except for:

- the difference in university entrance English examinations in the two places, the NMET for the mainland group and the UE for the Hong Kong one; and

- the availability of Chinese translations in the version for the mainland group. These translations included some English terminologies (e.g. bilingual, tone and style, and collocation) and sentences which some students might find difficult to understand.

The availability of these Chinese translations is attributable to the pilot study for the current research project. After the pilot study, two focus-group interviews (on the Hong Kong group and the mainland group) were held to identify possible problems in the questionnaire. The Hong Kong group seemed to have no problem with the questionnaire while some of their mainland counterparts expressed concerns about difficult words/phrases/sentences. Members in the mainland group were thus invited to identify these items and then suggest the best possible Chinese translations for the items (see Appendix G for the version for the Hong Kong group and Appendix H for the mainland group).

The questionnaire survey was intended to provide empirical analysis about possible differences in the results of the vocabulary tests and essay-writing between the two participating groups. When the analysis was conducted, different facets of the questionnaire were further divided according to theories of second

language acquisition. For example, aspects related to learning strategies were examined in line with Ehrman, Leaver and Oxford (2003) who categorised learning strategies into direct learning strategies (memory, cognitive and compensation strategies) and indirect learning strategies (metacognitive, affective and social strategies).

3.4.3. Focus-group interviews and case studies

After collection of the above quantitative data, qualitative information was also gathered via focus-group interviews and case studies. This part of the information was intended to shed light on the causes of mainland students' unusual language development patterns identified in the results of vocabulary tests and essay writing. Because of this, all the participants involved were from the mainland. Another reason for only including mainland learners was their misinterpretation of some questionnaire-survey questions (see Section 7.1 for more details). By conducting focus-group interviews, some of the information originally planned to be elicited from the questionnaire survey was obtained from the interviews with the mainland learners.

Three focus-group interviews were conducted at the beginning, in the middle and at the end of the first semester of the participants' university study, respectively. All the interviews lasted about one hour and were undertaken in English. To increase the variety of participants, each participant was invited to only one of the interviews, leading to a total number of 21 students involved in the interviews. All three interviews were recorded and transcribed. To verify information from students when discussing related but different topics, each interview focused on a specific topic (see Table 3-3).

Table 3-3. Topics, participants and timing of focus group interviews

Interview	Topic	Gender of participants	Timing
I	Experiences in vocabulary learning and training in writing in secondary school	4 males 3 females	Beginning of semester
II	Learning experiences of students from big and small cities in China	4 males 2 females	Mid-Semester
III	Training of writing experienced in secondary school and the university	4 males 4 females	End of semester

The focus-group interviews help the project draw a broad-brush landscape picture of the recurrent problems in lexical learning and essay-writing in mainland China and Hong Kong. To gain a deeper insight into the problems and exemplify the impact of contextual influences on learners from mainland China, five case studies were also conducted. These studies can serve as the "portraits" of learners from different economic and social development regions. The five learners were Karrie, Simon, Sam, Mecky and Abby. All five of them underwent the disadvantageous English learning context in the country. The five learners were chosen due to the representativeness of their experiences in mainland China. They were followed up by the researcher via formal meetings, informal conversations and email communications. The length of time for each case varied, from one semester (14 weeks) to three years.

3.4.4. Essay scoring and corpus building

Essay scoring for the present study consisted of two stages. Holistic scoring was adopted at the first stage. This method evaluates "a piece of writing in which the rater reads the paper without marking on it, then rates the paper as a whole (holistically) and assigns the paper a single score within a given range on scales" (Reid, 1993, p. 291). This evaluation method is often used for placement tests whose major objective is to separate learners into different levels according to their writing proficiency. Since the main objective of the writing test in the current study is similar, this assessment method should be appropriate. Another important reason for adopting this rating method was that it was not necessary for the present project to provide detailed feedback on participants' independent skills in their writing. To enhance the inter-rater reliability of the rating, two measures were taken:

- Two experienced language instructors rated each essay; and

- Both raters evaluated the essays according to the writing band descriptors for IELTS (Task 2), which have undergone careful research and piloting for their reliability and predictive validity (Shaw & Falvey, 2008; Shaw & Weir, 2007).

To further improve the inter-rater reliability in the rating process, the following measures were also taken:

- After all the essays were rated, an initial round of intra-class coefficient (ICC) analysis which measures the consistency between the raters' judgment was performed;

- The scores awarded by the two raters for each essay were compared. If they were the same or different by one scale (e.g. Rater A rates the essay 5 but Rater B scores it 6), they were acceptable. If the scores were diversified by two or more than two scales, they were graded for the second time by both raters; and this time, using analytic scoring, a method in which different aspects of an essay are evaluated independently and each given a score (Michieka, 2010). This detailed scoring procedure requires raters to "attend to the multidimensionality" of an essay and thus allows them to make "more valid judgments" about the writing (Hamp-Lyons & Kroll, 1996, p. 62).

Aligned with the IELTS band descriptors (Task 2), four independent aspects of an essay were assessed. All the four components weighted equally, i.e. 25% each. The four aspects were:

a. Task achievement (relevance to the topic, development and support of ideas)

b. Coherence and cohesion (organisation and sequencing of ideas)

c. Lexical resource (range, appropriateness and sophistication in lexical use)

d. Grammatical range and accuracy (range, appropriateness and correctness of sentence structures)

After the analytic scoring process, essays still awarded markedly different scores by the two raters (i.e. different by two scales or above) were discussed so that the two raters reached consensus about the final scores. After this process, a second round of intra-class coefficient analysis was performed to examine the improvement in inter-rater consistency.

Insofar as corpus-building was concerned, the following procedure was involved:

a. Separate the scripts into two groups, one written by the Hong Kong participants and the other by their mainland counterparts so that two learner corpora were built;

b. Type the scripts. After being typed, the scripts underwent the following modification steps. This procedure was intended to maximise the matching opportunities of words in the corpora with those in the two word lists built in the LFP, the General Service Word List (West, 1953) and the Academic Word List (Coxhead, 2000). After this procedure, the chances of the percentage of words in the "not-in-the-list" being pushed up in the LFP were minimised. These procedure included:

1. Correcting all words that were spelt incorrectly in the essays

2. Replacing contractions with the full form of the words. For example, *I've, doesn't, it's* and *let's* will be changed into *I have, does not, it is* and *let us*.

3. Removing items such as:

   - titles of all essays;

   - proper names, names of areas, people and organisations, such as *Guangzhou, Mr. Cheung,* and *WHO*. However, names of countries, such as China and nationality related terms such as *Chinese government* remained;

   - apostrophes (i.e. 's), for example, *teenagers'* and *city's* were changed to *teenagers* and *city*;

   - quotation marks, for example, "one-child policy" and "sex" were changed into *one-child policy* and *sex*;

- newly developed words such as *iPhone, iPad, iPod, iTunes, Tablet, Kindle, mp3, mp4, smartphone, YouTube, QQ, PowerPoint, internet, web, website, chatroom, 3G, Wi-Fi, wireless, laptop, Bluetooth, Skype, Apps, Talkbox, WhatsApp, online, hacker, download, software, login* and *password*;

- made-up English words, i.e. words that do not exist in English, for example, *handphone* (a direct translation from the Chinese term mobile phone);

- short forms (e.g. *hi-tech* and *flu*); and

- acronyms, for example, *CD, TV, PC, DIY, AIDS* and *DINK* (dual income, no kids) and special terms (e.g. *CO2* and *GDP*)

c. Building the two corpora, one from scripts by Hong Kong subjects and the other by the mainland participants.

The two built corpora were then employed for quantitative word analysis about the essays written by the two participating groups for the current study.

The analysis focused on lexical variation and lexical frequency profile, the two factors that determine lexical quality (Laufer, 1994). Lexical variation refers to the type-token ratio in a piece of writing; the greater the number of different words used, the better the writing. Lexical Frequency Profile displays the lexical richness, or "the relative proportion of words from different frequency levels" (Laufer & Nation, 1995, p. 311) in a piece of writing. The comparison of these two lexical items can help us to gain an insight into the variation and sophistication level of word-use in these two groups of learners.

Analyses were also performed to compare the occurrence rate of conjunction devices (sentence connectors), exclamative, interrogative and imperative sentences, use of personal pronouns and model verbs in the two corpora. These analyses were intended to measure the dialogic and hortatory (conversational) features in the writing by the two groups so as to enable comparison of the writers' awareness of genre features, register and style of different types of writing.

The analyses were performed via three computer programs:

- Wordsmith 6.0, a concordancing computer program;

- Vocabprofiler, an on-line computer program operated by Tom Cobb and available at http://www.lextutor.ca/vp/; and

- D Tools (version 2.0), a computer program developed by Paul Meara and Imma Miralpeix to calculate lexical variation, available at http://www.lognostics.co.uk/tools.

3.5. Data Analysis

Both quantitative and qualitative data were gathered for the current project. Details for analyses of the data are as follows.

3.5.1. Qualitative data

Qualitative data for the current project came from three focus-group interviews and five case studies. Each interview was recorded and transcribed. After all the interviews were completed, major findings from each interview were categorised and then summarised. Based on the summary, a report containing information from all the interviews was written.

For the case studies, no recording was made. This was to ensure that the participants were able to engage themselves comfortably in conversations at meetings with the researcher without worrying about negative consequences. Because of this, notes were taken immediately after each meeting with a participant. Unclear points identified during note-taking were clarified with the participant within two days after the meeting. After each case was closed, a detailed report was written.

3.5.2. Quantitative data

Analyses for the quantitative data were undertaken in four phases. The first phase was for the reliability of results from the two vocabulary tests and the writing task. The second, third and fourth phases were intended to generate data for Hypotheses 1, 2 and 3 of the current study, respectively.

3.5.2.1. Phase I

This phase was intended to ascertain the reliability of results from the vocabulary tests and the essay task. Two types of reliability test were performed. Cronbach's co-efficient alpha, a test often employed to assess internal consistency of continuous data (Streiner & Norman, 2002), was conducted on the scores derived from the two vocabulary tests. These analyses were performed for the full sample (N=150) as well as each individual sample (N=67, Hong Kong sample; N=83, mainland sample). Due to a disparity between the results from the main study and those from the pilot study, the mean, standard deviation and score-range from the two phases of the study were compared. This comparison was to provide explanations for the differences.

A two-way mixed intra-class correlation coefficient test, a measure often deployed to assess inter-rater reliability, was performed to identify the level of agreement between the two raters' judgment about the essay scores. Two such tests were conducted for the current project, one before the adjustment of scores and the other after the adjustment (see Section 3.4.4 for details).

3.5.2.2. Phase II

This phase was to test the first hypothesis of the current study. To achieve this purpose, Pearson product-moment correlation coefficient analyses were performed to measure the correlation between the participants' vocabulary size (VS), depth of vocabulary knowledge (DVK), and essay-writing (EW). This included three correlation tests:

1) VS and DVK;

2) VS and EW; and

3) DVK and EW.

These tests were conducted for the full sample of participants as well as for each participant group. It should be noted that from this chapter onward, the focus of the dissertation will be placed more on learners' breadth and depth of lexical knowledge rather than the two vocabulary tests themselves. Hence in most parts of the following sections, VS will be used to refer to the test results from the VLT, and DVK from the WAT.

3.5.2.3. Phase III

This phase was intended to test the second hypothesis. To achieve this purpose, a series of multiple regression tests were performed. The first one was on the VS,

DVK and EW, with the EW as the dependent variable. To measure the unique contribution of each independent variable, a forced entry operation was performed. This operation allowed manual control over the entry sequence of variables into the regression equation so as to determine if the depth dimension of learners' lexical knowledge could make a unique contribution to the prediction of writing proficiency on top of the contribution already afforded by vocabulary size. This was achieved by observing the R^2 change (see Section 5.4.1 for details). At the initial attempt, the VS was manually entered into the regression equation first. To further explore the roles of the two independent variables, the study conducted an additional multiple regression test, with the DVK being entered into the model at the first step this time.

The above multiple regression tests were also performed for each individual participant group, the first one for Hong Kong participants and the second one for their counterparts, the mainland students. Procedures of these tests were the same as those for the full sample.

3.5.2.4. Phase IV

This phase was performed to test the third hypothesis. To achieve this objective, a series of tests were conducted to identify if the language use in the writing of the two participant groups was significantly different. These include 1) t-tests to compare the lexical frequency profile and lexical sophistication (beyond 2,000 words) in the essays by the two groups, and 2) correlation tests to identify the relationship between the essay scores and the lexical variation (by calculating the TTR) in their writing. Due to the reliability concern expressed about the TTR (see Section 2.6.1 for details), the lexical diversity D, an alternative parameter to

measure lexical variation in learners' language production, was also calculated. A correlation test was performed afterwards to test the association between the lexical diversity D and the essay scores.

This phase also included the examination of language use in the essays by the two groups. This part focused mainly on computing the ratio of frequency (RF) of various language items in the two learner corpora. When sentence connectors, use of imperatives, exclamatives and interrogatives were examined, sentence was used as the basic unit for the calculation. To allow for comparison of very low occurrence figures, the frequencies were multiplied by 1,000 (see Section 7.5 for details). Since other analysed items, such as personal pronouns and modal verbs, may occur a number of times in one sentence, word was used as the basic unit for the calculation of their RF. To permit the computation of very low occurrence figures, the RF was calculated at one in every 10,000 words.

The analyses in this phase were further extended to the calculation and comparison of data from the three focus-group interviews. These analyses were largely on the learner/teacher beliefs and learner experiences in vocabulary learning and training of writing in English.

3.6. Summary

This chapter has provided detailed information about the instruments, subjects and research procedure of the project. The methods for the four-phase data analysis have also been described. The research procedure and the data analysis methods were strictly followed in the main study.

In the pilot study, however, owing to the limited number of the participants (24 in total, 11 from Hong Kong and 13 from mainland China), T-tests and regression

analysis were not performed. The main consideration was that the sample size was too small to yield any reliable results. The limited sample size also led to difficulties in identifying problematic questions in the pilot of the questionnaire survey. In the main study, three questions in the survey were misinterpreted by many mainland participants (see Section 7-6 for more details), but this problem did not seem to have appeared in the pilot study. Table 3-4 provides an overview of the instruments that were piloted and the statistical analyses that were performed in this phase of the study.

Table 3-4. Instruments delivered and statistical analyses performed in the pilot study

Instruments	Delivered	Statistical analyses performed	
VS, DVK & EW	✓	Reliability tests,	✓
		M, SD and score range	✓
		Correlation tests	✓
		T-tests	X
		Regression tests	X
Questionnaire	✓	✓	

CHAPTER FOUR

PILOT STUDY

4.1. Overview

Prior to the large-scale main study, a pilot was conducted. This chapter presents the research procedure and findings of the pilot study. Detailed information about the participants is also provided. The study mainly measures the reliability of the two vocabulary tests, examines the consistency between the raters' judgment concerning essays and tests the correlation between the participants' vocabulary size, depth of vocabulary knowledge and essay-writing.

4.2. Sample

The sample for the pilot study involved 25 participants at the initial stage, 13 from mainland China (four males and nine females) and 12 from Hong Kong (eight males and four females). However, one Hong Kong participant did not achieve the satisfactory level, i.e. scoring 67% or higher, in the VLT at the 3,000 word families; his scores were therefore excluded from the data analysis. The final number of learners counted in the pilot study was consequently 24.

The mainland group consisted of similar learners in terms of age, education level and English learning experiences. Of the 13 participants, 12 were fresh secondary-school graduates who had just been admitted into one of the universities in China and one was a year-one tertiary student there. This comparable education background means that they also have a similar age (18-20 years old), similar years of English learning experience (11-12 years) and similar English proficiency.

The subjects in the Hong Kong group were much more diversified. The 11 participants (six males and five females) were all students in the Hong Kong Polytechnic University, one higher-diploma, seven degree and four postgraduate students. This difference in educational background also indicates their dissimilarity in age (17-24 years old), years of English learning experience (11 to 23 years) and English proficiency. Details of the sample in the two groups are presented in Table 4-1.

Table 4-1. Pilot study, participants' profiles (N=24)

	Hong Kong	Mainland China
Age	17-24 years old	18-20 years old
Gender	5 males 6 females	4 males 9 females
Education level	11 tertiary students at 3 levels, higher diploma, degree and postgraduate	1 tertiary and 12 about-to-be tertiary degree students
Years of English-learning	11-23 years	11-12 years

4.3. Delivering procedure

Data collection for the pilot study took place in June 2011. To facilitate the process, and to obtain as many participants as possible for the pilot, data collection sessions were arranged in both Hong Kong and mainland China. In both places, several sessions were organised so as to meet the time-schedules of the various participants. Notwithstanding differences in time and location, all the participants followed the same procedure within a controlled time for each task. No dictionary was allowed in the whole process. This time and procedure control

was designed to ensure the reliability of the collected data. In each session, each subject was asked to:

- fill in a consent form which contained background information on the subject;

- complete the WAT and then the VLT; and

- write an essay.

After the main data-collection procedure, a questionnaire survey was administered to volunteers. Upon completing the questionnaire, the participants were also invited to comment on the questions in the survey and discuss possible problems they encountered or the participants in the main study might encounter while answering the questionnaire. The comments were later used for the questionnaire revision which took place soon after the data collection (see Section 3.4.2 for details).

4.4. Analysis and findings of the pilot study

After scores for the vocabulary tests and essays became available (see Section 3.4.4 for the essay-grading procedure), a series of statistical analyses were performed. The analysis was focused on the levels of reliability of the main instruments and the levels of intercorrelation between the participants' vocabulary size, depth of vocabulary knowledge and essay-writing. Due to the small sample size, t-tests and multiple regression tests, which will be performed in the main study, were not viable in the pilot study (see Section 3.6 for more details). This limited sample size suggests that the statistical results reported in this section might only provide preliminary profiles rather than serve as conclusive indicators.

4.4.1. Reliability analysis for vocabulary tests

Cronbach's co-efficient alpha analyses were performed on scores given to the two vocabulary tests. It was found that the α-value from the VS was .83 and that for the DVK was .93, indicating both tests were reliable (see Table 4-2).

Table 4-2. Pilot study, reliability of vocabulary tests (N=24)

Tests	MPS	Score range	Cronbach's alpha reliability
VS	60	42- 60	.83
DVK	160	76-139	.93

MPS: maximum possible score
VS: vocabulary size
DVK: depth of vocabulary knowledge

To gain a more detailed profile of each individual group, Cronbach's coefficient alpha analyses were also performed on the scores awarded to each subject group. The results showed that the score reliability of the Hong Kong group was higher than that of the mainland group. The α-value of the DVK from the Hong Kong participants, for example, was .97 while that from their mainland counterparts was only .79. To identify reasons for this difference, statistical data was also gathered to measure the mean (M) and standard deviation (SD) of the scores. From these analyses, it was found that the SD of the mainland subjects' scores was much lower than that of the Hong Kong participants. This could probably be attributed to the similar English proficiency level of the mainland subjects. This similarity could have resulted in a narrow range of scores in the two vocabulary tests, which possibly explains why the mainland group's score range of the DVK was only between 107 and 139 (a narrow range, most scores high) while that of their Hong Kong counterparts was between 76 and 147 (a wide range, some scores high but

some considerably lower). This narrow score range from the mainland participants   also helps to shed light on why they outperformed their Hong Kong counterparts on both the DVK (119 versus 116 in mean score, see Table 4-3) and the VS (54 versus 48, see Table 4-4) whilst in the main study, for which the sample size was substantially increased, the mainland group underperformed their Hong Kong counterparts on both of these two tests (see Section 5.2.1 for details).

Table 4-3. Pilot study, scores on DVK (Hong Kong versus mainland China)

Participants	M	SD	Score Range	Cronbach's alpha reliability
Hong Kong (N=11)	116	22.3	76-147	.97
Mainland China (N=13)	119	10.3	107-139	.79

Table 4-4. Pilot study, scores on VS (Hong Kong versus mainland China)

Participants	M	SD	Score Range	Cronbach's alpha reliability
Hong Kong (N=11)	48.	21.50	42-60	.85
Mainland China (N=13)	54	9.70	52-58	.70

Despite differences in reliability levels, the analyses indicated that the α-values derived from both vocabulary tests for each group were acceptable for the purpose of the current study. The instruments could thus be employed in the main study to compare the two target participating groups (see Section 3.3 for details).

4.4.2. Intra-class correlation coefficient analysis for essay scores

The first correlation test conducted in the pilot study was a two-way mixed intra-class correlation coefficient. This test measured the consistency between the raters' judgment about the essays. Before analytical scoring and discussions between the two raters took place (see Section 3.4 for details of the procedure),

the correlation coefficient was .76, which was already at a very acceptable level. After the adjustment, the coefficient reached .95, indicating that the grades awarded by the two raters were highly consistent (see Table 4-5). These results suggested the planned essay-scoring method for the main study was feasible.

Table 4-5. Pilot study, correlation between two essay-markers

	Intra-class correlation coefficients
Before adjustment	.76**
After adjustment	.95**

** Significant at .01

4.4.3. Intercorrelations between VS, DVK and EW

Several other correlation tests were conducted in the pilot study. This time Pearson product-moment correlations were calculated. These analyses were used in order to identify the association between scores on the VS, DVK and EW. Following procedures of the previously discussed reliability tests, an analysis was first performed on scores from all the subjects (see Table 4-6) and then separate tests were carried out on those from each individual group (see Table 4-7 & Table 4-8).

Table 4-6 indicates that there was a moderately strong positive correlation between scores from the two vocabulary tests ($r = .68$, $p < .01$), and a moderate positive correlation between scores on the essay-writing and the two vocabulary tests (r - value at .42 and .45 respectively). Notwithstanding the small sample size, these results seem to have lent an initial empirical support for Hypothesis 1 in the current study, which proposes that EFL learners' depth of word knowledge correlates with their writing proficiency.

Table 4-6. Pilot study, intercorrelations between VS, DVK and EW (full sample, N=24)

Tests	DVK	EW
VS	.68**	.42**
DVK		.45**

** Significant at .01

The analyses on scores from each individual group yielded different results (Table 4-7 and Table 4-8). The correlation between scores on the VS and DVK were still moderately strong (Hong Kong group at .61; mainland group at .70). However, the scores on the VS and the EW from the Hong Kong group were only weakly related (r = .39, p < .01). This situation seemed worse with the mainland group. In this group, the essay scores were not only weakly related with those on the VS (r = .35, p < .01) but also the DVK (r = .30, p < .01). This finding pointed to a difference in the level of association between depth of vocabulary knowledge and writing proficiency in these two groups. This difference could be accounted for by the divergent language learning contexts these groups experienced before they entered the university. If similar empirical data could be obtained from the main study, the third hypothesis of the current project, which proposes that language learning contexts affect the association between EFL learners' depth of vocabulary knowledge and their writing proficiency, would then be further supported.

Table 4-7. Pilot study, intercorrelations between VS, DVK and EW (Hong Kong, N=11)

Tests	DVK	EW
VS	.61**	.39**
DVK		.50**

** Significant at .01

Table 4-8. Pilot study, intercorrelations between VS, DVK and EW (mainland China, N=13)

Tests	DVK	EW
VS	.70**	.35**
DVK		.30**

** Significant at .01

4.5. Summary

The objective of the pilot study was to test the reliability of the instruments to be employed in the main study and to assess the effectiveness of the proposed research methods. The results suggest that the instruments are reliable and the proposed research methods are feasible, although minor changes were deemed to be necessary in the questionnaire survey. Details are as follows:

a. The essay prompts and proposed essay-scoring procedure were effective. No revision was necessary.

b. The two vocabulary tests were reliable (.83 from the VS and .93 from the DVK), although the reliability level from the main study may not be as high as that from the pilot study. This is largely because the expected participants in the main study will not be as diversified. Nevertheless, due to a great increase in the number of participants in the main study, the reliability level should at least remain at an acceptable level.

c. It was identified that minor changes were necessary for the questionnaire survey (in the version for the mainland group). The required changes were made before the main study was formally started (see Section 3.4.2 for details).

CHAPTER FIVE

THE MAIN STUDY

5.1. Overview

Data collection for the main study was conducted soon after the pilot study and the subsequent revision of the survey questionnaire (the version for mainland students only). After the data collection and the filtering of data that failed to meet the requirements of the present project (see Section 3.4.1 for details), grading of the test papers and essays took place (see Section 3.4.4 for details). When all the grades were available, a series of statistical tests were performed. The tests were focused mainly on two areas:

- analyses of reliability of the main instruments employed in the study, i.e. scores from the vocabulary tests and the inter-rater agreement in essay scoring; and

- investigations into answers to the three research questions

5.2. Reliability of vocabulary and essay scores

5.2.1. Vocabulary scores

As in the pilot study, Cronbach's co-efficient alpha analyses were performed on the scores derived from the two vocabulary tests which were the main instruments for the project. The results show that the α-value from the VS was .86 and that from the DVK was .89, which indicates that both tests were highly reliable (see Table 5-1).

Table 5-1. Reliability of vocabulary tests (N=150)

Tests	MPS	Score range	Cronbach's alpha reliability
VS	60	19-60	.86
DVK	160	57-142	.89

MPS: maximum possible score

VS: vocabulary size; DVK: depth of vocabulary knowledge

Cronbach's coefficient alpha analyses were also performed on the scores derived from each participant group. The objective was to identify whether there was a major difference between the reliability levels from the grades of the two groups as in the pilot study. The results suggested that such a disparity seemed to have diminished in the main study. The α-value of the DVK from the Hong Kong participants was .91, only slightly higher than that from their mainland counterparts, which was .88 (see Table 5-2). This situation could be seen more clearly from the VS results. Here, the α-value difference from the two groups almost disappeared, resulting in an alpha level of .86 for the Hong Kong group and .85 for the mainland group (see Table 5-3a).

Table 5-2. Scores on DVK (Hong Kong versus mainland China)

Participants	M	SD	Score Range	Cronbach's alpha reliability
Hong Kong (N=67)	119	15.90	74-142	.91
Mainland China (N=83)	112	14.16	57-136	.88

Maximum possible score: 160

Table 5-3a. Scores on VS (Hong Kong versus mainland China)

Participants	M	SD	Score Range	Cronbach's alpha reliability
Hong Kong (N=67)	44	7.57	23-60	.86
Mainland China (N=83)	42	7.36	19-58	.85

Maximum possible score: 60

This change is possibly the result of a significant increase in the sample size in the main study (150 in the main study versus 24 in the pilot study). The increase allowed for a considerably greater variety of participants, which in turn widened both groups' test-score ranges. Taking the mainland group's score range from the VS for example, the range was 52-58 in the pilot study but increased to 19-58 in the main study (see Table 4-4 and Table 5-3a). This change in score ranges narrowed the differences in the standard deviation (SD) of the two groups. The SD difference from the VS was 11.8 in the pilot study (21.5 from the Hong Kong group and 9.7 from its mainland counterparts, see Table 4-4); but was only 0.21 in the main study (7.57 from Hong Kong group and 7.36 from the mainland counterpart group, see Table 5-3a). The increased score ranges, in particular in the mainland group, also resulted in reversed DVK scores from the two subject groups: the mean score on the DVK from the mainland group (112) was lower than that from their Hong Kong counterparts (119) (see Table 5-2). The test results on the DVK in the pilot study were completely different, with the mean score from the mainland group (119) higher than that from their Hong Kong counterparts (116). See Section 4.4.1 for more details.

The high reliability of the test scores from both groups suggest that scores from each individual group can be analysed independently in the main study. Meanwhile, the diminished difference in the test reliability level from the two groups indicates that, when necessary, scores from the two groups can also be merged in the subsequent data analyses.

Analyses were also performed on the reliability of VS3,000 and VS5,000 from each participant group. The results showed that the tests at both levels were

reliable (see Table 5-3b and Table 5-3c). This indicates that scores from each individual level can be analysed independently when necessary.

Table 5-3b. Scores on VS3,000 (Hong Kong versus mainland China)

Participants	M	SD	Score Range	Cronbach's alpha reliability
Hong Kong (N=67)	25.21	3.86	11-30	.81
Mainland China (N=83)	24.36	3.30	16-30	.71

Maximum possible score: 30

Table 5-3c. Scores on VS5,000 (Hong Kong versus mainland China)

Participants	M	SD	Score Range	Cronbach's alpha reliability
Hong Kong (N=67)	19.67	4.46	10-30	.75
Mainland China (N=83)	17.14	4.79	2-28	.78

Maximum possible score: 30

5.2.2. Inter-rater agreement in essay scoring

A two-way mixed intra-class correlation coefficient test was performed to identify the inter-rater reliability of the raters' judgment about the essay scores. Two such tests were necessary for the current project, one before the adjustment of scores and the other after the adjustment (see Section 3.4.4 for details of the procedure). The objective was to measure the consistency between the two raters' judgment about the essays. The results indicate that before the adjustment (based on results from holistic scoring), the correlation coefficient was .75 (p < .01), a level of substantial agreement between the two raters. After the adjustment (based on results from analytical scoring and discussions between the two raters, see Table 5-4a), the coefficient surged to .93 (p <. 01), suggesting that the grades awarded by the two raters were highly consistent (see Table 5-4). These results reveal that

the essay scores awarded by the two markers agree strongly with each other, which indicates that the scores are highly reliable.

Table 5-4. Correlation between two essay-markers

	Intra-class correlation coefficients
Before adjustment	.75**
After adjustment	.93**

** Significant at .01

The high reliability of scores from both the vocabulary tests and essay writing provides a firm foundation for analyses related to the three research questions for the current project.

5.3. EFL learners' depth of word knowledge and their writing proficiency

This section addresses the first hypothesis of the study:

Hypothesis 1. EFL learners' depth of word knowledge correlates positively and significantly with their writing proficiency.

To test this hypothesis, the study performed Pearson product-moment correlation analyses to identify the association between scores on the DVK and the EW. Table 5-5 summarises the findings. The table shows that when all the 150 participants were considered, there was a positive and moderate correlation between the learner's depth of vocabulary knowledge and their essay proficiency. The *r*-value between the DVK and the EW was .43 (p < .01). To further investigate this relationship, correlation tests were also performed for each group of participants. The tests evinced a disparity between the participants from Hong Kong and mainland China. Hong Kong students' depth of word knowledge was still moderately related to their writing proficiency (r = .48, p < .01) but mainland

students' depth of lexical knowledge was only weakly related to their writing
ability (r = .28, p < .01). The r-value at this level is not very significant.

Table 5-5. Inter-correlation between DVK and EW

DVK	EW
All participants (N=150)	.43**
Hong Kong participants (N=67)	.48**
Mainland participants (N=83)	.28*

** Significant at .01

  * Significant at .05

The above findings have, to some extent, supported the first hypothesis of the
current study, i.e. EFL learners' depth of word knowledge correlates with their
writing proficiency. Meanwhile the results suggest that different language
learning contexts in which learners gain their vocabulary might affect the level of
association between the learners' depth of lexical knowledge and their writing
proficiency. A more in-depth investigation into this issue will be discussed in
Section 5.4.

5.4. Lexical knowledge as predictors of writing proficiency

This section addresses the second and third hypothesis of the present study:

Hypothesis 2: Compared with vocabulary size, depth of lexical knowledge
is a better predictor for the writing proficiency of EFL learners.

Hypothesis 3: Different language learning contexts affect the association
between EFL learners' depth of vocabulary knowledge and their writing
proficiency.

Hypothesis 2 is based on the assumption that vocabulary size is a predictor of EFL learners' writing ability. Although this assumption derives from the well-accepted theoretical stance that vocabulary size is closely associated with the four micro-language skills (see Stæhr (2008), Schoonen & Verhallen (2008) and Laufer & Ravenhorst-Kalovski (2010) for more details), testing the hypothesis in particular settings still requires empirical data. To achieve this purpose, the study performed a 2-tailed Pearson product-moment correlation test to identify the association between the scores on the VS, the DVK and the EW. As evinced by the figures in Table 5-6, both the participants' vocabulary size and their depth of vocabulary knowledge were moderately related to their writing proficiency (r = .40 and r = .43 respectively, p < .01). This indicates both dimensions of EFL learners' vocabulary knowledge were indicators of their writing proficiency. This table also demonstrates that the correlation level between the DVK and the EW was slightly higher (r = .43, p < .01), suggesting the DVK might be a better indicator of the essay writing by these students. Nevertheless, a calculation of the *z* value of these two correlation figures by using the Fisher r-to-z transformation indicated the difference was not statistically significant. As a result, this result will be further tested in the next phase of the study when a series of multiple regression tests are carried out.

