



**Michigan
Technological
University**

Michigan Technological University
Digital Commons @ Michigan Tech

Dissertations, Master's Theses and Master's Reports

2017

EFL EDUCATION IN MAINLAND CHINA: WORD MEMORIZATION AND ESSAY WRITING AMONG HIGH SCHOOL SOPHOMORES

Rehema Clarcken

Michigan Technological University, rmclarke@mtu.edu

Copyright 2017 Rehema Clarcken

Recommended Citation

Clarcken, Rehema, "EFL EDUCATION IN MAINLAND CHINA: WORD MEMORIZATION AND ESSAY WRITING AMONG HIGH SCHOOL SOPHOMORES", Open Access Dissertation, Michigan Technological University, 2017.

<https://doi.org/10.37099/mtu.dc.etr/493>

Follow this and additional works at: <https://digitalcommons.mtu.edu/etr>



Part of the [Language and Literacy Education Commons](#), [Rhetoric and Composition Commons](#), and the [Secondary Education Commons](#)

EFL EDUCATION IN MAINLAND CHINA: WORD MEMORIZATION AND ESSAY
WRITING AMONG HIGH SCHOOL SOPHOMORES

By

Rehema M. Clarcken

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

In Rhetoric, Theory and Culture

MICHIGAN TECHNOLOGICAL UNIVERSITY

2017

© 2017 Rehema M. Clarcken

This dissertation has been approved in partial fulfillment of the requirements for the Degree of DOCTOR OF PHILOSOPHY in Rhetoric, Theory and Culture.

Department of Humanities

Dissertation Advisor: *Karla Saari Kitalong*
Committee Member: *Beatrice Quarshie Smith*
Committee Member: *Ronald Strickland*
Committee Member: *Zhuang-Zhong Lehmberg*
Department Chair: *Ronald Strickland*

to my students, colleagues, friends, and family

Table of Contents

LIST OF FIGURES	IX
LIST OF TABLES	X
ACKNOWLEDGEMENTS	XI
DEFINITIONS AND ABBREVIATIONS.....	XIV
ABSTRACT.....	XVIII
CHAPTER 1: THE CHINESE EDUCATION SYSTEM: A NARRATIVE.....	1
ESSAY WRITING	13
<i>Why I Believe in Teaching the Essay</i>	13
<i>Difficulty of Writing Essays</i>	17
<i>Content with Essay Topics</i>	23
CHAPTER 2: A REVIEW OF COMPOSITION AND EFL STUDIES IN THE U.S. AND	
CHINA	31
AMERICA’S NINETEENTH CENTURY NEW UNIVERSITIES AND ENGLISH DEPARTMENTS	31
<i>Emergence of English as a Second Language alongside Composition</i>	33
<i>Pedagogy of Composition since the Establishment of CCCC</i>	35
<i>Culture and Language</i>	39
THE PEOPLE’S REPUBLIC OF CHINA	42
<i>Chinese Education: Overview of the Current System</i>	43
<i>Historical Development of English Education in the PRC since 1949</i>	46
CULTURAL EDUCATIONAL INFLUENCES.....	50
SUMMARY	52
CHAPTER 3: RESEARCHING VOCABULARY SIZE AND ESSAY WRITING.....	55
RESEARCH METHODOLOGY AND DESIGN.....	58
<i>Nation’s Vocabulary Size Test</i>	61
<i>Writing Assessment</i>	61
POPULATION AND SAMPLE SELECTION	63
INSTRUMENTS	65
<i>Nation’s Vocabulary Size Test A</i>	66
<i>Greek Myths Essay Exam</i>	67
<i>TOEFL iBT® Integrated Writing Rubrics</i>	67
VALIDITY AND RELIABILITY	68
<i>The Validity and Reliability of Nation’s VST</i>	68
<i>The Validity and Reliability of the TOEFL iBT® Rubrics</i>	69
DATA COLLECTION, MANAGEMENT, AND ANALYSIS	70
ETHICAL CONSIDERATIONS	72
LIMITATIONS AND DELIMITATIONS	73
<i>Nation’s Vocabulary Size Test</i>	73
<i>Essays</i>	73
<i>TOEFL iBT® Integrated Writing Rubrics</i>	74

SUMMARY	74
CHAPTER 4: ANALYZING THE RESULTS FROM THE VOCABULARY SIZE TEST AND THE ESSAY EXAM	75
POPULATION	76
<i>Nation's Vocabulary Size Test</i>	77
<i>Essay Exam</i>	78
<i>Accuracy of Marks for Essay Exam</i>	79
<i>VST & Essay Exam Are Not Normally Distributed</i>	80
AGE: TIGERS, RABBITS, AND DRAGONS	81
<i>Age and the Vocabulary Size Test</i>	83
<i>Age and the Essay Exam</i>	83
GENDER	85
<i>Males and Females on Essay Exam</i>	87
RELATIONSHIPS BETWEEN MEASURES	88
<i>Pearson's Correlation Coefficient</i>	89
<i>Regression Analysis</i>	89
SUMMARY	90
REFERENCE TABLES	92
CHAPTER 5: CHINGLISH TO STANDARD ACADEMIC PROSE—CHINESE HIGH SCHOOLERS BECOME PROS AT WRITING ESSAYS	93
SUMMARY OF THE STUDY	96
SUMMARY OF FINDINGS AND CONCLUSIONS	98
IMPLICATIONS	102
RECOMMENDATIONS FOR FUTURE RESEARCH	104
<i>Chinese English Education</i>	106
<i>Testing</i>	108
RECOMMENDATIONS FOR FUTURE PRACTICE	109
<i>Transparency of Vocabulary Used in Learning Materials</i>	109
<i>Explicitly Teaching How to Write a Diversity of Text Types</i>	111
SUMMARY	112
REFERENCES	113
APPENDIX A: CLASS HANDOUTS AND TEACHING MATERIALS	121
APPENDIX B: ESSAY EXAM WRITING PROMPT	124
APPENDIX C: TOEFL IBT® INTEGRATED WRITING RUBRICS	125
APPENDIX D: CHAPTER 4 REFERENCE TABLES	126

List of Figures

CHAPTER 4

FIGURE 4. 1 THE VOCABULARY SIZE TEST MEASURES FAMILIARITY WITH THE 20,000 MOST FREQUENT WORDS IN THE ENGLISH LANGUAGE.	78
FIGURE 4. 2 THE ESSAY EXAM ASSESSED WRITING ABILITY AMONG PARTICIPANTS.	79
FIGURE 4. 3 MANY OF THE STUDENTS WERE 16 YEARS OLD AT THE END OF GRADE 10.	81
FIGURE 4. 4 THIS SCATTERPLOT PRESENTS THE VOCABULARY SIZE TEST (VST) SCORES (VERTICAL AXIS) AND THE ESSAY EXAM SCORES (HORIZONTAL AXIS) FOR THE TIGERS (DARK DOTS), THE RABBITS (WHITE SPOTS), AND THE DRAGON (GRAY). ...	84
FIGURE 4. 5 THE STUDY POPULATION CONTAINED 12 FEMALES AND 26 MALES.	86
FIGURE 4. 6 THIS SCATTERPLOT PRESENTS THE VOCABULARY SIZE TEST SCORES (VERTICAL AXIS) AND THE ESSAY EXAM SCORES (HORIZONTAL AXIS) FOR THE MALES (DARK DOTS) AND THE FEMALES (WHITE DOTS).	87

CHAPTER 5

FIGURE 5. 1 THE VOCABULARY SIZE TEST MEASURES FAMILIARITY WITH THE 20,000 MOST FREQUENT WORDS IN THE ENGLISH LANGUAGE.	99
FIGURE 5. 2 THE ESSAY EXAM ASSESSED WRITING ABILITY AMONG PARTICIPANTS.	100
FIGURE 5. 3 THIS SCATTERPLOT PRESENTS THE VOCABULARY SIZE TEST SCORES (VERTICAL AXIS) AND THE ESSAY EXAM SCORES (HORIZONTAL AXIS) FOR THE MALES (DARK DOTS) AND THE FEMALES (WHITE DOTS).	101

APPENDICES

FIGURE 6. 1 THIS IS THE EXAMPLE COMPARE-AND-CONTRAST ESSAY TEMPLATE.	121
FIGURE 6. 2 THIS IS AN EXAMPLE GREEK MYTH COMPARE-AND-CONTRAST ESSAY ASSIGNMENT SHEET.	122
FIGURE 6. 3 THIS IS AN EXAMPLE VOCABULARY WORD LIST.	123
FIGURE 6. 4 THIS IS AN EXAMPLE GREEK MYTH COMPARE-AND-CONTRAST ESSAY EXAM QUESTION.	124
FIGURE 6. 5 THE TOEFL iBT® INTEGRATED WRITING RUBRICS FROM ETS (2014) ARE INCLUDED WITH PERMISSION.	125

List of Tables

CHAPTER 4

TABLE 4.1 ON THE VOCABULARY SIZE TESTS THERE IS NO SIGNIFICANT DIFFERENCE BETWEEN TIGERS' AND RABBITS' PERFORMANCE.	83
TABLE 4.2 ON THE ESSAY EXAM THERE IS NO SIGNIFICANT DIFFERENCE BETWEEN TIGERS' AND RABBITS' PERFORMANCE.	83
TABLE 4.3 ON THE VOCABULARY SIZE TEST THERE IS NO SIGNIFICANT DIFFERENCE BETWEEN FEMALES' AND MALES' PERFORMANCE.	87
TABLE 4.4 ON THE ESSAY EXAM, THERE IS A NUMERICAL AND STATISTICAL DIFFERENCE IN SCORES BETWEEN MALES AND FEMALES.	88
TABLE 4.5 THE PEARSON'S CORRELATION COEFFICIENT.....	89
TABLE 4.6 REGRESSION ANALYSIS	90

APPENDICES

TABLE 6.1 THE INTERRATER RELIABILITY FOR THE THREE ESSAY EXAM SCORES	126
TABLE 6.2 THIS IS THE TEST OF NORMALITY FOR VOCABULARY SIZE TEST AND ESSAY EXAM.....	126
TABLE 6.3 T-TEST RESULTS: TIGERS' AND RABBITS' VOCABULARY SIZE TEST.....	127
TABLE 6.4 T-TEST RESULTS: TIGERS' AND RABBITS' ESSAY EXAMS.	128
TABLE 6.5 T-TEST RESULTS: MALES AND FEMALES ON THE VOCABULARY SIZE TEST.	129
TABLE 6.6 T-TEST RESULTS: MALES AND FEMALES ON THE ESSAY EXAMS.....	130

Acknowledgements

First and foremost, I would like to express gratitude to all the educators who have invested so much time and energy in my education. At Michigan Technological University, I am indebted to Karla Saari Kitalong, who has consistently worked with me throughout my doctoral degree; she was my professor in seminar classes, a reader during my comprehensive exams, and my adviser while writing my dissertation. Karla's patience and understanding during this process has motivated and pushed me towards graduation. Next, I am obliged to Beatrice Quarshie Smith, who has consistently given me excellent feedback on everything from my dissertation proposal to my dissertation manuscripts; her keen understanding of both EFL studies and composition studies has been invaluable. I am very appreciative for Ronald Strickland stepping up to be my final committee member when other members could not continue in this role; his dedication to seeing this project to completion is heartwarming. Finally, I would like to thank my other professors at Michigan Technological University, who each greatly contributed to my understanding of the world and shaped my academic interests: Dieter Wolfgang Adolphs, Marilyn M. Cooper, Elizabeth A. Flynn, Randall, R. Freisinger, Nancy M. Grimm, Robert R. Johnson, Jinfang Ren, Patty Sotirin, and Craig Waddell.

In addition, I am extremely grateful to my professors at Northern Michigan University, who have continually encouraged me throughout my academic career as an undergraduate and graduate student at NMU, and even from afar when I moved on to Michigan Technological University. Most especially, I would like to acknowledge all that Zhuang-Zhong Lehmberg has done as my outside reader; she has been a mentor to me for

more than a decade, and her timely advice is always apt. Furthermore, I must thank Mark Smith and Paul Lehmborg for teaching me how to write. Without their creative tutelage, I would never have been able to string a sentence together let alone an essay or dissertation. Likewise, I am indebted to Laura Soldner, who invested innumerable hours in my development as a scholar and the late Raymond Ventre who was a courageous advocate for my education. However, my list of gratitude would not be complete without a sincere thankyou to Jon L. Saari: he has had been the single greatest influence over the destiny of my life. His love of China enabled my family to go to Sichuan through a university exchange program when I was a mere teenager, and his passion for Asian history, which he communicated in his college courses, encouraged me to return as a young adult. Without doubt, China would not have been such an important part of my life without his direction.

Moreover, I would like to thank my professional associates who have been supportive of my career in education and my educational research. Most importantly, I would like to thank the supervisors and principals who encouraged my research by creating learning environments that allows for critical inquiry among faculty and students alike. At the International Curriculum Center at the High School Affiliated to Renmin University, I am indebted to Ned Forney, Amy Maurer, Graham Turner and Shuyan Wang; while at Beijing Foreign Studies University, I must thank Wen Cao, Di Ma, and Peter Newell. These supervisors and principals have been supportive at every turn. Finally, no acknowledgement note is complete without a heartfelt shout-out to those near and dear. I could not have written this dissertation without the loving support of my

family and friends. I am grateful to my husband Jeffrey A. Thornton, who was with me every step of the way through this journey—including countless walks where I discussed the theories of numerous linguists and compositionists until he could recite their opinions back to me. I want to thank my parents, Rodney H. Clarken and Mary Jo Wilder, who continually encouraged me to become the first female Ph.D. in the family. Also, my siblings deserve much gratitude; Ziya Neema, Lua K. Weatherdon, and John V.A. Clarken have been—each in his or her own way—very supportive of my choice to pursue a doctorate. Finally, I want to thank my friends who have been tireless cheerleaders throughout—even when I could not see them because I needed to stay home to write. Without doubt, many individuals have encouraged me on my educational journey; I am appreciative for everything everyone has done to assist me. Thank you!

Definitions and Abbreviations

ACT: is one examination used in university admissions to measure high school students' college readiness. For more information visit <http://www.act.org>.

A Level (General Certificate of Education): refers to a series of courses and examinations with affiliations to the United Kingdom's education system. In China, an A Level program usually includes IGCSE, AS Level, and A Level coursework and examinations. For more information visit <http://cie.org.uk>.

AP (Advanced Placement): is a program is a series of courses and examinations run through the College Board that prepare students for US university coursework. For more information visit <https://aphighered.collegeboard.org>.

BDC (背单词 Beidanci): means recite vocabulary words. It is a way of quickly memorizing a vast number of English words and their Chinese translations. Resource guides or reference books (such as *TOEFL Vocabulary Root + Associative Memory Method* and *IELTS Vocabulary Root + Associative Memory Method*) are readily available throughout China. The most popular books among my students are the ones published by New Oriental, the largest test-prep center in the China if not the world.