Table 5-6. Inter-correlation between VS, DVK and EW (full sample, N=150)

Tests	DVK	EW
VS	.66**	.40**
DVK		.43**

** Significant at .01

5.4.1. Depth of lexical knowledge as a better indicator for writing proficiency

Having confirmed that both facets of EFL learners' vocabulary knowledge, vocabulary size and depth of vocabulary knowledge, are predictors of their writing ability, the study proceeded to test the second hypothesis, i.e. compared with vocabulary size, depth of vocabulary knowledge is a better indicator for EFL learners' writing proficiency. Testing this hypothesis required a series of multiple regression tests. The first one was on the scores from the full sample (Hong Kong and mainland students) with the EW as the dependent variable. As indicated in Table 5-7a, both variables together, i.e. VS and DVK, could significantly predict more than one-fifth (21%) of the variance in the writing scores. To measure the unique contribution of each independent variable, a force entry operation was performed. This operation allowed manual control over the entry sequence of variables into the regression equation so as to determine if the depth dimension of the learners' lexical knowledge could make a unique contribution to the prediction of writing proficiency on top of the contribution already afforded by vocabulary size. This was achieved by observing the R^2 change (see Table 5-7a).

Table 5-7a. Regression results with VS and DVK as independent variables (full sample, N=150)

Step	Procedure	Variable	Status	R^2	R^2 Change
	Both variables	VS & DVK		.21*	.21*
1	Forced entry	VS	In	.16*	.16*
2	Forced entry	DVK	In	.21*	.05*

*Significant at .05

At the initial attempt, the VS was manually entered into the regression equation first. This step showed that VS alone explained 16% of the variance in these

learners' writing efficiency (p < .01). When the DVK was entered into the model later, the R^2 changed to .21, suggesting that the depth dimension of the learners' lexical knowledge added 5% into the variance already accounted for by the learners' vocabulary size. This change, although statistically significant (at the .05 level), added only limited increase into the variance already explained by vocabulary size. To further explore the roles of the two independent variables, the study conducted an additional multiple regression test, with the DVK being entered into the model at the first step this time (see Table 5-7b). As evinced by the figures in the table, the DVK alone accounted for 19% of the variance of the learners' writing ability and their vocabulary size only added 2% of the variance already explained by the depth dimension of their lexical knowledge.

Table 5-7b. Additional analysis

Step	Procedure	Variable	Status	R^2	R^2 Change
1	Forced entry	DVK	In	.19*	
2	Forced entry	VS	In	.21*	.02*

*Significant at .05

The difference between the above 5% (DVK) and the 2% (VS) added-variance seems not very great. These results appear to have confirmed the conclusion from the aforementioned correlation test that the difference between vocabulary size and depth of vocabulary knowledge is statistically insignificant in predicting EFL learners' writing proficiency (see Table 5-6). Nevertheless, when multiple regression tests were performed for each of the individual sample groups, a very different picture emerged. The first one was performed for Hong Kong participants and the second one for their counterparts, the mainland students.

5.4.2. Language learning contexts and the predicting power of the DVK

The same operational procedure as above was employed for each individual group. The results from the Hong Kong group yielded results as follows (see Table 5-8a and Table 5-8b).

Table 5-8a. Regression results with VS and DVK as independent variables (Hong Kong sample, N=63)

Step	Procedure	Variable	Status	R^2	R^2 Change
1	Forced entry	VS	In	.14*	
2	Forced entry	DVK	In	.25*	.11*

*Significant at .05.

Table 5-8b. Additional analysis

Step	Procedure	Variable	Status	R^2	R^2 Change
1	Forced entry	DVK	In	.24*	
2	Forced entry	VS	In	.25*	.01

*Significant at .05

- Both variables together, i.e. the VS and DVK, could significantly predict a quarter (25%) of the variance in the writing scores.

- The VS alone accounted for 14% of the variance of the learners' writing scores. The DVK added a unique 11% of the variance already explained by the vocabulary size, i.e. it provides a unique prediction of the learners' writing scores (11%), over and above the contribution made by the breadth of lexical knowledge.

- The VS added only 1% of the variance already accounted for by their depth dimension of lexical knowledge, which is statistically insignificant. This suggests that, in predicting learners' writing proficiency, their vocabulary

size cannot significantly add to the contribution already made by the depth dimension of their lexical knowledge.

The results derived from the mainland group, however, were very different. Details are as follows (see Table 5-9a and Table 5-9b):

Table 5-9a. Regression results with VS and DVK as independent variables (mainland sample, N=87)

Step	Procedure	Variable	Status	R^2	R^2 Change
1	Forced entry	VS	In	.11*	
2	Forced entry	DVK	In	.12*	.01

* Significant at .05

Table 5-9b. Additional analysis

Step	Procedure	Variable	Status	R^2	R^2 Change
1	Forced entry	DVK	In	.08*	
2	Forced entry	VS	In	.14*	.04*

* Significant at .05

- The two variables of the VS and the DVK together could only predict 12% of the variance in the writing scores of these students.

- The VS alone explained 11% of the variance of the learners' writing scores. The DVK added very little (only 1%) to the variance already afforded by the vocabulary size scores. This increase is statistically insignificant.

- The DVK alone could only account for 8% of the variance of the learners' writing scores. The VS added 4% of variance already explained by the learners' depth dimension of lexical knowledge.

The above results indicate that, in the writing by the mainland learners, vocabulary size is a stronger predictor of their writing proficiency. However, the

depth dimension of their lexical knowledge cannot significantly add to the contribution already made by their vocabulary size.

A comparison of results from the above two sets of multiple regression tests provides clear indications within the following two perspectives.

- Mainland students' lexical knowledge, in particular the depth dimension, is not highly associated with their writing proficiency. Although some of them achieved a score as high as 136 (out of 160, i.e. 85% correct answers) on the DVK, many of them performed unsatisfactorily on the test with the lowest score at only 57 (see Table 5-2), which means less than 36% of their answers were correct. This possibly explains why the mean score of the mainland group (112) was lower than that of their Hong Kong counterparts (119). Given that the overwhelming majority of the mainland participants performed very successfully on the NMET, which assesses candidates' ability to recognise a great number of words for reading comprehension and multiple-choice questions, it is highly likely that the lexical knowledge of many of the participants remained largely at the receptive level. Some learners might have gained the depth lexical knowledge of a number of words, but their knowledge has not been overly translated into ability to use the lexical items fluently/automatically. This situation suggests that these students are learners who "know *little* about a *large* number of words" (Schmitt, 2014, p. 915). In Milton's explanation, they belong to those "with lots of words" in the mental lexicon but the words are in "poor organisation" (2009, p. 150). This lack of organisation probably explains why the mainland participants' depth of lexical knowledge was only weakly related

116

to their writing ability (r = .28), whereas the association level of these two items in the writing by the Hong Kong subjects was much higher (r = .48). This shortage of association between words in their lexicon could also shed light on the results of the comparative regression analysis. These results showed that, in the Hong Kong data, the DVK added a unique 11% of the variance already explained by the VS; whilst in the mainland data, the DVK added only 1% of the variance already accounted for by the DVK, which is not significant.

Since the two groups of participants differed mainly in the learning contexts in which they gained their lexical knowledge and other language skills, this divergence in the association level could be caused by the disparity in their language learning contexts. This deduction has begun to lend some support to the third hypothesis of the current study, i.e. different language learning contexts affect the association between EFL learners' depth dimension of lexical knowledge and their writing proficiency.

- The test results from the mainland group deviate from a widely-accepted theoretical understanding that vocabulary size is the basic dimension of learners' lexical knowledge and the depth of vocabulary knowledge further refines their vocabulary knowledge and thereby facilitates their automaticity in activating words for language production (Qian, 1999, 2002; Stæhr, 2009; Schmitt, 2008). This means the depth dimension of a learner' lexical knowledge should be better associated with productive language skills such as writing. This theoretical position is substantially supported by the test results (see Tables 5-8a & b) from the Hong Kong participants who built the

117

words in their mental lexicon in a more conducive language-learning context (see discussions in Section 2.10); and will be further supported with empirical data set out in the next two chapters.

5.5. Summary

The above two conclusions point to the need for the current study to depend on the Hong Kong samples when determining if the DVK makes a unique contribution to EFL learners' writing proficiency on the top of the variance already explained by the VS. This decision allows the present study to reach the final conclusions as to the association between EFL learners' lexical knowledge and their writing proficiency. These conclusions lead to answers to the second research question:

EFL learners' depth of vocabulary knowledge can provide a unique prediction of their writing scores (11%) on top of the contribution made by vocabulary size. This unique contribution is as great as that found in Qian's study (1999) on reading comprehension (11%), and much greater than that in Stæhr's study (2009) on listening comprehension (2%). This result is not particularly surprising given that learners' depth of lexical knowledge provides more information about the extent of word knowledge in their lexicon (Nation & Gu, 2007) and their level of automaticity in lexical production (Schmitt, 2008), both of which are directly related to their writing ability.

EFL learners' lexical knowledge accounts for a quarter (25%) of the explained variance in their writing scores. This figure, although lower than the 71% of the afforded variance in reading comprehension (Qian, 1999) and the 51% in listening comprehension (Stæhr, 2009), is substantial. This is because many other factors in

the process of writing , both linguistic ones, such as grammar, cohesive devices and register, and non-linguistic ones, such as content, organisation and even hand-writing, also play important roles in determining the quality of a piece of writing. A more important consideration is that both reading and listening require only retrieving the meaning of a word from one's mental lexicon (receptive vocabulary knowledge) but writing demands retrieving and producing an appropriate form of a word in a given context (both receptive and productive dimensions of one's lexical knowledge). A much higher command of learners' lexical knowledge is thereby needed for writing.

5.6. Additional remarks

Before ending this chapter, it might be worth noting that the 3,000 word families played a limited role in predicting writing scores of the participants in the present project. This finding, although beyond the scope of the three research questions for the current project, have significant implications for related future studies (see detailed discussions in Section 8.3.1.2). Empirical evidence for this finding can be found in the following three aspects:

1. In the correlation tests (all samples included, see Table 5-10), the VS at the 3,000 frequency level was only weakly associated with the participants' essay scores ($r = .31$), whilst the VS at the 5,000 level was moderately correlated with the essay scores ($r = .41$).

Table 5-10. Correlation between DVK, VS3,000, VS5,000 and EW (N=150)

	VS 3,000	VS 5,000	EW
VS 3,000		.65**	.31**
VS 5,000			.41**

** Significant at .01

2. In the T-tests, Hong Kong participants performed significantly better than their mainland counterparts on the DVK (p = .003), the VS at the 5,000 level (p =. 001) and the EW (p < .001). However, their performance on the VS at the 3,000 level did not follow the same trend. The mean performance level of Hong Kong students was higher but the difference between the two groups was not significant (p = .149). This is possibly because the 3,000 word families are all high-frequency words and therefore the difference in lexical knowledge at this level between the two groups of learners was insignificant (see Table 5-11).

Table 5-11. T-tests: performances of two sample groups

	Region	N	Mean	Sig. (2-tailed)
VS 3,000	Hong Kong	67	25.21	.149
	Mainland	83	24.36	
VS 5,000	Hong Kong	67	19.67	.001
	Mainland	83	17.14	
DVK	Hong Kong	67	119.04	.003
	Mainland	83	111.90	
EW	Hong Kong	67	5.21	.001
	Mainland	83	4.11	

*p < .005

3. In the regression analysis (see Table 5-12), the VS at the 3,000 frequency level was forced out from the equation, indicating its role in predicting the learners' writing scores was insignificant.

Table 5-12. Stepwise analysis: predicting power of DVK, VS 3,000 and VS 5,000 on essay scores

Step	Procedure	Variable	Status	R^2	R^2 Change
1	Forced entry	VS 3,000	In	.09*	
2	Forced entry	VS 5,000	In	.17*	.08*
3	Forced entry	DVK	In	.21*	.05*
4	Stepwise	VS3,000 removed	Model suggests: DVK and VS 5,000 second try		
1	Forced entry	DVK	In	.187*	
2	Forced entry	VS 5,000	In	.218*	.03*

* Significant at .05

CHAPTER SIX

EFFECTS OF LANGUAGE LEARNING CONTEXTS ON VOCABULARY STUDY

6.1. Overview

As Nation and Gu (2007) have pointed out, an appropriate size of mental lexicon and the extensiveness of knowledge about the words in that lexicon are both requisites for EFL learners' efficient and fluent lexical use. Both dimensions of lexical knowledge, in particular the depth dimension, as shown in Chapter 5, could be affected by the language learning context in which the learners gain their vocabulary knowledge. This chapter explores the extent of this effect by providing related empirical data. It firstly discusses the effect on the learners' vocabulary knowledge by examining the results of the vocabulary tests. It then explores the impact on their vocabulary use in writing by comparing the vocabulary profile, the lexical sophistication, the lexical variation and the D value (an alternative parameter to measure lexical variation in learners' language production), in the writing of the two groups of participants.

6.2. Effects of language learning contexts on vocabulary test results

Before examining the vocabulary test results, it is important to provide more detailed background information about the two groups of participants. As presented in Chapter 3, the gender ratio, age range and education level of these two groups are essentially similar (see Table 3-1). The only major difference between them lies in the contexts in which they received their education, in

particular the learning context in which they studied English before entering the university.

Another crucial factor to consider is that only elite secondary school students from the mainland (about the top 5%) may have an opportunity to enter a Hong Kong tertiary institute. Due to the practice of English as the MOI in Hong Kong universities, a requirement on English proficiency is imposed on candidates from mainland China (an overall mark of 120 out of 150 on the NMET, the English test in the national university entrance test battery in China). Language requirement for local students, on the other hand, is much more relaxed. The English test for the university entrance in Hong Kong was *Use of English* (UE). UE employed a 7-band scale marking scheme (from A to U). Based on the UE results, a large majority of Hong Kong participants for the current project received grades between D and F. One of the reasons is that the Hong Kong Polytechnic University, where the data for the present study was collected, is not one of the top universities in the territory, notwithstanding the fact that some faculties/disciplines in the university, such as Hotel and Tourist Management and Humanity Studies, enjoy a relatively high international academic ranking. This status means that the university can only attract a limited number of very top or elite local secondary school students. The corollary is that the UE grades of a large majority of Hong Kong participants for the current study (74.6%) fall between D and F. This is in marked contrast to the grades of the mainland subjects. When their NMET grades were converted according to the 7-band scale marking scheme used in UE, the grades for all the participants fell between A and B (see Table 6-1a). The conversion was calculated according to the 150 total scores on

the NMET. When divided by seven, the scores of 30 participants fell into the A-level of UE results and those of 53 fell into the B-level (see Table 6-1b).

Table 6-1a. Participants' English level before entering university

Region	English level based on Entrance Test	English level based on Entrance Test
% of grades received	A - B	D – F
Mainland China	100%	0%
Hong Kong	4.5%	74.6%

Table 6-1b. Score conversion on 7-band marking scale

Hong Kong group UE results	Number of participants	Mainland group NMET results	Number of participants
A	1	137-above	30
B	2	107-136	53
C	17	86-106	0
D	28	64-85	0
E	19	43-63	0
F	0	22-42	0
U	0	0-21	0
	Total=67		Total=83

If judged by the above grade ranges of the two groups of participants, the mainland participants should significantly outperform their Hong Kong counterparts on the vocabulary tests as well as on the writing task for the present project. However, an analysis of the collected data yielded very different results. As indicated in Table 6-2, Hong Kong participants received a higher average grade than their mainland counterparts on both the vocabulary size test and the depth test. The means from the two tests were both significantly different (2-tailed,

see Table 6-2). This suggests that the mainland students obtained significantly lower results than their Hong Kong counterparts on both vocabulary tests, which was rather surprising given their outstanding performances on the NMET. To further explore the differences between these two groups, the current study examined their results from the writing task.

Table 6-2. T-tests: DVK and VS

	Region	N	Mean	Sig. (2-tailed)
DVK	Hong Kong	67	119.04	.003
	Mainland China	83	111.90	
VS	Hong Kong	67	44.88	.007
	Mainland China	83	41.51	

6.3. Effects of language learning contexts on writing proficiency

The writing test for the current study, as illustrated in Section 3.3, required that all participants write an expository essay. The task was administered under controlled conditions regarding time, venue and accessibility to dictionaries. To maximise participants' opportunity to write on topics they are familiar with and comfortable to write about, a pool of eight to nine topics were provided for them to choose from. In addition, Hong Kong and mainland groups were given different essay topics so that each group was able to have topics that were "culturally accessible to intended interpretation" (Hamp-Lyons & Kroll, 1996, p. 60). All the completed essays were then scored in accordance with the 9-scale band descriptors used for the second writing task in IELTS.

To investigate the performances of the two groups, the mean and standard deviation of each group were computed. As shown in Table 6-3a, the mean grade from the Hong Kong subjects (5.21) was higher than that from their mainland counterparts (4.11). This indicates that on average the Hong Kong participants outperformed their mainland counterparts in the writing task. More detailed information can be seen in Figure 6-1.

Table 6-3a. Participants' performance on writing test (1)

Essay	Hong Kong				Mainland			
	Min	Max	**Mean**	SD	Min	Max	**Mean**	SD
Grades	2	8	5.21	1.54	1	7	4.11	1.32

* Maximum possible score is 9
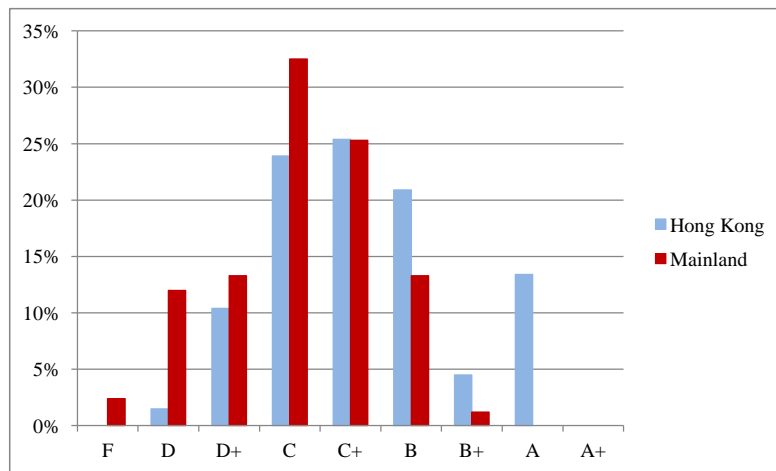
Figure 6-6-1. Essay grades from two groups of participants

Table 6-3b. T-tests: essay writing

	Region	N	Mean	Sig. (2-tailed)
Essay writing	Hong Kong	67	5.21	< .001
	Mainland China	83	4.11	

To obtain further insight into the situation, a t-test was conducted to compare the mean grades from these two groups. The result revealed that the difference in the means was statistically significant ($p<.001$) (see Table 6-3b). This result, again, suggests that Hong Kong subjects performed significantly better than their mainland counterparts on the writing task. It has hitherto become clear that the performances of the mainland participants on both the vocabulary tests and the writing task were significantly worse than their counterparts, notwithstanding their remarkable achievements on the NMET, the national university entrance English test.

Another point worth noting is that the standard deviation (SD) of grades from the mainland group (1.32) was lower than that from the Hong Kong group (1.54). This could mean that most mainland students received low grades and very few of them performed very satisfactorily on the writing test. To investigate this possibility, the grades of the two groups were further divided into *high, medium-high, medium-low* and *low* to identify which group received high grades and vice versa. The comparison indicates that a large majority of the mainland subjects (60.2%) received either low or medium-low grades whereas similar percentage of the Hong Kong participants (63.3%) achieved high or medium-high scores (Table 6-4). An even more important finding is that almost one-fifth of the Hong Kong subjects (18%) were awarded with high grades (A+ to B+) while very few mainland participants (1.2%) achieved this level of success.

Table 6-4. Participants' performance on writing test (2)

	Low	Medium low	Medium High	High
Grade Range	D - F	D+ - C	C+ - B	B+ - A+
Mainland China	14.4%	45.8%	38.6%	1.2%

Hong Kong	1.5%	34.3%	45.3%	18%

Given that a vast majority of the mainland participants were the elite students in China's secondary school student population, and scored high on the NMET, the above findings were surprising. A review of related literature (e.g. Jin & Cortazzi, 2006; Qi, 2007; Paltridge, 2007; Wu, 2008; You, 2010) points to two possible explanations for this situation:

1. The NMET, the English language test in the university entrance examination system in China, cannot effectively measure learners' writing proficiency. It is, to some extent, an invalid test; and

2. The English-learning context in China has negatively affected learners' vocabulary knowledge, in particular their productive lexical skills.

It is highly likely that both factors played an important role in the making of the above situation. It is also more likely that the first factor is imbedded in or a part of the second, i.e. the learning context is the principal cause and the NMET is a contributing factor to the learning context. The scale of the current project, however, does not allow room for substantial investigation into the first factor. An entirely different project would be required if such an investigation were to be conducted. The second factor which is related to the learning context, however, falls within the scope of the current project. Examining this factor will require an analysis of lexical richness in the participants' essays, including lexical frequency profile, lexical sophistication and lexical variation. Statistical data from these lexical measures not only provide indications of the number of words EFL learners have memorised but also of how the words are used in language production, which is often affected by the language learning context in which they

128

learn the words (Ellis, 2008b; Schmitt, 2008).

6.3.1. Lexical frequency profile in essays by participants

The first examined feature for lexical richness is the Lexical Frequency Profile (LFP) in the essays by the subjects. The LFP measures the free productive lexical knowledge or the word-usability of learners. It does so by identifying the lexical richness at four frequency levels: the first 1,000 words, the second 1,000 words, words from the AWL and words "not-in-the-list", which are normally low frequency words. In the LFP, a word is defined as a base form plus its inflected and derived forms, i.e. a word family. A free on-line program called Vocabprofiler (Cobb, 2011) was deployed to compute the lexical frequency profile for the current study. Since only consistent sample sizes of around 300 words can generate reliable results in the operation of the LFP (Smith, 2005), only the first 300 words of each essay were used in the computing process. Essays shorter than 300 words were excluded accordingly. This explains the change of participant number (66 from the Hong Kong group and 80 from their mainland counterparts) in each group (see Table 6-5a).

Table 6-5a. LFP in participants' essays (1)

	Region	N	Mean	Sig. (2-tailed)
K1	Hong Kong~~HK~~	66	81.4	.001
	Mainland China~~Mainland~~	80	85.9	
K2	Hong Kong	66	5.6	.016
	Mainland China~~Mainland~~	80	4.8	
AWL	Hong Kong~~HK~~	66	7.9	.002
	Mainland China~~Mainland~~	80	6.3	
Not-in-the-list	Hong Kong~~HK~~	66	5.1	.001
	Mainland China~~Mainland~~	80	2.9	

Significant at .05

As indicated in Table 6-5b, except for the first 1,000 word level, Hong Kong students employed higher proportions of words than their mainland counterparts at the other three levels and the difference at the each level is statistically significant (p < .05, see Table 6-5a). Their disparity in the use of "not-in-the-list" words, which are often low frequency and more sophisticated lexis, is particularly discernible. This possibly means that, compared with their Hong Kong counterparts, the mainland participants were more comfortable in using high frequency words, in particular the first 1,000 word level, but less confident in using not-in-the-list words, which are largely low frequency and more sophisticated words. This can be seen more clearly in Table 6-5b.

Table 6-5b. LFP in participants' essays (2)

	K1	K2	AWL	Not in the list
Hong Kong	81.4	5.6	7.9	5.1
Mainland China	85.9	4.8	6.3	2.9

6.3.2. Lexical sophistication in participants' writing

The second measure used to examine the lexical use in the essays was lexical sophistication. Lexical sophistication refers to the proportion of infrequent or advanced lexis in a text. In most computer programs of lexical analysis, such as Heatley *et al*.'s (2002) *Range,* words outside the most frequent 2,000 words in West's (1953) General Service Word List are categorised as infrequent/advanced words. This is why Laufer (1994) also uses "beyond 2,000" words to refer to advanced lexis in a learner's language production. This measure is built on the widely-accepted theoretical assumption that the higher a learner's proficiency is,

the more likely s/he is able to deploy infrequent words in language production (Nation & Laufer, 1995).

This measure, however, is also affected by the length of texts, as sample sizes too long (over 300 words) or too short (below 200 words) can trigger unreliable results (Meara & Bell, 2001). To attain reliable data, the current study used only the first 300 words for essays over 300 words long and excluded essays shorter than 200 words in length. This explains a further change of participant number (66 from the Hong Kong group and 82 from their mainland counterparts) in each group (see Table 6-6a and Table 6-6b).

Table 6-6a. Lexical sophistication in participants' essays

	N	Most frequent 2,000 words	Beyond 2,000 words
Hong Kong	66	86.5%	13.5%
		SD 4.1	SD 4.1
Mainland China	82	90.4%	9.7%
		SD 3.3	SD 3.3

Table 6-6b. T-test: beyond-2,000-words

	Region	N	Mean	Sig. (2-tailed)
Beyond 2,000 words	Hong Kong	66	13.5	.001
	Mainland China	82	9.7	

Significant at .05

As shown in the Table 6-6a, the mainland students employed more words from the most-frequent-2,000-word category (90.4%) than their Hong Kong counterparts (86.5%) but much fewer infrequent or advanced words (9.7%) than the Hong Kong participants (13.5%). According to the t-test shown in Table 6-6b, the difference in using the frequent/advanced words between the two groups is statistically significant ($p < .05$). This confirms the aforementioned supposition

that mainland students were more at ease than their Hong Kong counterparts in using high frequency words but less confident in using low frequency and more sophisticated words (see Section 6.3.1).

6.3.3. Lexical variation in essays by the participants

Another measure deployed to investigate the lexical use in the essays was lexical variation. Lexical variation, also referred to as lexical diversity, measures the variety of activated words in a learner's free language production. It calculates the type-token ratio (TTR) in a piece of language production. This parameter draws upon a theoretical assumption that the higher a learner's proficiency is in a language, the better s/he is able to activate a variety of words in his/her mental lexicon when speaking or writing (Milton, 2009). In accordance with this assumption, an essay by a learner of high proficiency should contain more diversified lexis and therefore the TTR derived from the essay should be higher than that from an essay written by a low-proficiency learner. This measure, notwithstanding its limitations, has been much deployed to measure learners' language proficiency. Data derived from the measure demonstrate a significant correlation between learners' lexical diversity and their language proficiency (see Daller & Phelan, 2007; Milton, 2002).

Due to the major limitation of the measure, its "extreme dependence" (Tidball & Treffers-Daller, 2007, p. 136) on length of texts (TTR curve falls with the increase of text size), it is now only used "with texts of equal length" (Milton, 2009, p. 126). In line with this practice, the current study only used the first 300 words of each essay and excluded essays shorter than 300 words when processing the texts. This explains why the participant number of each group has resumed to 66 from

Hong Kong and 80 from the mainland, the same figures used in the procedure of the lexical frequency profile in the students' writing (see Table 6-7).

Table 6-7. Correlation between essay scores and lexical variation

			Type/Token ratio
Hong Kong	N=66	Pearson Correlation	.31[*]
		Sig. (2-tailed)	.011
Mainland China	N=80	Pearson Correlation	-.07
		Sig. (2-tailed)	.527

* Significant at .05 (2-tailed)

As indicated in the table, the TTR derived from the essays by Hong Kong participants correlated significantly with their essay scores (r = .31, p < .05). Such a correlation, however, could not be found in the essays by their mainland counterparts. What is perhaps even more surprising is that a negative correlation (r = -.07), albeit insignificant, was found between the lexical variation and their essay grades.

These findings were rather surprising because lexical variation has been generally regarded as a highly sensitive predictor of learners' language proficiency (e.g. Zareva *et al*, 2005) and an important indicator of their writing ability (e.g. Laufer & Nation, 1995). This positive association has been evidenced in the results from the Hong Kong participants as indicated in Table 6-6. The findings from the mainland students, however, showed a striking contrast, a negative correlation between their ability to write and the lexical variation in their writing, which deviates markedly from the sufficiently researched conclusion. Even if one can argue that the TTR in learners' writing does not *always* correlate highly with their

writing proficiency as shown in several studies (e.g. Daller & Phelan, 2007; Yu, 2009), no study so far has reported a negative correlation between these two facets of learners' writing.

To resolve the reliability concern expressed about the TTR, it was decided to calculate the lexical diversity *D*, an alternative parameter to measure lexical variation in learners' language production. Although this concern has been largely addressed in the process by only using texts of equal length (the first 300 words in each essay) when the ratio was computed, deploying another measure could help to check or even further develop the finding. The advantage of lexical diversity *D* is that it allows for texts of various lengths so that a fuller picture of lexical variation can be identified. Owing to this feature, all words in the essays by participants in the study were used in the computing process. This explains why the participant number from Hong Kong has returned to 67 and that from the mainland to 84 (see Table 6-8).

Table 6-8. Correlation between D and writing scores

			D
Hong Kong	N=67	Pearson Correlation	.02
		Sig. (2-tailed)	.89
Mainland China	N=84	Pearson Correlation	-.19
		Sig. (2-tailed)	.08

The rationale of the parameter *D* rests on the assumption that "the higher the *D*, the greater the diversity of a text" (Yu, 2009, p. 239). After the D value from each essay was ascertained, the average D value from all essays by each participant group was computed and then a Pearson coefficient correlation test was

performed to identify the relationship between the essay scores and the *D* derived from them.

As evinced by the figures in Table 6-8, there was a positive association, albeit insignificant, between Hong Kong students' essay scores and the *D* from their essays. However, such a positive relationship diminished in the results from their mainland counterparts. Instead, a negative relationship between the items, although insignificant (r =-1.9), emerged.

This result has further confirmed the findings arising from the TTR parameter. The ratification then allowed the study to make the following two conclusions:

1. The lexical variety in the writing of the Hong Kong participants was closely related to their writing proficiency; however

2. There was a negative relationship between the diversity of words used by the mainland participants and the quality of their writing.

6.4. Summary

This chapter has examined the performances of the two participant groups in the vocabulary tests and essay writing. The results indicate that, notwithstanding the outstanding performances on the NMET, the mainland group were significantly outperformed by their Hong Kong counterparts in both the vocabulary tests and essay writing. They were also found to use much less sophisticated lexis in their writing than the Hong Kong students. This indicates that whilst the mainland participants may have the ability to recognise the forms of many English words and retrieve their meaning in Chinese, which allowed them to perform outstandingly on the NMET, they were not able to use many of these words in

their writing. Since the meaning of most words in their lexicon was memorised in Chinese, many participants were not able to achieve acceptable scores on the DVK, for which candidates need to know either the collocation or semantic meaning (including synonymy and polysemy) of a headword in English. Some participants were not even able to achieve a high score on the VS, for which candidates have to select the meaning of a tested word in English.

It was also identified that there was a negative relationship between the diversity of words used by the mainland participants and the quality of their writing. This negative relationship deviates severely from the normal language development patterns of EFL learners, in which one's writing ability is closely associated with the lexical diversity in his/her writing (Nation & Laufer, 1995). This deviation cannot simply be triggered by the limited language proficiency of these learners, given their outstanding performances on the NMET, but more likely by the approach through which they learned English, in particular, the way their vocabulary was built in their mental lexicon and the training they underwent in English writing. Since EFL learners' learning approaches are much shaped by their language learning contexts (Ellis, 2008b; Schmitt, 2008), it is highly likely that the identified deviation is a consequence of the English learning context these learners were placed in before coming to Hong Kong. Their Hong Kong counterparts, on the other hand, may not have such a problem due to their more conducive English language learning environment (Evans, 2011a; Lee, 2012). To test this theoretical deduction, data related to participants' experiences in writing and vocabulary learning would be necessary. This data should include both

qualitative and quantitative investigation. Discussions of such data will be presented in the next chapter.