China: In this dissertation China refers only to the mainland portion of the People's Republic of China which recognizes the capital in Beijing. I do not include Hong Kong, Macau or any other jurisdiction with different government administrations, currencies, and languages. These outside areas have unique histories with different international influences which make comparisons to the education system in mainland China problematic.

Chinese education system, The: is divided into kindergarten (幼儿园 you'eryuan: approximately 3-6 year olds), elementary school (小学 xiaoxue: grades 1-6), middle school (初中 chuzhong: grade 7-9), high school (高中 gaozhong: grade 10-12), and college (大学 daxue: bachelor's master's, doctoral education). Currently in China attendance is required through middle school (Grade 9), though a few decades ago only elementary school was mandatory. Academically poor students, or those who do not test well, and financially poor student, or those without money to pay tuition fees, may attend poor high schools, those lacking in physical and professional resources, but will seldom have the opportunity to attend colleges or universities.

college: is synonymous with university (大学 daxue) in American English. It refers to an institution that grants bachelor's degrees and other tertiary degrees. I do not use college to refer to high schools as it is used in other Englishes.

Common European Framework of Reference for Languages (CEFR): This is a mechanism to discuss and assess linguistic ability in forty different languages. It was developed in Europe and has established benchmarks to evaluate fluency, which are now being adopted around the world. For more information visit <http://www.coe.int>.

EFL (English as a Foreign Language): is distinctly different from ESL (English as a Second Language). EFL is taught in countries where English is not a local language to students who do not have regular exposure to the language outside of the classroom or school environment.

elementary school: is used synonymously with primary school and 小学 (xiaoxue). In China, elementary school includes grades 1-6.

ELL (English Language Learner): is someone who is learning English as a first, second or additional language. Usually, the individual is learning or is fluent in another language and is living in a multilingual community where other languages are regularly used.

ESL (English as a Second Language): is taught in situations where the legal or dominant language of the country is English, and where natural English environments are accessible to the wider community outside the classroom.

Gaokao(高考): is the Chinese university entrance examination. It is a standardized examination that tests students' abilities in Mandarin, English, mathematics, and humanities or science. It is administered once a year in June, and its scores determine the future of students' lives. Based on performance, students are accepted to different universities and programs of study throughout China. Compared to the US admissions process, very little choice is granted to the individual students about which university they will attend and which program they will study. It is common for a student to be forced to choose between a desired university in a major city and a desirable discipline of study at a less prestigious institution in smaller cities.

Huikao(会考): is the Chinese high school exit examination. It should not be confused with the Gaokao (高考), though both are taken by Chinese high school students, as they have considerably different purposes. From a students' and teachers' perspective, the Huikao is a simple yet important exam to complete high school; but the general feeling among most college-bound students is that it is not much more difficult than the Zhongkao, the high school entrance exam. For this reason, the different subject tests are often scheduled by the school to be completed in tenth or eleventh grade. However, to prevent early graduation, schools leave a few Huikao subject exams (such as PE and Chinese Literature) until the end of the Grade 12. Then, most of junior and senior year are spent preparing (cramming) for the Gaokao.

IB (International Baccalaureate): is an educational program offered in schools around the world. It is a standardized curriculum with external moderation and assessment that certifies quality of education and standardized results despite national, regional, and local educational requirements. For more information visit <http://ibo.org>.

IELTS (International English Language Testing System): is an English language examination to test proficiency in reading, writing, listening and speaking. It is mainly used in the United Kingdom and those English-speaking countries with strong ties to the UK. For more information visit <http://www.ielts.org>.

junior: is used synonymously with Grade 11 students

L2 (second language): For my purposes, an L2 is any language that is not the dominant or first language used at home. It is any additional language (third, fourth, fifth, etc.) language learned outside of the home usually in a school or professional setting.

middle school: is used synonymously with junior high school and 初中 (chuzhong). In China, middle school includes grades 7-9.

high school: is used synonymously with senior high school and 高中 (gaozhong). In China, high school includes grades 10-12.

preschool: is used synonymously with kindergarten and 幼儿园 (you'eryuan). In China, preschool includes children who are approximately 3-6 years old.

productive vocabulary knowledge: are the words that a language learner can appropriately use when writing or speaking

receptive vocabulary knowledge: are the words that a language learner can understand when reading or listening, but is unable to independently produce.

SAT: is an examination used in university admissions to measure high school students' college readiness. For more information visit <http://www.sat.org>.

senior: is used synonymously with Grade 12 students

sophomore: is used synonymously with Grade 10 students

TESOL: is an international association of teachers of English to speakers of other languages they have a variety of forums for discussing matters of interest to professionals such as publications and conferences. More information can be found on their website: www.tesol.org.

TOEFL: is an English language examination to test proficiency in reading, writing, listening and speaking. It was developed in the United States and tests US English.

However, it is now accepted around the world as proof of one's English-language level.
For more information visit <http://www.ets.org/toefl>.

Abstract

This dissertation explores English as a Foreign Language instruction within the context of the contemporary Chinese education system. Basic outlines chart the historical development of EFL studies in the United States and China framing the question of what each community values as important measures of success when assessing language learning. While traditional Chinese methods value strict memorization of vast word lists (背单词, BeiDanCi, BDC) the US educational community stresses essay writing—particularly on standardized tests such as the ACT, SAT, and TOEFL, which are required for university admissions. Therefore, this study investigates the relationship between vocabulary knowledge and writing ability among Grade 10 Chinese high schoolers in a megalopolis in mainland China. Students' vocabulary knowledge was measured with Nation's Vocabulary Size Test, and students' writing ability was assessed with an essay graded using the TOEFL iBT® Integrated Writing Rubrics. The results validate previous findings among different L2 populations by observing a moderate correlation between vocabulary knowledge and writing ability.

Chapter 1: The Chinese Education System: A Narrative

In the autumn of 2012, I moved with my husband to a megacity in China to teach English at a public high school with an American education program. This was my third time moving to China to live and work. As a teenager, I spent a year with my family in Sichuan; my father had been granted a sabbatical to research the Chinese education system, so we moved to Chengdu, Sichuan as part of a sister-university exchange program. After university when I was in my early twenties, I returned to China as a Peace Corps Volunteer; I lived, taught and studied in both Mianyang and Chengdu, Sichuan. Then in my thirties, after being married for a few years, my husband and I decided to try living and working in China again. In total, I have spent 10 years living, working, and studying in China—five years as a youth and five years continuing as an adult.

This new teaching job in a metropolis in China was very different from my former teaching experiences at high schools and universities in both China and the United States. My previous work had been among middle-class, rural and small-town students typical of where I had taught in Sichuan and Michigan; these students were interested in learning in my class in proportion to the usability of the information on their path to graduation and finding gainful employment. Generally, if I could show how the advanced reading and writing skills were necessary in future endeavors, they were good pupils.

However, my new classes of students in metropolitan China were full of individuals I did not immediately understand. Overwhelmingly, the school attracted affluent children of Chinese business people and bureaucrats who wanted to secure a better education for their offspring than they had had for themselves growing up in a

China that was less prosperous and developed and had fewer educational requirements and opportunities than today.

To give some background, students in China are only required to attend school until the end of middle school. Just over half of those who graduate from ninth grade are privileged enough to enter a regular high school where they will study for three more years; the remainder go on to technical schools or discontinue their education (Ministry of Education, 2015, “Number of Students”). Because a high school education is not required by the government, there are not enough facilities for everyone who wants a diploma. Thus, it is extremely competitive to attend publicly funded institutions and rather expensive to attend private ones.

When I started teaching in China in 2012, I taught many students whose parents were wildly successful (financially and in their careers) despite not having achieved accolades academically. This caused many tensions: most notably, the value of an education as a road to a prosperous future was disproven by the omnipresent examples of their family members and classmates’ relatives. Consequently, teaching methods which inspired a good work ethic in the socio-economically striving students I had taught in the past backfired in this new educational climate. This new work environment caused me to reevaluate how and what I taught as my former pedagogical approach did not seem to be as effective as before.

The students who did strive in my classes seemed to do so for very different reasons from those of the middle-class students I had taught in the past. These metropolitan Chinese often had parents who encouraged their education as a path to

higher and more distinguished levels of social status. Entering an elite international university, that their friends' children could not, would bring admiration in their parents' social circles. In essence, a child's academic achievements would be held up as accolades to broadcast the family's genetic strain of intellectual genius and the success of their enlightened parenting skills. The phenomenon of using education as a purchased gem in a crown—without any need to use what was learned to procure employment or financial success—was initially quite perplexing to my middle-class sensibilities that saw how hard my grandparents had worked on the farm and how my father's educational prowess provided him with an escape to an easier career path as an academic. However, I, eventually, realized that the only way my students would value their expensive schooling was if they had wise tutelage (hopefully, from my colleagues and myself) that pushed them to see the world differently.

The school in which I found myself in that fall of 2012 was one with students on the extremes of a continuum which went from those who did not want to learn because they believed that they did not need an education to succeed in the world and those who trusted how an education would bring social status associated with higher, more elite, and more exclusive levels of schooling. The first group could only be reached when I tricked them into having so much fun in the classroom that they dropped their guard long enough to learn something, while the second group had been trained through hours, days, months, and years of rigorous, boring drills that learning was difficult and painful and could not and should not be enjoyable, so they would become wary when the classroom activities became too much fun.

The complicated cultural milieu of apathetic students and educationally driven pupils created a great number of tensions in the classroom and school. As a Western-trained teacher, I tried to awaken the intellectual curiosity among the poorly performing but often quite bright students through engaging activities while holding the attention of the book-smart classmates who were uneasy with applying knowledge in creative ways. By using games, group work, music, posters, presentations, video clips and other activities, I tried to reach students through multiple senses and different learning styles.

It quickly became apparent that my pedagogical practices were at odds with the Chinese philosophy of education. This observation is supported in a study by Fang and Warschauer (2004) that found Chinese faculty avoided implementing new technology and projects in their courses because they “do not accord with norms of Chinese higher education” (p. 308). Even though the parents and students had chosen the school because they knew an American-style education would be refreshingly different (and hopefully better) from the traditional Chinese options, they did not completely understand the fundamental incongruences between the two educational models. Their invested tuition fees assumed a different wrapping on a familiar product—a high school education—so the Western-style of teaching and assessing brought unease.

To summarize what the traditional Chinese philosophy of education is could only be an oversimplification of nearly fifty centuries of history and culture. From before the Mohist until after the Maoist, there have been many diverse schools of thought. Briefly, the Mohist were established by Mo Tzu (470 BCE— 391 BCE) in the century following Confucius. They are credited with influencing the realms of science, religion, politics,

and philosophy (Needham, 1954) while the Maoist movement was established by Mao Zedong (1893-1976); he established a new kind of Marxist-Leninism for modern China (Ministry of Education, 2015a).

Despite the huge variety of influences throughout China's long history, there are some traditional ideas and practices still widely used today: teaching is most often done through lecturing; knowledge is acquired through reading, memorization, and recitation exercises and drills; and students are most often assessed through paper-based examinations with multiple-choice and gap-fill questions. To be sure, these practices are not ubiquitously implemented across all learning environments in China nor are they unquestionably applied; however, they are regularly seen and accepted in most schools in China. It should be noted that these teaching methods are not unique to China; they are used around the world and can be observed in part or whole in nearly every school on the planet.

Adams and Sargent (2012) describe the prevalent school learning environment found in China at the end of the millennium:

The traditional model of learning focuses on mastery, practice, rote memorization and discipline. It emphasized the role of the textbook and the teacher as the source and conveyor of knowledge. Students are expected to listen and learn the materials, to accept the information presented by their teachers, and reproduce the knowledge on high-stakes examinations. (p. 4)

More recently, the Chinese government recognized that this teacher-centered mode of education was not preparing students for the needs of the future, so the New Curriculum policy, which encouraged student-centered classrooms, was introduced. In the decades since that time, some progress has been made towards more inclusive learning

environments; however, much more teacher training needs to take place for the student-centered learning to be widely implemented throughout the country (Zhao, Zhu, and Liu, 2008).

One way to establish values in a culture or organization is to look at how time and resources are spent. Therefore, a description of important similarities seen in schools from Sichuan to Beijing will highlight commonalities across these educational environments. In highly competitive public institutions with local, if not national, reputations for excellence, there are similarities in the way they organize space and time, which hint at larger Chinese educational values. Furthermore, though teachers and schools across the country are making changes, most schools follow specific government guidelines and educational initiatives.

In terms of the physical design, classes are arranged to best accommodate teacher-centered, lecture-style courses. Classrooms have a very clear front indicated by a large blackboard or whiteboard and a small stage with a podium (industrial education model/proscenium classroom) and neat rows of forward facing desks and chairs for forty or more students. Because of the size of the rooms, the amount and kind of furniture, and the number of people occupying the space during a class period, it is nearly impossible to arrange the learning environment in a non-lecture format. At best, pair work and small group work can be organized, but there is neither enough space between the rows of desks nor enough time in a class period for a teacher to walk around and easily monitor each student's or group's progress.

Schools organize the school year according to a rigorous testing schedule. Each level of education—elementary school, middle school, high school, and college—begins and ends with an examination; admittance and graduation are determined by these tests. Therefore, yearly, semesterly, quarterly, and monthly exams are administered to check progress and assess preparedness. Students are ranked and ordered by the results with names and grades being posted in the classroom and new seating arrangements being assigned to place the highest-scoring students in the coveted front rows. Nevertheless, because seating is under the control of the teacher, a variety of methods for arranging students have been implemented in different classrooms.

In large schools with multiple classes at the same grade-level, students are tracked by their examination marks; students who score higher or lower can be removed from one class and reassigned to another as a reward or punishment for their performance. In such a situation, class numbers and classroom spaces are used to designate rank and prestige; for example, Class 1 is populated by the highest ranked students and is placed in the most conveniently located (closer to stairwells that lead to the cafeteria and sports fields) and well-equipped classroom (nicer furniture and better audio-visual equipment). Class 1 is encouraged to higher levels of achievements with lessons taught by the best (most highly-educated and experienced) teachers. The students in the lower ranked classes are expected to compete with fewer advantages (more students in the classroom, more inexperienced teachers, etc).

Such spacial and temporal arrangements have been designed as an efficient means for teaching masses of students a great deal of information in a short time period. With

approximately twenty-five million students in regular high schools each year (Ministry of Education, 2015c), it is only pragmatic for the school system to be arranged in such a way. However, this organization puts a great deal of pressure on students wanting to succeed. Information is devoured in preparation for exams; learning becomes synonymous with memorization; content need not be understood, but only internalized long enough to be used when answering a question on a test. To meet the high demands, teachers use a lecture format which most efficiently communicates information, and then the teachers test retention with a series of mock exams. Thus, the lecture-test cycle is perpetuated.

In China and in America, there are significant theoretical differences among educational practices. As a new foreign teacher in my school, I was asked to defend my pedagogical practices to irate parents who were upset that their child was suddenly not as good a student as he or she had been in middle school. Unwittingly, I had gone against expectations since my class assessments were not solely based on examination scores but included class participation, activities, and projects. One point of contention concerned the senior research projects, where students were assigned to write an original research paper in MLA or APA style. Improper citing, poor paraphrasing, and unabashed plagiarism were serious hurdles to passing. In an effort to turn errors into opportunities for learning, many students rewrote their papers multiple times to edit, revise and remove plagiarized portions. However, the students and I both underestimated the time, energy, and skill necessary to write a research paper—this failure of both teacher and student

gave ample opportunity for discussion and reflection on the requirements of academic writing.

Additionally, parents and students were furious that the poorer performing students in my classes were given so much of my attention in and out of class. Effectively, I violated the rules of the traditional Chinese education system where students are ranked by proven academic success which is measured by standardized testing administered regularly; moreover, a teacher's attention should follow predictable cycles of assessment where the top students won more time to be mentored. Culturally, it is acceptable for teacher's favoritism to be used as a reward or withheld as a punishment in response to students' grades. Therefore, my new assessment methods and egalitarian attention to all students was inappropriate and worthy of comment by senior administration.

Fortunately, my immediate supervisors were very protective of the international staff and their Western teaching philosophies, so they often provided us with support and encouragement in the face of frustrated parents. However, this only exacerbated the contradictions inherent in an American educational program being run in a Chinese high school. By failing to comprehend the theoretical gulf between the pedagogies, the various constituents—administrators, teachers, parents, and students—became locked in an interminable battle of wits, cultural expectations, and bilingual miscommunications.

As a teacher in the classroom, my more immediate concern was attending to the tensions between my apathetic students and the educationally driven pupils. When I conducted the class in a traditional lecture format with Socratic questions, I had issues

with apathetic students getting restless and talking or playing practical jokes. When I used more interesting activities where students played games or worked in groups, I would find my educationally driven pupils reading English books under their desks. To address the two forms of acting out, I implemented new classroom management systems; this greatly reduced disruptions by the apathetic crew, but my driven ones still insisted on studying their books under their desks.

I used several ideas from two different education manuals to regain more control over my classroom: *The First Days of School* (Wong, et.al., 2014) and *Tools for Teaching* (Jones, Jones, and Talbott Jones, 2007). For example, both books discussed how classroom arrangements could dramatically impact teacher control and student discipline. Basically, the organization of the physical space and the students within that area make disruptions to instruction more difficult. Because the classroom layout was set by the school, I created new seating charts to disperse hotspots of trouble and encourage learning partnerships and groups. The mostly random seating charts I had had before did not work well because of my lack of familiarity with the students' personalities. Therefore, I asked older teachers who had been at the school a few years longer and knew the students well to advise me on student placement. At least three teachers advised me on each class's new seating arrangement. This dramatically reduced disruptions.

However, an unintended consequence of my taking-back-the-classroom initiative was that the students who were performing poorly (often because they were too busy doing other things in class to pay attention) were near the front of the room and the students who were the best and brightest were at the back of the room. Furthermore, I

was regularly checking on the progress of the troublemakers to make sure they knew I was carefully watching them so that they would think twice about acting out. This arrangement—poorly performing students in the front of the room with extra teacher attention—was in direct opposition to what the students were used to.

Several exceptional students complained about my teaching because I was not rewarding them for their natural abilities and work ethic as they had come to expect. Truthfully, I struggled with the justice and ethics of this cross-cultural situation. Some of the questions I asked myself were: Does every student have equal right to learn the content of the course, or should I teach to the interested? Should teachers spend an equal amount of time with each student, or more time with students who need more help? Should teacher attention and seating arrangements be used as rewards or punishments? How can I explain, to an upset student or parent, that a smoothly functioning classroom will, eventually, work better for everyone involved because less time spent on discipline will be more time spent teaching and learning?

Now, turning to the other portion of my disgruntled student population: my bright pupils who were quietly and stealthily acting out by reading English dictionaries compiled for Chinese students cramming for the TOEFL and SAT examinations. These books from the New Oriental¹ conglomerate, printed with traffic-light red and green covers, were my warning flag, my siren, and finally my door to discovering the unarticulated expectations and motivations my students had for learning English. These vocabulary learning manuals were articulating an unspoken belief in the power of BeiDanCi (BDC, 背单词), or word memorization as an integral part of language

learning. My students were using their actions—reading the New Oriental study guides under their desks during class—to convey their fear that what I was teaching them in our English as a Foreign Language (EFL) course, was not adequate: vocabulary learning is important to them too.

And, what was it that I was trying so hard to teach them? Essay writing. For as much as I did not understand their desire to memorize words as a path to learning English, my students (and their parents too) did not understand my zeal for teaching composition. However, I strongly believed that for them to succeed in the international universities which they so coveted to attend, they must learn to write—if only to complete their ACT, SAT, and TOEFL exams and write their application essays for the admissions offer.

This process of reconciling the tensions in the classroom and in my school between the disruptive underachievers and the motivated overachievers led me to research vocabulary acquisition and composition practices among Chinese high school students studying in an international education program. This specific research focus developed because of my continual interaction with the Chinese education system as a teacher in the classroom, and it came out of my desire to understand the complexities of how and why the US and China differed in their assumptions and approaches to EFL education. My day-to-day and year-to-year teaching experience was the backdrop for my dissertation research. What I soon fell into was the gap: the small amount of research about English language acquisition among Chinese high schoolers in mainland China.

Therefore, I chose to conduct a study on vocabulary knowledge and writing abilities among Chinese EFL sophomores.

Essay Writing

This dissertation developed from my experience teaching English to high school students in China where two different ways of thinking about language learning crossed paths. As a compositionist, I believe in the importance of learning to write essays in English because it teaches one to put ideas together in a coherent, organized fashion; however, Chinese students insist on the importance of sight-memorizing English vocabulary because they want to express themselves with the most precise words even when those words are archaic and incongruent with the register of the text or when they have subtle emotional nuances that do not match the tone of the writing. Early on, I wondered how these two different ways of learning a language informed each other, so I endeavored to set up a research project to determine what kind of relationship vocabulary learning had to writing essays. What came out of this exploration was not only some statistically analyzed test results but a reflection on essay writing and vocabulary acquisition as components of a high school EFL curriculum in China.

Why I Believe in Teaching the Essay

For more than a decade I have been a composition instructor for novice and developing writers with native, L2, and EFL English capabilities, which has led me to appreciate the importance of teaching essay writing in all my English courses. Before continuing, I must address an essential question: what is the essay? I do not believe in

forcing topics into a 5-paragraph-essay mold, but I do believe that an essay is like an accordion that can grow and shrink to accommodate the phrasing of the music or the content of the prose. My definition of an essay is a nonfiction piece of writing with a beginning, middle and end that conveys an implicit or explicit message from the writer to the reader. Because of the flexibility of the form, it is a good basic writing style to teach and learn because it can be used to structure a multitude of everyday writing tasks such as cover letters for jobs, response essays on exams, emails to friends, descriptive writing in storytelling, and even sharing ideas through blogs.

Nonetheless, I must acknowledge that there is a regular debate among writing teachers about the place of the essay in a composition course. On TESOL's blogosphere, there was a discussion about the usefulness of teaching the essay sparked by Brian Sztabnik's (2015, December 13) article which is a call to authentic writing which he believes should not be confined by the essay form. Robert Sheppard (2016, January 4) replied with a piece defending the essay as one component or stage of a writing curriculum which brings learners from writing simple grammatical sentences to expressing ideas in lengthy works. However, the debate continues with Nigel Caplan and Luciana de Oliveira (2016, February 12) weighing in with another post criticizing all essay writing as generic, tone-deaf 5-paragraph essays written by novices without voice and agency.

This debate shows that there are clear divides in opinions which come from educators' different place in the hierarchy of education from primary through tertiary schools and the different purposes for teaching English writing from preliminary practice

with forms and styles through authentic writing situations. Though many educators dislike the strict structure of the 5-paragraph essay as shown by this recent blog-post debate, many instructors still teach this structure and other variations of the essay because they find it to be useful for students' development as writers. Without doubt, different circumstances require different writing forms: a creative writing class may teach narrative arcs, communication courses may favor journalistic styles, while academic subjects may teach research citations.

This debate shows that no matter which side you agree with, the essay holds a place in the consciousness of many English writing instructors, and whether or not you decide to teach it is a choice to be considered. One aspect of this debate which this blog discussion did not broach is the place of teaching essay formats in an international context. For example, in a nonnative English environment in a culture with completely different expectations for written expression, in general, and literature, in particular, such as China, it is essential to expose students to the essay to show a contrast to what they are expected to write in their Chinese-language composition courses.

This conversation is broached by Sullivan, Zhang and Zheng (2012) in "College Writing in China and America: A Modest and Humble Conversation, with Writing Samples" when they compare typical college-level writing samples from two university students, one Chinese and one American. The Chinese student wrote a *sanwen* or "free-topic lyrical prose text of approximately two thousand words" (310) focusing on her hometown while the American student wrote an essay about Thomas Jefferson as a historical figure as discussed in three different biographies.

The case study of two writing pieces completed by two college-level students on two continents details the very different cultural expectations behind the writing tasks. In summary Zheng explains:

American writing values logic, reasoning, and self-discovery, while Chinese writing values beauty of language and the moral message delivered in writing. This observation does not apply to all situations, of course; nevertheless, it does catch at least some critical characteristics of the two rhetorical traditions. (324)

This example illustrates the very real differences between Chinese-language-and-culture compositions and English-language and American-culture writing pieces. For novice EFL students, dissimilar writing expectations must be made explicit to enable students to intentionally use forms that are appropriate for and expected by their audience.

For EFL high school students, learning how to write an essay is an essential milestone in their English education. There are many excellent reasons to teach these students how flexibly to use the essay structure to meet the demands of multiple writing tasks which they will encounter during their high school years while preparing to enter top US colleges. First, they must prove their English proficiency by writing two essays on the TOEFL exam; next, they will prove their academic ability with the extended essay in the SAT exam or ACT exam; then, they will further impress a top university by taking several AP exams which often require one or more essay responses as part of the document based questions; finally, they will compose a number of essays as part of college admission applications.

These essays are assessed using a variety of criteria. Some institutions prefer a standard 5-paragraph essay in response to the question prompt (independent writing), while other institutions expect a longer essay with references to reading and listening material discussed in the content (integrated writing), so students must learn to write with flexibility. Plakans (2010) discusses the different processes individuals go through when given these two dissimilar writing tasks on the TOEFL exam and she concludes that students must be adequately prepared for both kinds of essays in order to successfully complete the required tasks and pass this required exam for international students intending to study in the United States.

Looking at the population of student participants in my study, I estimate that most will have written a minimum of 10 extracurricular (test and application) essays while preparing for university abroad. One student complaining on a WeChat post showed that he had written more than 78 essays for his Common Application submissions (Tiger, 2017). It should be remembered that all the exam essays will have been written in high-stakes, timed-test environments while the admissions essays would have been revised to near-native levels of English proficiency. Therefore, teaching essay structure and requiring EFL students' mastery is essential for them to surpass these hurdles and be admitted to college in America.

Difficulty of Writing Essays

In 2012 when I returned to China to teach high school students, I struggled to teach my students how to write adequate essays. I had been tasked with helping the juniors and seniors prepare for the 2013 SAT writing exam, which entailed writing a 250-

word essay in 25 minutes in response to a question prompt. It should be noted that the SAT changed the format of their essay section of the exam in January 2016 to make it more challenging; furthermore, the new format makes it more difficult for students to get points for prepared or memorized essay responses.

My experience of teaching English in a Chinese high school at times seemed only one step removed from teaching at a test-prep center. Many Chinese parents put pressure on the schools to add preparation materials for extracurricular English exams such as the TOEFL, IELTS, SAT, and ACT to the high school curriculum because of the pervasive anxiety about international university entrance requirements. Many ICC English Departments, at one time or another, must explain how their course materials prepare students for these exams. As the English Department Head, I regularly had to discuss the English curriculum with the principals to assure them that the students would be ready for the external exams; in addition, I presented this same information to parents at the annual meetings and individually met with parents who had concerns about their child's English abilities and performance. Whenever suitable, the English Department chose textbooks with some connection to these exams; for example, one book had another supplemental teachers' text with additional activities to further tie each chapter with the TOEFL exam. When the publishing companies did not have readily available materials like this to consult, the English team worked to create lesson plans with explicit links between the coursework and these examinations' content to soothe the students' and parents' test anxieties.

While teaching juniors and seniors SAT writing (for the former, discontinued exam), I discovered that besides having the normal EFL struggles with grammar, vocabulary and fluency, they had few ideas about what to write about in their essays. Many students were reticent to use relatable stories from their daily lives as examples in their essays; moreover, they had little ability to discuss current events or cultural and historical events with any eloquence or depth. Consequently, they relied on a series of short nonfiction narratives about important world political and business leaders provided in an easily purchased book written by a major Chinese test-prep company (Dan & Liu, 2014). My students would read, memorize, and regurgitate the books' example essays for the in-class mock exams.

Unfortunately, their recall of the scenarios was so awkward and formulaic it led me to question several students about copying each other's work and plagiarizing the book. However, this book had another drawback. The material was written by a Chinese author for a Chinese audience which cast the information in a particular way that was uncomfortable for Western readers with sensitivities to historical accuracy or inequality in society—especially when discussing gender, race, and class.

Strikingly, my colleagues and I read far too many essay that praised Adolph Hitler as a great leader. Yes, you read that correctly: many students' essays considered Hitler to be famous not infamous. Time and time again, they would recount his ability to transform a nation, not his proclivities for genocide. Students would argue he rebuilt his country after WWI and despite losing WWII has been influential over the second half of the 20th Century. The problem with all of these essays that mentioned Hitler was their severe lack

of historical and cultural context. It seemed students had only learned about Hitler in English from these short biographies in the aforementioned test-prep book, so they were unable to critically discuss his place within the larger context of WWII, Europe, and the world. For example, they had little understanding of how Germany and Japan cooperated during the war nor the parallels between their countrymen's suffering in concentration camps and the Jewish holocaust.

Horribly, students often wrote rigid 5-paragraph essays with one of each of the three body paragraphs dedicated to a different "great" leader. They followed a formula where each body paragraph was a discrete entity unto itself that only had to loosely meet the criteria of the prompt, which they often forced into the theme of successful leadership whether the prompt had anything to do with these two concepts. Consequently, it was not uncommon to see a paragraph about Hitler placed into an essay with paragraphs about Martin Luther King Jr and Steve Jobs. This was not an ironic compare-and-contrast essay, but it was a testament to the serious lack of cultural understanding about the historical and cultural significances of these men. Gradually, the sheer number of times I had to explain to an otherwise good English student why this kind of pairing in an essay was inappropriate led me to become much more specific in my instructions on essay writing and my information about the cultural significance of topics discussed in class.