CHAPTER SEVEN

ENGLISH LEARNING EXPERIENCES IN CHINA

7.1. Overview

This chapter presents data on the learning experiences of the participants for the current study, in particular those from the mainland before they came to study in Hong Kong. It aims to ascertain the causes of the deviant language development patterns identified in the mainland students (see Chapter 6).

The data include a large scale questionnaire survey (150 respondents), three focus-group interviews (21 participants in total) and five case studies. The questionnaire survey was administered immediately after the participants completed the two vocabulary tests and the writing task. When the survey data was analysed, it was found that the participants from the mainland and Hong Kong seemed to have considerably different interpretations of some questions in the survey. For example, one of the questions asked about the sufficiency of their training in English writing in the secondary school. The results indicated that more Hong Kong participants (41.4%) thought they received insufficient training than their mainland counterparts (33.7%) (see Table 7-1 below).

Table 7-1. Training in English writing

Question: I did not have much training in English writing.		
	Hong Kong Participants	Mainland Participants
Very true	13.6%	3.6%
Almost true	27.8%	30.1%
Total	41.4%	33.7%

This finding was not in line either with the literature in the related field or the information already gathered for the present project. As a consequence, some participants were invited to discuss their answers. The discussion revealed that these two groups of students had very different interpretations of "training in English writing". To Hong Kong students, this term referred only to writing compositions such as letters, descriptive essays and expository essays; whilst to most of the mainland participants, this term seemed to mean much more, ranging from translating a sentence from Chinese to English to joining two English sentences to make a new sentence. The mainland participants' interpretation confirmed the observation of Qi (2007), in which she points out:

> *Writing practice is a vague term in China's ELT circle. It can mean constructing sentences, re-writing what one has read, or writing an essay. Even copying two simple sentences and join them with an adverbial clause to make a complex sentence is treated as writing practice* (p. 59).

Qi's comments suggest that the misinterpretation of the survey questions by the mainland students was a corollary of the learning context in China. This context also caused misinterpretations of several other questions in the survey. To deal with this situation, it was decided to conduct focus-group interviews to collect qualitative data, which could help to better understand the answers to the survey questions by the mainland participants and gain more detailed information about their learning experiences in their secondary schools. The questionnaire survey and focus-group interviews could result in a broad-brush explanation of the recurrent problems the present project intends to investigate. To obtain a deeper understanding of the problems, five students were identified for case studies.

These cases could serve as the "portraits" of learners who underwent difficulties caused by the disadvantageous English learning context in China. The description of these portraits provided one part of the qualitative data for Hypothesis 3 of the study, i.e. language learning contexts affect the association between EFL learners' depth of vocabulary knowledge and their writing proficiency.

7.2. Case studies

Five students were identified for the case studies based on the representativeness of their experiences in mainland China. Of the five students, two were males, Simon and Sam, and three were females, Karrie, Mecky and Abby. All the names presented in this chapter are pseudonyms and written consent was obtained from all five students prior to their participation.

Given that all five cases were followed up by the current researcher who was teaching many of the project participants in their first-year English subject, *English for University Studies*, personal pronouns such as *I* and *my* will be used in the following few sections to allow for accurate and precise description and analysis of the five cases. All communication with the five candidates, such as meetings and email correspondence, took place in English.

7.2.1. Case study I

The first case was a female student, Karrie, who was from Beijing. Karrie attained a very high grade on the NMET (142 out of 150), particularly in writing (27 out of 30)*. She was not my student during the time when the case studies were undertaken and was referred to me by a language instructor who is a native speaker of English. (*Note: The total score given to the writing section of the*

NMET varies, ranging from 35 to 25, in different provinces/capital-cities that set their own NMET papers. How many scores to allocate to this section depends on the decision of each local government.)

Because of the current study, I emailed the teaching staff members in the English Language Centre where I teach in an attempt to identify students whose writing was encumbered with overly complex sentences and lexis but makes little sense to readers. A considerable number of scripts were received from colleagues, of which Karrie's was the most illuminative (see Appendix D). I invited Karrie to discuss the causes of her writing problems with me so as to help her write better in the future. She accepted my invitation and we had several meetings afterwards. In our meetings, we discussed a number of issues related to her experiences in English learning, particularly in writing.

Karrie believed the grades a learner received from high-stakes English tests were the most reliable indicators of his/her English proficiency. To obtain top grades on such tests, Karrie thought two factors were essential: memorising as many low-frequency words as possible and composing long complex sentences in essay writing. To achieve the vocabulary goal, she memorised word lists and set-phrases in her textbooks as well as in TOEFL and even GRE (Graduate Record Examinations) preparation books. To accomplish her mission of using long and complex sentences in writing, she diligently worked on grammar and used many sentence-connectors (mainly conjunctions) in her essays.

Karrie had not taken any TOEFL test by the time we met and was not prepared to take one in the near future despite the fact that she attended several TOEFL preparation tutorial schools. At one of our meetings, she showed me the notes she

took in these schools. The notes mainly listed words and phrases that are supposed to help her obtain a high score on a TOEFL test.

Karrie's strategies seemed to have worked for her considering her outstanding performance on the NMET. Nevertheless, she was still much concerned about her examination skills and asked me many questions in this regard during the meetings with me. One of the consequences of her focus on test-strategy training was her reluctance to spend time reading or listening to English in daily life because such activities "take too much time". Whenever she engaged herself in independent language learning, she read an examination-preparation book. A much more serious consequence was her tense and ponderous style of writing, which was full of low-frequency words and overly complex sentences but often made little sense to her readers.

With the assistance from me and possibly also other people in the university, her writing showed some improvement at the end of the semester but the progress was "very limited" according to the instructor for her class. Karrie claimed this was because she did not even know "how to compose simple sentences anymore".

7.2.2. Case Study II

The second case study was Simon, a student from Chang Chun, the capital city of Jilin province in northeast of China. He came to my attention because of his unusual performance in the subject I was teaching. He appeared to be a very attentive student in class but seemed not able to follow what was going on in class sometimes. The first time he delivered an oral presentation, many students including myself could understand very little of his content despite the number of PowerPoint slides he deployed during the presentation. His writing skills seemed

not much better. This situation caused me concern and I therefore approached him a number of times during the semester with the intention of providing special assistance for him.

During our meetings, Simon told me about his English learning experiences before coming to Hong Kong, which he believed were the main cause of the problems he was facing at the time. In his secondary school time, textbooks were the focus of students' English learning. Supplemented with the textbooks were usually exercises on discrete-item knowledge of grammar and vocabulary. Most of such exercises were in multiple choice (MC) format. Students were also required to do many reading comprehension exercises (usually in MC as well). These exercises, according to Simon, helped students to gain high grades on the NMET because around 85% of the NMET questions were in the format of MC.

Simon confessed he paid little attention to speaking skills in his secondary school time because speaking was not tested (and is still not being tested currently) on the NMET. Students in his secondary school also rarely practiced writing. When they occasionally had a practice, they only focused on descriptive writing, for example, describing a person or a place, and attached little importance to expository essays. This was largely because such a genre of writing was usually given as guided writing on the NMET, i.e. all the main points for the essays were given in the task prompts which were normally in Chinese. Students only needed to translate the main points and then link the points with the phrases they had memorised, of which many were transitional devices (mainly sentence connectors), such as *furthermore, moreover, first* and *second*.

Simon commented that clear and tidy handwriting was extremely important on the NMET. According to what he was told, NMET raters had very little time for each piece of writing and were not able to read the details. They therefore rated candidates' compositions according to their "impression". This impression, according to Simon, was largely determined by the following three factors:

- coverage of main points given in the task prompts;

- use of complex sentences and difficult words; and

- clear and tidy handwriting

Simon appeared to have satisfied the above three criteria and attained 22 (out of 25) on the writing task of the NMET. However, he was not able to cope with the writing tasks required in Hong Kong universities. With my assistance and his determined effort, he had made some progress by the end of the semester. However, I believe that he still has a long and arduous journey before he can meet the writing standard of the university.

7.2.3. Case Study III

The third student was Sam, a student from Dong Ying, a small city in Shandong Province. Similar to Simon, Sam was also in one of my seminar groups for the subject of *English for University Studies*. He came to my attention due to his unusual passivity in class. He seldom spoke up in class. Whenever asked to answer a question, he could hardly utter a complete sentence for his answer. He received a D for his first in-class written assignment, a grade markedly low in his group. His situation captured my attention from the very beginning of the semester. I therefore observed him carefully and communicated with him on a

number of occasions with the intention of helping him perform better in the subject.

From our conversations, I learnt that his passivity in class was caused by his low confidence in his productive language skills. Similar to Simon, Sam attached little attention to his speaking since this skill was not tested on the NMET. His training in writing did not formally start until a few months before he sat the NMET. Nevertheless, with diligence and perseverance, Sam attained a very high score on the NMET. His overall score was 136 (out of 150) and his score on the writing task was 26 (out of 30). He attributed his success to his high command of English grammar, his large vocabulary size and his teacher's strategic training to help him manage the writing task on the NMET.

Sam's teacher demanded students' particular attention to two factors in writing: tidy handwriting and correct grammar. According to the teacher, inasmuch as a composition satisfied these two criteria, even if sentences in the composition made little sense of the given composition topic, a mark of at least 21 (out of 30) could be given. His teacher also trained him to memorise exemplary essays for the NMET and, in particular, the formulaic expressions in these essays. By the time Sam sat the NMET, he had memorised at least 500 such sentences. These sentences were for different parts of an essay, i.e. some were particularly for writing introductions, some for conclusions and others for body paragraphs. Below are examples of sentences that Sam memorised for an argumentative essay:

To be included in an introduction:

    1. Recently, the problem of … has aroused people's concern.

2. There are different opinions among people as to … Some people suggest that…

3. Everything has two sides and … is not an exception, it has both advantages and disadvantages.

To be included in a conclusion:

1. There is no doubt that … has its drawbacks as well as merits.

2. Taking into account all these factors, we may reasonably come to the conclusion that…

3. Hence/Therefore, we'd better come to the conclusion that…

Equipped with these sentences, Sam only needed to activate some of them, changing words here and there based on the composition topic, and then adding a few other sentences to link the memorised sentences together so as to achieve a high grade on the writing tasks of the NMET.

Before the NMET, Sam and his classmates took many mock NMET tests, sometimes as many as four sets of tests per day. For each paper, they had to write a composition. This practice helped Sam and his classmates obtain high grades on the NMET. However, Sam believed that most of his classmates, including himself, lacked productive skills in English. One of the telling examples he cited was his classmate George. This student received 148 marks (out of 150), almost a full mark on the NMET. Nevertheless, he could neither speak, write nor even read well. What he was competent with were examination skills, particularly in managing reading comprehension questions.

George could not write well but he used the same strategies Sam employed and attained nearly a full mark on the NMET. Because of this, he received an offer from Peking University, one of the very top universities in China. When his university life started, however, he found it difficult to manage the English learning on the campus where the medium of instruction for most English subjects is English and the teaching approach is relatively more communicative. Both Sam and George believed that in their secondary school, "the only purpose of learning English is to achieve high grades on the NMET".

With my assistance and his assiduous effort, Sam's speaking and writing skills showed some improvement at the end of the semester. Nevertheless, I believed his journey of becoming a successful language-learner would still be a long and onerous one.

7.2.4. Case Study IV

Mecky's English learning experiences and her English writing were examined for the fourth case study. Mecky is from Guangzhou, one of the most economically and politically developed cities in China. She came to my attention because of the unnecessarily long sentences she often deployed in her writing and also because of her repeated claims that she did not know how to write short sentences in English.

Mecky believed that her secondary school English teacher, the only one she had in the four years at school, was the major cause of her problems in English writing. She studied in one of the most prestigious secondary schools in Guangzhou. The English teacher, although only in her early 40s, adopted very traditional approaches in teaching. She attached much attention to grammar and strongly

encouraged students to use various subordinate clauses, in particular attributive clauses, which are commonly used to modify, describe and add meaning to a noun, a phrase or an idea (Halliday, 1967). This function can be seen from Sentences 1 and 2 below (both adapted from Watkins and Biggs, 1996). The noun *students* in Sentence 1 and the noun phrase *one aspect of the paradox of the Chinese learners* in Sentence 2 are modified by an attributive clause.

1. *Students who adopt deep and achieving approaches to learning would be more successful than those who adopt a surface learning.*

2. *There is one aspect of the paradox of the Chinese learners that has been inadequately explored in previous studies.*

To Mecky's English teacher, however, the main purpose of using attributive clauses was to lengthen sentences. Long sentences, according to this teacher, are often written by advanced learners to improve the level of sophistication of their writing. Based on this belief, she strongly encouraged her students to use complex sentences in their compositions. This encouragement became stronger as the NMET approached, and then evolved into a "*one paragraph, one sentence*" request. She explained her request this way: "If you are a good writer, you should be able to use only one long sentence to fill the space of a whole paragraph".

Mecky's teacher expressed strong discontent with students who paid little attention to grammar and often punished such students. Mecky disliked grammar and was therefore constantly disciplined for this reason. Her parents were even summoned to the school many times because of her lack of interest in English grammar. As a consequence, Mecky hated English learning for a considerably long time.

148

Under the pressure of the coming NMET, and more importantly to avoid further disciplinary actions taken against her by the teacher, Mecky finally made concessions and began to follow the teacher's instructions. She started writing unnecessarily long and overly complex sentences, which gradually became a habit and then a major stylistic feature of her writing. According to Mecky, using long, complex sentences and "high-level" words (complex and low frequency lexis) were the teacher's key requirements for students' writing.

By the time Mecky and her classmates started to prepare for the NMET, they had to write a composition every single day. The compositions, however, were normally not commented on and often not even read by the teacher. Instead, the teacher provided the whole class with an exemplary essay for each of the given written task. If a student went to see the teacher for feedback, the teacher usually commented only on obvious grammar problems and checked if the sentences in the composition were long enough and if the composition contained a large number of low frequency words. These experiences largely shaped Mecky and her classmates' views of what counts as good writing.

To help students obtain a high grade on the NMET, the teacher required that all the exemplary essays be included in the class's morning reading, a common practice in China's secondary school system, in which the whole class read aloud texts together under the lead of a class monitor in early mornings. By the time Mecky and her classmates sat the NMET, most of them were able to memorise every single exemplary essay provided by the teacher.

The intensive training by Mecky's English teacher apparently yielded a very mixed impact on Mecky. On the one hand, she successfully obtained a score of

142 (out of 150) on her NMET and her writing score was as high as 24 (out of 25), almost a full score for the task. On the other hand, her understanding of what counts as good writing was by and large distorted. According to my observation, she had very little awareness of English writing conventions, and in particular in register (i.e. tone and style), and often deployed overly long and complex sentences and unnecessarily difficult words, of which many were inappropriately used. This is evident in a detailed analysis of her lexical use in Section 7.4. The analysis was conducted on a paragraph in the second version of a submitted essay. She submitted three versions of the essay (see Appendix I).

The main objective of my requiring Mecky's submission of three versions of the essay was to engage her in process writing, which is one of the most effective approaches to improve learners' writing skills (Sandmel & Graham, 2012). The second purpose was to boost her confidence in her writing ability because the process allowed her to witness her own progress each time she submitted a new version. The third objective was to test her claim that she did not know how to write short English sentences. To achieve this purpose, I carefully explained the problems caused by the overly long and complex sentences in her essays and then advised before each submission that she should use complex sentences only when it was very necessary.

Table 7-2. Three versions of Mecky's essay

	Version 1	Version 2	Version 3
Average word length (by letter)	4.76	4.90	5.07
AWL words (by %)	7.08	8.71	8.97
Average sentence length (by token)	27.44	16.57	16.11

Table 7-3. Word and sentence length in essays by two groups of participants

	Mainland China	Hong Kong
Average word length (by letter)	4.60	4.87
Average sentence length (by token)	17.13	17.12

The empirical evidence derived from the three versions indicate that her claim was rather questionable. The average token number of her sentences in the first version was 27.4 (see Table 7-2), which means her average sentence length was more than 60% longer than that in essays by other participants (around 17.2, see Table 7-3). The average, however, was dramatically reduced to 16.6 in the second version and then further decreased to 16 in the third version. Despite a limited decline from the second to the third version, her average remained similar to that of essays by other participants. Another noticeable feature is that, as her average sentence length declined in each version, her average word length in the essay gradually increased (from 4.76 to 4.90 and then to 5.05); and the percentage of AWL words followed a similar trend (from 7.08% to 8.71% and then to 8.97%). More importantly, the quality of the essay was much improved.

This trend, however, did not continue. When another essay of hers was submitted at the end of the semester, the average sentence length ascended to over 24, much higher than the average of 17 in the writing of other participants.

Mecky believed intelligent students in China should have little difficulty achieving high scores on the NMET if they had sufficient test-taking strategies. Proficiency of the language, on the other hand, was often not the primary concern for students like herself. Motivated by these beliefs, she tried her luck with TOEFL several times despite her relatively low English proficiency. By the time my case study started, she had taken TOEFL twice. The first time she attained 102 (out of 120) but the second time she only received 82 (out of 120), a much lower

score which she blamed on her "bad luck". She was scheduled to take TOEFL for the third time in a week following the day of my last interview with her.

7.2.5. Case Study V

This case study was closely associated with the previous one. It was conducted because of the re-appearance of unnecessarily-long and over-complex sentences in Mecky's essays after her conscious efforts to tackle these problems. Owing to this result, I decided to extend my observation of the fifth case to three years instead of one semester as with the practice for the previous four cases.

This case studied a learner called Abby. She came to my attention because of a message she posted on the forum of my course at the beginning of the semester. One part of the message expressed her frustration at her unsatisfactory participation in the seminar. She attributed this problem to her secondary school English learning experience. The original text is as follows:

> *In high school, we study for the college entrance examination, so we mostly do multiple-choice questions, either right or wrong. And though I may have a large vocabulary, I can hardly think out one when we have to answer questions, for teachers in high school only require us to know the meaning of the words, not think out one when we talk with others (that's why I may hesitate for a long time when having a talk, my brain just can't find a suitable word).*

After reading this forum post, I invited her to tell me more about her secondary English learning experience. She is from Guiyang, a relatively underdeveloped province in Southwest China. Similar to the English teaching in many other underdeveloped areas in China, most English teachers in Guiyang adopted very

traditional teaching approaches. Because of this, Abby experienced very similar training to that of Sam in Case Study III: intensive grammar and vocabulary drilling, exhaustive exercises in the format of multiple-choice and exhausting vocabulary dictations. The productive facets of the language, speaking and writing, were seldom attended to by the teacher. As a consequence, Abby received very little training on speaking and writing in her entire secondary school education. Fortunately, Abby's parents were relatively wealthy and were determined that Abby receive a good education, preferably at the postgraduate level, in the United States. To prepare for her future overseas studies as well as for the intimidating NMET, the parents hired one of the most well-known secondary school teachers in the area to tutor her privately. The main objective of the tutoring was to enhance her productive skills, in particular in writing because writing was tested on the NMET while speaking was not. This tutor, similar to many other experienced secondary school teachers in China, trained Abby to use difficult words and overly complex sentences in writing. This training was clearly manifested in her essays for my course, in which a number of sentences contained more than 50 words (tokens) and were difficult to understand. To raise the awareness of such problems in the essays by Abby and several other mainland students in the class, I took a series of measures. One of them was that I selected overly long and complex sentences from student essays and asked the whole class to use several cohesively connected short sentences to replace them, an approach taken by one of my native-English-speaker colleagues to address the same issue in his class. Another was my introduction to the class of several online programs that assist learners with their lexical use, for example, *Corpus Concordance English* (Cobb, 2011) and *Just the Word* (Edmonds, 2011). I followed this up by requiring

that Abby and her classmates use these online programs to check words they were not sure of before submitting their essays to me. After my determined effort, Abby started consciously reducing overly complex sentences and unnecessarily difficult words in her essays. By the end of the semester, the quality of her essays had greatly improved.

I was the subject lecturer for Abby for only one semester in 2011. In the following two years, I maintained regular contact with Abby but did not ask for any of her writing. This was part of my plan to see how Abby's writing would change before her graduation from the university. In the first semester of her graduation year, i.e. 2014, and elicited by my suggestions, Abby sent me several pieces of her writing, including internship reports, assignments for subjects she was taking and application letters/forms for postgraduate studies. Upon receiving these texts, I immediately examined them and found that unnecessarily difficult lexis and overly complex sentences reappeared in the writing. To verify my initial observation, I performed an in-depth comparative analysis of her writing. To conduct this analysis, the following procedure was followed:

1. Choose two texts written by Abby, one in 2011 after my determined effort to reduce her use of unnecessarily difficult words and over-complex sentences in her writing (Text A) and the other in 2014, the year of her graduation (Text B). Texts A and B were chosen because they contained a similar number of words (around 800) and they both were on generic topics, and hence no technical lexis was necessary (see Appendix J for both texts);

154

2. Ask two native speakers of English (both are academic staff members teaching in the Hong Kong Polytechnic University) to highlight all the words that were used inappropriately in the two texts;

3. Mark the words highlighted by *both* native speakers and then calculate the percentage of the marked words (words used inappropriately) in each text;

4. Identify the *average word length*, *average sentence length* and *max sentence length* of each text by using the Wordsmith tools; and

5. Compare figures derived from Steps 3 and 4 above (see Table 7-4).

Table 7-4. Comparison of two texts written by Abby

	Text A (written in 2011)	Text B (written in 2014)
Average word length (by letter)	5.7	5.3
Inappropriately used words (by %)	5.2	9.4
Average sentence length (by token)	21	25
Max sentence length (by token)	35	55

Figures in the table suggest that Text B contained much longer sentences and more lexical errors than Text A. The average sentence length in Text B was 25, with the longest sentence containing 55 words whilst the average of Text A was only 21 with the max sentence length being 35 words. Although the average word length in Texts A and B was considerably similar (5.7 and 5.3 respectively), the percentage of the lexical errors in Text B (9.4%) was much higher than that in Text A (5.2%). When asked for overall observations about the quality of the two texts, both native speakers commented that Test A was better composed than B. This was not only attributable to the much fewer lexical errors but also the overall shorter sentence length in Text A. These features are evinced in the two sentences

from Text A (one of them being the longest sentence in the text; all lexical errors underlined):

> *A number of <u>research</u> has presented <u>evidences</u> to show that government regulations are helpful to <u>improve</u> the problem of deceptive internet advertising. As numerous innovative and persuasive advertisements with <u>delicate</u> illustrations and inspiring words spread <u>out</u> rapidly through the internet, customers are more likely to be misled by their false descriptions <u>for</u> the products.*

Notwithstanding the errors both in lexical choices (e.g. *improve* and *research)* and morphological forms (e.g. *delicate* and *evidences*), these two sentences are readily comprehensible to most people. In Text B, however, a number of sentences are unintelligible to many readers due to lexical errors and the complexity of the sentences. This can be seen from the two sentences below (all lexical errors underlined):

> *For example, the courses at Columbia SIPA MPA Program will help <u>push</u> my learning in macroeconomics, economic sociology theories to a higher level, while laying solid foundation for further learning such as Budgeting for non-profits, which are particularly crucial since economic and finance knowledge is the base for <u>proposing</u> to solve any <u>fields</u> of issue. <u>Meanwhile</u> those <u>economics</u> <u>rules of thumb</u> will help understand the course of Politics of Policy-making and <u>uplift</u> my <u>visions</u> to view <u>conflict interests</u> at the <u>commanding point</u>.*

The first sentence consists of 55 tokens (words) and is the longest sentence in the text. The following one, although not very long (27 tokens), contains a number of

severe lexical errors. Owing to these errors and the complexity of the structures, it is rather difficult to fully understand the meaning of these two sentences. Both native speakers of English who read the text pointed out the incomprehensibility of the sentences.

The comparative analysis of Text A and Test B indicates that the habit of using unnecessarily difficult lexis and overly complex sentences is deeply rooted in many mainland learners. Efforts to eliminate its impact have to be continuous. A one-semester training session such as was provided for Abby and her classmates seems to have been insufficient. More determined and continuous efforts from teachers and researchers are needed to address this issue.

7.2.6. Summary

The section painted five "portraits" of students who are high achievers according to the results of the NMET but are low achievers insomuch as communicative language skills are concerned. Their problems were caused by the NMET and other high-stakes English tests which form parts of the language learning context in China. These five cases exemplify many students who have undergone and/or are still enduring the disadvantageous language learning environment in the country. The results from the focus-group interviews and the questionnaire survey in the following two sections will enable a broader and deeper understanding of this problem.

7.3. Focus-group interviews

Three focus-group interviews were conducted, respectively at the beginning, in the middle and at the end of the first semester of the participants' university study.

Each interview was focused on a specific topic (see Section 3.4.3 for more details). Although on different topics, most data derived from the three interviews was either confirming or supplementary to the information from the five case studies regarding learners' vocabulary learning and training in writing. Differences were also identified. These differences lie mainly in the learning contexts including teaching and learning resources and teaching approaches in different regions in China.

7.3.1. Lexical learning

All participants at the interviews admitted that the ability to recognise the form of a word and retrieve its meaning in Chinese is the most important aspect of word knowledge they attached attention to. Other types of lexical knowledge pointed out in Nation (2001), such as morphological forms, syntactical and semantic features, constraints in use, collocations and frequency, were largely ignored. The participants also reported that they mainly learnt words from word lists. These lists could be from their own textbooks and teachers' notes as well as NMET and TOEFL/IELTS preparation books. A key factor leading to the popularity of these lists was that they contain complex lexis which are regarded as "high-level" words by many students. Learning complex words, according to the participants, was crucial for achieving a high NMET score, and, in particular, in the writing section. The reason was explained by a participant:

An important criteria to judge our writing ... in the college entrance
examination is ... you should have some ... so called ... high-level words. I
mean, if you can use "significance", you should not use "importance". You

should use words ... words that look longer and more complicated and ...

something like that. It works! ... It give (gives) more marks to your writing.

Because of this, some students memorised 100 words per day and their teachers helped them by delivering frequent dictations in class. These findings add support to the five case studies and also the observations of Qi (2004).

7.3.2. Training in writing

The most common experience of the participants was that their training in writing did not start until the last few months of their secondary school education, to be exact, one to three months before the NMET. The so-called training was essentially test-taking strategy training which, according to Cohen (2006), includes (1) language learning strategies; (2) test management strategies; and (3) test wiseness strategies. Due to the task requirements, task prompts and adopted assessment criteria of the writing task on the NMET (see Section 2.9.2.2), the students at all three interviews reported that test wiseness strategies were especially used to train them to manage the writing task. This is because these strategies enable them to take advantage of "the characteristics and formats of the test" to attain a high grade (Millman, Bishop & Ebel, 1965, p. 707). To achieve this purpose, teachers asked their students to do the writing tasks in past NMET papers and mock-test papers and then trained them in skills necessary to manage similar tasks. This finding echoes the experience described by Sam in Case Study III and accords with the observation of Kennedy and Lui (2013).

According to the participants, an overarching area accentuated by all teachers during the training was the discrete-point knowledge of English grammar and vocabulary. Most teachers believed these two facets of knowledge were the

cornerstone of effective writing. Another area that was centred on was students'
ability in using complex sentences and "beautiful structures" (You, 2010), also
known as "shining phrases" by the participants (See Section 2.9.2.2 for more
details). Examples of "beautiful structures"/ "shining phrases" and complex
sentences that the teachers expected their students to use in examinations can be
seen in a pre-determined composition template (see next page). It is one of the
five templates the participants highlighted for the researcher to show the use of
these language items. All the templates were from the teachers of the participants.

It is well known to us that the proverb: "_____ (*state the proverb*)" has a
profound significance and value not only in our job but also in our study. It means
_____ (*explain the proverb*). The saying can be **illustrated** through a series of
examples as follows. A case in point is _____ (*give your example*). Another
case is _____ (*give your second example*). Therefore, it goes without saying that
it is of great **importance** to practice the proverb _____ (*restate the proverb*).
With this rapid **development** of _____ (*state the essay topic*), an increasing
number of people come to realise that it is _____ (*link the proverb to* the *essay
topic*). The more we are aware of the significance of this famous saying, the more **benefits**
we will get in our daily job and life.

This is a template for writing expository essays. The italicised parts (in brackets)
were in Chinese in the original template and were translated into English by this
researcher for the purpose of the current study. In this template, there are
altogether eight sentences, of which seven are complex sentences (only the third
one being a simple sentence). These eight sentences contain six complex words
(in bold) and seven "beautiful structures"/"shining phrases" (underlined).

160

Reasons for the prominence of these structures, phrases and words, the participants explained, are two-fold. According to the "insiders", i.e. raters for the NMET, both the content and organisation of the writing tasks on the NMET were pre-determined by the task prompts and therefore most compositions looked similar. A composition with complex words/sentences and "shining phrases" could easily please the eye of raters and was therefore more likely to receive a high grade.

The second reason is psychological. Many raters who were in-service teachers felt a sense of achievement when seeing the complex words, sentence structures and "shining phrases" they taught in class appear in the NMET writing; and they therefore tended to give higher scores for compositions with such structures and phrases.

Since complex words/sentences and "beautiful structures" play such an important role in helping students achieve high grades on the NMET, most students worked diligently on these items. A corollary of this practice was that a considerable number of the students believed that they did not know how to write simple sentences anymore. This echoes the descriptions given by Karrie in Case Study I and Mecky in Case Study IV when they discussed their own writing problems. Of course, one could argue that there is nothing wrong for a composition or an essay to contain complex sentences. The problem is that some learners have gone so far in this direction that their writing becomes cumbersome and laboured. This could cause difficulties in readers' comprehension of the written texts. Below are two unnecessarily long and overly complex sentences extracted from an essay by a mainland learner. The essay is about the balance between press freedom and

161

protection of privacy, and is attached as an appendix (Appendix K) in this dissertation.

1. *Society should attach more importance to the freedom of expressing opinions and comments reasonably by the media rather than the freedom of unveiling the private lives of citizens which is regarded as offensive and impolite.*

2. *Furthermore, exposing the private lives of celebrities to the public for entertainment purposes is actually barely constructive but only damage the reputation of the victims and raise concern about privacy which are significantly threatened.*

The third factor highlighted by most participants was tidy and neat handwriting. This, again, was based on information from the "insiders" who claimed that raters could only spend around 20 seconds on each composition. Under such a time constraint, most raters could not read the details of each composition. Instead, they skimmed through each piece of writing with attention largely paid to the candidate's ability (or lack thereof) in achieving the following points:

- inclusion of main points listed in the prompts;

- use of correct grammar and complex sentences;

- employment of complex lexical items and "shining phrases"; and

- tidy and neat handwriting.

This finding adds support to the information reported by Simon in Case Study II and Sam in Case Study III respectively, and are also in line with the observations of Paltridge (2007), Qi (2007), Wu (2008), and, Xu and Wu (2012). Wu (2008)

points out that these four items have in practice become the four key assessment criteria, albeit unofficial, in assessing candidates' writing on the NMET.

For many candidates, one way to satisfy the above criteria is to imitate sample essays for the NMET. Memorising exemplary texts is part of the learning process for many students in China because students can imitate the structure and/or style of the memorised model essays in their own writing (Singh & Fu, 2008). Because of this, when training students to write for the NMET, some teachers never give feedback on students' compositions. Instead, they provide students with exemplary essays and ask them to memorise these essays (Zhao, 2009). The problem is that sample texts for the NMET, according to the interviewees, came from different sources, ranging from essays in NMET preparation books to student compositions recommended by teachers. Some of these texts were not well-composed. Using these texts as models in one's writing could cause unexpected consequences, such as incorrect lexical use, overuse of sentence connectors and inappropriate tone and style. Nevertheless, most interviewees strongly believed in the value of memorising these exemplary essays because these texts could equip them with (1) a pool of complex words, formulaic expressions and "shining phrases" that were useful in obtaining high grades on the NMET; and (2) a number of composition structural templates they needed for different genres of writing. This finding accords with the experience described by Mecky in Case Study IV and also conforms to the finding in Xu and Wu (2012), who report on a pre-determined template that many students follow when completing the writing task in one of the past NMET papers.