After recovering from the shock of exactly what my students (who were planning to attend Western universities) did and did not know, I became one of many to question English education within their school, their country, and the world. In particular, I

noticed the multiple ways in which the Chinese education system in which I worked seemed to value getting into university more than learning. This tension between stated goals and assumed outcomes influenced the design of all future courses within the English department which I ran as the department head. Therefore, each year the courses within my purview changed in ways that I hoped would better prepare the students linguistically and culturally for an international university experience.

As for the question of historical context and interpretations, I am not quite sure why there was such a positive analysis of Adolf Hitler among my Chinese students' essays. There are many possible reasons the students expressed these notions such as an inability to express complex ideas about war and society in English as a Foreign Language or a lack of knowledge about historical events. For example, a simple linguistic concern is that students lacked an appreciation for the difference between "famous" and "infamous." However, for a historical explanation, Breitman (1990) quotes a portion of a speech given by Adolf Hitler in 1939 which may speak to the heart of the difference in interpretation: "Genghis Khan had millions of women and children killed by his own will and with a gay heart. *History* [author's emphasis] sees only in him a great state builder..." (p. 337). This explanation of Genghis Khan by Hitler would also seem to apply to the Chinese students' interpretation of Hitler in their essays. I very much wish more research to be conducted in China about Chinese views of Hitler to clarify the different perspectives taught in school. It would also be interesting to understand how the Chinese view what happened to the Jews in Europe in comparison to the use of similar genocidal practices that were used by the Japanese against the Chinese.

Honestly, my immediate response to reading the in-praise-of-Hitler essays was “make it stop!” Therefore, before I could implement extensive curricular changes, I issued an ultimatum that all students who mentioned Hitler in an in-class SAT mock essay would automatically fail the assignment. Then, I had a few difficult conversations with classes about my ultimatum explaining how their current uncritical representation of Hitler may negatively sway the SAT reader, and consequently, their grade on the essay writing section of the exam, whether or not that was actually true. This emergency decision only bought time during that first year of teaching, so I could implement new course materials to address my students’ need for cultural understanding. One such change was the implementation of a unit which addressed WWII and the holocaust with texts such as *Night* (Wiesel, 1960) and *Maus* (Spiegelman, 1986).

Now, let me take a moment to say that I do not believe in the supremacy of the US education model in which I was raised and trained, nor do I think that any one person or group of people should export their views of the world to the world through imperialist practices. Nonetheless, as an American English teacher of Chinese students who had chosen to work in a high school program that grooms individuals for universities abroad, it became my contracted obligation to prepare my students by instructing them in the language and philosophy of the country in which they wished to study. I found myself as an unwitting colonizer of minds. I grappled with the question of how I could prepare my students for their American examinations (TOEFL, SAT, & ACT) and their university experiences abroad without negatively influencing their Chinese identity. This is an on-

going discussion which cannot be resolved within my dissertation research, but which I hope many more people will deem worthy of research and writing.

Content with Essay Topics

It became abundantly clear that my students needed just as much help choosing what to write about in their essays as they needed guidance on the mechanics of how to write an essay. Without a common frame of reference and a shared body of knowledge, it would be difficult for students and teacher to get on the same page. That is how my quest began to find material to form the new curriculum for the incoming sophomore class.

It is difficult to teach or study material that is neither interesting nor exciting. Choosing engaging books to use with sophomore-level EFL students is demanding because so many of the resources suitable for students' language level are not suitable for their developmental sophistication. Finding the "Goldilocks Zone" (just right curriculum) for both language level and intellectual content requires diligence and creativity. From the students' perspective, books which match their reading level may seem quite childish, while books that are intellectually and academically engaging contain challenging vocabulary and grammar which can be overwhelming; therefore, I endeavored to create coursework which was sufficiently linguistically challenging as well as educationally viable. The final hurdle was finding a book that could be purchased in China and used in a Chinese high school.

This entailed endless discussions with the textbook distributor who solicited publishers and negotiated with the government to ensure compliance—meaning the book was not banned in part or whole and it could either be printed in China or shipped to

China before the start of the new school year. Several textbooks were either unavailable for publication in China because of copyright law violations or distribution difficulties. When books were available, they had to first be cleared by the government for use. Several of the school's textbooks arrived with pages missing or stickers covering "inappropriate" content. Sometimes pictures were covered while other times entire sections of chapters containing information deemed sensitive by the Chinese government were removed. Usually, the banned content was historical or political information about China that was not believed to be accurate or appropriate by the Chinese government agencies.

It should be noted that the school had adopted Pearson Longman's *NorthStar* Series² as the basic EFL textbooks for the language courses. This textbook covered the typical reading, writing, listening and speaking activities and was adequate for teaching many of the basic skills necessary to become proficient in English. For writing practice, the book had sections on sentence-level grammar, paragraph development, and essay structure that provided admirable scaffolding along the way. However, this series of textbooks were not enough to fill the hearts, minds and time of the students. They needed extended, fictional reading materials to supplement the nonfiction articles predominating the *NorthStar Reading and Writing* textbook.

Before discussing the particulars of the supplemental text chosen, it is important to discuss the particulars of the content. Why did I choose to study Greek mythology over other possible topics such as Chinese mythology, a collection of fairy tales, or

contemporary novels? Many of these other possibilities did cross my mind; some were adopted while others were shelved for one reason or another.

First, I reviewed the book *Chinese Fairy Tales and Fantasies* (Roberts, 1979) which contains a good collection of Chinese mythology. Unfortunately, after some consideration, I found the topic was not suitable because the Chinese pantheon is too vast and varied with many regional and local difference in deities and stories. Furthermore, Chinese myths do not form good examples to be used in response to generic TOEFL and SAT essay prompts because each story requires too much background explanation for a small point to be made. Western readers are not familiar enough with the stories so a simple name and a few identifying details of the deity and myth are insufficient to add support or prove a point in an essay.

Next, I reviewed a collection of Western fairy tales, but this idea was dismissed when no suitable book presented itself. Working with the publishers and distributors, it was difficult to find something that was appropriate and obtainable in China. Lastly, a range of novels were considered, and the English faculty agreed that multiple nonfiction and fiction books should be used throughout the sophomore, junior, and senior curriculum as supplemental materials to the EFL textbooks.

Thus, the sophomore curriculum adopted *The Short Stories of Sherlock Holmes* (Doyle, 1992) and *The No. 1 Ladies Detective Agency* (McCall Smith, 2002). These books were chosen because of students' familiarity and unfamiliarity with the content respectively. They were also chosen because they can be easily used to help students write descriptive, process and compare and contrast essays. Detective stories have rich

detail that allows for ample elaboration in descriptive essays. Whodunits lend themselves to analysis of events as they happened and were discovered in process essays. And finally, these two sets of stories have primary detectives and assistant detectives who can easily be compared and contrasted among and between each other. An added bonus is that both fictions have TV series that closely follow the books' storylines, so students can watch the adaptations and see the pages come to life on the screen which accommodates the multiple learning styles in any given class. However, the sophomore students still needed reading material when learning how to write basic essays.

After disregarding Chinese mythology as a viable topic, being unable to find a suitable collection of fairy tales in a single book, and incorporating novels and memoirs in other ways throughout the curriculum, it was clear there was still a place for Greek mythology in the sophomore reading and writing course. So, why teach Greek mythology to EFL students?

Greek myth has a special place in Western education because it is considered to be a component of the well-educated or classically-educated elite (Taggart, 2009). Though most primary and secondary schools, nowadays, have done away with Greek myth curriculum, Western society is full of references to gods, goddesses, heroes, heroines and legends. For example, many names of companies and common English words directly reference Greek mythology, so a basic familiarity with the characters and stories is useful for an EFL student who intends to study in a Western university. A short list of company names includes Amazon, Midas, Nike, OKCupid, and Pandora while

common words derived from Greek myth could not leave out aphrodisiac, chaos, herculean, nemesis, and odyssey.

Learning Greek mythology, though often overlooked in the L2 community, is an important component of becoming literate in English. At some point, ELLs can read all of the words on the page but completely misunderstand a piece of writing. Many English literary masterpieces often obliquely reference the myths of old. For example, a sentence about Jupiter could mean the planet, the god Zeus or both. Without an introduction to the Greek myths and the characters, many EFL student would be completely blind to the complexity of multiple readings of a text. Therefore, a unit on mythology among foreign language and culture students is practical for introducing sophisticated sociolinguistic concepts.

Furthermore, Greek mythology is an excellent reservoir of ideas and examples to be used in the essays required by the TOEFL and old SAT exams. Since many of the prompts are open-ended, speculative and vague, students need to be quick with ideas to match the overarching theme. With an adequate understanding of the personalities of the gods and goddesses and a memory for the intriguing adventures of the heroes and heroines, students could write an essay peppered with references to Greek myth, which would add color and depth to a piece of writing or even add substantial content to become the thrust of the essay.

Fortuitously, after a few attempts, I found a book: *D'Aulaires' Book of Greek Myths* (1962) met all of the criteria required by the curriculum. Most importantly, the universe of Greek mythology provided abundant material for a variety of essay writing

assignments. Furthermore, the stories of love, war, and adventure captured the imagination of the students which encouraged them to push through the unknown vocabulary and grammar to understand the intriguing stories. It seemed to inhabit the Goldilocks Zone of curricular content and language difficulty. Finally, a sufficient number of copies were actually available in China and could be procured before the start of the school year through the school's publisher and distributor network.

Ultimately, *D'Aulaires' Book of Greek Myths* (1962) helped students gain an in-depth understanding of Greek mythology which enabled them use the material in their essay writing. The intricate relationships described among gods, goddesses, humans, and animals; brothers and sisters; husbands, wives, and mistresses; and friends and enemies sparked age-appropriate dialogs about customs, morals and personal beliefs concerning relationships and societal expectations for behavior. Most notably, students commented on the narrow roles of females and males in society described in the myths—ruler, lover, wife, mistress, warrior, creator—and compared them to life in twenty-first-century Beijing. Sometimes class conversations even ventured into the mores of the Greek gods and goddesses and how such behavior would be accepted among contemporary mortals today.

These discussions led into excellent essay topics. However, enabling students to write well-structured, well-supported, well-written essays took some time. Previous experience in Chinese high schools had shown that when I just assigned an essay, students would produce a page of sentences loosely centered on the assigned topic. The

students with some TOEFL or SAT training would divide these sentences into chunks resembling paragraphs, but seldom would they have much more organization than that.

Without doubt, these texts contained excellent ideas and articulate points, but the overall message was often lost in the stream-of-conscious style. It is clear from further research into Chinese essay conventions (Sullivan, Zhang, & Zheng, 2012) that my students were attempting to mimic traditional Chinese works they had grown up reading in their literature courses. However, this did not translate well in English as they did not yet have enough mastery of the language, nor did this satisfy the academic expectations of the courses and examinations for which they were preparing.

Therefore, I taught formulaic essay structures in line with the academic conventions required to pass their courses and exams. The initial work on essay writing came from the *NorthStar* textbook which elaborated on the basic form of an essay (introduction, body, and conclusion) and drew attention to important parts of an essay (thesis statements, transition phrases, topic sentences, and supporting details). Undeniably, this was a useful foundation for the students' essay writing, but it was hardly enough as they were trying to mimic a style in which they had little experience. My students needed much more explicit instruction and extensive practice to write better essays.

This realization led me to create example essays and detailed outlines for new assignments. In addition, we spent days in class talking about main ideas and supporting details as in the beginning students were unable to see or explain how a specific example related to their main point. Then, we spent days crafting well-written sentences that

expressed their ideas clearly and concisely. Finally, we practiced peer-editing texts to learn how to give and receive effective feedback. This extensive process meant it took several weeks to write one well-crafted essay.³ Without doubt, it was worth it as students could then write each new essay with less supervision and in less time.

Notes

1. New Oriental sells a series of study guides that tout the wonders learning vast amounts of data through the associative memory technique. Translated titles of these resource books on Amazon.com read *New Oriental TOEFL Word Root + Associative Memory* or *New Oriental: SAT Vocabulary + Associative Memory*.
2. The school curriculum used the third edition of *NorthStar 3 Reading and Writing* (Barton & Dupaquier Sardinas, 2008) and *NorthStar 3 Listening and Speaking* (Solorzano & Schmidt, 2008) for the sophomores while the juniors used level 4 and the seniors used level 5 of the *NorthStar* Series.
3. Look at Appendix A to see an example essay templet and other support materials.

Chapter 2: A Review of Composition and EFL Studies in the U.S. and China

How I learned to teach essay writing to EFL students is a result of my education in the American field of composition studies. Therefore, a brief overview of the history of where composition studies come from, in general, and EFL studies, in particular, is needed. Though strongly influenced by the development of English Departments, linguistics and rhetoric also cast strong shadows on my preparation as an EFL instructor.

I will begin with an outline of the major historical events which solidified and professionalized the studies of English, Composition, and ESL within the system of American higher education. This grid of dates and happenings will then be filled in with the theories, ideologies and pedagogies which make up the rich scholarship of composition. Finally, some ideas will be offered about the future relationships between ESL and composition studies in the twenty-first century.

America's Nineteenth Century New Universities and English Departments

One hundred and fifty years ago the U.S. Congress passed the Morrill Act of 1862 which brought into existence the land grant university. This investment in higher education opened the doors of universities to the middle class who needed to be trained to be professionals and managers in the businesses of the industrial revolution. The immediate increase in possible places to get advanced training caused the old universities

to change. The competition for students and faculty required reorganization. From this restructuring, the English department, with its responsibility for teaching good writing and good reading, was born.

Discovering the exact reason why these two fields—literature and composition—came to be paired within the American university is a more difficult proposal. The explanation that reading and writing have always been joined together because they are united by text is too simple. Susan Miller (1991) in *Textual Carnivals: The Politics of Composition* sets forth several other plausible explanations. The two subjects were paired for pragmatic reasons; composition was an immensely practical subject and would inevitably provide both a financial base and the educational need for an English department to exist. They were also united for political reasons. With the development of the university system in the late 19th Century, a social compromise between the aristocratic elite and the industrial elite, who were both taking advanced studies, needed to be brokered. Good training in the classics as well as proper grammar were symbolic marks of culture and class. The old money expected the new money to honor the traditions of class to which they aspired, so English coursework took on the challenge of correcting "unsophisticated" grammars and indoctrinating the classics (Miller, 1991).

However, James Berlin (1987) offers a different explanation in *Rhetoric and Reality: Writing Instruction in American Colleges, 1900--1985*. From the beginning, literature and composition were the purview of one department, so reading and writing courses were taught by the English faculty. However, from the start they were not given equal rank as academic disciplines worthy of advanced study and research. Literature and

the classics established the great canon which became the heart of the field of English. It is not completely clear why this happened; two possible reasons are proposed. Perhaps, a preexisting prejudice about the practicality of teaching writing already existed and a precedent for low pay and disrespect was already in place, or the early professors who were most influential in English departments preferred reading and literary criticism and disliked grading hundreds of undergraduate essays (Berlin, 1987).

After the establishment of English departments in universities, the field began to organize and professionalize itself. In 1883, the Modern Language Association (MLA) was established; this organization's main objective throughout the last century has been to elevate the level of scholarship in literature. In 1911, the National Council of Teachers of English (NCTE) was created as a forum for teachers of English to discuss pedagogy. MLA became the outlet for the literature professors of universities while NCTE became a gathering which emphasized the pre-college English teaching experience and the non-literature English professionals (Berlin, 1987).

Emergence of English as a Second Language alongside Composition

Paul Kei Matsuda (2006) gives a succinct description of the history of second language writing instruction in the second half of the twentieth century in a piece entitled "Second-Language Writing in the Twentieth Century: A Situated Historical Perspective." Matsuda starts his history in the 1940s and 1950s when the university system had just experienced a fivefold increase in international student enrollments. In 1941, the University of Michigan started its English Language Institute (ELI) under the direction of Charles C. Fries, who was heavily influenced by developments in the field of

linguistics—not English, which was dominated by the study of literature at that time.

Fries “assumed that students would be able to write once they mastered the structure and sounds of language” (Matsuda, 2006, p. 16).

In 1949, the Conference on College Composition and Communication (CCCC) emerged from NCTE as a space for the growing group of compositionists working at the college level. The second half of the twentieth century was a time of empowerment for writing professionals within English departments and universities due to the almost universal adoption of the general education curriculum, which included freshmen writing courses. For the first fifteen years of the CCCC, teaching writing to second language students was a topic of conference workshops and panels. At the same time in the 1950s, the professionalization of the field of Teaching English as a Second Language (TESL) developed in direct relation to the needs of the ELI programs which were being opened at schools with large international student populations, such as top research universities. Post-baccalaureate programs and graduate programs in TESL were offered through applied linguistic departments which had concentrations in second-language acquisition. In 1966, a new organization called Teachers of English to Speakers of Other Languages (TESOL) was founded (Matsuda, 2006).

The professionalization of ESL through graduate degrees and organization under the auspices of linguistics departments, not English faculty, created a divide which would not be bridged for years to come. Communication increased in times of need, for example, when the number of international students enrolling in American higher education continued to increase. In 1940, less than 7,000 international students were

studying in the US while in 2016 more than one million had entered US higher education (Matsuda 2006; IIE Open Doors Report, 2016). These students had difficulties writing in their freshman writing classes and in the basic writing classes which were under the purview of English departments, so ESL writing concerns began to arise in composition conferences and publications. At the same time within the field of TESL, concerns about the quality of writing education were being raised. The general interest in L2 writing encouraged the second-language and composition professionals to share across fields eventually creating the need for an interdisciplinary field of second-language composition. It is my assertion that when there was a sudden increase in the number of international students (during 1940-1950, 1975-1980, and 2005-present) discussions about second language learners erupted in composition studies.

Pedagogy of Composition since the Establishment of CCCC

Trends in teaching composition have been quite fluid during the last sixty years. No one pedagogy was dominant in the classroom even when the teaching methods and ideology inundated the literature. That said, some consensus about the history of teaching theory has been formed. In order, to see how the different ideas worked together as a whole, I have ordered the ideas according to the decades which they emerged in the literature; however, traces of all the different kinds of composition can be found in the literature at different times. Many old arguments, which have not been conclusively resolved, still throw shadows over the scholarship today.

1950s—1960s Current-traditional

1960s—1970s Expressivism, rhetoric, linguistics

1970s—1980s Process pedagogy

1980s—1990s Maturation of the field (histories published)

1990s—2000s Cultural studies (special interest groups: computers, diversity, ESL)

2000s—2017 Multiliteracies, world Englishes, cross-language

Current-traditional pedagogies have been a part of how English is taught throughout our history. What exactly is this rhetoric? It has been given many names. Fulkerson classifies it as a "formalist theory" and defines it as a pedagogy which over stresses spelling, grammar, punctuation, and form of the entire piece including style, layout and length (Fulkerson, 1979, p. 344). Basically, it is a way of teaching and thinking about language that gives preference to form and structure, and, many would argue, at the expense of expression and meaning. Berlin's chronicle states, "the rhetoric that appeared in the English department in the late nineteenth century has come to be called current-traditional rhetoric" (p. 26). Unfortunately, current-traditional methodologies and ideologies, with their objective or formalist view of the language, have held sway in English departments from the beginning, and despite their unpopularity in published forums, traces of their influence can still be seen in many classrooms today.

The place of rhetoric is a theme which keeps appearing in composition studies throughout the decades. I would not consider it a dominant theory at any one time, but its rich theoretical and pedagogical approaches leave a mark on all who seriously study them. Therefore, its influence regularly emerges within the field of composition. Rhetoric is not a monolithic subject; it has many iterations and philosophies. Fulkerson names

three that he believes were influential within composition studies in the last few decades: argumentation, genre or theme based, and dialectical discourse (Fulkerson, 2005).

Linguistics made a significant contribution to the fields of composition and English as a Second Language studies. "The appearance of structural linguistics in the fifties created a stir in English departments that affected teachers of literature as well as writing teachers.... This theory of language, moreover, promised to provide the writing course with its own subject matter, offering a new grammar that was to be the counterpart to literature's New Criticism" (Berlin, 1987, p. 111). While linguistics' influence on composition was not the great unifying theory it was hoped to be, linguistics did significantly impact English as a Second Language development as applied linguistics adopted ESL as a sub-field.

Expressivism is a direct outcome of the current-traditional rhetoric that came before it, for it seems that after a time of tight rule-bound thinking, order will be eschewed. This new approach to teaching was also an outcome of the political turmoil within the United States during the 1960s and 1970s. Berlin (1987) defines Expressivism (which he terms expressionism), as a subjective rhetoric which encouraged uniquely individual metaphors, regular journaling and peer workshoping. As this Expressivist view of writing developed, it became paired with a new process approach to writing which encouraged interchangeable stages of writing such as idea generation, organization, drafting, revising and editing. However, these two methodologies should not be confused. Fulkerson explains: "It is simply inaccurate to equate 'process' pedagogy with an expressive axiology, although the two were entwined in the influential early work

of Don Murray" (Fulkerson, 2005, p. 669). It is easy to see why they are confused when many instructors use journaling for idea generation and peer groups for revising and editing. As a field, composition should look at expressivism and process for what they were: a break from the old traditions of teaching. Miller (1991) compares process pedagogy to current-traditional pedagogies by saying these "pedagogies that focus on methods of writing clearly improve on either emphasis on correctness and grammar" (p. 97), but may not necessarily lead to the production of better written texts by students.

The 1980s and 1990s saw the maturation of the field of composition; it was a time of coming into its own. This opinion is supported by Fulkerson (2005) who notes that during the twenty-year period between 1980 and 2000 there was a "growing 'scholarizing' of the field" (p. 657). Berlin's (1987) history *Rhetoric and Reality* and Miller's (1991) political commentary *Textual Carnivals* both attest to the excellent discussions which were occurring in the field. This scholarship shows a need to look back to see where we have come from to inform future developments in the field.

The last twenty years of composition scholarship has been inundated by what Fulkerson calls "critical and cultural studies" (Fulkerson, 2005). He defines it as a pedagogy which highlights "systemic cultural injustices inflicted by dominant societal groups and dominant discourses on those with less power, and upon the empowering possibilities of rhetoric if students are educated to 'read' carefully and 'resist' the social texts that help keep some groups subordinated" (Fulkerson, 2005, p. 659). This era has been the time of special interest groups. This would include all the diversity issues related to the liberation of subjected peoples and attention to their needs in the composition

classroom. Topics that also arose from this discussion were the status of ESL or EFL students, World Englishes and cross-language communication.

Culture and Language

There is a relationship between culture and language. Things really are lost in translation. Pithy expressions, proverbs or clichés in one language often require background to explain in another language. We assume when we explain something to someone else that if they understand the words, they can also understand the concept and deeper meaning. We forget that there are many traditions which underlie our words. These un-verbalized norms, unstated standards, or hidden expectations are very difficult to know and understand, unless one can read between the lines and decipher the code.

To complicate matters, each culture has specific ways of communicating. Most cultures have different vocabularies and mannerisms for speaking with elders, minors, bosses, and subordinates. These differences in language and gestures are extremely difficult to put into written form and to translate into another language and culture.

Some research has been done about various rhetorical styles of presenting information. These contrastive rhetorics (Robert Kaplan discussed in Fox, 1994) only begin to identify the different possible constructions for formal writing. By overemphasizing the thought patterns and writing traditions of the English-speaking Anglo-American—the norms of a small subset of the world population that currently has considerable power influencing the world culture—other styles of communication are undervalued and underappreciated.

Anyone who has picked up an advanced book in a field of study they are unfamiliar with will feel like they are reading a foreign language even when they understand each and every word by itself. Language is a flexible system. Within different contexts, words take new connotations and meanings. Complex ideas are represented with referential phrases which are easy codes to understand for the insider. However, the context-rich nature of advanced study in an unfamiliar discipline is a good example of the code cracking required between language and culture.

Helen Fox's (1994) *Listening to the World: Cultural Issues in Academic Writing* is an excellent book to introduce the complex relationship between language and culture as it relates to international students who are trying to write academic papers within the US university system. As a writing tutor and instructor, Fox worked with hundreds of international students over approximately ten years and documented their stories of struggling with writing within academic institutions. There are two themes which run throughout her book: first, self-expression through words is a very personal endeavor, so it is hard to find voice in another language because the ideas you have do not fit easily into that language's grammar. Second, each student has different cultural expectations for what is required by academic writing at the university level, and they have learned different rhetorical and organizational strategies to compose research scholarship (Fox, 1994).

An article that illustrates Fox's second point is "Toward a Writing Pedagogy of Shuttling between Languages: Learning from Multilingual Writers" written by Suresh A. Canagarajah (2006). In this piece, the multilingual, multigenre writing of Professor K.

Sivatamby is analyzed to shed light on the complex relationship language and culture expectations of different audiences. Canagarajah looks at three articles in three different publications written by the author Sivatomby on the same subject. Canagarajah shows the different rhetorical moves that Sivatomby employs, his deft understanding of both English and Tamil, and his experience with academic writing in Western and Eastern traditions—skills which had been honed over decades of work in an international higher educational setting. This enables the reader, who neither speaks Tamil nor has access to Sri Lankan publications, to see just how different academic writing can be across languages and cultures (Canagarajah, 2006).

When teaching English as the Lingua Franca, Celce-Murcia (2013b) explores the difficult questions that surround World Englishes and how to ethically teach language. With so many regional and cultural groups adopting English as a means for communicating within and across cultures, it is important to consider the learners' identities in relation to the intelligibility of their variety of English. Her chapter, entitled "Teaching English in the context of World Englishes" (Celce-Murcia, 2013b), encourages more questions than it answers: should educators focus on communication over correctness; should we encourage standardized pronunciation for ease of communication or social benefits; and when, where, and why should academic written English be taught instead of or alongside conversational English.

The subject of English language acquisition education is fraught with many interesting questions and concerns. These debates are not settled in the field now, nor will they likely be fully resolved anytime soon. Studying English as an additional language is

a political as well as practical concern and as such it has many diverse peoples adding their opinions to how the fields (ESL, EFL, L2) should develop and change. Before we look to the future, we must look to the past and reflect on what has occurred in different parts of the world. In particular, more attention should be given to English language studies in China because this country is exporting a large number of students who are flooding educational institutions around the world.

The People's Republic of China

Xiaoye You (2010) argues in *Writing in the Devil's Tongue* that composition studies as a field has focused most of its attention on US university students with marginal attention being paid to writing among students of diverse backgrounds (1.5 generation, second language, and foreign language learners) or professionals influenced by other ethnic and national discourse communities. Compositionists make assumptions about research and pedagogy based on norms which have been set on Western soil. Since much of the English language teaching taking place on the planet happens outside of English-as-a-primary-language nations, it is crucial that an assessment of diverse international localities takes place. This dissertation will do this by showing the experiences of one group of Chinese students in an International Curriculum Center in a Chinese high school. By looking at some of their ways of learning English as they try to master the TOEFL writing exam, new insights into how to better to teach these students can arise.

Chinese Education: Overview of the Current System

The People's Republic of China has organized its education system as follows: six years of elementary school (小学 *xiaoxue*), six years of secondary education (中学 *zhongxue*), and two or more years of tertiary instruction, (大学 *daxue*), depending on the degree sought. Secondary schooling is divided into three years of middle school (初中 *chuzhong*, junior high school) and three years of high school (高中 *gaozhong*, senior high school). In 1986, the Compulsory Education Law of the People's Republic of China (2009) went into effect stipulating all children and young adults should be educated until the end of middle school—grade nine or age fifteen. Before that time, only primary school was prescribed. However, among the social, political, and financial elite in developed areas of China, students are expected to complete high school and go on to tertiary education.

It is commonly known that Chinese students are good test takers; though a stereotype, this impression is bolstered by centuries of institutionalized strenuous examinations for governmental power positions. More than a century ago in 1905, the imperial examination system was abolished with the ending of imperial reign. Many assumed that this signaled the end of this model of using tests for the stratification of the masses. However, Huang and Brown (2009), when commenting on contemporary China have said, “Chinese education is an examination-driven system” and it is quite competitive throughout primary, secondary and tertiary levels (p. 645). To enter and exit each school level (elementary, middle school, high school, and college) very rigorous

testing occurs. The Zhongkao (中考) is administered to all grade nine students to determine if and where they can attend high school. Scores on this exam decide futures—tracking individuals to enter key schools, which prepare for college admissions, or other less rigorous vocational or training schools, which train for other professions.

Examinations control both graduation from high school (会考 Huikao) and admission to university (高考 Gaokao). The first of these exams, the Huikao, is not considered very difficult, allowing many students to pass. However, the Gaokao is one of the most competitive examinations in the world with millions of students (Ash, 2016) jockeying for admissions to approximately 2,500 institutions of higher learning in mainland China (Ministry of Education, 2015b). Elite students in key schools will spend the entire senior year of high school cramming 18 hours a day for this three-day-long exam offered only once a year in June. If they fail to achieve their objective—entrance into the university of their choice—students may retake their senior year and the Gaokao, or they will try to gain admissions to private colleges in China or other institutions abroad. The importance of English language study in China is underscored by the fact that approximately one third of this crucial Gaokao exam tests English skills (Ash, 2016).