163

To understand the importance of pre-determined templates for NMET candidates, it might be worth examining one such template. Students and teachers often obtain these templates from NMET preparation books. These books are normally written by people who are authoritative in NMET training. Advice in these books is therefore accepted more or less unconditionally by many teachers, students and parents. The one below is adapted from an NMET writing preparation book (Wang, 2013, p. 20). The italicised parts (in brackets) were in Chinese in the original template and were translated into English by this researcher for the purpose of the present project. This template is for argumentative essays.

Nowadays, there is a widespread concern over the issue that _____ (*essay topic*). In fact, there are both advantages and disadvantages in _____ (*the main issue*). Generally speaking, it is widely believed that there are several positive aspects as follows. First (*first advantage*). And secondly_____ (*second advantage*).

Just as a popular saying goes, "Every coin has two sides", _____ (*the main issue*) is no exception, and in another word, it still has negative aspects. To begin with, (*first disadvantage*). In addition, _____ (*second disadvantage*).

To sum up, we should try to bring the advantages of_____ (*the main issue*) into full play, and reduce the disadvantages to the minimum at the same time. In that case, we will definitely make a better use of the _____ (*the main issue*).

There are 110 words in this template and the total word limit for NMET essays is normally 150. The template contains the elements required for top-grade compositions: coverage of main points of the task, use of advanced lexis, complex sentence structures, "beautiful structures" and sentence connectors. It is not difficult to envisage that a candidate who has memorised this template will obtain

164

a high score because all s/he needs to do is to translate the italicised parts into English and some of the translations (e.g. main issue of the topic) may have been given in the test prompts already. This is why memorising pre-determined templates is regarded as an important means of training English writing by many teachers and students in China (Zhao, 2009).

One area that was not mentioned in the case studies, but was remarked on frequently at the interviews, was the dialogic and hortatory features in the writing of the mainland students (see examples of sentences with such features in Section 7.5.2.). These unique features in the writing of Chinese learners of English are reported in Major (2006). When asked about the causes of these features, the interviewees pointed out two main factors: (1) the influence of Chinese writing, which is much shaped by the historical and political development in China, and (2) the prominence of "shining phrases" in the writing for high-stakes national English tests in China.

For historical reasons, in particular as a result of political developments after the establishment of the People's Republic of China in 1949, the public were taught to be positive about social issues. Expressing negative views about the country in public places, including examination papers, could have many unexpected consequences (Xu & Wu, 2012), such as deduction of one's grades or even failure in an examination. A further step along the "being positive" continuum is to be assertive about one's pro-communist political orientation. This, according to the interviewees, makes it necessary to deploy sentences with strong tone and style, such as exclamative, interrogative and even imperative sentences, personal pronouns, such as *I* and *you,* and assertive words, such as *surely, undoubtedly* and

definitely, in one's writing. The use of exclamative and interrogative sentences and personal pronouns makes one's writing dialogic, and the employment of imperative sentences and assertive words/phrases causes a hortatory tone in one's writing. The strong promotion of "beautiful structures" or "shining phrases" by teachers, tutorial schools and high-stakes test preparation books has made the hortatory features in the writing of these learners more pronounced. This is because many of the phrases/structures are very strong in register. For instance, Wang's (2013) NMET preparation book provides 12 examples of "classical" beautiful structures (p. 100). Five of them are very strong in register (see below).

1. *There is no doubt that…* (Example: There is no doubt that our football team will win the game.)

2. *It is … that…* (Example: It is his selfishness that causes pain to the whole family.)

3. *It is high time somebody did …* (Example: It is high time the government introduced education reforms.)

4. *Only in this way can somebody ….* (Example: Only in this way can we resolve the problem properly.)

5. *Something is so… that….* (Example: The weather condition was so undesirable that we had to cancel our planned trip.)

7.3.3. Regional differences in ELT in China

The issue of regional differences in English language teaching (ELT) in China is discussed in Hu (2003; 2005a). This disparity was also observed from the results of the focus-group interviews. Although this topic is beyond the scope of the

166

current study, it is worth noting because of the implications it may have on the findings of the study.

Of the 21 participants, the vast majority were from developed areas, usually metropolitan cities. This is most likely attributed to the English proficiency requirement of the university (an overall mark of 120 out of 150 on the NMET). Few students from underdeveloped areas were able to achieve this level due to contextual factors, such as a lack of qualified teachers as well as a shortage of input and output opportunities, the availability of language learning materials and chances to use the target language (Hu, 2005). In general, students from developed areas enjoyed better language learning contexts in school and those from underdeveloped areas suffered from more contextual influences. Nevertheless, the experiences of the vast majority of the participants were by and large similar in lexical learning and training in writing as discussed in the aforementioned sections.

A girl from Beijing, however, recounted a very different learning context in her school. Her English teacher was well-qualified and had received training in an English speaking country. English was the MOI of all her English classes and the teaching was interactive and communicative. In class, the teacher normally started by playing an audio-recorded passage and then asked students questions about the passage. This was followed by the students reading on the same topic and then writing a summary of the passage as homework. Although this student also received intensive training for the NMET in the last year of her secondary school studies, her weekly writing-class started long before the test, and it was taught by a native speaker of English.

Another two participants, one from Shanghai and the other from Hang Zhou, also reported the existence of a better than average English-learning environment in their schools. Their teachers were well-qualified and overseas-trained. Apart from training in grammar and vocabulary, their teachers encouraged students to listen to BBC (British Broadcasting Corporation) news and watch international television programmes. Sometimes their teachers even asked them to recite in class news items and short passages on the BBC website. Nevertheless, English teaching in their classrooms was far from being interactive and communicative. Much of the teaching was still centred on discrete items of linguistic knowledge of grammar and vocabulary, and the MOI in their English classes was switched completely into Chinese in the last year of their secondary school schooling. The student in Hang Zhou reported that his teacher had to change the MOI under the requests from the students.

On the contrary, a very different account from that of the Beijing girl was given by a student from a small city, Pan Ji (Liao Ning Province). It might be more appropriate to describe his learning context as a military camp rather than a school. Here is how he described his English learning experience.

My classmates and I went to school seven days a week. We started at 7 am and finished …normally by 9 pm every day. That was … all through our six years in the secondary school. We had English classes in the morning and afternoon. The teacher normally taught new content (a new text)… in the afternoon. He first of all went through the vocabulary list, then explained the text … and then new grammar … in Chinese ... After that, we read aloud the text, the whole class together. For homework … we often do [did]

multiple-choice exercises, normally 100 MC questions ... every day. On the second morning, the teacher checked the answers to the MC questions ... and then gave us dictations to see if we have [had] remembered the new words in the lesson.

Writing? No, we had no writing exercises ... until one month before our college entrance examination…

The above cases exemplified the regional differences in ELT in China. The contextual influences caused by this regional disparity could have a serious impact on the students' language ability. Some learners' learning environment in school, such as the one described by the Beijing girl, is as communicative as that found in most Hong Kong schools, if not better. Nevertheless, most participants' learning experiences were shaped by the overall English learning context in China. Inside mainland China, the number of students who are undergoing the experiences depicted by the student from Pan Ji should be rather high due to the unbalanced economic and social developments in China (Hu, 2005a).

The implication of this finding is that a small number of participants from the mainland may not have the problems regarding lexical learning and essay writing that are examined in the present project. However, since the number of such learners is very limited, their impact on the results of the study should also be rather limited. Meanwhile, due to the high number of students who are still experiencing the language learning context similar to that recounted by the student from Pan Ji, the problems in vocabulary learning and essay-writing in the mainland could persist into the future.

7.4. Questionnaire survey

The questionnaire was intended to provide quantitative data about the differences between Hong Kong and mainland learners in their experiences of vocabulary learning and English writing in their secondary schools. Because of this, two slightly different versions were produced for these two groups of participants (67 Hong Kong students and 83 mainland students, see 3.3.2 for more details). The questionnaire contained 50 multiple-choice questions and was composed of two main sections: lexical learning and training in writing. The following section first analyses the results of the respondents' vocabulary learning and then discusses those of their training in writing.

7.4.1. Lexical learning

A number of similarities in the lexical learning of these two groups of participants have been identified. These are mainly related to their learning beliefs. The vast majority of them (90% of Hong Kong participants and 85% of their mainland counterparts) believed that vocabulary was the most important aspect of English learning. When asked about word knowledge, around half of the participants from both groups believed the association between the form of a word and its meaning in Chinese was the most important (see Table 7-5).

Table 7-5. Similarities in student strategies

	Hong Kong	Mainland China
Vocabulary learning	90%	85%
Link between word form and meaning in Chinese	46%	47%

However, when responding to questions about specific lexical learning strategies, the results from these two groups were considerably different. One major difference lies in the number of new words they learnt per week. As evinced by the figures in Table 7-6, only very few Hong Kong students learnt 20 to 30 words each week while nearly half of their mainland counterparts memorised that many words per week. The difference in learning more than 30 words per week was similar between the two groups.

Table 7-6. Student strategies in lexical learning (1)

Number of words learnt per week	Hong Kong	Mainland China
Between 20 and 30	5%	43%
Above 30	3%	17%

This distinct difference in the number of words learnt each week is probably attributed to the methods with which they learnt new words. Some empirical evidence related to this can be found in Table 7-7. According to this table, nearly half of the mainland subjects learnt new words from word lists whilst only a very small percentage of their Hong Kong counterparts used word lists. When asked about types of word lists, a considerable number of mainland respondents (30%) admitted using word lists from high-stakes examination preparation books (such as IELTS and TOEFL), whereas a large majority of Hong Kong students (67%) used their own word lists. It is highly likely that these lists were generated from their contextual language learning because most Hong Kong subjects (78%) acquired their words from language use such as reading and listening.

Table 7-7. Student strategies in lexical learning (2)

Items	Differences	Hong Kong	Mainland China

Sources from which students learnt words	Word lists	15%	42%
	Reading & listening	78%	50%
Types of word lists students used	IELTS and/or TOEFL books	8%	30%
	Students' own lists	67%	24%

The divergence in lexical learning strategies between these two groups could be associated with the difference in which words were taught in school. This could be evinced by the figures in Table 7-8. The table shows that a substantial percentage of mainland teachers (39%) used traditional approaches in lexical teaching. They normally went over the glossary of a text with students before teaching the text. The glossary was taught mainly by asking students to read aloud each word after them (57%) and explaining words in Chinese (53%). On the contrary, very few Hong Kong teachers (8%) felt the need to go over the glossary with their students. If they encountered words that required clarification in class, the majority of them (61%) explained the words in English.

Table 7-8. Teacher strategies in lexical teaching (1)

Teacher Input in class	Hong Kong	Mainland China
Go over glossary with students before teaching a text	8%	39%
Ask students to read each word after the teacher	8%	57%
Explain glossary in Chinese	10%	53%
Explain glossary in English	61%	39%

Apart from glossary teaching, mainland teachers were also closely involved in monitoring their students' progress in their word learning (see Table 7-9). Many of them (51%) performed this role by delivering dictations in class. Some Hong

Kong teachers also checked their students' progress in lexical learning; however, the majority of them (53%) did so by providing an environment in which their students could use the newly-learnt words in speaking and writing (e.g. conducting a class discussion about a topic related to recently-learnt words).

Table 7-9. Teacher strategies in lexical teaching (2)

Teachers' role in monitoring student progress	Hong Kong	Mainland China
Deliver dictations	16%	51%
Provide an environment for students to use newly-learnt words in class	53%	9%

It could be concluded from the above findings that lexical learning in mainland China was regarded as a discrete-point of knowledge accumulation in English learning. The link between the form of a word and its Chinese meaning is the most important aspect of the knowledge. Because of this, words needed to be explained in Chinese by many teachers and learnt in isolation by most students mainly in the form of words lists. Many teachers monitored student progress by delivering frequent dictations. These findings add empirical support to the results from the five case studies and the focus-group interviews, and are in line with the findings in related literature, for example, Qi (2004), Jin and Cortazzi (2006) and You (2010).

On the contrary, and as previously mentioned, the vast majority of Hong Kong teachers felt little need to go over the glossaries in their text books. This is possibly because they believed words should be acquired from contextual language use. This could explain why, when they monitored student progress in lexical learning, they provided an environment for students to use the

newly-learnt words in speaking and writing instead of delivering constant dictations. Because of this, a majority of Hong Kong students (78%) acquired their words by means of reading and listening. These results shed light on the finding that Hong Kong participants performed significantly better on the two vocabulary texts and, in particular, on the WAT which tests several facets of the learners' depth of lexical knowledge (see Section 6.2 for details).

The above different strategies in lexical learning and teaching also exerted a different impact upon the writing competence of the two groups of participants. Many mainland students may have a large vocabulary size but their depth dimension of vocabulary knowledge is still rather limited. This could result in lexical errors of various kinds occurring in their language production, particularly in word choices and morphological forms. The seriousness of this problem is manifested in the short paragraph below. This paragraph is from an essay written by Mecky, one of the participants in the case studies of the current project.

> *As well as the <u>unfair</u> in education, the wealth gap may also lead to social instability. The poor people <u>would</u> <u>resentment</u> the wealthy people and may even <u>abuse</u> <u>violent</u> to protest against the social <u>unfair</u>. According to the research of Clifford and Pau (2001), 96% <u>publics</u> hate the wealthy and <u>myriad</u> feel discontented <u>toward</u> the government. In 2011, a <u>plethora</u> of people <u>protest</u> the low income and high cost violently.*

Mecky's essay was about the impact of an increasing wealth gap in Hong Kong. What she intended to argue was that the growing wealth gap causes much resentment among the poor towards the rich. This resentment also extends to the government and the whole society. To express the discontent with their worsening

174

situation, many poor people stage mass protests which could become violent sometimes. However, this intended argument was not very clearly presented, largely due to the lexical errors in the paragraph (underlined). These errors are mainly in the following three categories (intended words italicised and placed in parentheses).

a. word choices: such as would (*could*), abuse (*use*) and toward (*with*)

b. morphological forms: such as unfair (*unfairness*), resentment (*resent*) and publics (*the public*); and

c. unnecessary use of complex (low frequency) words, such as myriad (*a countless or extremely great number of*) and a plethora of (*a large or excessive amount of*)

The above lexical errors suggest that Mecky may have sufficient lexical knowledge (a large vocabulary size) but her lexical competence, the ability to use lexis appropriately in communication, is rather limited. This insufficiency in lexical competence, however, seemed to have very little impact on her goal of achieving a high score on the NMET. According to what she reported in the case study, she received an overall score of 142 (out of 150) on the NMET and her score for the writing task in the test was 24 (out of 25). This raises questions concerning the validity of the writing task in the NMET, including the extent to which it tests the candidates' competence in written communication.

7.4.2. Training in writing

The questionnaire survey yielded results of both similarities and differences between the two groups of respondents about their training in writing. As stated in

the aforementioned discussion (see Section 7.1 for details), some survey questions were interpreted differently by the two groups of students due to the diversity of their learning contexts. After checking with both groups of students, the study decided to exclude the results from any questions that might be read differently by the two groups.

Table 7-10. Similarities in training of writing

Student beliefs	Hong Kong	Mainland China
Vocabulary and grammar are the most important considerations in writing	50%	48%
Using connectors is the most important strategy in achieving cohesion and coherence	67%	69%

There are two major similarities. The first one was in the importance of vocabulary and grammar in writing. Around half of the respondents from both groups (50% and 48% respectively) believed that these two facets of knowledge were most important in writing (see Table 7-10). The other similar view was found in the use of sentence connectors, such as *however, moreover* and *furthermore*. A large majority of both Hong Kong and mainland students (67% and 69% respectively) considered using connectors as the most important means of achieving cohesion and coherence in writing.

Regarding the differences, the major ones are in the strategies recommended by the teachers of the two groups of subjects. As is evident in Table 7-11, a substantial number of mainland teachers (39%) strongly suggested that their students use proverbs and set phrases, which are also known as "beautiful structures" or "shining phrases" in mainland China, but only a small number of Hong Kong teachers (13%) recommended the use of such language elements in

176

their students' writing. Another major difference was in the teachers' views on long and complex sentences. A large majority of mainland teachers (67%) strongly suggested that their students use such sentences in essays but much fewer Hong Kong teachers (39%) seemed to think it was necessary to provide such an advice for their students.

Table 7-11. Writing strategies recommended by teachers

Strategies recommended by teachers	Hong Kong	Mainland China
Using proverbs and set expressions	13%	39%
Writing long and complex sentences	39%	67%

The teachers' training in writing seemed to have affected the strategies students adopted (see Table 7-12). To achieve a higher grade on the writing task of the NMET, over half of the mainland respondents chose to deploy beautiful structures/shining phrases in their writing. These structures/phrases mainly consist of clichés, proverbs and set phrases which were strongly recommended by their teachers.

Table 7-12. Student strategies in writing

Strategies students used to achieve a higher grade in university entrance examinations	Hong Kong	Mainland China
Enhancing the tidiness of my writing	17%	68%
Using beautiful structures	36%	51%
Memorising sample essays	29%	43%
Reciting exemplary texts	35%	58%

The above findings about beautiful/shining phrases are consistent with the information provided in Wang's (2013) NMET preparation book. To encourage candidates to use clichés, proverbs and set phrases in their writing, the book lists 209 proverbs and clichés, and 43 famous quotes and sayings at the end of the book (pp. 296-309). These items are divided into categories according to themes (10 themes in proverbs/clichés and 24 themes in famous quotes/sayings). However, there is no explanation as to how to use these in contexts and, in particular, what types of writing genres the different proverbs/clichés/quotes might best be used in. This is possibly because many people in China believe that these structures and phrases are a "panacea" (You, 2010) and therefore can be employed in any genre of writing. After all, "it is these beautiful and native-sounding sentences that will make your writing stand out" (p. 154).

What is even more surprising is that some of the proverbs/clichés are not correctly listed in the book, which can be evinced by the two examples below (adopted from p. 298). The italicised parts (in brackets) show the proverbs/clichés that are better known to native speakers of English.

1. One cannot put back the clock (*Life waits for no man, cannot turn back the hands*).

2. Life is but a span (*The life of man is but a span*).

Another interesting finding is that many mainland students memorised sample essays (43%) and exemplary texts (58%) because such essays contain words, phrases and structural templates their teachers expected them to use on the NMET (see Section 7.3.1.2). On the contrary, not many Hong Kong participants appeared to feel the need to take such measures for their UE.

The most distinct discrepancy between the two groups of respondents was their views on the tidiness in writing. A large majority of mainland respondents (68%) took tidiness as a priority in the writing but much fewer Hong Kong students (17%) seemed to feel this way. This finding accords with Mei (2008) who reports that many teachers and students regard maintaining tidiness as the most important strategy in obtaining high marks in the writing component of the NMET. The finding also adds strong empirical evidence to the qualitative data collected from the case studies and the focus-group interviews for the current study.

Table 7-13. Student experience in writing

Student experience in writing	Hong Kong	Mainland China
Writing was only to prepare for exams	32%	52%
Training was largely paragraph writing rather than essay writing	20%	31%

When asked about experiences in writing, findings from the two groups of subjects were also much diversified. As is evident in Table 7-13, more than half of the mainland respondents had little writing experience until they had to prepare for the NMET. Many of those (31%) who had some experience only practised writing at the paragraph level, meaning they seldom had to write a complete essay. The situation in this regard was different with Hong Kong participants. The great majority of them (78%) received training in essay-writing long before attending the UE. Most of them (80%) had to write complete essays in their secondary schools.

The above results indicate that, notwithstanding the similar views of the two groups' beliefs on the roles that vocabulary and grammar played in producing effective writing, and the roles that conjunction devices performed in achieving

179

cohesion and coherence, their experiences in writing were considerably different. Most mainland participants had little training in writing before attending the NMET and the training for the NMET was largely based on test-taking strategies such as using beautiful structures and complex sentences, following structural templates and maintaining tidiness. The more conducive learning context in Hong Kong, on the other hand, made it unnecessary for the students to undergo such training for the UE. Instead, writing has always been part of their English learning experience (Lee, 2012, p. 5). Because of this, these students could have learnt some basic concepts of rhetorical devices and genre features, including style and register of different types of writing.

As to the similar beliefs the two groups shared, it is not difficult to understand the importance that both groups attached to vocabulary and grammar because these two facets of knowledge have been regarded as the foundations of a language (Barani & Seyyedrezaie, 2013; Delmonte, 2008). Lexical knowledge is even considered as a "precondition" for other language skills (Roche & Harrington, 2013, p. 2). It was, however, rather surprising that the percentage of the two groups' views on the role of sentence connectors in writing was virtually the same, given that the mainland participants seemed to consider the using of connectors as the principal device to achieve cohesion and coherence. Although it is well-documented that Hong Kong learners overuse connectors (Field & Yip, 1992; Milton & Tsang, 1993; Bolton, Nelson & Hung, 2002), a deep understanding of this situation requires systematic analysis of the writing by these two groups. Such analysis may also help to shed light on the aforementioned supposition that Hong

Kong participants had gained a better understanding of genre features of different types of writing before entering the university.

7.5. Corpus analysis of participants' writing

This part of the analysis was performed on the learner corpora built from essays written by the two groups of participants. Due to the limited number of participants for the current study, both corpora are relatively small. The number of essays and tokens in each corpus can be found in Table 7-14. To maximise the accuracy and reliability of the data, both corpora underwent strict modification procedures (see Section 3.3.3 for details). The analysis of the corpora was focused on the use of connectors and the dialogic and hortatory features in the two groups' writing.

Table 7-14. Essay and token number in two learner corpora

Corpora	Hong Kong	Mainland China
Total number of essays	67	83
Total number of tokens	33,832	39,577

7.5.1. Use of connectors

As indicated in the aforementioned discussion, Hong Kong learners have been found to overuse and misuse connectors. If mainland learners use connectors even more frequently, the quality of their writing could be seriously questioned because overuse of such cohesive devices is often related to the writers' lack of essay writing experience (Field & Yip, 1992) and is a way of "disguising poor writing" Crewe (1990, p. 321).

According to Halliday and Hasan (1976), connectors can be divided into four main categories: additive, such as *also, moreover* and *furthermore*; adversative, such as *but, however* and *nevertheless*; causal, such as *because, therefore* and *thus;* and temporal, such as *firstly, secondly* and *finally*. Following this classification, the present project computed the ratio of frequency (RF) of conjunctions used in the two groups' writing.

The calculation of the RF was only on connectors used to join sentences. When connectors, such as *and, also, but* and *so,* were employed to link words or phrases, their occurrences were not counted. Because of this, the sentence was the basic unit for analysis of the RF for connectors. To allow for comparison of very low occurrence figures, the frequencies were multiplied by 1,000. This means the figure in each column of the related tables shows the occurrence of the item in every 1,000 sentences. The figure in brackets is the total occurrences of the item in the corpus. Since other analysed items, such as personal pronouns and modal verbs, may occur a number of times in one sentence, word was used as the basic unit for the calculation of their RF. To permit the computation of very low occurrence figures, the RF was calculated at one in every 10,000 words.

Table 7-15 demonstrates the overall use of connectors by the two groups. The table clearly shows that the RF in the writing of the mainland participants was higher. They employed 36 more connectives in every 1,000 sentences than their Hong Kong counterparts. Insofar as the different types of connectors is concerned, it was found the RF derived from the mainland learners was higher in adversative, causal and temporal connectors (14.2, 21.6 and 8.5 respectively), but was slightly lower (4.1) in additive connectors.

Table 7-15. Overall use of sentence connectors in participants' writing

	Mainland China per 1,000 sentences		Hong Kong per 1,000 sentences		Ratio difference (mainland China versus Hong Kong)
Additive	84.4	(153)	92.5	(139)	−4.1
Adversative	92.1	(167)	77.9	(117)	+14.2
Causal	124.1	(225)	102.5	(154)	+21.6
Temporal	52.4	(95)	43.9	(66)	+8.5
Total	353.0	(640)	316.9	(476)	+36.1

Note: Figures in brackets indicate the total number of item occurrences.

When the detailed use of connectors in each category in the two corpora was compared (via the ratio change, see Table 7-16a), it was found that the most overused additive device in the mainland corpus was *and* (22) and the one in the Hong Kong data was *also* (16.4). In the adversative category (see Table 7-16b), the highest RF in the mainland corpus was *but* (21) and the most frequently used connector applied by their Hong Kong counterparts was *however* (11.6). The category that caused the most striking difference in RF was that of causal devices (Table 7-16c). In this category, *therefore* was used much more frequently by Hong Kong participants, with a corresponding difference of +28.5; and the RF of *so* was much higher in the mainland data, with a difference of +43.3 from that of the Hong Kong corpus. In the study by Bolton, Nelson and Hung (2002), *so* was found to be the most overused connector by the Hong Kong learners, 32 more in every 1,000 sentences than the figure derived from academic papers. Now that the mainland learners were found to have used this connector much more frequently than their Hong Kong counterparts, the RF in the mainland data should be much higher than the academic norm.

Table 7-16a. Additive connectors in participants' writing

	Mainland per 1,000 sentences		Hong Kong per 1,000 sentences		Ratio difference (mainland China versus Hong Kong)
and	51.3	(93)	29.3	(44)	+22.0
also	8.2	(15)	22.6	(34)	−16.4
furthermore	2.8	(5)	8.7	(13)	−5.8
moreover	3.9	(7)	17.3	(26)	−13.4
besides	12.7	(23)	12.6	(19)	+0.1
what is more	5.5	(10)	1.9	(3)	+3.6
Total	84.4	(153)	92.5	(139)	−4.1

Note: Figures in brackets indicate the total number of item occurrences.

Table 7-16b. Adversative connectors in participants' writing

	Mainland China per 1,000 sentences		Hong Kong per 1,000 sentences		Ratio difference (mainland China versus Hong Kong)
but	36.9	(67)	15.9	(24)	+21
however	36.3	(64)	47.9	(72)	−11.6
on the one hand	7.2	(13)	1.9	(3)	+5.3
on the other hand	9.9	(18)	5.9	(9)	+4
nevertheless	2.8	(5)	6.0	(9)	−3.2
Total	92.1	(167)	77.9	(117)	+14.2

Note: Figures in brackets indicate the total number of item occurrences.

Apart from a conclusion that the overall RF of conjunction devices used by the mainland learners was high, the above findings do not appear to allow further conclusions to be drawn regarding the use of connectors by these two groups.

However, one further conclusion that can be drawn concerns the use of temporals. The occurrence of all the connectors in this category (except for the use of *finally*) was higher in the mainland corpus than in the Hong Kong one (Table 7-16d). This

is possibly a consequence of the mainland students' employment of structural templates and memorised exemplary essays in their writing (see Section 7.3.1.1). In these templates and essays, temporals often appear as devices used to join paragraphs (see Appendix M).

Table 7-16c. Causal connectors/conjunctions in participants' writing

	Mainland China per 1,000 sentences		Hong Kong per 1,000 sentences		Ratio difference (mainland China versus Hong Kong)
thus	9.4	(17)	12.6	(19)	−3.2
so	55.2	(100)	11.9	(18)	+43.3
therefore	9.4	(17)	37.9	(57)	−28.5
because	50.2	(91)	39.9	(60)	+10.3
Total	124.1	(225)	102.5	(154)	+21.6

Note: Figures in brackets indicate the total number of item occurrences.

Table 7-16d. Temporal connectives/conjunctions in participants' writing

	Mainland China per 1,000 sentences		Hong Kong per 1,000 sentences		Ratio difference (mainland China versus Hong Kong)
first	15.4	(28)	16.6	(25)	+1.2
first of all	8.8	(16)	7.3	(11)	+1.5
firstly	9.9	(18)	7.3	(11)	+2.6
secondly	13.2	(24)	5.3	(8)	+7.9
thirdly	3.9	(7)	2	(3)	+1.9
finally	1.1	(2)	5.3	(8)	−4.2
Total	52.4	(95)	43.9	(66)	+8.5

Note: Figures in brackets indicate the total number of item occurrences.

Another important finding was that an overwhelming number of connectors in the data from both groups were fronted to the sentence initial position. This finding accords with the observations of Field and Yip (1992) who studied the use of

connectors by Hong Kong learners. This situation seemed more severe in the mainland data. Table 7-17 lists connectors that were frequently used both inside sentences and at the initial position by the two groups. The shows that, except for *also*, all other listed connectors including *and, but, however, so* and *thus* were used more frequently at the initial position by the mainland participants. Of these five conjunctions, the most distinct difference was on the use of *so*. The mainland learners used 22.5% more of this word at the initial position than their Hong Kong counterparts. Frequent use of connectors at sentence initial position could be problematic because a large majority of the overuse and misuse of connectors by Hong Kong learners occur at this position (Field & Yip, 1992; Bolton, Nelson & Hung, 2002).

Table 7-17. Connectives at sentence initial position

	Mainland China	At initial position	Hong Kong	At initial position	Changes to mainland data
and	93　(972)	9.6%	44　(908)	4.4%	+5.2%
also	15　(168)	8.9%	34　(140)	24.2%	−15.3%
but	67　(209)	32%	24　(125)	19.2%	+12.8%
however	56　(64)	87.5%	53　(72)	73.6%	+13.9%
so	100　(228)	43.9%	18　(83)	21.7%	+22.2%
thus	17　(22)	72.3%	19　(28)	67.9%	+4.4%

Note: Figures in brackets indicate the total number of item occurrences.

The above discussion indicates that, in general, the mainland participants used sentence connectors much more frequently than their counterparts, and their frequency of conjunctions used at the initial position of sentences is also much higher than that in the writing of their Hong Kong counterparts. This could have caused overuse and/or misuse of conjunctions in their writing since Hong Kong

students themselves have been reported to overuse and misuse this rhetorical device (Bolton, Nelson & Hung, 2002). An example of essays containing such problems and written by a mainland participant can be seen in Appendix L. This essay contains only 18 sentences. Seventeen sentence connectors were placed in the text (all in bold in the appendix). Most of the connectors were inappropriately used.

7.5.2. Dialogic and hortatory features in the writing by the two groups

A lack of awareness of genre features, style and register in the writing by mainland learners was discussed in the focus-group interviews (see Section 7.4.2.). This problem is hardly new, though. Mayor (2006) reported this problem after examining the writing by Chinese candidates for IELTS tests (Task 2, expository writing). She believes dialogic and hortatory features are the two major manifestations of this problem. Given that essays for the current project were also expository in nature, Mayor's framework was followed when comparing the use of these language items in the two corpora.

Dialogic (or conversational) features are caused by writers' inappropriate choices of grammatical forms and lexical items, including employment of exlamative and interrogative sentences, use of first and second person pronouns, such as *I* and *you*, and other informal language use, for example contractions such as *don't, won't* and *it's*. Examples of such sentences can be seen in the next page. These sentences were extracted from essays written by participants of the current project. One essay has been placed as an appendix (Appendix N). The essay contains a number of sentences with dialogic features (see below).

1. *Excuse my impolite comment, I think it is nothing serious to experience sex in early age just when it is in a safe and right way.*

2. *Just tell yourself that you are the asset of your family and don't worry too much about other's words.*

3. *Let's attempt to be someone with wisdom and enthusiasm!!!*

4. *So what can we, a small little person in the big society do, to help the government achieve this goal?*

The comparative corpus analysis revealed that the mainland participants employed 16 more exclamative sentences, 27 more interrogative sentences (Table 7-18a) and 61 more contractions (Table 7-18b) in every 1,000 sentences than their Hong Kong counterparts. The disparity of RF between the two groups in their use of personal pronouns was even wider (Table 7-18c). More than 88 personal pronouns appeared in the mainland corpus in every 10,000 words than in the Hong Kong data. What is also worth noting is that the RF for every single personal pronoun in the mainland data is higher, with the greatest difference of +42 (the use of first person pronoun, *I*) and the second greatest difference of +18 (use of second person pronoun, *you*) from the Hong Kong corpus. These learners' frequent employment of *I* and *you* would have greatly increased the conversational tone of their writing.