The development of international sections (国际部 guojibu) or International Curriculum Centers (ICCs) within public high schools is a relatively recent phenomenon in the history of Chinese education. International schools, which provide international education in a school setting, are authorized to admit only foreign students, so Chinese nationals are unable to participate without acquiring additional citizenships. One of the first international sections in a public high school was opened in 2002 (The High School,

2010). This was a significant shift in the Chinese education system because it allowed Chinese nationals an opportunity to study in an international school environment while enabling them to take the Huikao exam and graduate from a public high school. These international-section students usually do not take the Gaokao, the exam for admissions into Chinese tertiary institutions, because they are intending to study at universities abroad. Since 2002, many other schools across China have opened international sections within their public schools. The popularity of international sections in public high schools is directly related to the competitive nature of the Gaokao exam. For families who can afford the tuition and plan on sending their child (and in rare cases children) abroad for college, they circumvent the Gaokao and the related stress endemic to the traditional examination system.

However, the international education paradigm has not escaped China's competitive culture—instead, it has become big business fueled by parents' lack of knowledge about international universities' educational requirements and students' anxiety about the battery of tests, such as TOEFL, IELTS, SAT, ACT, GRE, and GMAT, which purport to determine one's future, much like the detested Gaokao.

In this context, an industry of English training centers and admissions counseling services has emerged in China. Very quickly it has become a billion-dollar industry. New Oriental—the largest competitor in the market with nearly 28 million students and 790 learning centers and schools across the country (New Oriental, 2015)—expounds one method for teaching English which encourages vocabulary memorization of individual words. Each year they produce several books which contain vocabulary relevant to

TOEFL, SAT, GRE and any other standardized test that millions of Chinese citizens are studying for and preparing to take. Many Chinese students claim to 背单词 (BeiDanCi) BDC, memorize thousands of vocabulary words in a short span of time, and believe it helps them to learn English. Therefore, within the context of Chinese high school students, it is important to interrogate the ubiquity of this study method and its usefulness.

Historical Development of English Education in the PRC since 1949

The history of English language teaching in China has not followed the same progression in development as those documented by Western scholars studying in U.S. institutions; this is because cultural and political traditions within China have strongly influenced its evolution.

Adamson and Morris (1997) summarize the development of English language teaching in China since 1949. Initially, the predominant foreign language learned and valued in China was Russian, which was considered more patriotic to learn than English because of its importance within the communist world; this was a political choice to facilitate communication between other communist countries and to show an affinity to the USSR. During the Cultural Revolution (mid 1960s to mid 1970s), all schooling was interrupted and foreign language teachers, whatever language they taught, were accused of being spies of other nations and persecuted. Despite these historical facts, Mao Zedong is noted to have said he prefers English to Russian and encouraged students to learn it. After Mao's death in 1976, English was chosen over Russian as the foreign language of choice to be taught in junior high and senior high schools across China.

Since the reintroduction of foreign language lessons into secondary schools in the late 1970s, many English teachers in China lacked adequate training. Some were the Russian foreign language teachers who were obliged to learn English, the new foreign language of choice, immediately or lose their jobs as foreign language teachers. Others had studied English informally and clandestinely during the Cultural Revolution by reading contraband books and listening to illegal international radio broadcasts such as the BBC radio transmissions (Z. Fu, personal communication, April 27, 2017). Therefore, a generation of the Chinese English teachers who taught throughout the 1980s and 1990s suffered from a lack of adequate training as professionals, which has impacted future generations of ELLs in mainland China.

Consequently, it is not surprising that in 1978 when Cowan, Light, Mathews, and Tucker (1979) conducted an educational tour of China they “observed an impressive commitment to the teaching of English in the People’s Republic of China,” they added, however, that “the practices we observed are not those advocated by all TESOL specialists” (1979, p. 477). Arguably, the teachers and students they encountered were politically eager to implement English education as a new government mandate, but the country did not yet have mechanisms in place to support this goal.

In the last three decades, English language instruction has gradually become more professionalized. Now, high school and university teachers are usually graduates from tertiary institutions with degrees in English; however, many still do not hold degrees in education nor do they have the extensive training in pedagogy and curriculum development routinely required by Western educational institutions. According to the

“Teachers Law of the People’s Republic of China” adopted in 1993, Chinese teachers must graduate with a degree from at least one level of education beyond what they teach, pass China’s national teacher’s qualification exam, and then receive further on-the-job training within their school district (Ministry of Education, 2015d).

The progression in language teaching theories in Western academies is documented by Zimmerman (1997, 2013). Her assessment starts more than two hundred years ago with the Grammar -translation approach and steadily moves through the Reform approach, Direct method, Reading approach, Audiolingualism, Communicative Language Teaching, the Natural Approach, and ends with Lexicographical Research at the end of the millennium. Each new theory of teaching did not completely replace existing approaches but slowly moved the conversation about best practices in new directions often swinging in a pendulum between the importance of reading comprehension and verbal fluency. Celce-Murcia (2013a) continues the list of different movements in the field of English language teaching through the end of the last century by adding the following to Zimmerman’s initial list—the Cognitive Approach, the Affective-Humanistic Approach, the Comprehension-Based Approach, the Communicative Approach and Designer Methods. For the new millennium, Celce-Murcia (2013a) believes the field of language teaching has entered a post-methods era where no one way of teaching holds sway over theory or practice; instead, pedagogy is influenced by strategies that encourage teachers and learners to critically and ethically engage with the process of acquiring a language.

English teaching in the People's Republic of China has a different development and timeline. A few language researchers have completed historical surveys of the development of English teaching in China through the analysis of textbooks and curriculum. The article "English Curriculum in the Peoples' Republic of China" (1997) documents how politics often trumped pedagogy in the publication of English textbooks; "teachers complained that the books were too difficult to teach, primarily because of the emphasis on political content to the detriment of language instruction" (Adamson & Morris, p. 18). That said the methods of education did progress as new textbooks were written and adopted. The Chinese-published English textbooks moved from using grammar-translation methods during the 1950s-1960s to "a blend of audio-lingualism and grammar-translation method" during 1977-1993 (Adamson & Morris, 1997, p. 20), with a break in language teaching altogether during the Cultural Revolution.

Compared with the development of Western English language pedagogy, China's English education followed a different progression and timeline. This is further elucidated in You's study of textbook production in the Chinese mainland. In his analysis of two editions of *A College Handbook of Composition* published in 1984 and 1994, You (2010) asserts that both have distinct current-traditional tendencies despite the incorporation of some elements of process pedagogy in the latter publication. Additionally, he shows how "Explicit political teaching, including morals and ethics, was found in more than foreign language classes during the 1980s; it was a vital part of one's college education" (You, 2010, p. 639).

Cultural Educational Influences

Without doubt, diverse cultures have different methods for educating future generations; the American and Chinese models are based on distinctive values held by the different societies. China's unique history and political system has led to a novel blend of elements which shape beliefs about how individuals learn and how subjects should be taught.

The basis of education in China has its foundations in Confucian thought and teaching practices; imperial examinations, which ruled China for more than a millennium and selected the officials of historic China, were based on the curriculum developed in the Confucian private schools. These examinations stressed memorization and attention to detail while original thought and individualistic ideas were not rewarded. "The fundamental principle of Confucian education was to train talent loyal to the government" (Huang & Brown, 2009, p. 645).

Because of China's unique history, there are many fundamental differences between Western and Chinese education, so a brief comparison between the two may illuminate sites of tension and misunderstanding.

Both Western and Chinese cultures believe teachers must have sound and extensive knowledge of the content area to educate students well (Flowerdew, Levis, Zhang, Watkins, 2007, p. 784). Western students accept that a teacher does not know everything about a subject, but the instructor knows enough to teach the course. Chinese students believe that instructors should be "masters" (师傅 shifu) of the material;

therefore, they often ask extremely obtuse questions to test the knowledge of an educator and will lose faith with an instructor who admits to limited understanding.

Another discrepancy arises in beliefs about educational methods. Western pedagogy values interesting classroom activities and quality resources, while Chinese instruction emphasizes the ability of an educator to transmit knowledge to pupils and be a moral example (Flowerdew, Levis, Zhang, Watkins, 2007). Therefore, Western students have become accustomed to being entertained with a variety of well-prepared activities which help engage them in the content of the curriculum. Chinese students expect fast-paced lectures with key-points written on the blackboard (Huang & Brown, 2009), and they believe a good teacher will be a beacon of virtue which they can emulate.

There are further differences when one discusses writing—both its content and organization; the Western tradition and the Chinese tradition can barely agree on what should be written about and how it should be put down on paper. “American writing values logic, reasoning, and self- discovery, while Chinese writing values beauty of language and the moral message delivered in writing. This observation does not apply to all situations, of course; nevertheless, it does catch at least some critical characteristics of the two rhetorical traditions” (Sullivan, Zhang, & Zheng, 2012, p. 324).

When describing typical vocabulary acquisition techniques used in Chinese English language pedagogy, it is apparent that different methods and values are being implemented from those currently accepted as best practices in the West. “The popular model for college teachers to teach vocabulary is asking their students to read after them the whole word list in the text followed by the teacher’s translation of each word or the

teacher chooses some basic words and just gives the Chinese equivalents. Students are really passive in this process and they just commit the new words into memory by rote” (Zhao, 2009, p. 123).

This account is comparable to the grammar translation method as described by Zimmerman (1997): “Students were provided detailed explanations of grammar in their native languages, paradigms to memorize, and bilingual vocabulary list to learn” (p. 5). These standards, reciting vocabulary words and literary texts by heart, are still used in numerous places in present-day China to judge students’ English language comprehension.

Therefore, it is not surprising that Chinese foreign language pedagogy has strong overtones of the Grammar Translation method because Confucius thought encouraged the memorization and internalization of texts (Huang & Brown, 2009) and many textbooks borrowed heavily from the USSR (Adamson & Morris, 1997), which had a history of this tradition (Zimmerman, 1997).

Though many Western scholars view grammar translation as an outdated education model, an analysis of the research on vocabulary learning can contextualize and validate some of the practices within English language methodology in China.

Summary

The US and China have English education programs that arise out of completely disparate histories, so it is no surprise that their current educational pedagogies are dissimilar in make and purpose. The US created a vast network of public universities

more than a century ago and experienced accelerating influxes of international students each year since WWII; therefore, the academies have not only pragmatically needed to undertake the challenge of teaching English to non-native speakers, but they have had the opportunity to build masters and doctoral programs which study language acquisition. These trends in conjunction with the desire to understand and welcome the diversity of cultures and languages in the US education system has made a unique English language curriculum.

China's history with English education has been a switchback path with government policies directly influencing classroom curriculum. With the early strong influence of Russian communist thinking to the need to educate millions of citizens every year for the global economy, everything from political correctness to educational mandates have played a role in the formation of the current teaching methods widely use across the country. It is certain that the future of English education in China will continue to evolve as the universities increase their focus on research and the government reevaluates educational approaches to better prepare its citizens for the 21st Century. This will make the future study of English language acquisition and teaching in China and the US quite fruitful as their two separate courses will continue to produce different outcomes that require honest assessment.

Chapter 3: Researching Vocabulary Size and Essay Writing

For more than a decade, EFL and ESL researchers have been assessing the significance of vocabulary teaching and learning. It is standard practice to teach vocabulary when developing reading, writing, listening, and speaking skills (Hinkel, 2006). To quantify just how important receptive vocabulary is for understanding authentic English materials, Nation (2006) determined that English learners should understand more than 6,000 words to watch children's movies, 8,000 words to read English-language newspapers and 9,000 words to read novels (Nation, 2006).

In addition to knowledge of vocabulary, L1 and L2 compositionists agree that writing proficiency is critical for success in educational environments where standardized tests such as the SAT, ACT, and TOEFL exams, which all have writing components, “weigh heavily in college admission decisions” (Jeffrey, Kieffer, & Matsuda, 2013). Furthermore, because the writing requirements on tests such as the TOEFL exam have similarities to university academic writing, they can be used as a valid assessment of students' college readiness (Riazi, 2016); therefore, preparing for university coursework and university entrance exams do not need to be mutually exclusive activities.

Without a doubt, vocabulary words are to communication like bricks are to a house—a sought-after building material; however, much more than organization is required to turn words into an essay or bricks into a house. Clearly vocabulary words are necessary for understanding English as a Foreign Language and good writing skills are required for admission to university and success in university. Thus, it is logical to ask: to what extent can a correlation be found between English vocabulary knowledge and the

ability to write an essay for the TOEFL exam? Is there a relationship between vocabulary knowledge and writing skills?

Many different researchers have studied this question in different ways among different populations. A generation ago, Koda (1993) researched American college students learning Japanese as a foreign language by assessing writing, through a narrative and descriptive task, and linguistic knowledge, through a vocabulary and grammar task. She found: “Vocabulary knowledge was correlated highly with the overall quality of ratings in both tasks (Descriptive, $r = .75$; Narrative, $r = .70$)” (Koda, 1993, p. 334). This research established a strong link between vocabulary knowledge and writing skills among students learning Japanese as a foreign language.

More recently Baba (2009) investigated Japanese college students learning EFL through eight tasks: five English language measurements, one self-assessment, and two Japanese language assessments. His extensive data yielded different relationships between vocabulary and writing. He found: “Two lexical proficiency indices—vocabulary size and word definition ability—had moderate correlations of .40 and .51, respectively, with summary writing performance” (Baba, 2009, p. 199). He concluded that there is some correlation between participants’ ability to write definitions of English words in English and their ability to write good summaries.

Baba’s (2009) results were not as conclusive as Koda’s (1993) earlier work. It is unclear what affected the differences between the two studies because there are too many diverse variables at play—foreign languages being studied, vocabulary assessment measures, and writing tasks assigned, to name a few of the different procedures used in

the two different studies. When reviewing the research, attention to detail is required; divergent results may stem from a plethora of possibilities which all require consideration.

Though several studies show positive correlations between vocabulary knowledge and writing proficiencies, there are large discrepancies in significance among studies. Unfortunately, it is difficult to determine if the divergences arise because of the differences in populations or the incongruent research instruments used and the variety of study designs. Therefore, more research needs to be conducted. This doctoral study assesses the extent to which English vocabulary size correlates with the essay writing proficiencies among Chinese high school students studying EFL (English as a Foreign Language) in mainland China. By using a standard English Size Test developed by Paul Nation (Beglar & Nation, 2007; Nation, 2012; Coxhead, Nation, & Sims, 2014), which has been used in multiple contexts around the world (Nguyen & Nation, 2011; Coxhead, Nation, & Sims, 2015), and the widely-accepted TOEFL grading rubrics, which has been used by Baba (2009) and Cumming et al. (2005), this study will add to other research in the field by looking specifically at the population of Chinese high school students in International Curricular Centers.