Table 7-18a. Exclamative and interrogative sentences in participants' writing

Sentence types	Mainland China per 1,000 sentences	Hong Kong per 1,000 sentences	Ratio difference (mainland China versus Hong Kong)
exclamative	17.7 (32)	2.0 (3)	+ 15.7
interrogative	45.8 (83)	18.6 (28)	+ 27.2

Note: Figures in brackets indicate the total number of item occurrences.

Table 7-18b. Contractions in participants' writing

	Mainland per 1,000 sentences	Hong Kong per 1,000 sentences	Ratio difference (mainland China versus Hong Kong)
didn't	3.3 (6)	0	+3.3
don't	22.6 (41)	8.7 (13)	+13.9
can't	22.1 (40)	0.6 (1)	+21.5
it's	23.1 (42)	0.6 (1)	+22.5
Total	71.1 (129)	9.9 (15)	+61.2

Note: Figures in brackets indicate the total number of item occurrences.

Table 7-18c. Personal pronouns in participants' writing

	Mainland China per 10,000 words	Hong Kong per 10,000 words	Ratio difference (mainland China versus Hong Kong)
I	56 (220)	14 (46)	+42
Me	6 (22)	1 (5)	+7
My	22 (88)	4 (13)	+16
mine	1 (2)	0 (0)	+1
you	29 (114)	11 (38)	+18
your	6 (24)	2 (8)	+4
Total	120 (470)	32 (97)	+88

Note: Figures in brackets indicate the total number of item occurrences.

Hortatory features are caused by an overly assertive tone in one's writing. The most salient manifestation of these features is a frequent use of imperative

sentences but a lack of tentative formulations, such as model verbs (e.g. *might, could* and *would*) in one's writing. Examples of sentences with hortatory features are as follows. An essay containing such sentences can be found in Appendix N. The essay was written by one of the project participants from the mainland.

1. *Don't let the outcomes of our knowledge destroy our wisdom.*

2. *Do not object to try it just because of some negative aspects it might lead to.*

3. *Take actions right now.*

4. *Just remember, as a saying goes "as you sow, so shall you reap".*

The disparities in the use of these two linguistic devices are presented in Table 7-19a and Table 7-19b. The figures in these two tables demonstrate clearly that the mainland learners used seven more imperative sentences in every 1,000 sentences but 22 fewer modal verbs in every 10,000 words than their Hong Kong counterparts. This means that the RF of sentences with an overly assertive tone was much higher, but that with a tentative tone was much lower in the mainland data.

Table 7-19a. Imperative sentences in participants' writing

hortatory features	Mainland China per 1,000 sentences	Hong Kong per 1,000 sentences	Ratio difference (mainland China vs Hong Kong)
imperative	11.6　(21)	4.7　(7)	+ 6.9

Note: Figures in brackets indicate the total number of item occurrences.

Table 7-19b. Modal verbs in participants' writing

Model verbs	Mainland China per 10,000 words	Hong Kong per 10,000 words	Ratio difference (mainland China vs Hong Kong)
could	7 (28)	7 (28)	+/−0
would	12 (48)	21 (71)	−9
may	32 (130)	44 (148)	−11
might	3 (10)	5 (17)	−3
Total	54 (216)	77 (259)	−22

Note: Figures in brackets indicate the total number of item occurrences.

7.5.3. Set-phrases, idioms and clichés

As discussed in the aforementioned sections, set-phrases, idioms and clichés are known as "beautiful structures" or "shining phrases" amongst mainland students and are strongly recommended for use in essay writing by their teachers (see Section 7.4.2). Empirical support for the results of teachers' recommendations can be evinced by the figures in Table 7-20. As is evident in the table, the mainland learners used 37 more such phrases in every 1,000 sentences than their Hong Kong counterparts. The tone and style of many of these phrases, such as, *as a saying goes*, are informal and can cause inappropriate register in academic writing.

Table 7-20. Set phrases/idioms/clichés in participants' writing

	Mainland China per 1,000 sentences	Hong Kong per 1,000 sentences	Ratio difference (mainland China vs Hong Kong)
*(1) every coin has two sides (2) every sword has two edges	15.4 (28)	1.3 (2)	+11.1

as the saying goes	2.8 (5)	0	+2.8
in a word	7.2 (13)	0	+7.2
it is well-known	4.9 (9)	2.6 (4)	+2.3
last but not least	6.6 (12)	4.7 (7)	+1.9
as far as… is concerned	3.8 (7)	0	+3.8
Total	45.2 (82)	8.7 (13)	+36.5

Note: 1. Figures in brackets indicate the total number of item occurrences.

* These sentences were directly collected from the learner corpus. Although containing similar meanings, these two sentences are more likely to be written in a slightly different way by a native speaker of English: (1). There are two sides to every coin.

(2). It (or something) is a two-edged sword.

7.5.4. Average word and sentence length

Both the quantitative and qualitative data from aforementioned sections including the case studies, focus-group interviews and questionnaire survey also suggest that mainland students employed overly long sentences and unnecessarily complex lexis. A corpus analysis was therefore performed to compare the average word length and average sentence length in the two corpora. The results, however, could not add further empirical evidence to this observation, as both the average word and sentence length were very similar in the corpora (see Table 7-21).

Table 7-21. Word and sentence length in two learner corpora

Corpus	Mainland China	Hong Kong
average word length	4.60	4.87
average sentence length	17.13	17.12

This result seems rather surprising considering mainland learners' employment of unnecessary long sentences and overly complex words appear to be a serious

problem in accordance with both the aforementioned data from the current project and related literature (e.g. Qi, 2004; Paltridge, 2007; Wu, 2008). Further evidence of the existence of this problem can also be found in Wang's (2013) NMET preparation book. This book particularly lists the official band-descriptors for the highest level of compositions (Band 5) on the NMET. The purpose is to help candidates understand the factors contributing to top-grade essays. Four criteria were listed on this band in Chinese and they were translated into English by this researcher for the purpose of the current project (see below).

1. *Addresses all the main points of the task*

2. *Uses a relatively wider range of sentence structures and vocabulary*

3. *May contain errors in grammar and vocabulary but the errors were caused by candidates' effort in using complex sentences and advanced/infrequent lexis; and*

4. *Effectively uses sentence connectors.*

Of these four marking criteria, two of them directly address the issue of sentence structures and vocabulary (Points 2 and 3). Point 2 is straightforward and easy to understand. However, few students may understand Point 3 due to the implicit message it contains. The author of the book particularly interpreted this criterion so as to draw readers' attention to the importance of using complex sentence structures and advanced lexis (i.e. low frequency words) on the NMET. The interpretation of the author possibly represents a large majority of teachers' views in this connection considering the data the current project has collected from the case studies, focus-group interviews and questionnaire survey. Below is the English translation of the author's interpretation (by the researcher). The original

193

sentences are in Chinese and appear on Pages 17 and 18 in the book.

> *This is to say, candidates who merely use simple words and sentence structures cannot demonstrate their language abilities, and therefore are less likely to attain high scores even if their compositions are error-free. On the contrary, those who purposely deploy complex sentence structures and advanced lexis are considered stronger candidates with a higher command of the English language. These candidates will be rewarded with high scores on the NMET.*

After receiving instructions such as this, many candidates "deliberately employed unnecessarily complex sentences and low-frequency lexis in their NMET compositions" (Zhao, 2009, p. 24). It is highly likely that many learners will also use unnecessarily long and difficult sentences/lexis in English writing for other purposes because to them such sentences/words are closely associated with high quality of writing.

This being the case, the only possible explanation for the similar average word and sentence lengths in the writing of the two sample groups is the higher number of dialogic and hortatory sentences used in mainland students' writing. Sentences with informal register, such as dialogic and hortatory ones, are normally short and contain non-complex lexis. Frequent employment of such sentences could have evened out the sentence and word lengths caused by unnecessarily long sentences and overly complex words in the mainland data.

7.5.5. Major lexical problems

It is widely acknowledged that lexical errors fall mainly into two categories, inappropriate word choices and morphological forms. Of the two types, the

former causes more communication problems than the latter (Ellis, 2008a). This is particularly evident in the writing of the mainland learners. Apart from general lexical errors in word choices as indicated in the discussion about Mecky's writing (see Section 7.4.1), the problems in the areas of connotation and collocation seem especially apparent. The connotation of a word refers to the implied or suggested meaning of the word in addition to its explicit meaning. The implied/suggested meaning can be extracted and acquired by learners if the word is learnt in contextual language use. Since the large majority of mainland learners tend to learn words in isolation by memorising word lists, in which the semantic meaning of a word is often given in Chinese translations, many learners fail to fully understand the connotations of some words. This could have led to lexical errors in the two sentences below (All sentences are from essays written by the participants in the current study. Intended words are italicised and in parentheses). These two sentences seem to be closely associated with contextual influences, including social and political ones, involved in language learning in China.

1. Propaganda (*education*) is one of the most important ways to help people avoid getting cancer.

2. To further protect our environment, the government should increase the scale of *propaganda* (*education in environmental protection*)

3. What wealth can bring to us is not only material enjoyment but also spiritual (*psychological*) satisfaction.

4. Many parents pay much attention to material provision instead of spiritual nourishment (*mental* or *psychological support*).

In mainland China, propaganda is one of the major forms of government communication aimed at the public. Learning the word *propaganda* in such a political context, many learners fail to extract the negative connotation of the word, i.e. the communication is often one-sided and politically motivated with the intention to influence people's opinions (*Cambridge Advanced Learner's Dictionary*, 2013). Relying largely on the Chinese translations of this word, many learners even think the connotations associated with the word are positive. This can be seen in Sentences 1 and 2. In addition, many of the propaganda messages promote the developing and/or raising of the "spirit" of the society, which is a political agenda of the government. This may have also caused these learners' lack of awareness of the connotation of *spiritual*, which is often associated with religious beliefs. (*Cambridge Advanced Learner's Dictionary*, 2013). This unawareness is very likely the cause of the error in Sentences 3 and 4.

Another type of lexical error which appears in many EFL learners' language production, and is also very apparent in the writing of learners from mainland China, is related to collocation. This problem is particularly evident in semantic prosody, a term used to describe a word's semantic preference to "co-occur with items that can be described as bad, unfavourable or unpleasant, or as good, favourable or pleasant" (Partington, 2004, p. 149). For example, the word *undergo*, as discussed in Stubbs (2001), often collocates with nouns related to unpleasant experiences that most people are forced into but would rather not endure, such as medical treatments, tests and dramatic changes in a person's life (e.g. a war, a divorce or an accident). Largely relying on word lists in lexical

learning, most mainland learners are unaware of the collocation preferences of words like *undergo*. This is very likely the cause of the lexical errors in Sentences 3 and 4 below (intended words italicised and in parentheses). All four sentences are from essays written by the mainland participants in the current study.

1. The government should adopt strict regulations to improve (*reduce* or *control*) air pollution in Hong Kong.

2. The government is working hard to improve (*eliminate*) the gap between the rich and the poor, which is a very popular social initiative in developed countries.

3. This new policy has caused (*produced* or *contributed to*) the improvement of the city's financial situation.

4. Generally speaking, human cloning can cause (*provide*) great help with medical science and improve family relationships.

The above analysis focuses only on the lexical errors in the writing of the mainland learners. This is, however, not to suggest that there is no lexical error in the language production of Hong Kong learners. Lexical errors, sometimes rather severe ones, for example, those caused by a lack of awareness of the semantic prosody of some words as discussed above, do appear in Hong Kong learners writing. Nevertheless, errors related to word connotation, as in the use of *propaganda* and *spiritual*, rarely occur in their writing. More importantly, the lexical errors in the language production of Hong Kong learners seem not to obstruct their communication as much as those made by mainland learners. To provide empirical evidence for this claim, however, requires a detailed

197

comparison between the scales of obstruction of lexical errors in the writing by these two groups of learners. This statistical comparison is beyond the scope of the current study and therefore requires another large scale project to achieve such an outcome of comparison.

7.6. Summary

The discussion in this chapter suggests clearly that the mainland learners used many more sentences with dialogic and hortatory features. These sentences caused conversational and often also overly assertive tone, which is inappropriate for expository writing. Adoption of such a tone indicates the writers' low awareness of genre conventions and is "counter-productive" to the quality of their writing (Thompson, 2001, p. 74). When the dialogic and hortatory sentences were mixed with complex lexis, long sentences and so-called "beautiful phrases", a cumbersome and laboured writing style and, more seriously, an inappropriate and/or even confusing register, were likely to become manifest. Another problem lay in their use of cohesion and cohesive devices. They deployed conjunction devices much more frequently than their Hong Kong counterparts notwithstanding the fact that Hong Kong learners themselves overuse such connectors (Field & Yip, 1992, Milton & Tsang, 1993; Bolton, Nelson & Hung, 2002). Their overuse of connectors could have generated more problems in cohesion and coherence than their Hong Kong counterparts because such use of conjunction devices "at best clutters up the text unnecessarily, and at worst causes the thread of the argument to zigzag about" (Crewe, 1990, p. 321). The situation could also have been compounded by a third factor - the learners' inappropriate choices of lexis. Most mainland students only attended to the association between

the form of a word and its meaning in Chinese, but largely ignored other facets of lexical knowledge listed in Nation (2001), such as collocations, syntactical and semantic features and constraints in use (see Section 7.4.1). This problem could have affected the readability of their writing because lexical items are the main carriers of the message in writing.

The above analysis has lent explanation to the deviant language development patterns identified in the mainland students which were manifested in the following two research results:

- There was a low level of association between the depth dimension of mainland learners' lexical knowledge and their writing skills (Chapter 5); and

- There was negative correlation between the lexical variation in their writing and their essay grades (Chapter 6).

The analysis indicates that the deviant language development was caused by a lack of a conducive English-learning environment in China. This disadvantageous learning environment shaped the approach through which students learnt English. The consequence of such a learning approach was that many of the learners were able to employ a variety of lexis in their writing. However, their inappropriate choices of lexical items, their lack of cohesion and coherence, their ponderous writing style and their inappropriate and/or even confusing register, have had a negative impact upon the quality of their writing.

This chapter has hitherto provided both quantitative and qualitative support to the supposition made in Section 6.3.3: that the deviant language development identified in the mainland students' writing was caused by their English-learning

approach, and in particular, the way words were built into their mental lexicon and the way they studied English writing. Considering the close association between EFL learners' learning approaches and their language learning contexts (Ellis, 2008b; Schmitt, 2008), it can be concluded that the identified deviation is a consequence of the previously mentioned unconducive English-learning context these learners experienced before entering university in Hong Kong. Their Hong Kong counterparts, on the other hand, had fewer such problems due to their more favourable English language learning environment.

Before moving to Chapter 8, the concluding chapter of the dissertation, it might be useful to provide an overview of the data which was excluded when conclusions for the three research questions of the study were drawn. This data was discussed in Chapters 5 and 7 respectively and is summarised in Table 7-22.

Table 7-22. Data not used for drawing conclusions of the three research questions

Chapter 5	Regression Tests	Table Number
	Regression tests with VS and DVK as independent variables (full sample, N=150)	5-7a (p. 104)
		5-7b (p. 105)
	Regression tests with VS and DVK as independent variables (mainland sample, N=87)	5-9a (p. 107)
		5-9a (p. 107)
Chapter 7	Survey questions	Question Number
	Cohesion and coherence in writing (mainland data)	B3 (p. 246)
	Amount of training in writing (mainland data)	C1 (p. 247)
	Frequency of training in writing (mainland data)	C3 (p. 247)

The data in Chapter 5 were mainly from the regression analysis for examining the power of the DVK in predicting the mainland participants' writing proficiency. Table 5-7a and Table 5-7b were for the full sample whereas Table 5-9a and Table

5-9b were for the mainland sample only. See Sections 5.4.1 and 5.4.2 for more details.

The data in Chapter 7 included three survey questions. All three questions were related to the writing experience of the mainland participants, the amount and frequency of the training they received (Questions C1 & C3) and their training on cohesion and coherence of essay writing (Question B3). The exclusion of B3 was attributed to the misunderstanding of many mainland participants as to this writing skill. To them, enhancing cohesion and coherence in one's writing simply means using more sentence connectors such as *moreover, however* and *therefore*. The reason for excluding C1 and C3 are detailed in Section 7.1. The questionnaire containing the three questions can be seen in Appendix H.

CHAPTER EIGHT

CONCLUSIONS, IMPLICATIONS AND LIMITATIONS

8.1. Overview

This chapter provides a brief summary of the major findings from the present project. In accordance with the findings, implications the project may have for future studies as well as pedagogical practices of EFL teaching are set out. Finally, the limitations of the project are also discussed.

8.2. Summary of major findings from the present project

8.2.1. Vocabulary knowledge as an indicator of EFL learners' writing proficiency

Before presenting the summary, it is necessary to restate a major finding of the present project: the language development of many mainland participants has deviated from the language development pattern that researchers have previously identified in EFL learners (see Chapter 5 and 6). This unusual development made it necessary for the project to exclude data from these learners when conducting analyses for Research Questions 1 and 2. This means that the computation was based on data from Hong Kong learners only. In accordance with this, the main findings for the current study are summarised as below:

1. Both the breadth and depth dimensions of EFL learners' vocabulary knowledge are indicators of their writing proficiency. This can be evidenced by the positive correlation between the participants' two facets of vocabulary knowledge and their writing ability. Another important empirical

support to this conclusion is the overall figure of 25% predicting power of the writing scores by the VS and the DVK.

2. The depth dimension of vocabulary knowledge is a better indicator of the writing proficiency of EFL learners. This is evinced by the distinctive predicting power of the DVK. The DVK could provide a unique prediction of the writing scores (11%), over and above the contribution made by the breadth of lexical knowledge. Meanwhile, the VS could only add 1% of the variance already accounted for by the learners' depth of lexical knowledge, a figure which is statistically insignificant.

3. Different language learning contexts affect the association between EFL learners' depth of vocabulary knowledge and their writing proficiency. Empirical support in this regard can be found in the results from the vocabulary tests and the measures of lexical richness in the learners' writing, i.e. Lexical Diversity, Lexical Sophistication and Lexical Frequency Profile.

Results from the vocabulary tests demonstrate that, due to different language learning contexts the two groups came from, the association between the DVK and the EW of the Hong Kong participants was moderate; whilst the association level of these two items found in the writing of their mainland counterparts was weak. More importantly, in the Hong Kong data, the DVK added a unique 11% of the variance already explained by the VS, whereas in the mainland data, the DVK added very little (only 1%) to the variance already afforded the VS.

The differences between the two groups are also evinced by the figures from the measures of lexical richness. In the analysis of the Lexical Frequency Profile it

was found that, except for the first 1,000 word frequency level, Hong Kong students employed significantly more words than their mainland counterparts at the other three levels, indicating that, compared with their Hong Kong counterparts, the mainland participants were more comfortable in using high frequency words but less confident in using low frequency or more sophisticated words. This observation was confirmed from the lexical sophistication analysis, which revealed that the mainland subjects employed significantly more words from the most-frequent-2,000-word category than their Hong Kong counterparts but much fewer infrequent or advanced words than the Hong Kong participants. In the lexical variation analysis, the TTR derived from essays by the Hong Kong participants correlated significantly with their essay scores. Such a correlation, however, could not be found in the essays by the mainland students. Instead, a negative correlation, albeit insignificant, was found between the lexical variation and their essay grades. This negative correlation emerged again when the lexical diversity *D,* an alternative parameter to measure lexical variation*,* was calculated. The negative correlation pointed to a deviation in the language development of the mainland learners.

8.2.2. Impacts of the unconducive language learning context in China

The English learning context in China is unconducive to learners' language development (see Section 2.9). Due to a lack of input and output opportunities, the main purpose for most learners to learn English is to attain a high score on the high-stakes English tests in the country. In order to achieve this purpose, teachers teach to the test and students learn for the test. The corollary of this practice can be seen from the five "portraits" which emerged from the case studies and the

"landscape pictures" formed from the focus-group interviews, questionnaire survey and the analysis of the learner corpus built from the writing by the two groups of participants.

These portraits and landscape pictures mainly illustrated two categories of problems associated with the mainland learners: lexical learning and training in essay writing. Lexical learning in mainland China was regarded as accumulation of discrete-point knowledge. The association between the word form and its meaning in Chinese was the most important facet of the knowledge gained. As a result, words needed to be drilled and explained in Chinese by many teachers in class and learnt in isolation by most students mainly in the form of word lists. Many teachers ascertained student progress by delivering frequent dictations on newly learnt words.

Regarding essay writing, the main focus was on learners' test-taking strategies rather than their written communication skills. Due to the test requirements, test prompts and assessment criteria, in particular the implicit ones of the NMET, many students used unnecessarily long and complex lexis and sentences, causing a cumbersome and laboured writing style. In many learners' writing, these items were mixed with sentences with dialogic and hortatory features, which generated an inappropriate and/or even confusing register. This problem in register was often compounded by two other factors: overuse and misuse of sentence connectors and inappropriate use of lexical items. The former could lead to problems in cohesion and coherence, and the latter in readability. These three factors explain the negative association between the lexical diversity derived from

the writing of the mainland students and their writing proficiency, and this is a strong indicator of the irregular language development of the learners.

8.3. Implications of the present project

The major implications of this study are twofold, and relate to future research directions and EFL pedagogical practices.

8.3.1. Implications for future research

8.3.1.1. Depth dimension of vocabulary knowledge as a predictor for EFL learners' speaking proficiency

Studies have been hitherto undertaken to identify the role of depth of vocabulary knowledge in predicting EFL learners' reading (Qian, 1999), listening (Stæhr, 2009) and writing (current project). These studies found that the predicting power of the depth dimension of vocabulary knowledge is much higher in reading (71%) and listening (51%) than in writing (25%). This could mean that the depth of lexical vocabulary knowledge plays a more important role in predicting receptive skills (reading and listening) than productive skills (speaking and writing). Future studies could test this assumption by identifying the predicting power of the depth of lexical vocabulary knowledge in speaking. Reliable data in this connection should help researchers to gain a deeper and more profound insight into the role of depth of lexical knowledge as an indicator of EFL learners' language proficiency.

Although a study by Koizumi and In'nami (2013) has made an initial attempt to address this issue, two factors in that study make it difficult to compare their findings with those of Qian (1999; 2002), Qian and Schedl (2004) and Stæhr

206

(2009). The first factor is the difference in the employment of instruments, as neither Nation's Vocabulary Levels Tests (2001) nor Read's Word Associates Test (1998) was used in the Koizumi and In'nami study. The second one is the conflicting results evident in the two parts of their research, caused largely by the fact that the first part focused on learners at the novice level, whilst the second on those at the intermediate level.

8.3.1.2. Role of the 3,000 word families in predicting EFL learners' writing proficiency

In the study by Qian (1999), the most frequent 3,000 word families were set as the threshold for assessing reading comprehension. In several other studies, the 3,000 word families also performed an important role in measuring EFL learners' language ability. For instance, Schmitt *et al* (2003) deployed the Vocabulary Levels Test at the 3,000 and 5,000 frequency levels to investigate the formulaic language used by the EFL students in the University of Nottingham. Nevertheless, the 3,000 word families played a limited role in predicting the writing scores of the participants in the present project (see detailed discussions in Section 5.6).

This limited role could be attributed to either the students' high cognitive level or the genre requirement of the writing. Given that participants for the present project were all relatively advanced learners (university students) and the writing is expository in nature, future studies could test whether the 3,000 word families have a stronger predicting power in the writing:

1. of lower level students, for example, at primary and/or secondary school levels; and

2. of different genre type, for example, descriptive or narrative writing.

207

Another indication is that further research into the predicting power of vocabulary knowledge on academic writing written by tertiary students should possibly use word families at the 4,000 frequency level or above. Since the predicting power of the 3,000 word families is rather limited, words at lower frequency levels should be deployed. Testing words at some frequency levels (e.g. the 4,000, the 6,000 and the 7,000) was not possible in the previous studies due to the unavailability of the *Vocabulary Size Test* (Nation & Beglar, 2007; Beglar, 2010), which consists of 14 levels and allows researchers to measure words at each of the first 14,000 frequency levels. Now that this test is available and has been increasingly used as a research tool (Nation, 2012), measuring word families at any of the 14,000 frequency levels is viable. Another reason is that there is an increasing agreement in the literature that the 5,000 word families constitute the minimum threshold required to undertake studies in a university where the MOI is English (Roche & Harrington, 2013). Applying this threshold, the 3,000 word families are far below the required level.

8.3.1.3. Validity of the NMET

The data gathered from the current study has pointed to a validity issue in relation to the NMET, i.e. to what extent this high-stakes examination measures what it intends to measure. The NMET is an English proficiency test used for admission decisions for university entrance in China (Cheng, 2008). However, both the qualitative and quantitative data from the current project seem to suggest that the test, to some extent, measures candidates' test-taking skills and even memorising abilities. This is particularly evident from the following four aspects:

1. Prominence of sample essays and pre-organised writing templates for the NMET

   Responses from the survey and case studies suggest that a candidate who can memorise many sample essays and strategically use the writing templates stands a greater chance of achieving a high score in the writing task of the NMET, regardless of his/her writing competence in English. Many learners, such as Sam and his classmates in Case Study III, have attained high scores in the writing task but writing in English for real communication purpose remains far beyond their ability.

2. Importance of neat and tidy handwriting in essays written for the NMET

   According to the official NMET writing construct, "main idea, coherence, grammar and vocabulary, writing purpose, authorship and readership" should be raters' foci in marking (Mei & Cheng, 2014, p. 180), but many of these aspects have been given little attention by the raters. Instead, handwriting, an item that is not part of the intended writing construct, has become a major consideration for most raters of the NMET (see Section 7.1).

3. Raters' own interpretation of the official marking criteria

   As indicated in the aforementioned discussions, cohesion and coherence are important considerations in the marking criteria of the NMET. The cohesion and coherence of an essay is achieved when there is a smooth flow of ideas within and between paragraphs (Fowler & Aaron, 2007). These criteria, however, were interpreted by most raters in terms of the inclusion of - or a lack of, by contradistinction - a large number of

209

conjunctions, such as *moreover, furthermore* and *besides,* in an essay (see Sections 7.2, 7.3 & 7.4). This misunderstanding has led to the overuse of such cohesive devices in the writing of many mainland learners (see Section 7.5.1).

4. Overemphasis on accuracy at the expense of communication skills

The NMET test developers intended to measure candidates' written communication ability (Qi, 2004; Qi, 2007, Mei & Cheng, 2014), but data from the current study strongly indicates that most raters paid more attention to discrete items in grammar and lexical use (see Sections 7.2, 7.3 & 7.4). This practice has distorted the intention of the text developers and thereby reduced the validity of the test.

The above points suggest that there exist two sets of marking criteria in the scoring process of the NMET, the *de jure* criteria, the intended ones by the test developers, and the *de facto* criteria, the enacted ones by many of the NMET raters. These two sets of inconsistent criteria help explain the "polluted scores" reported in Cheng and Curtis (2010, p. 270). They also help shed light on cases such as that of Mecky in Case Study IV, who managed to achieve an almost full score (24 out of 25) in the writing task of the NMET though her writing proficiency, particularly her lexical competence, remains very low (see Sections 7.2.3 & 7.4.1).

The inconsistency between the intended and enacted marking criteria could also, at least partially, help resolve questions derived from a statistical analysis conducted for the current study. The results of the analysis were not reported in the main study because they are beyond the scope of the current project. However,

they have significant implications for future studies. The analysis indicates that the NMET scores had no correlation with either the scores of the vocabulary tests, both the DVK and the VS, or the essay scores. When the same analysis was performed with the HKALE scores, a very different yield was produced: the scores correlated with those of all the tested items, the DVK ($r = .67$, $p < .001$), the VS ($r = .61$, $p < .001$) and the EW ($r = .46$, $p < .001$). This suggests that the NMET scores cannot serve as a reliable indicator of the candidates' English language ability whilst the HKALE scores can (see Table 8-1).

Table 8-1. Correlation between NMET/HKALE and quality of essay writing

		DVK	VS	EW
NMET	Pearson Correlation	.114	.215	.074
	N	83	83	83
HKALE	Pearson Correlation	.669**	.609**	.461**
	N	67	67	67

**Significant at .001

The above analysis has called into question whether the NMET truly tests what it intends to measure. Nevertheless, since the validity issue of the NMET is beyond the scope of the current study, more systematic studies on this topic are necessary.

8.3.1.4. Impact of change in language learning contexts

Data from the current project, both qualitative and quantitative, have also pointed to a problem in the language development of the mainland learners. After changing to a more conducive language environment, these learners might progress quickly since most of them already had a large vocabulary size and a high command of English grammar. Future studies could investigate the progress of these learners by identifying how fast they advance in their language

proficiency and which skill(s), (e.g. reading, listening, speaking or writing), improves more quickly. The studies could also investigate what language problems are still persistent notwithstanding the improvement in their language learning environment. For instance, Mecky and Abby (the two students who were studied in Case Studies IV & V in the current project) reverted back to their habit of using overly difficult words and unnecessarily long sentences in their writing only a few months after undergoing extensive training on avoidance of using such language forms in writing.

8.4. Implications for pedagogical practices in EFL

8.4.1. Education reforms in China

The disadvantageous English learning context in mainland China has had a severe impact on the pedagogical practices in the country, which in turn has negatively affected the communication skills of the learners as manifested in the findings of the current study. The most serious impact is perhaps caused by the disparity between the intended and enacted curricula. While the official or *de jure* curriculum aims at developing learners' communicative abilities, the enacted or *de facto* curriculum focuses on strategies for students to obtain high scores at high-stakes tests. A corollary of this discrepancy is that the implementation of education reforms "seldom touches the classroom ground" (Gu, 2014, p. 298) and the *de jure* curriculum has therefore become by and large a façade for the education system. A more serious consequence of this discrepancy is that many learners from mainland China manage to achieve high scores in international high-stakes examinations, such as IELTS and TOEFL, but experience serious

anxieties when encountering demands to use English communicatively in their overseas studies (Edward, Ran, & Li, 2007).

The most straightforward method to tackle the problem is perhaps to reinforce CLT/TBLT in instructional practices. However, research suggests that top-down implementation of the policies in China has failed to achieve the original intentions of CLT/TBLT, in particular in under-developed areas. Instead of imposing more top-down policies to ensure teachers teach communicatively, the authorities might need to examine the contextual factors discussed in Butler (2011), in particular the examination system. This is to say that making the official curriculum more congruent with the *de facto* curriculum in China would require fundamental reform of the examination system. This change should take place together with other necessary pedagogical reforms, for example, teacher training and the status of teachers, which is currently closely connected to the examination scores of their students in China (Cheng, 2008). Pedagogical research meanwhile could focus on helping teachers to adapt communicative language instruction in their pedagogical practices under the current contextual influences in the country.

The ultimate goal of the education reform is perhaps to lower the selection function of the NMET and other high-stakes examinations in the country. Along with this change, the dual roles of the current assessment system, role of admission decisions (i.e. the selection function) and that of promoting the official or *de jure* curriculum, could be more balanced. In this connection, the examination authorities and the test developers in China could draw inspiration from the school-based assessment initiated in Hong Kong. This initiative takes

into account teachers' classroom assessments in their students' final scores in the university admission procedure and thereby reduces the selection role of the final university entrance English test. Since the system promotes "assessment *for* learning" instead of "assessment *of* learning" (Cheng & Curtis, 2010, p. 269), the second role of high-stakes examinations, the role of promoting the intended curriculum, could be enhanced. The balance of the two roles of high-stakes examinations could then help researchers and educationalists design curricula that better promote students' communicative competence in language learning. So long as the prominent function of admission decisions in the high-stakes examinations remain unchanged, no matter what vision, goals and even directions the educationalists and education authorities may have, they could be easily "distorted" before landing on the classroom ground in China (Gu, 2014, p. 298).

8.4.2. Marking criteria and rater-training for high-stakes English tests in China

Lowering the dominant role of selection of high-stakes examinations, such as the NMET, the CET and the GSEEE, may not be achieved in China in a short time. Making such a change is to move "from an exam culture to a learning culture" (Hamp-Lyons, 2006, p. 487), which could be difficult to implement due to the historical and social context in the country (Cheng, 2010). However, immediate measures could be taken to address problems in marking criteria and rater training for these examinations. As reported by the respondents of the current study, raters were only allowed about 20 seconds for a composition on the NMET. This makes it extremely difficult, if possible at all, for the raters to carefully read and understand the language and structure of each composition. Being constrained by

such a serious time limit, many raters merely looked for cues that were indicators of effective essays to them, mainly including:

- neat and tidy handwriting;

- difficult (low-frequency) words;

- "beautiful structures" or "shining phrases" (including clichés);

- long and complex sentences; and

- sentence connectors/conjunctions, such as furthermore, moreover and besides.

This practice, to some extent, has shaped the candidates' views of what counts as good writing and impedes the advancement of written communication skills of many students.

Another immediate measure that could be taken is to provide more specific and detailed instructions in the marking criteria. These instructions may include guidelines for dealing with essays with chunks of texts from exemplary essays and/or even those completely regurgitated from sample essays (with minor changes in some cases). These instructions could reduce the situation reported in Paltridge (2007), where inclusion of such texts is approved by some CET-4 testing centres but not by others. More importantly, these guidelines could eliminate the rise in circumstances reported in He (2010) where some raters award very high scores (e.g. 19 out of 20) but others give punishing scores (e.g. 2 out of 20) for scripts containing regurgitated texts from sample essays.

Instructions as above would be important because they provide more detailed guidelines for raters. Meanwhile more comprehensive rater training should also

be provided. Such training could focus on, among other skills discussed in Fulcher and Davidson (2013), strengthening raters' understanding of the test rubrics and rating scales. This training can further eliminate the "polluted scores" pointed out in Cheng and Curtis (2010, p. 270), and hence reduce the unfairness in scoring scripts in high-stakes English tests in China (Mei & Cheng, 2014). In the long run, these guidelines and rater training could facilitate the teaching and learning of learners' written communication skills.

8.4.3. Assistance to EFL students from less conducive language learning contexts

Findings from the current project allow teachers and language specialists to gain further insight into the problems that students from a disadvantageous language learning context could experience. Such insight could then help teachers and language specialists better understand students from such learning contexts, and thus provide more relevant assistance for them. Apart from measures to help students with their intercultural communication skills, as indicated in Myles and Cheng (2003), another important measure is perhaps to raise their awareness of the potential language problems they may have, for example, the ones discussed in the current study. This is to make the invisible problems visible to them so that these learners may consciously avoid such problems in their language production.

With the increasing interest in corpus studies, in particular in learner corpora, this measure is becoming more feasible. Paquot (2010), for example, identifies the overuse of imperative structures in the academic writing of French learners of English (e.g. *let us take more measures to deal with the problem*). One of the reasons for this overuse, according to the study, is that this imperative is often

216

used as an organisational marker to enhance cohesion and coherence in French academic texts.

For Chinese learners of English, a number of studies have investigated the common language errors they are especially prone to. For example, Lin (2002) explores the overuse, underuse and misuse of the pronoun "it". Bolton, Nelson and Hung (2002) examine Chinese students' inappropriate use of sentence connectors. Lin (2003) looks into the Chinese learners' language problems caused by the typology difference between Chinese and English; and Li (2014) explores the overuse and misuse of first and second English personal pronouns (i.e. *I* and *You*) by Chinese learners of English, in particular learners from the mainland. If the results of these studies are integrated into learning and teaching materials for these learners, their awareness of related language problems may be developed. This could make the struggle with their studies overseas less onerous. Some researchers have started their investigation into this area. One of the examples is Chuang and Nesi (2006) who study the errors that frequently appear in academic texts by Chinese learners of English. Built on their analysis, an online self-study learning package, GrammarTalk, was designed to raise learners' awareness of these common errors so that they can avoid similar errors in their language production. This study, however, focuses mainly on the writing of secondary school students in China. Research on Chinese speaking students in tertiary education might be even more important considering that many of them study in universities where English is used as the medium of instruction.

8.4.4. Components in writing assessment criteria

The current project has identified a negative correlation between the mainland students' writing scores and the lexical diversity in their writing. These learners' strong belief in using unnecessarily long and complex sentences also indicates that the sentence variety in their writing may not be highly related to their writing quality. The implications of these findings are that markers/raters of writing tests should exercise caution in awarding scores to the range or variety of lexical items and sentence structures, which are both important components in the writing assessment criteria for high-takes international proficiency tests of English, such as IELTS and TOEFL. Possibly more consideration should be given to how appropriately these lexical items and sentence structures are used in contexts, rather than the variety of them.

8.5. Limitations of the study

The limitations of the study are three-fold. They are related to the sample size and timing of data collection and also the choices of instruments.

8.5.1. Sample size and representation

The overall sample size for the current study was 150. However, later on it was found that many mainland learners' language development deviates severely from the language development patterns of EFL learners identified in previous studies. This deviation made it necessary for the present project to depend upon data from the Hong Kong sample when determining the predicting power of vocabulary knowledge in writing. Because of this, the sample size reduced to 67 in the related analyses. This size, although close to that of 74 in Qian (1999), was considerably

smaller than that of 217 in Qian (2002), 207 in Qian and Schedle (2004) and 113 in Stæhr (2009). A larger sample size could have made the current study more comparable with these studies.

Another consideration is that, as a variable control measure, the present project only recruited learners who were first-year university students and had the same mother tongue. This control measure helped to eliminate many of the variables that may affect the results of the research, in particular, the factors of language background and cognitive capacity. However, it, at the same time, also reduced the scope of the sample representation. Considering the limitations of sample size and representation, any future studies focusing on different research populations should make reference to the present study with caution.

8.5.2. Level inclusion of the Vocabulary Levels Test

With the purpose of recruiting as many participants as possible for the current project, the researcher took measures to limit the procedure of data collection, which included the completion of two vocabulary tests, an essay (400-500 words), a questionnaire survey and an informed consent form, to two and half hours in total duration. To accomplish this task, the researcher decided to employ the VLT only at the 3,000 and 5,000 frequency levels (see Section 3.2 for the rationale of this decision). This measure successfully controlled and placed limits upon the time required from the project participants, but at the same time limited the project scope in conducting more comparisons with related studies.

8.5.3. Timing for data collection

Data collection for the present project was undertaken in 2010. That was the year when the Foundation Year Programme for the mainland students in the Hong Kong Polytechnic University where the data were collected, was still being operated. The researcher of the present project was coordinating the subjects offered for the programme by the English Language Centre and was also teaching on the subjects. This role allowed the researcher to have more direct contact with the mainland students in the university, which helped to recruit project participants and identify learners for the case studies. The drawback of this timing was that in 2009, a new three-year senior secondary curriculum was implemented in Hong Kong. Students under the new curriculum take the HKDSE instead of HKALE, and the HKALE ceased operation in 2013. This examination reform is beneficial to students (Carless, 2013; Lee, 2013; Carless & Harfitt, 2013); but unfortunately it may limit the possibilities of any interested parties in conducting a replica study of the current project.

8.5.4. Data from questionnaire survey

The questionnaire survey was piloted and revisions were made accordingly afterwards (see Section 3.4.2). However, due to the limited number of participants in the pilot study, it was impossible to conduct an in-depth statistical analysis of the pilot results. After the survey was formally administered in the main study it was found that some mainland students misinterpreted a number of questions in the survey. Further investigation indicated that their misinterpretation was caused by the English-learning environment they experienced before coming to Hong Kong (see Section 7.1). The corollary of this was that the results from some

questions had to be excluded from the study. This matter was resolved by conducting focus-group interviews to collect qualitative data for the related questions. However, it would have been more persuasive if both qualitative and quantitative data had been collected for these questions.

8.6. Concluding remarks

This study aimed to examine the association between EFL learners' depth of lexical knowledge and their writing proficiency and the role that language learning contexts play in determining this relationship. With sufficient empirical support, the main study of the project yielded results leading to the following answers to the first and second research question set in Chapter 1:

- EFL learners' depth or extensiveness of vocabulary knowledge is closely related to their writing ability; and

- The depth dimension of word knowledge can provide a unique and distinctive prediction of EFL learners' writing scores (11%), over and above the contribution made by the breadth of lexical knowledge.

This unique contribution is as great as that found in Qian's study (1999) on reading comprehension (11%) and much greater than that in Stæhr's study (2009) on listening comprehension (2%). The result is little surprising given that learners' depth of lexical knowledge provides more information about the quality of word knowledge in their mental lexicon and the level of automaticity in their lexical use. Both of these factors are directly associated with learners' writing ability.

In addition, the study has also found that EFL learners' word knowledge accounts for a quarter (25%) of the explained variance in their writing scores. This figure,

although lower than the 71% of the afforded variance in reading comprehension (Qian, 1999) and the 51% in listening comprehension (Stæhr, 2009), is substantial, considering that many other linguistic and non-linguistic factors come into play when a learner composes a piece of writing. More importantly, both reading and listening require only retrieving the meaning of a word from one's mental lexicon (receptive vocabulary knowledge) but writing demands retrieving and producing an appropriate form of a word in a given context (both receptive and productive dimensions of one's lexical knowledge). A much higher command of learners' lexical knowledge is therefore needed for writing.

In the subsequent study, both qualitative and quantitative analyses were conducted to ascertain the answer to the third research question. The case studies helped to paint five "portraits" of students who underwent the less conducive language learning context in China. The questionnaire survey and focus-group interviews then provided a broad-brush "landscape painting" of the recurring problems caused by the same factor. The corpus analysis on the writing of these learners further confirmed the causes of the problems. Results from all these analyses led to the answer to the third research question:

- EFL learners' language learning contexts have a significant impact on the predicting power of the depth of lexical knowledge on their writing proficiency.

The significance of this current project is two-fold. Theoretically, it has filled a gap in the existing literature appertaining to the relationship between EFL learners' depth of word knowledge and their language proficiency. A more significant contribution is that the study has drawn researchers' attention to the

factor of language learning contexts when determining the predicting power of lexical knowledge and, in particular, the depth dimension of vocabulary knowledge on learners' writing proficiency. Pedagogically, the study has further confirmed the observations in previous studies, which suggest that the NMET cannot objectively measure candidates' writing ability, and thus reasserts the pressing need for fundamental examination reforms in China. At the same time, the study has ascertained the imperative importance of cultural and language assistance universities should provide for students from relatively unconducive English-learning contexts.

APPENDICES

Appendix A. Receptive Vocabulary Test (3,000 word level)

Choose a right word on the left hand side to match its meaning on the right hand side. Write the **number** of that word next to its meaning. One example has been done for you (30 marks).

Example	
1. business	<u>6</u> part of house
2. clock	<u>3</u> animal with four legs
3. horse	<u>4</u> something used for writing
4. pencil	
5. shoe	
6. wall	

Source:

Nation, I. S. P. (2001). Learning vocabulary in another language. Cambridge: Cambridge University Press.

Version: Test B

Now you can start your test.	
1. bull	__ formal and serious manner
2. champion	__ winner of a sporting event

3. dignity 4. hell 5. museum 6. solution	__ building where valuable objects are shown
1. blanket 2. contest 3. generation 4. merit 5. plot 6. vacation	__ holiday __ good quality __ wool covering used on beds
1. comment 2. gown 3. import 4. nerve 5. pasture 6. tradition	__ long formal dress __ goods from a foreign country __ part of the body which carries feeling
1. administration 2. angel	__ group of animals __ spirit who serves God

3. frost 4. herd 5. fort 6. pond	__ managing business and affairs
1. atmosphere 2. counsel 3. factor 4. hen 5. lawn 6. muscle	__ advice __ a place covered with grass __ female chicken
1. abandon 2. dwell 3. oblige 4. pursue 5. quote 6. resolve	__ live in a place __ follow in order to catch __ leave something permanently
1. assemble 2. attach	__ look closely __ stop doing something

3. peer 4. quit 5. scream 6. toss	__ cry out loudly in fear
1. drift 2. endure 3. grasp 4. knit 5. register 6. tumble	__ suffer patiently __ join wool threads together __ hold firmly with your hands
1. brilliant 2. distinct 3. magic 4. naked 5. slender 6. stable	__ thin __ steady __ without clothes
1. aware 2. blank 3. desperate	__ usual __ best or most important __ knowing what is happening

4. normal	
5. striking	
6. supreme	

Receptive Vocabulary Test (5,000 word level)

Choose a right word on the left hand side to match its meaning on the right hand side. Write the **number** of that word next to its meaning. One example has been done for you (30 marks).

Example	
1. business	<u>6</u> part of house
2. clock	<u>3</u> animal with four legs
3. horse	<u>4</u> something used for writing
4. pencil	
5. shoe	
6. wall	

Now you can start your test.	
1. analysis	__ eagerness
2. curb	__ loan to buy a house
3. gravel	__ small stones mixed with sand
4. mortgage	
5. scar	
6. zeal	

1. cavalry	__ small hill
2. eve	__ day or night before a holiday
3. ham	__ soldiers who fight from horses
4. mound	
5. steak	
6. switch	
1. circus	__ musical instrument
2. jungle	__ seat without a back or arms
3. nomination	__ speech given by a priest in a church
4. sermon	
5. stool	
6. trumpet	
1. artillery	__ a kind of tree
2. creed	__ system of belief
3. hydrogen	__ large gun on wheels
4. maple	
5. pork	
6. streak	
1. chart	__ map
2. forge	__ large beautiful house
3. mansion	__ place where metals are made and shaped

4. outfit 5. sample 6. volunteer	
1. contemplate 2. extract 3. gamble 4. launch 5. provoke 6. revive	__ think about deeply __ bring back to health __ make someone angry
1. demonstrate 2. embarrass 3. heave 4. obscure 5. relax 6. shatter	__ have a rest __ break suddenly into small pieces __ make someone feel shy or nervous
1. correspond 2. embroider 3. lurk 4. penetrate 5. prescribe 6. resent	__ exchange letters __ hide and wait for someone __ feel angry about something

1. decent	__ weak
2. frail	__ concerning a city
3. harsh	__ difficult to believe
4. incredible	
5. municipal	
6. specific	
1. adequate	__ enough
2. internal	__ fully grown
3. mature	__ alone away from other things
4. profound	
5. solitary	
6. tragic	

Appendix B. Word Associates Test

Word Associates Test

Instructions:
This is a test of how well you know the meaning of adjectives that are commonly used in English. There are eight words in the two boxes (left & right boxes). The words on the left side may help to explain the meaning of the stimulus word. The words on the right side are nouns that may come after the stimulus word in a phrase or a sentence. Take a look at the following example with sudden as the stimulus word:

sudden

beautiful	(quick)	(surprising)	(change)	doctor	(noise)
thirsty			school		

From the two boxes, circle **four words** that you think are relevant to the stimulus word (i.e. *sudden* in this example), according to the criteria mentioned above.

1. **beautiful**

enjoyable	expensive	free	education	face	music
loud			weather		

2. **bright**

clever	famous	happy	colour	hand	poem
shining			taste		

3. **calm**

open	quiet	smooth	cloth	day	light	person
tired						

4. **natural**

expected	helpful	real	foods	neighbours	parents
short			songs		

Source: http://www.lextutor.ca/tests/associates/

233

5. **fresh**

another	cool	easy	raw	cotton	heat	language
				water		

6. **general**

closed	different	usual	country	idea	reader
whole			street		

7. **bare**

empty	heavy	uncovered	cupboard	feet	school
useful			tool		

8. **acute**

hidden	often	rich	sharp	angle	hearing	illness
				stones		

9. **common**

complete	light	ordinary	boundary	circle	name
shared			party		

10. **complex**

angry	difficult	necessary	argument		passengers
sudden			patterns	problem	

11. **broad**

full	moving	quiet	wide	night	river	shoulders
				smile		

12. **conscious**

awake	healthy	knowing	face	decision	effort
laughing			student		

13. **convenient**

easy	fresh	near	experience	sound	time
suitable			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. favorable

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	accident	examination	illness
unplanned			shortage		

24. **compact**

effective	small	solid	group	kitchen	medicine
useful			string		

25. **crude**

clever	fair	rough	behaviour	drawing	oil
valuable			trade		

26. **domestic**

home	national	regular	animal	movement	policy
smooth			speed		

27. **profound**

bright	deep	exact	great	effect	machine	taste
				thought		

28. **fertile**

dark	growing	private	business	egg	mind	soil
special						

29. **formal**

fast	loud	organised	bomb	education	growth
serious			statement		

30. **independent**

changed	equal	important	child	country	ideas
separate			prices		

31. **original**

careful	closed	first	condition	mind	plan
proud			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

cautious hopeful traditional warm	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 C. Sample of the NMET writing prompts*

2010 年普通高等学校招生全国统一考试（全国二卷）

第三节 书面表达（满分 30 分）

假设你是李华，你的美国笔友 Peter 层表示希望来中国教书。你校现在需招聘外教，请给他写封信，告知招聘信息。内容主要包括：

1.教授课程：英语口语、英语写作、今日美国、今日英语等

2.授课对象：高中生（至少三年英语基础）

3.工作量：

—每周 12 学时，任选三门课

—担任学生英语俱乐部或英语校报顾问（advisor）

注意：

1.字数 100 左右；

2.可以适当增加细节，以使行文连贯；

3.开头语已为你写好，请将完整的回信书写在答题卡上。

****

Dear Peter,

I remember you told me you were interested in teaching in China.

Best,

Li Hua

*Note: My translation of this task is on the next page.

Source: http://www.rr365.com/Article/gkmess/gkxx/201006/23628.htm

National Matriculation English Test (Paper II)
2010

Part III.   Writing (30 marks)

You are Li Hua. Your school is looking for a native-speaker of English to teach for the school. They know that you have a pen friend, Peter, who is interested in teaching in China and have asked you to write an invitation letter to him. In your letter, you need to provide details of the post for Peter (see below).

1. Subjects to teach: Oral English, Writing, American Today and Contemporary English
2. Level : senior high school
3. Workload: three subjects per week (12 hours). The successful candidate should also serve as an advisor to the English Club or the school English newspaper.

Please note that
1. Your letter should be around 100 words in length.
2. You can add details to make your letter more coherent.
3. The beginning sentence of the letter has been written for you (see below). Please write the full letter on the answer sheet.

****

Dear Peter,

I remember you told me you were interested in teaching in China.

Best,

Li Hua

Appendix D. Sample of Karrie's writing

City Life and Country Life

With the development of economy, leading a stable life in cities has been playing an increasingly important role in pursuing and accomplishing most people's objectives and realizing their value, especially in some big cities. However, the others may have recognized the wisdom of having a good time which is free from restraint as well. While it is argued that the different advantages of life in alternative places vary a lot in numerous aspects, they both share the advanced society's resources, such as living conditions, industrial demands, working attitudes and lifestyle.

Giving consideration to the quality of life, working hard and being prepared for challenges would benefit both citizens and peasants. An examination of those who enjoy high returns confirms a fact that hardship should be a strong determinant of achievements associated with life standard no matter where it is.

Besides, city is not alone in emphasizing the vital factors of technology in the life. As with the information-based system in cities, agriculture requires contemporary knowledge and skills as well. In addition, science may afford the ability of improving the life standard for people either in cities or countries.

In spite of the fact that both of these lifestyle have several similar elements in reality, the pressure of living in the country is not as much as that of city. People enjoy more pleasure in life itself and are less busy with their work and business. Owing to the healthier life manner, the group of people who are fascinated by living in the village suffer less from many diseases caused by tension, for instance, cancers, even though citizens care more on health and are provided with excellent

medical conditions and techniques. Nowadays, faced with the high rates of death, an increasing number of citizens hold the opinion that the importance of health can never be over estimated albeit under too much surviving stress.

On the other hand, city life is superior to country life in living conditions. The majority of people favor cosy and comfortable conditions as opposed to simple facilities. Take one kind of transports, the subway which creates much convenience as an example. The rapid pace of individuals' life requires the necessity of easy approach to daily purchases, eatings and vehicles. Nevertheless country life can not catch up with the people's various demands and always only specializes in its less polluted environment and comparatively fresh air. It should be one of the obvious reason that city life is valued more than country life to some extent by a great number of people.

Whereas city life may provide easier access to advanced technology, t will not be easy for most people to find satisfactory shelter to stay in. On account of the high price of housing ,the attempt to occupy an own apartment last for their whole working process. Moreover, it can be inferred that housing price is not the only stress the poor can not tolerate. In another word, prices of the bulk of goods have the same effects. On the contrary, a sharp rise in prices may not occur in the countryside. The lower expense benefits individuals' pleasure and bring much satisfaction with life.

In summary, city life differs a lot, such as living conditions and people's attitudes from country life, although they should both depend on the development of civilization for science and technology. Yet the attempt to obtain a high standard

of life by working hard and being engaged in own jobs can always be affirmed as a tremendous success wherever it is.

Appendix E. Consent form

Consent Form

Information about Participants

Name   _____   Student ID _____

Age   _____   Sex   _____

Name of programme (e.g. BSE, ITC)   _____

Level of study (circle one from the 3 choices below)

Foundation Year

Year 1 (degree)

Year 1 (Higher diploma)

Place where you received secondary school education:

City _____   Province _____ (if you are from the mainland)

Score at the NMET (高考) in 2011_____ (if you are from the mainland)

Grade at HKALE in 2011 _____ (if you are from Hong Kong)

I understand that the above information will be for research purposes **only** and will be kept strictly confidential.

By signing this form, I am giving my permission for the results of my vocabulary tests and essay writing to be used for research purposes.

Signature _____   Date _____

Appendix F. Topics for writing test

1. Increasing concerns have been expressed in Hong Kong and mainland China about youngsters starting to experience sex at an early age, some even before ten years old. Discuss the causes of this problem and suggest ways to deal with it.

2. A private school in London has recently divided its students at the same year level into three streams according to their academic abilities. Would you like to see such divisions taking place in the schools of Hong Kong or/and mainland China? Please give reasons.

3. Which is more important for a country/region, economic development or environmental protection? Please give reasons.

4. High housing price in China has caused serious social concerns. Discuss the causes and impacts of such social problem.

5. The life span in many places such as Hong Kong, China and Europe is much longer than before, resulting in an ageing population in these places. Discuss the causes and effects of an ageing population.

6. Discuss the advantages and disadvantages of rapid development of e-communication devices, such as iPad, iPhone and Kindle.

7. The 'one-child policy' has been practiced in China for many years. Many people have voiced their different opinions about this issue. Discuss the positive and negative effects of this policy.

8. Many children in Hong Kong and China are overweight nowadays. Discuss the causes and effects of the overweight problem among youngsters.

9. Many teenagers take celebrities as their idols. Discuss the positive and negative effects of such social phenomenon.

10. Discuss the advantages and disadvantages of conducting sex education in secondary schools.

Appendix G. Questionnaire Survey (version for Hong Kong students)

This questionnaire is to discover your experience on vocabulary learning and writing practice before you came to the Hong Kong Polytechnic University.

1. Some people seem to think vocabulary is the most important part of English learning.   Do you agree?
    a. I completely agree
    b. I partially agree
    c. I don't agree
    d. I don't know

2. How many new words did you learn each week on average?
    a. above 30
    b. 30-20
    c. 20-10
    d. below 10

3. Of the following vocabulary learning strategies, which one did you use most often?
    a. memorising word lists
    b. reading, e.g. novels, newspapers, magazines
    c. listening to, e.g. news on TV/radio, recorded stories
    d. reciting selected texts

4. When using word lists to memorise words, which of the following lists did you use most?
    a. vocabulary lists in your English textbooks
    b. vocabulary lists for IELTS or TOEFL
    c. vocabulary lists you compiled yourself
    d. others, please specify
        _____

5. How often did you revise your word lists?
    a. every day
    b. every week
    c. every fortnight
    d. others, please specify

    _____

6. When learning a word, which of the following aspects was the most important to
   you?
    a. meaning in Chinese
    b. meaning in English
    c. spelling
    d. pronunciation

7. When learning a word, did you pay attention to its collocations?
    a. always
    b. often
    c. sometimes
    d. rarely

8. When learning a word, did you study the sentences given as examples of the use
   of this word?
    a. always
    b. often
    c. sometimes
    d. very rarely

9. If no example sentence was given for a new word, did you look it up in a
   dictionary to find some examples?
    a. yes, always
    b. yes, sometimes
    c. yes, but very rarely
    d. never

10. Which type of dictionary did you use most often (including both paper and electronic dictionaries)?
    a. English-Chinese dictionary (word meaning & examples only in Chinese)
    b. English-English dictionary (word meaning & examples only in English)
    c. bilingual dictionary (word meaning & examples both in Chinese and English)
    d. others, please specify

    _____

11. Some of you liked to learn "beautiful" words/phrases. What kind of words/phrases was beautiful for you?
    a. words/phrases that make your writing look more formal
    b. words/phrases that strongly express your feelings and emotions
    c. words/phrases that did not very often appear in texts you read
    d. set phrases and proverbs

12. Did your teacher go through the new words with the class before teaching a text?
    a. always
    b. sometimes
    c. rarely
    d. never

13. When teaching new words, did your teacher ask you to read aloud after him/her?
    a. always
    b. sometimes
    c. rarely
    d. never

14. When teaching new words, did your teacher explain their meanings in Chinese?
    a. always
    b. sometimes
    c. rarely

d. never

15. When teaching new words, did your teacher explain their meanings in English?
        a. always
        b. sometimes
        c. rarely
        d. never

16. When teaching new words, did your teacher provide example uses of these
     words?
        a. Yes, for every word
        b. Yes, for most words
        c. Yes, for important or difficult words only
        d. No

17. When teaching new words, did your teacher draw your attention to the
    collocations of these words
        a. always
        b. sometimes
        c. rarely
        d. never

18. What were the methods your teacher often used to check your progress with
    vocabulary?
        a. by giving dictations in class
        b. by having quizzes in class (e.g. fill in blanks with provided words)
        c. by asking you to make sentences for words you had learnt in class
        d. by checking how well you used words you had learnt in writing

19. How often did your teacher have a dictation and/or a quiz on vocabulary?
        a. everyday
        b. every 3-4 days
        c. every week

d. others, please specify

_____

A. Teaching input

Which of the following strategies did your teacher recommend to you to help you cope with the writing tasks in the HKCEE/HKALE?   Choose the answer that suits your situation most.

1. Provide a lot of background information in the introduction.

    a. *always*   b. *sometimes*   *c. rarely*   *d. never*

2. Address the topic indirectly because this is part of Chinese culture.

    a. *always*   b. *sometimes*   *c. rarely*   *d. never*

3. Draw an outline before writing.

    a. *always*   b. *sometimes*   *c. rarely*   *d. never*

4. Start with a topic sentence for each paragraph and develop the topic sentence logically.

    a. *always*   b. *sometimes*   *c. rarely*   *d. never*

5. Use connectors/conjunctions to connect sentences and paragraphs.

    a. *always*   b. *sometimes*   *c. rarely*   *d. never*

6. Use logic to make sure the whole writing connects well.

    a. *always*   b. *sometimes*   *c. rarely*   *d. never*

7. Use proverbs and set expressions.

   a. *always*   b. *sometimes*   c. *rarely*   d. *never*

8. Use low-frequency and/or difficult words.

   a. *always*   b. *sometimes*   c. *rarely*   d. *never*

9. Think in Chinese logically and then translate the thoughts into English.

   a. *always*   b. *sometimes*   c. *rarely*   d. *never*

10. Use long and complex sentences.

   a. *always*   b. *sometimes*   c. *rarely*   d. *never*

11. Use very strong tone to indicate your personal feelings and opinions.

   a. *always*   b. *sometimes*   c. *rarely*   d. *never*

12. Decide on the tone and style according to the audience.

   a. *always*   b. *sometimes*   c. *rarely*   d. *never*

B. Your beliefs and practice

Are the following statements *true* in your case?

1. I did not have much training in English.

   a. *very true*   b. *almost true*   c.   *true to some extent*   d. *not true at all*

251

2. My classmates and I did not write unless we had to prepare for the HKCEE/HKALE.

   a. *very true*   b. *almost true*   c.   *true to some extent*   d. *not true at all*

3. I believed tidiness was crucial in getting a higher grade in the HKCEE/HKALE.

   a. *very true*   b. *almost true*   c.   *true to some extent*   d. *not true at all*

4. I believed writing long and complex sentences helped me to obtain a higher grade in the HKCEE/HKALE.

   a. *very true*   b. *almost true*   c.   *true to some extent*   d. *not true at all*

5. While writing a composition, I paid much attention to its cohesion and coherence.

   a. *very true*   b. *almost true*   c.   *true to some extent*   d. *not true at all*

6. I believed the most efficient way to achieve cohesion and coherence in a composition was to use connectors/conjunctions.

   a. *very true*   b. *almost true*   c.   *true to some extent*   d. *not true at all*

7. I believed memorising example essays is crucial for improving my writing.

   a. *very true*   b. *almost true*   c.   *true to some extent*   d. *not true at all*

8. I believed reciting selected texts was also very important in improving my writing.

   a. *very true*   b. *almost true*   c.   *true to some extent*   d. *not true at all*

9. After completing a piece of writing, I always proofread it if I had time.

252

a. *very true*   b. *almost true*   *c. true to some extent*   *d. not true at all*

10. I spent most of my English study time memorising new words and phrases and trying to understand grammar because this is the most efficient way to improve my English.

    a. *very true*   b. *almost true*   *c. true to some extent*   *d. not true at all*

11. I memorised many low-frequency words in my writing because they are useful for exams.

    a. *very true*   b. *almost true*   *c. true to some extent*   *d. not true at all*

12. I memorised many "beautiful" words and phrases because use of them can increase my   grade in the HKCEE/HKALE.

    a. *very true*   b. *almost true*   *c. true to some extent*   *d. not true at all*

C. Others

1. How often were you asked to write in school?
    a. every week
    b. every fortnight
    c. every month
    d. others, please specify

    _____

2. When asked to write, did you usually have to write a paragraph, a few paragraphs or a completed composition in school?
    a. a paragraph
    b. a few paragraphs
    c. a completed composition
    d. others, please specify

    _____

3. What type of writing did you often write in school?
   a. narrative
   b. descriptive
   c. argumentative
   d. correspondence, e.g. letters, emails

4. Were you often given a task to write at class within limited time or out of class where you could spent as much time as you would like to on the writing?
   a. always at class
   b. most at class and occasionally out of class
   c. most out of class and occasionally at class
   d. always out of class

5. After finishing marking compositions of your class, did your teacher discuss the problems in your compositions in class?
   a. always
   b. sometimes
   c. rarely
   d. never

6. When discussing your compositions in class, what did your teacher comment on most?
   a. content
   b. organisation
   c. vocabulary and grammar
   d. others, please specify
      _____

7. Rank the following items from 1 to 4 (from the most important to the least important), according to what you were told about marking of writing in the HKCEE/HKALE.
   a. content
   b. organization
   c. vocabulary and grammar

254

d. tidiness (including handwriting)

Appendix H. Questionnaire Survey (version for mainland students)

This questionnaire is to identify your beliefs and strategies in learning vocabulary and writing before you came to the Hong Kong Polytechnic University.