This dissertation statistically analyzes two measures to establish the validity of the claim that both Koda's (1993) and Baba's (2009) research make: that vocabulary knowledge impacts writing ability. To assess this claim among Chinese EFL students, two tests were given to the participants: Nation's Vocabulary Size Test (VST) and a compare-and contrast essay marked according to the TOEFL Integrated Writing Rubrics.

The research statistically analyzes the test results, explores outcomes, and compares results with past studies.

The fundamental hypothesis:

H₀: Receptive vocabulary size, as measured by Nation's VST, has no effect on writing proficiencies, as measured by a compare-and-contrast essay marked using the TOEFL Integrated Writing Rubrics.

H₁: Receptive vocabulary size, as measured by Nation's VST, positively correlates with writing proficiencies, as measured by a compare-and-contrast essay marked using the TOEFL Integrated Writing Rubrics.

If it is determined through analyzing the VST score and the essay score that there is a statistically significant correlation between receptive vocabulary knowledge and writing proficiencies, then further questions will be asked of the data—such as, what is the quantifiable relationship, and are other factors of statistical significance?

Research Methodology and Design

After much consideration about the nature of the site and the age of the participants, I determined to use a quantitative study evaluating performance measures that were already in place within the school where I worked. Prior experience had revealed logistical concerns with implementing a qualitative or mixed method design in high schools in China as the schools do not usually conduct research which requires Institutional Review Board (IRB) approval; instead, they have other procedures in place, which are acceptable within the local context, for approving and conducting research.

Therefore, the site did not have an IRB, nor did it have significant experience with the processes required for approving faculty research, facilitating consent and assent forms for guardians and participants, and other necessary measures to ensure the safety of the participants and the data that is required by a typical US university Internal Review Board (IRB). Thus, a straightforward request for test data was made to the school administration; after several explanatory meetings, the principals signed a letter agreeing to release the necessary research materials at the end of the school year.

In addition to the logistical concerns of the population and site to be studied, some consideration was given to the intended audience of the research. This research was undertaken to better understand how Chinese high school students learn English language skills and to determine how vocabulary acquisition impacted their writing abilities. Whatever the results, the research was meant to be shared with the school community. Therefore, empirical data, which could be expressed in charts and graphs, would be easily translated for the primary audience—Chinese colleagues, students, and parents.

The research was designed to take advantage of the kind of data that could be collected from minors in a high school setting in China: test scores. Because it was only possible to get IRB approval to collect data from examinations normally conducted within the school, an exploratory quantitative study with a population of convenience was chosen—a population of students who were conveniently located within the school where I taught. Then, an analysis of the curriculum through the lens of other studies conducted in the fields of language learning, in general, and EFL, in particular, was required. Finally, a design which resembles the studies conducted by Koda (1993) and Baba (2009)

was selected; they each collected and evaluated some form of a vocabulary assessment and writing assessment from participants. Their basic method could be replicated, with modifications, by using two instruments, one vocabulary measure and one writing measure, and piloted during final-exam week.

This research design is similar to other studies where the assessments appear determined by local particularities. For example, several articles concerned students studying languages other than English; Koda (1993) investigated American college students learning Japanese as a foreign language while Hazenberg (1996) examined how many words are needed to study in a Dutch-language, first-year university program. These specific learning environments required specific measures for the languages being researched; unfortunately, their measures are difficult to transfer across languages and communities. This is a recurring theme in language research—there are not freely available, standard measures to assess language learning in different contexts. That said, other researchers are attempting to find parallel tools to use to measure language learning around the world among different languages and different educational communities (Schmitt, 2010; Milton, 2009).

Like other studies in the field, my design has two parts to the assessment: a vocabulary test to measure the number of words known in English and a writing prompt to measure participants' compare-and-contrast writing abilities. Paul Nation's Vocabulary Size Test B was used to measure student's vocabulary size while a writing prompt based on curricular content was used for the writing assessment.

Nation's Vocabulary Size Test

Nation's Vocabulary Size Tests had been previously implemented in the school as an instrument to determine students' language learning over the course of their time in high school. Multiple versions of the test were employed throughout the academic cycle to measure incremental language development. The exams were not part of the students' grades but were used to better inform the development of the English Language program as it related to course planning and textbook choices. The tests were chosen for their relative ease of use and their minimal cost to the school.

Writing Assessment

Deciding which writing sample to use required consideration. Finally, it was determined that the final examination, which is kept on file for an extended period of time, was the best option because the papers could be made available for longitudinal research. Several concerns should be addressed about the writing prompt: First, the prompt is not a standard one that had been used in other testing or research settings. This makes it difficult to determine if writing prompts are comparable across testing situations and populations. However, it is clear from the literature that many other studies use writing prompts which are specific to the courses' content and the participants. Furthermore, regardless of content or level of the prompt, participants will respond in all manners; therefore, another way must be found to analyze writing responses for comparability across settings.

One way to minimize the discrepancies in research results because of the variations in prompts and responses, is to use a standard grading rubrics to assess the

work such as the TOEFL iBT® Test Independent and Integrated Writing Rubrics. This assessment tool is recognized by students as they learn about it when preparing to take the TOEFL exam as a first step to studying abroad in college. Furthermore, I followed Baba's (2009) example to holistically assess students' writing with the TOEFL rubrics because, as he explains, it looks at multiple levels of writing such as ideas and concepts, organization, linguistic structures and forms, and text from the source material when considering the quality of the piece of writing. Finally, after submitting an online application, TOEFL gave me permission to use their grading rubrics in this doctoral study.

TOEFL has two different grading rubrics for the two different essays required on the exam. The independent essay, a personal experience essay, is assessed using the TOEFL iBT® Test Independent Writing Rubrics while the integrated essay, an academic essay, is assessed using the TOEFL iBT® Test Integrated Writing Rubrics. The major difference between the two essay types is that the independent essay prompt encourages test takers to write about a topic using ideas and examples from daily life experiences while the integrated essay prompt requires examinees to include material from a specific lecture and reading (Collins, 2012). Therefore, in this study, the participants were assessed using the TOEFL iBT® Test Integrated Writing Rubrics because they were required to use examples from *D'Aulaire's Book of Greek Myths* (1962) in their essay responses. In essence, they were asked to write an academic essay which required the examples to come from course materials instead of from personal experience.

Population and Sample Selection

There is significant research among the general population of international high school students studying English as a Foreign Language. Nation (2013) wrote a general teacher's guide to help highlight the most useful pedagogy in the field of EFL studies. Nevertheless, little research has been conducted among the target population of Grade 10 mainland Chinese high school students studying English as a Foreign Language though related studies have been conducted among three similar populations: Chinese undergraduate students in China (Zhou, 2010; Li, 2011), in Hong Kong (Ma, 2013), in Taiwan (Chen, Warden, & Chang, 2005), and in the United States (Sullivan, Zhang, & Zheng, 2012). Undoubtedly, some parallels can be drawn between these populations and Chinese high school students studying in mainland China, but many differences make the comparisons incomplete or problematic.

First, although the Chinese undergraduate students cited in the literature might seem to be the most similar population to those I studied, they are normally several years older than grade 10 high school students; subsequently, a great deal of emotional and intellectual development occurs in that time. Second, Hong Kong students have the additional concern of living in a multilingual community with a recent history of British colonization with English language content in public spaces and the community; consequently, it cannot be considered a true EFL environment because English is spoken as a native language by many residents. Third, Taiwanese students have a different education system than mainland Chinese students; since 1949, the two places' school curricula have developed along separate lines with distinctive political and social

influences. Finally, the Chinese students studying in the US are often ethnically Chinese but hail from many different geo-cultural communities with diverse educational backgrounds. Therefore, although these communities share many similarities, they also have significant distinctions that must be considered when drawing meaningful comparisons.

The study sample consisted of mainland Chinese high school students at an International Curriculum Center in a mega city in eastern China who are studying English as a Foreign Language to prepare for enrollment in international universities abroad. The population consists of one class of 38 high school Grade 10 EFL students studying in an International Curriculum Center in Beijing. The participants are 26 males and 12 females between 15 and 17 years old. The participants were a convenience sample selected from among the classes to which I had access within the school. Originally, two classes were selected for study (approximately 80 participants), but because of different pedagogies employed by two different instructors, it was determined that the test results from only one class be assessed in this study. Otherwise, differences in how essay writing was taught could account for differing writing proficiency scores.

The participants in the study come from affluent Chinese families who enrolled their children in this school and the ICC in preparation for a university education abroad. This means that the participants' families can afford the above average high school fees in China and the cost of tuition at colleges in the United States, Canada, and the United Kingdom, among other countries. In addition, all participants take rigorous high school courses that prepare them for a university education abroad. For example, all students

were concurrently enrolled in a college-level physics course that was conducted completely in English. This kind of demanding coursework taught in English enabled many students to increase their English vocabulary dramatically outside of the English classroom because they had other native English teachers as instructors for different subjects.

Arguably, this immersive English-language curriculum within the ICC would enable students to feel like they were English as a Second Language learners instead of English as a Foreign Language learners. However, two things make this not so. First, Mandarin speaking students and teachers greatly outnumbered the native English speaking teachers; second, more than half of the time at school individuals would speak Mandarin in class (specifically in national Chinese curriculum courses such as Literature, History, Politics, and PE) and outside of class (during passing periods, at lunch, and at home). Thus, the English Language Learners are firmly in the EFL community and not ESL speakers.

Instruments

Two instruments were used to measure participants' vocabulary and writing, respectively. The Vocabulary Size Test (VST) gauged how many words a participant recognized, while an essay exam marked using the TOFEL iBT ® rubrics measured writing ability.

Nation's Vocabulary Size Test A

Nation's Vocabulary Size Test A is a 100-question, multiple-choice examination created to measure the number of words known by an ELL (Nation, 2012). The test was created following the logic that individuals learn (recognize) words by exposure, so ELLs are more likely to learn the most common, or frequently seen and heard, words before they learn the specialized, or infrequent, words. Therefore, a test could be created that samples words in bands (levels) of difficulty to determine mastery of each band of 1000 words.

Before going on, two points must be addressed: first, what counts as a word, and second, how is a frequency list created? To begin, Bauer and Nation (1993) organize words into families with headwords, so by having familiarity with one of the members of that group a learner could work out the meaning of the other members with a basic understanding of prefixes and suffixes. For example, the headword *learn* includes *learned*, *learner*, *learners*, *learning*, *learns*, and *learnt* (Cobb, 2017); altered forms of the headword that function differently grammatically within a sentence but have a shared basic meaning. Next, the VST uses a combination of the written and spoken British National Corpus to sequence the words into frequency order. This method was determined when it became clear that common words such as *hello* are seldom written though often said (Beglar, 2010).

The first version of the VST assessed the first 14,000 most frequent words in English (Beglar & Nation, 2007); however, a second version of the test was then created to assess recognition of the 20,000 most common words (Coxhead, Nation, & Sim,

2014). This second version was chosen for the final exam because the 14,000-level test was used as one, among several, intake assessments during the first month of Grade 10 to determine the background knowledge of incoming students.

Greek Myths Essay Exam

The writing sample was the final exam essay for the Grade 10 English course. The last quarter of the year focused on a unit where the students read *D'Aulaires' Book of Greek Myths* (D'Aulaire & D'Aulaire, 1962) and learned about the pantheon of Greek gods, goddesses, heroes and heroines. This material exposed Chinese students to characters, stories and themes much referenced in Western literature in general; furthermore, it primed students for the Grade 11 and 12 curricula where they would read widely in preparation for the AP English Language and the AP English Literature exams.

The test prompt encouraged participants to write compare-and-contrast essays using direct examples from *D'Aulaires' Book of Greek Myths* (D'Aulaire & D'Aulaire, 1962) about the Greek gods, goddesses, heroes and heroines. In particular, students were asked to look for parallels and peculiarities among the pantheon of characters and catalogue these details into a well-reasoned argument which drew informed associations that showed understanding of the material and society (Appendix B).

TOEFL iBT® Integrated Writing Rubrics

Once the topic and format of the essay examination had been determined, an objective method for analyzing and marking the written responses was required. After much consideration, the TOEFL iBT® Integrated Writing Rubrics (Appendix C) was chosen because it is well understood by the participants in the study who intended to take

the TOEFL exam in preparation for (as a requirement of) their international university admission applications. Moreover, the TOEFL iBT® Integrated Rubrics has been assessed by other researchers (Cumming et al., 2005; Riazi, 2016) and used in other studies and found to be a reasonable tool for measuring writing performance (Baba, 2009).

Validity and Reliability

The Validity and Reliability of Nation's VST

Nation (2006) has determined that, in a given piece of writing, the 14,000 most frequent words in the English language (plus proper nouns) account for more than 99% of the running text; therefore, a test like the VST, which assesses the number of known words, would evaluate approximately how much of the vocabulary in a given reading a student could recognize. Beglar (2010) assessed the 140-question version of the VST under a Rasch framework and determined the examination was valid and reliable for interindividual and intraindividual measurements and “displaying adequate fit to the Rasch model” (Beglar, 2010, p. 116).

Even critics of the VST agree that the test “has value as a measure of a more general form of written receptive language proficiency” (Stewart, 2014, 277) though it may not be a precise enough measure to estimate a test-takers' exact written vocabulary size. Originally, it had been hoped when the test was designed that it would be able to give examiners a reliable measure of how many English words (within 100 to 200 words) a test-taker knows—through receptive recognition of written vocabulary. However, the

test may be best used to determine “true score variance [where] the VST likely serves to distinguish learners with larger vocabulary sizes from learners with smaller ones” (Stewart, 2014, p. 280). That means that a student who responds to 38 questions correctly recognizes fewer English words than a student who answers 46 questions correctly; although not as catchy as saying an individual knows 7,600 words and another knows 9,200 words, it is clear that there is a difference between the two results which is useful when choosing textbooks and creating curriculum.

The Validity and Reliability of the TOEFL iBT® Rubrics

First, it must be asserted that the writing required by the TOEFL exam has similarities to typical writing assignments given in school settings. Riazi (2016) endorses this claim “that the textual features of the texts produced in the [TOEFL iBT® writing section] test situation are not significantly different from those produced in the real-life academic writing” (p. 25). Thus, using the TOEFL grading rubrics is an adequate assessment tool to measure writing in a school setting—especially with a population preparing for this language exam.

This idea is supported by Cumming et al. (2005), who worked on the prototype for the revised version of the current TOEFL. They explain, “the integrated writing tasks require examinees to write compositions that summarize ideas coherently that have appeared in academic-type source texts. The integrated tasks require complex cognitive, literate, and language abilities for comprehension as well as to produce written compositions that display appropriate and meaningful uses of and orientations to source evidence” (Cumming et al., 2005, p. 34). Basically, the examination is designed to

evaluate a test taker's ability to read, understand, summarize, synthesize, and write an essay that shows what he or she has learned. This goal matches how I wanted to assess the Greek Myth essays.