1. Some people seem to think vocabulary is the most important part of English learning.   Do you agree?
    a. I completely agree
    b. I partially agree
    c. I don't agree
    d. I don't know

2. How many new words did you learn each week on average?
    a. above 30
    b. 20-30
    c. 10-20
    d. below 10

3. Of the following vocabulary learning strategies, which one did you use most often?
    a. memorising word lists
    b. reading, e.g. novels, newspapers, magazines
    c. listening to, e.g. news on TV/radio, recorded stories
    d. reciting selected texts

4. When using word lists to memorise words, which of the following lists did you use most?
    a. vocabulary lists in your English textbooks
    b. vocabulary lists for IELTS (雅思) or TOEFL (托福)
    c. vocabulary lists you compiled yourself
    d. vocabulary lists for NMET (高考)

5. How often did you revise (複習) your word lists?
    a. every day
    b. every week

c. every fortnight (兩週一次)

d. others, please specify

   _____

6. When learning a word, which of the following aspects was the most important to you?
   a. meaning in Chinese
   b. meaning in English
   c. spelling
   d. pronunciation

7. When learning a word, did you pay attention to its collocations (語境；搭配)?
   a. always
   b. often
   c. sometimes
   d. rarely

8. When learning a word, did you study the sentences given as examples of the use of this word?
   a. always
   b. often
   c. sometimes
   d. very rarely

9. If no example sentence was given for a new word, did you look it up in a dictionary to find some examples?
   a. yes, always
   b. yes, sometimes
   c. yes, but very rarely
   d. never

10. Which type of dictionary did you use most often (including both paper and electronic dictionaries)?

a. English-Chinese dictionary (word meaning & examples only in Chinese)

b. English-English dictionary (word meaning & examples only in English)

c. bilingual (雙語) dictionary (word meaning & examples both in Chinese and English)

d. others, please specify

_____

11. Some of you liked to learn "beautiful" words/phrases. What kind of words/phrases was beautiful for you?

a. words/phrases that make your writing look more formal

b. words/phrases that strongly express your feelings and emotions (情感)

c. words/phrases that did not very often appear in texts you read

d. set phrases and proverbs

12. Did your teacher go through the new words with the class before teaching a text?

a. always

b. sometimes

c. rarely

d. never

13. When teaching new words, did your teacher ask you to read aloud after him/her?

a. always

b. sometimes

c. rarely

d. never

14. When teaching new words, did your teacher explain their meanings in Chinese?

a. always

b. sometimes

c. rarely

d. never

15. When teaching new words, did your teacher explain their meanings in English?
    a. always
    b. sometimes
    c. rarely
    d. never

16. When teaching new words, did your teacher provide example uses of these words?
    a. Yes, for every word
    b. Yes, for most words
    c. Yes, for important or difficult words only
    d. No

17. When teaching new words, did your teacher draw your attention to the collocations (語境；搭配) of these words
    a. always
    b. sometimes
    c. rarely
    d. never

18. What were the methods your teacher often used to check your progress with vocabulary?
    a. by giving dictations (聽寫) in class
    b. by having quizzes (小測試) in class (e.g. fill in blanks with provided words)
    c. by asking you to make sentences for words you had learnt in class
    d. by checking how well you used words you had learnt in writing

19. How often did your teacher have a dictation (聽寫) and/or a quiz (小測試) on vocabulary?
    a. everyday
    b. every 3-4 days
    c. every week
    d. others, please specify

_____

A. Teacher input

Which of the following strategies did your teacher recommend to you to help you cope with the writing tasks in the NMET (高考)?   Choose the answer that suits your situation most.

1. Provide a lot of background information in the introduction.

   a. *always*   b. *sometimes*   c. *rarely*   d. *never*

2. Address the topic indirectly because this is part of Chinese culture (在寫文章時 遵循中文習慣，間接地引出話題).

   a. *always*   b. *sometimes*   c. *rarely*   d. *never*

3. Draw an outline before writing.

   a. *always*   b. *sometimes*   c. *rarely*   d. *never*

4. Start with a topic sentence for each paragraph and develop the topic sentence logically.

   a. *always*   b. *sometimes*   c. *rarely*   d. *never*

5. Use connectors/conjunctions (連接詞) to connect sentences and paragraphs.

   a. *always*   b. *sometimes*   c. *rarely*   d. *never*

6. Use logic to make sure the whole writing connects well.

   a. *always*   b. *sometimes*   c. *rarely*   d. *never*

7. Use proverbs and set expressions.

a. *always*　　b. *sometimes*　　*c. rarely*　　*d. never*

8. Use low-frequency (不常用) and/or difficult words.

　　a. *always*　　b. *sometimes*　　*c. rarely*　　*d. never*

9. Think in Chinese logically and then translate the thoughts into English.

　　a. *always*　　b. *sometimes*　　*c. rarely*　　*d. never*

10. Use long and complex sentences.

　　a. *always*　　b. *sometimes*　　*c. rarely*　　*d. never*

11. Use very strong tone (語氣；語調) to indicate your personal feelings and opinions.

　　a. *always*　　b. *sometimes*　　*c. rarely*　　*d. never*

12. Decide on the tone and style (風格) according to the audience.

　　a. *always*　　b. *sometimes*　　*c. rarely*　　*d. never*

B. Your beliefs and practices

Are the following statements *true* in your case?

1. I believed tidiness (整潔) was crucial (非常重要) in getting a higher grade in the NMET
(高考).

　　a. *very true*　　b. *almost true*　　c. *true to some extent*　　d. *not true at all*

261

2. I believed writing long and complex sentences helped me to obtain a higher grade in the NMET (高考).

    a. *very true*   b. *almost true*   c. *true to some extent*   d. *not true at all*

3. While writing a composition (文章), I paid much attention to its cohesion and coherence (邏輯和連貫).

    a. *very true*   b. *almost true*   c. *true to some extent*   d. *not true at all*

4. I believed the most efficient way to achieve cohesion and coherence (邏輯和連貫) in a composition (文章) was to use connectors/conjunctions.

    a. *very true*   b. *almost true*   c. *true to some extent*   d. *not true at all*

5. I believed memorising example essays is crucial (非常重要) for improving my writing.

    a. *very true*   b. *almost true*   c. *true to some extent*   d. *not true at all*

6. I believed reciting selected texts was also very important in improving my writing.

    a. *very true*   b. *almost true*   c. *true to some extent*   d. *not true at all*

7. After completing a piece of writing, I always proofread (修改) it if I had time.

    a. *very true*   b. *almost true*   c. *true to some extent*   d. *not true at all*

8. I spent most of my English study time memorising new words and phrases and trying to understand grammar because this is the most efficient way to improve my English.

262

a. *very true*　　b. *almost true*　　c. *true to some extent*　　d. *not true at all*

9. I memorised many low-frequency (不常用) words in my writing because they are useful for exams.

   a. *very true*　　b. *almost true*　　c. *true to some extent*　　d. *not true at all*

10. I memorised many "beautiful" words and phrases because use of them can increase my 　grade in the NMET (高考).

   a. *very true*　　b. *almost true*　　c. *true to some extent*　　d. *not true at all*

C. Your English writing experience in secondary school

1. I did not have much training in English.

   a. very true

   b. almost true

   c. true to some extent

   d. not true at all

2. My classmates and I did not write unless we had to prepare for the NMET (高考).

   a. very true

   b. almost true

   c. true to some extent

   d. not true at all

3. How often were you asked to write in school?

   a. every week

   b. every fortnight (兩週一次)

   c. every month

d. others, please specify

_____

4. When asked to write, did you usually have to write a paragraph, a few paragraphs or a completed composition (文章) in school?

    a. a paragraph

    b. a few paragraphs

    c. a completed composition (文章)

    d. others, please specify

    _____

5. What type of writing did you often write in school?

    a. narrative (敍述文)

    b. descriptive (描寫文)

    c. argumentative (議論文)

    d. correspondence, e.g. letters, emails

6. Were you often given a task to write at class within limited time or out of class where you could spent as much time as you would like to on the writing?

    a. always at class

    b. most at class and occasionally out of class

    c. most out of class and occasionally at class

    d. always out of class

7. After finishing marking compositions of your class, did your teacher discuss the problems in your compositions in class?

    a. always

    b. sometimes

    c. rarely

    d. never

8. When discussing your compositions in class, what did your teacher comment on most?

    a. content

    b.   organisation

    c.   vocabulary and grammar

    d.   others, please specify

    _____

9. Rank the following items from 1 to 4 (from the most important to the least important), according to what you were told about marking of writing in the NMET (高考).

    a.   content

    a.   organization

    b.   vocabulary and grammar

    c.   tidiness (整潔) (including handwriting)

Appendix I. Three versions of an essay written by Mecky

Essay topic: Discuss the impacts caused by Hong Kong's widening wealth gap and suggest solutions to address the problem.

Version 1

With a flourishing economy, Hong Kong is one of the most prosperous cities of the world. However, because the wealth isn't shared by all the citizens equally, Hong Kong's widening wealth gap now becomes a problem that the public cannot afford to ignore. The influence caused by this problem and the probable solution will be given in this essay.

The widening wealth gap causes a kind of psychology call 'chou fu' [1] in Chinese. Poor people turn against rich people because the distribution of wealth is not equal. The poor may tend to believe the government have ignored that they probably not be able to enjoy the social welfare as the riches do. The violence in New Delhi and Shanghai [1] indicate that this abnormal feeling will turn into some event which might out of control. Lately, the little shop owners and small sellers turned the demonstration called 'occupy central' into anti-government protests. Life is difficult for them because the high price level and the movement almost cut all their income. This situation shows the significant wealth gap and the discontent of the public.

The inequality of wealth distribution also causes the discontent of the educational system. Nowadays, the DDS (Direct Subsidy Scheme) might enjoy more educational resources than the normal school do [2].The high fee set a barrier for

266

the poor families. DDS always charge nearly 40000HKD a year more than the normal school, which almost make it impossible for poor students to purse academic success. In addition, Y.K. Poon and Y.C. Wong point out that the majority teachers tend to believe that study will be more attractive and students can be given more independence if they study online [3]. Taking POLYU's (The Hong Kong Polytechnic University) e-learning system as an example, students have to finish some tasks online or they will lost credit, which would increase the possibility of failing the subjects. However, the Internet access is not available for all families in Hong Kong. The society also requires students to have a more outstanding comprehensive quality. That is to say, students will be more easier accepted if they have more talent such as dancing and playing the musical instrument .Perhaps, most of these skills may need parents to spend a significant amount of money. This situation puts poor family in trouble because their child may not gain the same possibility to be well-educated or get a well-paid work as the ones from rich family do.

According to the rules in Hong Kong, the highest rate of income tax is 17% [4], which means the rich people with considerable income will not pay much more tax than the poor people do. The Hong Kong government does not provide the benefits for the people without jobs or retire, which also led to the increase of the inequality of the wealth distribution.

To address the problem, the government could increase the opportunities for the public to decrease the ratio of unemployment. The social security system should be developed further and the tax system have to change the tax policy in order to redistribute the sources of the society. The government can hackle the educational

problem by means of advancing elementary education [5]. To keep the growth of economics, Hong Kong government has to take measures to stop the wealth gap from widening.

Version 2

The widening wealth gap in Hong Kong is causing problems in many aspects. This problem, to some extent, is the main obstacle that obstructs Hong Kong's economic developing and leads society instability. Addressing this problem requires polices setting by the government as well as the co-operation of the companies. This essay will discuss the impact and some solutions to this problem.

The major impact of the wealth gap in Hong Kong is the unfair in education. A study conducted the enrolment rates of university among 19- and 20- year-olds reveals that 48.2 per cent students are from rich families and 11 per cent from the poor families in 2011 [1]. The poor may feel hard to afford the valuable tuition fee of the university education. As Poon and Wong [2] indict that except the curriculum reform some new education forms have emerged such as online learning. Those new form can make learning more effective, however, required more family support. Numbers of lower- income families have little money and experience to support those new learning form [2].The impact of the wealth gap in education can be even serious. Wang *et al* [3] point out that "the relationship between average elementary education and wealth gap is negative". According to this result a conclusion may be drawn that the poor families may be even poorer, the wealthy may be even wealthier.

As well as the unfair in education, the wealth gap may also lead to social instability. The poor people would resentment the wealthy people and may even abuse violent to protest against the social unfair. According to the research of Clifford and Pau [4] 96% publics hate the wealthy and myriad feel discontented

toward the government. In 2011 a plethora of people protest the low income and high cost violently [4].

To solve the problem of unfair in education and social instability, the government should take immediately measures. Wang and Xia [5] believe that Hong Kong's tax policy is inequality. The income tax rate for people whose income is different is the same. The government should adjust the tax policy to redistribute the wealth between the poor and the wealthy.

In addition to adjust the tax rate, the government should also strength the social welfare system. Numerous European societies have much completed social welfare systems than Hong Kong, however, they have little income gap than Hong Kong [5]. That means to complete the social welfare can adjust the wealth gap effectively.

Apart from the tax policy and social welfare, the adjusting of employment could be another major factor contributing to this problem. Wang *et al* [3] point out that "employment is the basis for livelihood". The poor are always people who are unemployment. Unemployment always means low-income. Creating more opportunities for employment can be effective in limiting wealth gap.

The widening wealth gap in Hong Kong has created unfair in education and social instability. To address this problem the government could adjust the tax rate to redistribution the wealth and complete the social welfare system. Companies could create more employment opportunities. Only when both of them act immediately can this problem to be reduced.

Version 3

With a flourishing economy, Hong Kong is one of the most prosperous cities of the world. However, the wealth is not shared by all the citizens equally. It led to the serious problem such as demonstration and widening gap in education. Hong Kong's wealth gap now becomes a problem that the public can not afford to ignore. The influence caused by this problem and the probable solution will be given in this essay.

The widening wealth gap causes instability of the society and decreases the public trust in government. Poor people turn against rich people because the distribution of wealth is not equal. The poor may tend to believe the government have ignored them. Because they probably not be able to enjoy the social welfare as the riches do. This kind of psychology called 'chou fu' might increases the possibility of demonstration. Discontent with the government is also growing. According to Clifford and Pau, the protests in some other place in Asia turned into violent incident ultimately. If the situation keep deteriorating, the possibility of more protests and even violence might increase. The dissatisfaction to the government might also deal a significant blow to the authority's credibility.

Besides the social conflict might happens, the inequality of wealth distribution also causes the disparity of the education. Nowadays, the DDS (Direct Subsidy Scheme) might enjoy more educational resources than the normal school do [2]. The high fee set a barrier for the poor families. DDS always charge nearly 40000HKD a year more than the normal school. It almost makes it impossible for poor students to pursue academic success. As well as the gap in high school, there is also a educational disparity exists in University. Poon and Wong point out that

271

the majority of teachers attach great importance to online learning [3]. Taking POLYU's (The Hong Kong Polytechnic University) e-learning system as an example, students have to finish some tasks online or they will lose credit. The loss of credit increases the possibility of failing the subjects. However, the Internet access is not available for all families in Hong Kong. It puts the poor family at the disadvantage. Apart from the outstanding learning skills, the society also requires students to have a more outstanding comprehensive quality. Students might distinguish themselves easily if they have more skills such as dancing and playing the musical instrument. Perhaps, most of these skills need parents to spend a significant amount of money. The lack of money puts poor family in trouble because their child may not gain the same possibility to be well-educated or be outstanding as the ones from rich family do.

Concerning the frequent occurrence of social protests over the unequal wealth distribution, some measures have to be taken. Li *et al.* suggest that the tax structure should reform. The tax should be in proportion to the income. To improve the living quality of the poor, citizens with higher income have to take the responsibility to improve the social welfare system. The Hong Kong government can provide benefits for the people without jobs by creating more employment opportunities. When reallocate the resource of the society, the authority should redistribute more to enhance the social welfare.

To limit the educational gap, the government need to set up policies to ensure the possibility of obtaining the quality education is the same. Providing a amount of money to general schools. Schools can afford scholarship to students to provide them with financial support. Students can spend these money on buying electronic

devices to study online or attend extra-curricular participation to equip themselves with comprehensive skills.

Social restlessness and educational inequality have been created because of the widening wealth gap. By means of reforming in the tax system, settling more social resources to social welfare and providing financial support to general schools, the problem can be tackled. Only when the Hong Kong government realize how serious the problem is and take measures to address it will the wealth gap stop growing.

Appendix J. Two texts written by Abby

Text A

The recent decades have witnessed a fundamental change in traditional ways of advertising. With millions of people regularly assessing the internet nowadays, online advertising tends to be accepted as one of the most powerful media for product promotion. However, increasing concerns about deceptive internet advertising and the privacy of the customers have been raised. To alleviate this problem, some people expect the government to take more responsibilities for internet advertising while others doubt the potential troubles which would be caused by enhanced regulations. This essay examines both the arguments for and against regulations on internet advertising and provides suggestions based on these arguments.

A number of research has presented evidences to show that government regulations are helpful to improve the problem of deceptive internet advertising. As numerous innovative and persuasive advertisements with delicate illustrations and inspiring words spread out rapidly through the internet, customers are more likely to be misled by their false descriptions for the products. In order to protect the interests of the consumers, different kinds of legislations have recently been issued in many western countries. One of the successful examples is the implementation of the Federal Trade Commission Act in the USA which prohibits deceptive advertising in any medium and encourages advertisers to self-regulate efficiently. The introduction of the act strengthens both the awareness of the

public and the advertisers towards deceptive internet advertising and it is found out that there was a large decrease in customers' complaints after that.

However, some people would argue that regulations from the government may violate one's freedom of speech. Since online advertising consists of the expression of ideas and is created to be understood in the cognitive sense, it could be deemed as a specific form of speech. The freedom of the advertisers should therefore be protected and even the state has no rights to intervene in the name of regulations. According to a study, censorship should not be allowed to inhibit the free flow of online information as the freedom of expression has been a part of basic human rights for centuries.

The next consideration is that appropriate regulations may avert the problem of privacy invasion. It is reported that a growing number of advertisers have adopted the strategy of collecting customers' online data automatically for advertising purpose. In this way, companies could simply distribute advertisements to the customers who recently browse similar products through the internet. Expenses of the companies on displaying internet advertisements could be effectively curtailed, however, customers gradually realized that their privacy of personal data was seriously invaded. They began appealing to privacy protection and legislations like Privacy and Electronic Communications Directive was then enacted in Europe. Goldfarb and Tucker have investigated that the number of European exposed to distributive online advertisements approximately dropped by 65% after the new law took its effect. Great amount of studies like this have emphasized the significance of government regulations to protect the privacy of the consumers.

Another argument against regulations on internet advertising considers the great burden which they may bring to the government. The implementation of enhanced regulations such as censorship requires a considerable amount of time and resources. Various types of expensive equipments and software are combined to set up a filter or firewall. In addition to this, a number of professionals with sophisticated skills should be employed to frequently maintain and upgrade the system. There are also limitations in relevant laws as some of the advertisers are continuously able to find ways to bypass the legislations with the fast development of internet technology.

This essay has discussed the benefits and drawbacks of the government regulations on internet advertising. Advantages include effective control on deceptive advertisements online and illegal invasion to customers' personal information. Disadvantages include the possibility of violating one's freedom of speech and large burden that regulations may impose on the government. Taking these arguments into consideration, I believe the government should enhance the regulations on advertisements online. However, benefits can only be maximized through sustainable ways taking expenditure and the freedom of speech into account. One of the practical methods is to establish an online green platform monitored by the government. Selected advertisers would be permitted to promote their advertisements to the public through the platform and the criteria for selection are in accordance to the sales volume and complaint rate of the companies. Consumers can also report the companies which deliver false advertisements or invade privacy on the platform. Once illegal commercial activities is verified, those companies should assume corresponding legal

responsibilities. Apart from this, the government could reduce deceptive advertising by increasing the taxation for each advertisement. In this way, a fair balance will be achieved between human rights and the interests of the consumers.

Text B

Social sciences have always enthralled me as I enjoy exploring social structures within multiplex cultural backgrounds. My undergraduate study in Social Policy and Administration has greatly polished my skills to think critically, to draw comparisons between distinctive social institutions, and utilize my understanding to comprehend social issues. I am determined to pursue an MPA degree in SIPA, selecting the international finance and economic policy as my concentration and management as specialization.

My career interests were not confirmed until I attended the 26th European International Model United Nations held in the Netherlands in 2013, where university students from all over the world acted as diplomatic officers in the UN and dealt with complex issues happening in the world. At that time, I delegated the Netherlands in the Economic and Social Council. One of the major topics is about The Resource Curse and Under-development in Sub-Saharan Africa, which refers to the phenomenon that although a large quantity of Sub-Saharan Africa countries possesses abundant natural resources, they are not able to gain a competitive position in the global market. During the literature review, I was inspired by how the Netherland managed to deal with its over-dependency on export of resources, and proposed that Sub-Saharans try focusing on decentralized private manufacturing to gain competitive advantage. By working with delegates of other countries, we finally completed a draft resolution paper to address the problem.

The sense of achievement as well as my huge interests in other international issues inspire me to pursue a career in the United Nations and I believe SIPA

would equip me with the needed economic knowledge and management skills. For example, the courses at Columbia SIPA MPA Program will help push my learning in macroeconomics, economic sociology theories to a higher level, while laying solid foundation for further learning such as Budgeting for non-profits, which are particularly crucial since economic and finance knowledge is the base for proposing to solve any fields of issue. Meanwhile, those economics rules of thumb will help understand the course of Politics of Policy-making and uplift my visions to view conflict interests at the commanding point. Except for those rigorously designed courses, the intellectual depth of SIPA is remarkable, driven by world-class faculty. Moreover, the Center for Global Economic Governance in SIPA has produced a new wave of policy-oriented research, making SIPA the ideal School to shape me into a qualified United Nations contributor.

Being an open-minded and independent individual who live off passion, I am convinced that a man is the architect of his own fate. Although I am an only a fresh graduate, my proactive characteristics and fast-learning ability can always help me adapt to the new environment quickly. I am also a highly disciplined person able to work well under pressure. Perfectionism complex motivates me to keep making plans and trying to accomplish tasks perfectly. Years of work in the Students' Union also trained my problem solving and emergency capacity and I have successfully led the committee and organized several big occasions within limited time and budget. My perfectionism do sometimes make me a very demanding leader, but never a fastidious one. I love sharing with others because learning from other people can also broaden my vision and also come up with the best solutions. Besides, I think work life balance is crucial in enhance both

productivity and all around development of the person. Last but not least, I am a creative girl willing to try new things, I sometimes even design my own clothes and ask tailors to customize. I believe this would bring the most benefit for the organization.

My desire to return to US for further education reached fever pitch when I finished my internship in Chicago this summer. I led a group of three and conducted a social research about the News Literacy of Chinese American High School Students at Chinese American Service League. After comprehensive literature review, data was collected from 34 participants through questionnaire survey, focus group discussion and individual interviews. A 47 pages dissertation was completed based on the results of data analysis and my research competency was greatly enhanced through the experience.

Honestly speaking, I was quite anxious at the beginning of the internship facing the brand new environment. Gradually fitting into the proactive American culture, I began to realize the importance of taking initiatives and now I am assertive to convey my opinions whenever I wish to. I am very certain that I can maintain this kind of learning attitude and be proactive in all walks of my life in the future.

Besides, I also learned how to balance work and life during the internship. On weekdays I had to work hard enough so that I could enjoy the fantastic musicals, beautiful sceneries and sunshine outside on weekends. For example, the Blues festival and Motown musical shows offered me vivid exposure of different types of music and better understanding of American history. Additionally, my determination of proceeding to graduate study in the U.S. was further confirmed after the campus visit to University of Chicago and Northwest University. I really

appreciate the opportunity to undertake the internship in Chicago. The unforgettable experience will inspire me all along the way to become a more mature and resilient person with a strong sense of social responsibility.

Appendix K. A student essay containing unnecessarily long and overly complex sentences

Recently, an issue that a famous actress's inappropriate behavior on a beach was spread by the media raised a debate whether the freedom of the press should be compromised to the privacy of citizens.

Which is more important, freedom of the press or protect of privacy of citizens?

Unless in particular situation, the privacy of citizens is undoubtedly worth more protection than the present freedom of the press. Society should attach more importance to the freedom of expressing opinions and comments reasonably by the media rather than the freedom of unveiling the private lives of citizens which is regarded as offensive and impolite. Only on the condition that the issue such as corruptions has great impacts on society excluding entertainment should be privacy be revealed. People can not live in the society where they are constantly concerned about the ruin of their names by the press.

A community is based citizens and ensures its operation and democracy by means of the press. Citizenship is superior to the freedom of the press because if citizenship can not be secured, freedom will lose its basic and be controlled by those who have wicked purpose. Furthermore, exposing the private lives of celebrities to the public for entertainment purposes is actually barely constructive but only damage the reputation of the victims and raise concern about privacy which are significantly threatened. There is no such thing as absolute freedom. Everything has some restriction aiming to protect the freedom of others. Freedom of the press should not be the excuse for offending the citizenship of Hong Kong.

Some could argue that without the freedom of the press, the society will lost conscience and opportunities to approach the truth. It is evident that news whose representatives are Watergate and Wukan Issue brings the media courage and motivation to reveal corruption and the unfair of the society. However, the purpose of this exposure was to protect the citizenship of other people whilst it destroyed the names of those who had behaved out of law. The restrictions should be posed on the justice but on those who wickedly reveal the privacy of innocent citizens for ridiculous entertainment purposes.

The media are supposed to tell freedom from abuse. Self-discipline is much more important, and flexible than restriction from government. Now that they do not want to be limited, they should behave themselves without publishing news which on harms individual but contributes little to society. But if the media isn't on their free, government should intervene via law on claim to protect the citizenship.

No one wants to live without privacy which means that he loses freedom of his life. The media and the government should soon take action until the society fails to protect citizenships.

Appendix L. A student essay containing overused and misused sentence connectors

With the development of the society, we put more and more energy into the construction of the world. At the same time, there's a heated discussion about whether economic development is the first or environmental protection is the first. I think that environmental protection should be taken more seriously. Below are my reasons.

Firstly, as we all know, economic development is based on the environment. If the lands are all deserts and there is no water, we will not be able to construct a beautiful city. **Also**, assume that we can construct a beautiful city, the environment is not good enough for people to live in. **So**, environmental protection plays an important part in the modern society.

Secondly, economic development should be pursued, **but** use shall not pursue it by destroying the environment. As someone puts it, protecting the environment is protecting ourselves. By protecting the environment, **not only** can we live in a more beautiful city, **but also** we can have the enough sources to pursue the economic development. We can kill two birds with one stone. So why don't we do that?

Thirdly, the fact is that more and more countries are paying debts of the environment, like the unstable climate, floods, hurricanes, volcanoes etc. these disasters cost more and more people their lives and destroy many famous buildings and even the precious things of the world. **And** that's all because that we had done so many bad things to the environment only to develop economics

faster. **So**, environmental protection is obviously more important than economic development.

What's more, economic development is just good to the people in short periods, but the environmental protection is good to the citizens for long periods of time. Developing economics is obviously not wrong, **but** if we develop economics by dint of the environment, we can't get the biggest profit from the environment **and** we are even in danger if we pay no attention to the environment.

Above all, I think there's only one solution to solving this problem. We can't only do one thing because if we only develop economics, the environment will be destroyed **but, on the contrast**, if we only protect the environment, we can't make the whole city developed. **So**, as far as I'm concerned, we should put environmental protection on the first hand and then without destroying the environment, we can develop the economics to improve our living quality. What I mean is that we should develop economics and protect the environment at the same time **but** environmental protection plays a more important part in the process. We can develop economics and at the same time we won't destroy the environment.

In summary, environmental protection needs more attention and should be laid more emphasize. In this way, we can live in a more and more beautiful world while enjoying a higher and higher living standard. I hope this day will come soon.

Appendix M. Temporals used in a pre-determined template*

<div align="center">Which would you give up: TV, cell or Web?</div>

As we know, nowadays T.V, cell and web become three important methods for us to get information or communicate with others in our daily life. They are almost equally important to us. But now if I have to give one up, I choose to give up TV. And I have the following reasons.

Firstly, many times when I watch TB, I just want to spend my free time and get some entertainments from it. But I can't have a talk with the people in the TV plays. But with a cell I can talk with others and share my ideas.

Secondly, web can offer me as much as information I need. But when I watch TV, I just can gain some information which the TV programs give. Thirdly, whatever news or the programs the TV offers are, I can easy get them on the Internet.

All in all, either the function or the practicality of the TV could be replaced by cell and computer, so I give it up.

---

* The essay was listed as a "full-mark" or exemplary composition in Wang (2013). The comments on the essay include the following:

*The use of **firstly, secondly** and **thirdly** has improved the cohesion and coherence of the essay and consequently helps audience better understand the author's views.*

Source: Wang, Y. F. (2013). *Full-mark Compositions on the NMET*. Beijing: Beijing Education Press.

Appendix N. A student essay with dialogic and hortatory features

The new iPhone 5 will be released in late October!' Hearing this news, hundreds of thousands of Apple fans burst into screaming. It didn't take a long time for those kind of Apple products to become a part of our daily life. Thus, e-communication devices and social networking have become two of the most popular phrases.

Those communication devices seem to be perfect and even invincible. But are they really what they seem to be?

Nothing can be more obvious than the advantages of communication devices and we can never deny the change they have brought to our life. Ten years ago you may have never thought about video call and wireless network even though mobile phones had become common in our life. Yes, without those devices, the distance between you and me would never be so close and our communication would never be so easy like it is now. If we look further into this problem, those devices do play a crucial role in cases which lead the world to be more united and accelerate the democratic process in some Muslim countries. Those e-communication devices change the world in their own unique way.

Now that e-communication devices are so perfect, why not buy everyone a product like this? Imagine you step into a classroom of a primary school where there are only 10-year-old children. Is it a little surprising to see they are all using e-communication devices? No doubt, using those devices at an early age can never do good to children's development. We have seen countless cases where teenagers are addicted to e-communicating devices and social networking and if

this kind of cases become more and more common, how can teenagers focus on their academic performance and future development? Needless to say, those devices also do harm to our physical health, which is always ignored by the public. In a recent case, a lawyer even use e-communication devices and social networking to collect evidence, which is completely ridiculous. This case is also a warning sign for every e-communication devices user—your personal information and secrets can be easily given away while using them. If you're not careful enough, you will have to face the music.

No matter how colorful the life is in the virtual world, our life is definitely in reality. Find yourself an apple, have a bite. Is it as fast as that APPLE which controls your social life?

REFERENCES

Adamson, B., & Tong, S. Y. A. (2008). Leadership and collaboration in implementing curriculum change in Hong Kong secondary schools. *Asia Pacific Education Review*, *9* (2), 181-190.

Baddeley, A. D. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience*, *4*, 829–839.

Barani, S., & Seyyedrezaie, S. H. (2013). Language Learning and Vocabulary: (A review).*International Journal of Basic Sciences & Applied Research*, *2*(1), 87-90.

Beglar, D. (2010). A Rasch-based validation of the vocabulary size test. *Language Testing*, *27*(1), 101-118.

Benson, P., & Lor, W. (1999). Conceptions of language and language learning. *System, 27*(4), 459-72.

Biggs, J. B. (1996). Western misconceptions of the Confucian-heritage learning culture. In D. A. Watkins & J. B. Biggs (Eds.), *The Chinese Learner: Cultural, Psychological and Contextual Influences* (pp. 45-67). Hong Kong: CERC and ACER.

Biggs**, J.** & Watkins, D. (1996). *The Chinese Learner: Cultural, Psychological and Contextual Influences***.** Hong Kong: CERC and ACER.

Bolton, K., Nelson, G., & Hung. J. (2002)**.** A corpus-based study of connectors in student writing. *International Journal of Corpus Linguistics*, *7*(2), 165-182.

Braine, G., & McNaught, C. (2007). Adaptation of the 'Writing across Curriculum' model to the Hong Kong context. In J. Liu (Ed.), *English*

language leaching in China: New approaches, perspectives and standards (pp. 311-328). London: Continuum International Publishing Ltd.

Butler, Y. G. (2014). Parental factors and early English education as a foreign language: A case study in mainland China. *Research Papers in Education*, *29*(4), 410-437.

Cambridge Advanced Learner's Dictionary (4th ed.). (2013). Cambridge: Cambridge University Press.

Carless, D. (2012). TBLT in EFL settings: looking back and moving forward. In A. Shehadeh & C. Coombe (Eds.), *Task-based Language Learning and Teaching in EFL Contexts: Research and Implementation.* Amsterdam/Philadelphia: John Benjamins.

Carless, D. (2013). Sustainable feedback and the development of student self-evaluative capacities. In S. Merry, M. Price, D. Carless & M. Taras (Eds.), *Reconceptualising feedback in higher education: developing dialogue with students* (pp. 595-608). London: Routledge.

Carless, D. & Harfitt, G. (2013). Innovation in secondary education: A case of curriculum reform in Hong Kong. In K. Hyland & L. Wong (Eds.), *Innovation and Change in English Language Education* (pp. 174-185). London: Routledge.

Chai, C. S., Teo, T., & Lee, C. B. (2009). The change in epistemological beliefs and beliefs about teaching and learning: a study among pre-service teachers. *Asia-Pacific Journal of Teacher Education, 37*(4), 351–362.

Cheng, L. (2008). The key to success: English language testing in China. *Language Testing*, *25*(1), 15-37.

Cheng, L. (2010). The History of Examinations: Why, How, What and Whom to Select? In L. Cheng & A. Curtis (Eds.), *English Language Assessment and the Chinese Learner (*pp. 13-26). New York: Routledge.

Cheng, L. & Curtis, A. (2010). The Impact of English Language Assessment and the Chinese Learner in China and Beyond. In L. Cheng & A. Curtis (Eds.), *English Language Assessment and the Chinese Learner* (pp. 267-273). New York: Routledge.

Cheng, L., & Qi, L. (2006). Description and examination of the National Matriculation English Test. *Language Assessment Quarterly*, *2*, 53-70.

Chu, H. C. J., Swaffar, J., & Charney, D. H. (2002). Cultural representations of rhetorical conventions: The effects on reading recall. *Tesol Quarterly*, *36*(4), 511-541.

Chuang, F. Y & Nesi, H., (2006). An analysis of formal errors in a corpus of Chinese student writing. *Corpora, 1* (2), 251 – 271.

Cobb. T. (2011). *Vocabprofiler*. Last retrieved in Feb 2011 on the World Wide Web. http://www.lextutor.ca/eng.

Cobb. T. (2011). *Corpus Concordance English*. Last retrieved in May 2011 on the World Wide Web. http://lextutor.ca/conc/eng.

Cohen, A. D. (2006). The coming of age of research on test-taking strategies. *Language Assessment Quarterly, 3*(4), 307-331.

Cortazzi, M., & Jin. L. (1996). English teaching and learning in China. Language Teaching: *The International Abstracting Journal for Language teachers and Applied Linguists, 29*, 61-80.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*, 213-238.

Crossley, S. A., & McNamara, D. S. (2012). Predicting second language writing proficiency: The role of cohesion, readability, and lexical difficulty. *Journal of Research in Reading, 35*(2), 115-135.

Dallar, H., Milton, J., & Treffers-Daller, J. (2007). Editor's introduction. In J. J. Daller, J. Milton & J. Treffers-Daller (Eds.), *Modelling and Assessing Vocabulary Knowledge* (pp. 133-149). Cambridge: Cambridge University Press.

Daller, H., & Phelan, D. (2007). What is in a teacher mind? Teacher ratings of EFL essays and different aspects of lexical richness. In J. J. Daller, J. Milton & J. Treffers-Daller (Eds.), *Modelling and Assessing Vocabulary Knowledge* (pp. 133-149). Cambridge: Cambridge University Press.

Daller, H., van Hout, R. & Treffers-Daller, J. (2003). Lexical richness in spontaneous speech of bilinguals. *Applied Linguistics*, *24* (2), 197-222.

Delmonte, R. (2008). *Computational Linguistic Text Processing-Lexicon, Grammar, Parsing and Anaphora Resolution*. Nova Science Publishers-New York.

Deng, C. R., & Carless, D. (2010). Examination preparation or effective teaching: Conflicting priorities in the implementation of pedagogic innovation. *Language Assessment Quarterly, 7*(4), 285-302.

Djokic, D. (1999). Lexical Errors in L2 learning and communication. *Rassegna Italina di Lingustistica Aplicata*, *1*, 123-135.

Edmonds, P. (2011). *Just the Word*. Retrieved in Feb 2011 on the World Wide Web. http://www.just-the-word.com.

Ehrman, M. E., Leaver, B. L., & Oxford, R. L. (2003). A brief overview of individual differences in second language learning. *System*, 31, 313-330.

Ellis, R. (2008a). *The study of second language acquisition*. Oxford: Oxford University Press.

Ellis, R. (2008b). Learner beliefs and language learning. *Asian EFL Journal, 10*(4), 7-25.

Ellis, R. & Shintani, N. (2013). *Exploring Language Pedagogy through Second Language Acquisition Researc*h. Florence, KY, USA: Routledge.

Evans, S. (2009). The medium of instruction in Hong Kong revisited: Policy and practice in the reformed Chinese and English streams. *Research Papers in Education, 24*(3), 287-309.

Evans, S. (2010). Language in transitional Hong Kong: Perspectives from the public and private sectors. *Journal of Multilingual and Multicultural Development, 31*(4), 347-363.

Evans, S. (2011a). Historical and comparative perspectives on the medium of instruction in Hong Kong. *Language Policy, 10*(1), 19-36.

Evans, S. (2011b). Hong Kong English and the professional world. *World Englishes, 30*(3), 293-316.

Evans, S. (2011c). Hong Kong English: The growing pains of a new variety. *Asian Englishes, 14*(1), 22-45.

Evans, S. (2013). The long march to biliteracy and trilingualism: Language policy in Hong Kong education since the handover. *Annual Review of Applied Linguistics, 33*, 302-324.

Færch, C., Haastrup, K., & Phillipson, R. (1984). *Learner language and language learning*. Clevedon, Avon, UK: Multilingual Matters.

Fan, M. (2001). An investigation into the vocabulary needs of university students in Hong Kong. *Asian Journal of English Language Teaching*, *11*, 69-85.

Field, Y., & Yip, L. M. O. (1992). A comparison of internal conjunctive cohesion in the English essay writing of Cantonese speakers and native speakers of English. *RELC Journal, 23*(1), 15–28.

Fowler, H. R., & Aaron, J. E. (2007). *The little, brown handbook*. Pearson: Longman.

Fox, J., & Curtis A. (2010). International English Language Testing System. In L. Cheng, & A. Curtis (Eds.). *English Language Assessment and the Chinese Learner* (pp. 113-120). New York: Routledge.

Fulcher, G., & Davidson, F. (2013). *The Routledge handbook of language testing*. New York: Routledge.

Gan, Z. (2009). Asian learners' re-examined: An empirical study of language learning attitudes, strategies and motivation among mainland Chinese and Hong Kong students. *Journal of Multilingual and Multicultural Development*, *30*(1), 41-58.

Gan, Z. (2013). Understanding English speaking difficulties: An investigation of two Chinese populations. *Journal of Multilingual and Multicultural Development*, *34*(4), 1-19.

Gan, Z., Humphreys, G. & Hamp-Lyons, L. (2004), Understanding Successful and Unsuccessful EFL Students in Chinese Universities. *The Modern Language Journal*, *88*, 229-244.

Gass, S. (1998). Second language vocabulary acquisition. *Annual Review of Applied Linguistics*, 9. 92-106.

Gass, S., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd Ed.). New York: Taylor & Francis.

Gathercole, S. E., & Alloway, T. P. (2008). Working memory and classroom learning. In K. Thurman & K. Fiorello (Eds.), *Cognitive development in K-3 classroom learning. Research applications*. Mahwah, NJ: Lawrence Erlbaum.

Granger, S. (1996). From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora, in K. Aijmer, B. Altenburg, and M. Johansson (Eds.), *Languages in Contrast* (pp. 37-51). Lund, Sweden: Lund University Press.

Granger, S. (2012). Learner corpora. In Chappelle, C.A. (Ed.), *The Encyclopaedia of Applied Linguistics* (pp. 3235–3242). Oxford: Wiley-Blackwell.

Greidanus, T., Beks, B., & Wakely, R. (2005).Testing the Development of French Word Knowledge by Advanced Dutch- and English-Speaking Learners and Native Speakers. *Modern Language Journal*, *85*, 303-317.

Gu, Y. (2003a). Vocabulary learning in a second language: person, task, context and strategies. *TESL-EJ*, *7*(2), 1-4.

Gu, Y. (2003b). Fine brush and freehand: the vocabulary learning art of two successful Chinese EFL learners. *TESOL Quarterly, 37*(1), 73-104.

Gu, Y. (2005). *Vocabulary learning strategies in the Chinese EFL context*. Singapore: Marshall Cavendish Academic.

Gu, Y. (2014). The unbearable lightness of the curriculum: what drives the assessment practices of a teacher of English as a Foreign Language in a Chinese secondary school? *Assessment in Education: Principles, Policy & Practice*, *21*(3), 286-305.

Gu, Y., & Johnson, R. K. (1996). Vocabulary Learning Strategies and Language Learning Outcomes. *Language Learning, 46*(4)*,* 643-679.

Halliday, M. A. K. (1967). Notes on transitivity and theme in English Part I. *Journal of linguistics*, *3*(01), 37-81.

Halliday, M. A. K. (2013). *Halliday's Introduction to Functional Grammar* (4th ed), revised by Christian M. I. M. Matthiessen. New York: Routledge.

Halliday, M. A. K., & Hasan, R. (1976): *Cohesion in English*. London: Longman.

Hamp-Lyons, L. (2006). The impact of testing practices on teaching: ideologies and alternatives. In J. Cummins & C. Davison (Eds.), *International handbook of English language teaching* (pp. 474-490). Norwell, MA: Springer.

Hamp-Lyons, L., & Kroll, B. (1996). Issues in writing assessment: An overview. *College ESL*, *6*, 52-72.

He, L. (2010). The Graduate School Entrance English Examination. In L. Cheng & A. Curtis (Eds.), *English Language Assessment and the Chinese Learner* (pp. 145-157). New York: Routledge.

Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). *Range*. Retrieved in August 2012, from http://www.victoria.ac.nz/lals/staff/paul-nation/nation.aspx.

Hong Kong Examination Authority. (2010). *Longitudinal study on the school-based assessment of the 2007 Hong Kong Certificate of Education (HKCE) English language examination*. Hong Kong: Government Press.

Hong Kong Examination and Assessment Authorities. (2013). Retrieved in August 2013, from http://www.hkeaa.edu.hk/en/hkale.

Hu, G. W. (2003). English language teaching in China: Regional differences and contributing factors. *Journal of Multilingual and Multicultural Development, 24*(4), 290-318.

Hu, G. W. (2005a). Contextual influences on instructional practices: A Chinese case for an ecological approach to ELT. *TESOL Quarterly*, *39*, 635-660.

Hu, G. W. (2005b). English language education in China: Policies, progress, and problems. *Language Policy*, *4*(1), 5-24.

Hu, Y. (2008). China's English language policy for primary schools. *World Englishes*, *27*, 516–534.

Jensen, J. V. (1998). Teaching East Asian rhetoric. *The Rhetoric Society Quarterly, 18,*136 -149.

Johnson, R. K. & Ngor, A. Y. S. (1996). Coping with second language texts: the

development of lexically-based reading strategies. In D. A. Watkins & J. B. Biggs (ed.), *The Chinese Learner: Cultural, Psychological and Contextual Influences* (pp. 474-490). Hong Kong: CERC and ACER.

Jiang, N. (2000). Lexical representation and development in a second language. *Applied linguistics*, *21*(1), 47-77.

Jiang, N. (2002). Form–meaning mapping in vocabulary acquisition in a second language. *Studies in Second Language Acquisition*, *24*(04), 617-637.

Kane, R., Sandretto, S., & Heath, C. (2002). Telling half the story. A critical review of research on the teaching beliefs and practices of the university academics. *Review of Educational Research, 72*(2), 177–227.

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: theories, methods, individual differences and applications* (pp. 57-71). Mahwah, NJ: Lawrence Erlbaum.

Kennedy, S., & Lui, R. (2013). Washback of a high-stakes English test in China: Student and teacher perceptions. *Concorda Working Papers in Applied Linguistics*, *4*, 22-29.

Koizumi, R. & In'nami, Y. (2013). Vocabulary Knowledge and Speaking Proficiency among Second Language Learners from Novice to Intermediate Levels. *Journal of Language Teaching and Research*, *4*(5), 900-913.

Kormos, J. (2012). The role of individual differences in L2 writing. *Journal of Second Language Writing, 21*(4), 390-403.

Kachru, Y. (1998). Culture and argumentative writing in world Englishes. In L. Smith & R. Forman (Eds.), *World English 2000* (pp. 48-67). Honolulu: University of Hawaii Press.

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. J. L. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 126-132). London: Macmillan.

Laufer, B. (1994). The lexical profile of second language writing: does it change over time? *RELC Journal*, *25*(2), 21-33.

Laufer, B. (1997). The lexical plight in second language reading: words you don't know, words you think you know and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second Language Vocabulary Acquisition: a Rationale for Pedagogy* (pp. 20-34). Cambridge: Cambridge University Press.

Laufer, B. (2013). Vocabulary and writing. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Boston: Wiley-Blackwell.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: size, strength, and computer adaptiveness. *Language Learning*, *54*, 469-523.

Laufer, B., & Nation, P. (1995). Vocabulary size and use: lexical richness in written production. *Applied Linguistics*, *16*, 307-31.

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language, 22,* 15-30.

Lee, I. (2010). Writing teacher education and teacher learning: Testimonies of four EFL teachers. *Journal of Second Language Writing, 19*(3), 143-157.

Lee, I. (2012). Becoming a writing teacher: Using "identity" as an analytic lens to understand EFL writing teachers' development. *Journal of Second Language Writing*, *22*(3), 330-345.

Lee, S. H. (2003). ESL learners' vocabulary use in writing and the effects of explicit vocabulary. *System, 31*, 537-561.

Leki, I., & Carson, J. G. (1994). Students' perceptions of EAP writing instruction and writing needs across the disciplines. *TESOL Quarterly*, 28, 81-101.

Li, L. (2014). Contextual and cultural influence on the use of first person pronouns by Chinese learners of English. In D. Qian, & L. Li (Eds.), *Teaching and Learning English in East Asian Universities: Global Visions and Local Practices* (pp. 200-212). Cambridge: Cambridge Scholars Publishing.

Lin, L. H. F. (2002). Overuse, underuse and misuse: using concordancing to analyse the use of it in the writing of Chinese Learners of English. In M. Tan (Ed.), *Corpus Studies in Language Education (pp.* 63-80). Bangkok: IELS Press.

Lin, L. H. F. (2003). Corpus Evidence: Existential Constructions by Chinese Learners of English. *The English Teacher- An International Journal*, *6*(3), 279-280.

Lin, L. H. F. (2009). Status of English in Hong Kong. In S. Granger (ed.), *Status of English in the World* (pp. 180-185). Belgium: Université de Louvain Press.

Lin, L. H. F., & Morrison, B. (2010). The impact of the medium of instruction in Hong Kong secondary schools on tertiary students' vocabulary. *Journal of English for Academic Purposes*, *9*, 255-266.

Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. *Language Teaching*, *40,* 243-249.

Llach, M. P. (2005). The Relationship of Lexical Error and their Types to the Quality of ESL Compositions: An Empirical Study. *Porta Linguarum*, *3*, 45-57.

Llach, M. P. (2011). *Lexical Errors and Accuracy in Foreign Language Writing*. Bristol: Multilingual Matters.

Lo, Y. Y., & Lo, E. S. C. (2014). A meta-analysis of the effectiveness of English-medium education in Hong Kong. *Review of Educational Research*, *84*(1), 47-73.

Lo, Y. Y., & Murphy, V. A. (2010). Vocabulary knowledge and growth in immersion and regular language learning programmes in Hong Kong. *Language and Education, 24*, 215-238.

Louw, B. (1993). Irony in the text or insincerity in the writer? In M. Baker *et al*. (Eds.), *Text and Technology*. Amsterdam: John Benjamin's.

Luk, J. C. M. (2010). Talking to score: Impression management in L2 oral assessment and the co-construction of a test discourse genre. *Language Assessment Quarterly, 7*(1), 25-53.

Luk, J. C. M. (2012). Teachers' ambivalence in integrating culture with EFL teaching in Hong Kong. *Language, Culture and Curriculum*, *25*(3), 249-264.

Malvern, D. D., & Richards, B. J. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing*, *19*, 85-104.

Mayor, B. M. (2006). Dialogic and hortatory features in the writing of Chinese candidates for the IELTS test. *Language, Culture and Curriculum*, *19*(1), 104-121.

McCarthy, M. J. (1990). *Vocabulary*. Oxford: Oxford University Press.

McKeown, M. G., & Curtis, M. E. (1987). *The nature of vocabulary acquisition*. Hillsdale, NJ: Erlbaum.

Meara, P. M. (1982). Word association in a foreign language: A report from the Birkbeck vocabulary project. *Nottingham Linguistic Circular, 11, 29-37.*

Meara, P. M. (1992). *EFL vocabulary tests*. University of Wales, Swansea: Centre for Applied Language Studies.

Meara, P. M. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer, & J. Williams (Eds.), *Competence and performance in language learning* (pp. 35-53). New York: Cambridge University Press.

Meara, P. M., & Alcoy, J. C. O. (2010). Words as Species: An Alternative Approach to Estimating Productive Vocabulary Size. *Reading in a Foreign Language*, *22*(1), 222-236.

Meara, P. M., & Jones, G. (1990). *Eurocentres vocabulary size test 10 KA*. Zurich: Eurocentres Learning Service.

Meara, P. M., & Milton, J. L. (2003). *X-Lex: Swansea Vocabulary Levels Test* (Version 2.02). (Computer software). Swansea: Lognostics.

Mei, Y. (2008). *An attentive study on ethical issues in rating the NMET writing performance*. Unpublished dissertation, Shanxi University, Taiyuan, China.

Mei, Y., & Cheng, L. (2014). Scoring Fairness in Large-Scale High-Stakes English Language Testing: An Examination of the National Matriculation English Test. In D. Coniam (Ed.), *English Language Education and Assessment: Recent Developments in Hong Kong and the Chinese Mainland* (pp. 171-187). Springer: Singapore.

Melka, F. (1997). Receptive vs. productive aspects of vocabulary. In N. Schmitt, & M. McCarthy (Eds.), *Vocabulary: description, acquisition and pedagogy* (pp. 82-102). Cambridge: Cambridge University Press.

Michieka, M. (2010). Holistic or Analytic? Issues in Grading ESL Writing. *TNTESOL Journal. 3*, 75-83.

Millman, J., Bishop, C. H., & Ebel, R. (1965). An analysis of test-wiseness. *Educational and Psychological Measurement*, *25*(3), 707-726.

Milton, J. (2009). *Measuring Second Language Vocabulary Acquisition*. Bristol: Multilingual Matters.

Milton, J., & Tsang, E. S. C. (1993). A corpus-based study of logical connectors in EFL students' writing: directions for future research. In R. Pemberton & E. S. C. Tsang (Eds.), *Studies in Lexis* (pp. 215–246). Hong Kong: The Hong Kong University of Science and Technology Language Centre.

Myles, J., & Cheng, L. (2003). The social and cultural life of non-native English speaking international graduate students at a Canadian university. *Journal of English for Academic Purposes, 2*(3), 247-263.

Nagy, S., & Herman, P. A. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. G. McLean & M. E. Curtis

(Eds.), *The Nature of Vocabulary Acquisition* (pp. 19-35). Hillsdale, NJ: Erlbaum.

Nassaji, H. (2004). The relationship between depth of vocabulary knowledge and L2 learners' lexical inferencing strategy use and success. *Canadian Modern Language Review*, *61*(1), 107-134.

Nation. I. S. P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.

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*, *63*(1), 59-82.

Nation, I. S. P. (2008). *Teaching vocabulary and techniques*. Boston, MA: Heinle.

Nation, I. S. P. (2012). *The Vocabulary Size Test*. Retrieved in July 2014 from http://www.victoria.ac.nz/lals/about/staff/publications/paul-nation/Vocabulary-Size-Test-information-and-specifications.pdf.

Nation, I. S. P. (2014). How much input do you need to learn the most frequent 9,000 words? Reading in a Foreign Language, 26(2), 1-16.

Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, *31*(7), 9-13

Nation, I. S. P., & Gu, Y. (2007). *Focus on vocabulary*. Sydney: National Centre for English Language Teaching and Research, Macquarie University.

Normant, N. (1986). Organizational structure of Chinese subjects writing in Chinese and in ESL. *Journal of the Chinese Language Teachers Association, 21*(3), 49 -72.

Ouyang, H. H. (2000). One-way ticket: a story of an innovative teacher in mainland China. *Anthropology and Education Quarterly, 31*, 397–425.

Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Boston: Heinle and Heinle.

Özmen, K. (2012). Exploring student teachers' beliefs about language learning and teaching: A longitudinal study. *Current Issues in Education*, *15*(1), 1-16.

Paquot, M. (2010). *Academic vocabulary in learner writing: From extraction to analysis*. New York: Continuum.

Paribakht, T. S., & Wesche, M. B. (1997). Vocabulary enhancement activities and reading for meaning in a second language vocabulary acquisition. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 174-200). Cambridge: Cambridge University Press.

Parrington, A. (2004). "Utterly content in each other's company": Semantic prosody and semantic preference. *International Journal of Corpus Linguistics, 9*, 131-156.

Parris-Kidd, H., & Barnett, J. (2011). Cultures of Learning and Student Participation: Learners in Multicultural English Class in Australia. In L. Jin & M. Cortazzi (Eds.), *Researching Chinese Learners: skills, perceptions, intercultural adaptations* (pp. 169-187). Houndmills: Palgrave, Macmillan.

Philipson, R. (1984). *Learner language and language learning*. Clevedon, England: Multilingual Matters.

Pilcher, N., Cortazzi, M., & Jin, L. (2011). Different waves crashing into different coastlines? Chinese learners doing postgraduate dissertations in the UK. In: L. Jin & M. Cortazzi (Eds.), *Researching Chinese Learners: Skills, Perceptions and Intercultural Adaptations* (pp. 292-313). Basingstoke, U.K.: Palgrave Macmillan Ltd.

Poon, A. Y. K. (2013). Will the new fine-tuning medium-of-instruction policy alleviate the threats of dominance of English-medium instruction in Hong Kong? *Current Issues in Language Planning*, *14*(1), 34-51.

Qi, L. (2004). Washback Effects of the NMET. *Teaching and Researching in Foreign Language Studies*, *36*(5), 357-363.

Qi, L. (2007). Is testing an efficient agent for pedagogical change? Examining the intended washback of the writing task in a high-stakes English test in China. *Assessment in Education: Principles, Policy & Practice, 14*(1), 51-74.

Qian, D. D. (1998). Depth of vocabulary knowledge: assessing its role in adults' reading comprehension in English as a second language. Unpublished doctoral dissertation, University of Toronto, Toronto, ON.

Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge. *Canadian Modern Language Review*, *56*(2), 282-307.

Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, *52*, 53-536.

Qian, D. D. (2008). English language assessment in Hong Kong: A survey of practices, developments and issues. *Language Testing, 25*(1), 85-110.

Qian, D. D. (2014). School-based English Language Assessment as a high stakes examination component in Hong Kong: Perceptions and insights of frontline assessors. *Assessment in Education: Principle, Policy & Practice*, *21* (3), 251-270.

Qian, D. D., & Schedl, M. (2004). Evaluation of an in-depth vocabulary knowledge measure for assessing reading performance. *Language Testing*, *21*(1), 28-52.

Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing*, *10*, 355-371.

Read, J. (1998). Validating a test to measure depth of vocabulary knowledge. *Language Testing*, *10*, 355-371.

Read, J. (2000). *Assessing Vocabulary*. Cambridge: Cambridge University Press.

Read, J. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined? In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition and testing* (pp. 209-227). Amsterdam: John Benjamins.

Read, J. (2007). Second language vocabulary assessment: current practices and new directions. *International Journal of English Studies*, *7*(2), 105-125.

Read, J., & Nation, P. (2006). An investigation of the lexical dimension of the IELTS speaking test. In P. McGovern & S. Walsh (Eds.), *IELTS research reports* (pp. 207-231). Canberra: IELTS Australia.

Reid, J. (1993). *Teaching ESL writing*. New York: Prentice Hall.

Richards, J. C. (1976). The role of vocabulary testing. *TESOL Quarterly*, *10*, 77-89.

Roche, T., & Harrington, M. (2013). Recognition vocabulary knowledge as a predictor of academic performance in an English as a foreign language setting. *Language Testing in Asia*, *3*(1), 1-13.

Sandmel, K., & Graham, S. (2012). The Process Writing Approach: A Meta-analysis. *Journal of Educational Research*, *105*(1), 78-78.

Santos, T. (1998). Professor's reactions to the academic writing to non-native speaking students. *TESOL Quarterly*, *22*, 69-90.

Sasaki, M., & Hirose, K. (1996). Explanatory variables for EFL students' expository writing. *Language Learning*, *46*(1), 137-174.

Saville-Troike, M. (1984). What really matters in second language learning for academic achievement. *TESOL Quarterly*, *16*(2), 189-216.

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

Schmitt, N. (2010). Key issues in teaching and learning vocabulary. In R. Chacón-Beltrán, C. Abello-Contesse, and M. Torreblanca-López (Eds.), *Insights into Non-native Vocabulary Teaching and Learning* (pp. 28-40). Bristol: Multilingual Matters.

Schmitt, N. (2014). Size and Depth of Vocabulary Knowledge: What the Research Shows. *Language Learning, 64 (4)*. 913–951.

Schmitt, N., Dornyei, Z., Adolphs, S., & Durrow, V. (2004). Knowledge and acquisition of formulaic sequences: A longitudinal study. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing, and use* (pp. 55–86). Philadelphia: John Benjamins.

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), 1-7.

Schoonen, R., & Verhallen, M. (2008). The assessment of deep word knowledge in young first and second language learners. *Language Testing*, *25*, 211-236.

Schoonen, R., Gelderen, A. V., Glopper, K. D., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2003). First language and second language writing: The role of linguistic knowledge, speed of processing, and metacognitive knowledge. *Language learning*, *53*(1), 165-202.

Scott, M. (2012). *WordSmith Tools version 6*. Stroud: Lexical Analysis Software.

Shaw, S., & Falvey, P. (2008). *The IELTS Writing Assessment Revision Project: Towards a revised rating scale*. Cambridge: University of Cambridge ESOL Examinations.

Shaw, S., & Weir, C. J. (2007). *Examining writing: Research and practice in assessing second language writing*. Cambridge: Cambridge University Press.

Shi, L., & Qian, D. D. (2012). How does vocabulary knowledge affect Chinese EFL learners' writing quality in web-based settings? Evaluating the relationships among three dimensions of vocabulary knowledge and writing quality. *Chinese Journal of Applied Linguistics, 35*(1), 117-127.

309

Singh, M., & Fu, D. (2008). Flowery inductive rhetoric meets creative deductive arguments: Becoming transnational researcher-writers. *International Journal of Asia Pacific Studies*, *4*(1), 121-137.

Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal, 36*(2), 139-152.

Stæhr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition, 31*(4), 577-607.

Streiner, D. L., & Norman G. R. (2002). *Health Measurement Scales: A Practical Guide to their Development and Use*. Oxford: Oxford Medical Publications.

Stubbs, M. (2001). *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell.

Tang, E., & Nesi, H. (2003). Teaching vocabulary in two Chinese classrooms: schoolchildren's exposure to English words in Hong Kong and Guangzhou. *Language Teaching Research*, *7*(1), 65-97.

Thompson, G. (2001). Interaction in academic writing: Learning to argue with the reader. *Applied Linguistics, 22*(1), 58-78.

Tidball, F., & Treffers-Daller, J. (2007). Exploring Measures of Vocabulary Richness in Semi-Spontaneous French Speech. In J. J. Daller, J. Milton & J. Treffers-Daller (Eds.), *Modelling and Assessing Vocabulary Knowledge* (pp. 133-149). Cambridge: Cambridge University Press.

Tsang, W. K. (2008). Evaluation research on the implementation of the medium of instruction guidance for secondary schools. *Hong Kong Institute of Educational Research Newsletter*, *24*, 1-7.

Tsui, A. B. M. (2007). Language Policy and the Social Construction of Identify: The case of Hong Kong. In A. B. M. Tsui & J. W. Tollefson (Eds.), *Language Policy, Culture and Identity in Asian Contexts* (pp. 121-142). Mahwah, NJ: Erlbaum.

Vermeer, A. (1992). Exploring the second language learner lexicon. In L. Verhoeven & J. H. A. L. de Jong (Eds.), *The Construct of Language Proficiency* (pp. 147-162). Amsterdam: John Benjamin's.

Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics, 22*, 217-234.

Victori, M. (1999). An analysis of writing knowledge in EFL composing: a case study of two effective and two less effective writers. *System*, *27*, 537-555.

Walters, J., & Wolf, Y. (1996). Language awareness in non-native writers: metalinguistic judgments of need for revision. *Language Awareness*, *5*(1), 3-25.

Wang, L. (2003). Switching to first language among writers with differing second-language proficiency. *Journal of Second Language Writing*, *12*(4), 347-375.

Wang, Y. F. (2013). *Full-mark Compositions on the NMET*. Beijing: Beijing Education Press.

Watkins, D. (1996). Hong Kong secondary school learners: a developmental perspective. In D. A. Watkins & J. B. Biggs (ed.), *The Chinese Learner: Cultural, Psychological and Contextual Influences* (pp. 474-490). Hong Kong: CERC and ACER.

Wesche, M., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. *The Canadian Modern Language Review*, *53*, 13-40. West, M. (1953). *A general service list of English words*. London: Longman.

Wu, Q. (2008). Washback of the Writing Part in NMET. *Foreign Language Studies*, *28*, 565-566.

Xu, Y., & Wu, Z. (2012). Test-taking strategies for a high-stakes writing test: An exploratory study of 12 Chinese EFL learners. *Assessing Writing*, *17*(3), 174-190.

Yamamoto, Y. M. (2011). Bridging the gap between receptive and productive vocabulary size through extensive reading. *The Reading Matrix*, *11*(3), 226-242.

You, X. (2004). "The Choice Made from No Choice": English Writing Instruction in a Chinese University. *Journal of Second Language Writing***, **13*(2), 97-110.

You, X. (2010). *Writing in the devil's tongue: A history of English composition in China*. Carbondale: Southern Illinois University Press.

Yu, X. (2012). *Text Memorisation in Chinese Foreign Language Education*. Bern: Switzerland.

Yuan, W. (2011). Academic and Cultural Experiences of Chinese Students at an American University: A Qualitative Study. *Intercultural Communication Studies*, *1*, 141-157.

Zareva, A., Schwanenflugel, P., Nikolova, Y. (2005). Relationship between lexical competence and language proficiency: Variable sensitivity. *Studies in Second Language Acquisition*, 27 (4), 567-595

Zhang, Y. F., & Hu, G. W. (2010). Between intended and enacted curricula: Three teachers and a mandated curricular reform in mainland China. In K. Menken and O. Garcia (Eds.), *Negotiating language policies in schools: Educators as policymakers* (pp. 123-142). New York: Rutledge.

Zhao, J. H. (2009). *Error analysis on students' English writing in senior high school and pedagogical suggestions*. Unpublished dissertation, Northeast Formal University, Changchun, China.

Zimmerman, C. B. (1997). Do reading and interactive vocabulary instruction make a difference? An empirical Study. *TESOL Quarterly*, *31*, 121-139.

Zughoul, M. R. (1991). Lexical choice: towards writing problematic word lists. *IRAL*, *29*, 45-59.