Data Collection, Management, and Analysis

There were five basic steps to the research process. First, I asked for and was granted permission to conduct research within the school. Second, the English Department administered the two final exams. Third, the VST scores and written essays were given to me for use in this study. Fourth, the essays were rescored by three individuals per the TOEFL iBT® Integrated Writing Rubrics. Fifth, the data was analyzed using SPSS.

I was granted access to school data in May 2015 after several discussions with the school principals and administrators about the nature of my research project. This allowed me time to coordinate with the class teachers to collect data from the final examination period in June 2015.

The two examinations were administered during one examination period. Students sat a one-hour-and-thirty-minute examination which consisted of three parts. The first 40 minutes was allocated for Paul Nation's Vocabulary Size Test; next, students took a multiple-choice test covering content from the Greek myth unit, which is not included in this study; finally, the students spent the last 30-40 minutes writing the essay. This multiple section, multiple choice, and essay examination was not untypical of other

exams given over the course of the school year during midterm and semester assessments.

It should be noted: Chinese students have a great deal of experience sitting long examinations. Throughout their academic career, they are regularly tested in each subject. In high school Grade 10 alone, this was the fourth major English exam of the school year; they had already taken an exam each quarter which was similar in length and content consisting of multiple choice, vocabulary and writing sections.

Once the examinations were completed, they were marked for class grades; then, the documents were scrubbed of identifying details, and the VST scores and essay papers were released for external assessment. The VST was graded using a computerized scantron machine within the high school, which allowed for quick, reliable scores to be generated. The essays were marked by three independent readers who each have more than 5 years of experience teaching English to students abroad.

Finally, the numerical data was explored using SPSS, a statistical analysis software package. First, descriptive statistics were generated concerning the demographics of the population, including their gender and age and the mean and range of scores for the VST and the essay. Second, correlative statistics were generated to determine the relationship between the two exams; a Pearson correlation coefficient and regression analysis were generated to tease out which results were more than a chance occurrence and indeed statistically significant.

Ethical Considerations

Ethical considerations were the main determinant when designing and implementing the research plan. Because of the difficulty of working with minors and parents in a multi-lingual environment, where IRB oversight is not well understood within the local community, I was limited to analyzing data that had already been captured in standard examination papers.

Previous attempts to work with the same community on another study never progressed past the planning stages because of cross-cultural communication concerns. After getting permission to present my research plan to approximately 100 parents at an official parents' meeting, only two consent forms were returned. A confidant within the school explained the lack of support: within China, research has traditionally just been approved by the school or only carried out in special experimental classes, so any attempt to get assent and consent from students and parents makes the research seem, from the parents' perspective, extremely unusual at best and risky at worst.

Furthermore, the parents' reluctance to allow their children to participate might have been a function of pure lack of understanding since the PowerPoint presentation and consent forms were translated documents, which could have lacked the necessary nuance to persuade parents that the research was valuable, that their children would be adequately protected from harm, or that they would benefit from participating.

Therefore, a non-experimental correlational design was chosen. With this, the school provided consent to use specific data from specific courses, and then participant

data were furnished for research purposes. This arrangement was more culturally acceptable.

Limitations and Delimitations

Nation's Vocabulary Size Test

Without doubt, the number one limitation with this research is that it is an exploratory pilot study completed with a relatively small number of participants. Another limitation is that the two tests were given in a classroom setting instead of a one-on-one testing environment; for the vocabulary size test, Nation advised individual testing (Nation, 2006). Third, questions have been raised about the precision of parallel versions of the VST; therefore, it may be difficult to estimate the precise number of words known receptively by a student because guessing on the exam may inflate scores (Coxhead et al., 2014); however, students who have more correct answers on the VST know more words than those who get fewer answers right, so scores can still differentiate between test takers' word recognition skills but may not be able to exactly measure the precise number of words known because of the general fallibility of test designs calculating student knowledge.

Essays

The Greek myth essay prompts are uniquely original to this study and have not been vetted by other researchers, so an objective measure of their difficulty is not available. This is an unfortunate but common problem in EFL research where the dictates of the curriculum often impose limitations on the scope of the study

TOEFL iBT® Integrated Writing Rubrics

The TOEFL iBT® Integrated Writing rubrics is an externally accepted assessment tool used around the world to differentiate the quality of writing among English learners. Consequently, as long as the use of the rubrics is consistent among the three scorers, which was ensured through training and norming before the scoring session and statistical analysis afterward, then the quality of the scores can be determined.

Summary

This project was designed to look at the relationship between receptive vocabulary knowledge and productive writing proficiency among EFL Grade 10 students in a high school in mainland China. Participants took Nation's VST and wrote an essay exam, and their scores were then analyzed using SPSS to determine to what extent these two language skills are correlated.

This research is another attempt to look at what others have previously studied in different language learning contexts. In the 1990s, Koda (1993) observed US students studying Japanese as a foreign language and showed the clear relationship between vocabulary knowledge and writing ability. Then in the 2000s, Baba (2009) assessed Japanese EFL students and found a moderate link between words and writing. These two studies amply demonstrate that more research is needed in foreign language studies to determine how learning words impacts using words. This current study among Chinese high school students is another attempt to verify past research results and therefore to better understand how particular learning situations may change the statistical significance of word knowledge on essay production.

Chapter 4: Analyzing the Results from the Vocabulary Size

Test and the Essay Exam

Travel is a good metaphor for research as they share many things in common. For me with both, I set out on a quest to discover new things, challenge my assumptions, and learn how reality may appear as I imagined yet be unbelievably counter to my expectations. When I began researching English language learning within the Chinese education system, I naïvely thought: how different can it be as they attend schools, go to classes, read books, and take tests just like American students? It is clear such thinking is a trap that oversimplifies the philosophical differences, so the variety of beliefs should not be underestimated.

As I tried to relate with my personal experiences of living and working in China in the first chapter, there are many differences between the U.S. and China in what, how, and why we teach English. Much more time should be spent researching cultural assumptions that may or may not be beneficial when teaching and learning. Furthermore, more energy should be invested in understanding how differences which emerge from the complexities of historical circumstances, cultural upbringing, and personal learning differences influence individuals and societies. Clearly, the plethora of questions yet to be asked overwhelm what can be explored in this dissertation.

This study is one step towards discovering reliable theories and best-practices in EFL studies. It is evident with the specificity of my own research question: does English vocabulary knowledge correlate with writing ability? A comparative analysis of the Vocabulary Size Test, which measures vocabulary knowledge, and the Essay Exam,

which measures writing ability, shows that there is a relationship between vocabulary learning and essay writing. The results of this study are similar to past research among different populations of language learners. However, though the research results are not unanticipated, the correlation between words known and writing quality compositions is weaker than many Chinese learners would have anticipated and stronger than I expected. Language learning is a complex process, so how two different components work together, such as words and writing, is still in need of much more exploration.

This chapter is organized as follows. First, there is a brief introduction to the population and the two different measures used with general statistics about the students' overall performance on the exams. Second, there is an analysis of how the population did when considerations for age and gender differences are included. Third, an exploration of the relationship between performance scores on both exams reveals some links that can be made between the two scores. Finally, the chapter will conclude with a summary of the most salient results from the study and suggestions for future research.

Population

The participants in the study are high school students in a megacity in mainland China. They attend a prestigious public school and are enrolled in a special section of the high school, called an International Curriculum Center (ICC), which prepares students for universities abroad; they take the required high school courses and exams to pass the Huikao exam (the high school graduation test) in addition to a full regimen of globally accredited coursework (such as Advanced Placement, A Level, or International

Baccalaureate programs) that are useful for attending international institutions of higher education overseas. However, at the end of their senior year, they opt out of taking the Gaokao exam (for admittance into Chinese tertiary programs) because they do not believe it necessary to take this difficult exam when going abroad for study, and because it is considered one of the most competitive and stressful university entrance exams in the world (Ash, 2016). Yet this decision then hinders them from applying to and enrolling in public colleges in China. At the time of the study, the participants were taking their final exams to complete Grade 10 in an ICC program.

This sample population took two different examinations; the first measured their vocabulary knowledge, while the second assessed their writing ability. Nation's Vocabulary Size Test measured students' familiarity with the 20,000 most common words in the English language (Coxhead, Nation, & Sim, 2014), while the Essay Exam graded with the TOEFL iBT® Integrated Writing Rubrics assessed students' performance on a compare-and-contrast essay (Baba, 2009; Cumming et al., 2005; Riazi, 2016).

Nation's Vocabulary Size Test

In the convenience sample of Grade 10 English participants ($N = 38$), the Vocabulary Size Test (VST) scores range from 34 to 85 questions answered correctly this means on average participants answer 48 questions correctly, which, according to Nation (2012), shows that test takers know approximately 9,600 English word families (questions answered correctly x 200). This is impressive for an EFL Grade 10 student since an average 16-year-old, native-English-speaking high school students in a New Zealand study answered 65.46 questions correctly which is equivalent to knowing

approximately 13,092 English word families (Coxhead, Nation, & Sim, 2015). Knowing more than 10,000 words indicates a vocabulary level that is college ready—especially for EFL learners (Beglar & Nation, 2007).

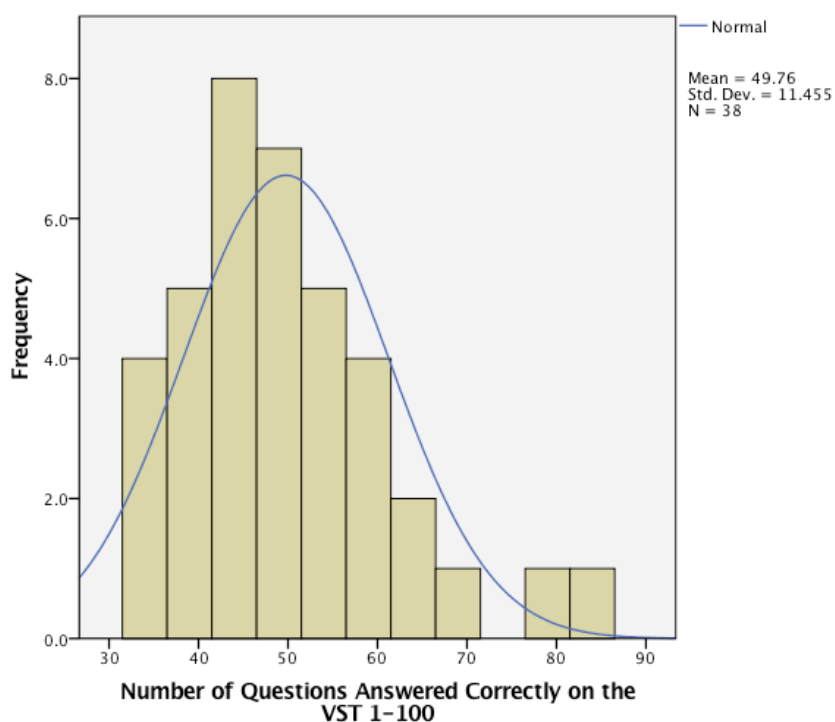


Figure 4. 1 The Vocabulary Size Test measures familiarity with the 20,000 most frequent words in the English language. The median score for the Chinese EFL participants is 48 questions correct out of 100. This score could be interpreted that on average participants recognize 9,600 English words and their meanings. ($M = 49.76$, $SD = 11.46$). The median score for the VST is 48 ($M = 49.76$, $SD = 11.46$);

Essay Exam

In the convenience sample of Grade 10 English participants ($N = 38$), the Essay Exam scores ranged from 2.00 to 5.00. The median score for the Essay Exam is 4.33 ($M = 4.06$, $SD = 0.77$), so on average participants wrote strong essays that would give them a good score on the essay section of the TOEFL examination and would, all things being

equal (all other sections of the TOEFL being comparably strong), enable them to enter a good university abroad. On the Essay Exam, the participants scored much closer to the high end of the exam. Since the exam grading rubrics was designed as an undergraduate college entrance measure, the Grade 10 participants were much closer to reaching the ideal score since they were well-prepared for advanced high school material.

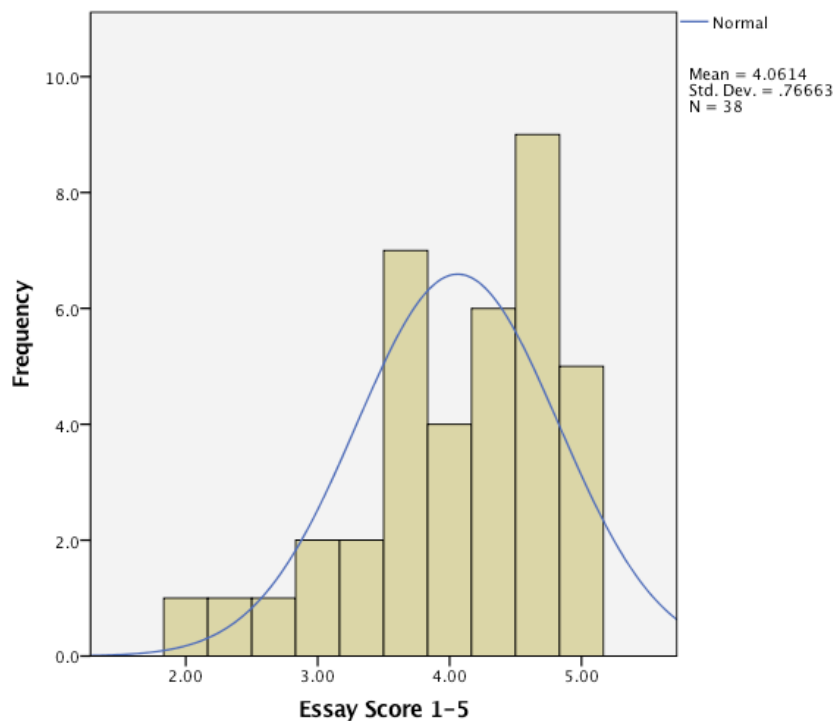


Figure 4. 2 The Essay Exam assessed writing ability among participants. It was graded on a scale of 1-5 using the TOEFL iBT® Integrated Writing Rubric. The mean score for the Chinese EFL participants is 4.06 out of a total of 5 while the median score was 4.33 out of 5. This shows that the quality of the essays is generally good for high-school, foreign-language writers.

Accuracy of Marks for Essay Exam

The Essay Exams were holistically scored by three experienced native-English-speaking teachers using the TOEFL iBT® Integrated Writing Rubrics (Appendix C).

Each essay was evaluated on a five-point scale independently by these three graders. These scores were then assessed using Cronbach's Alpha reliability statistics; there was an inter-rater reliability of 0.736 (Table 6.1), which is satisfactory for exploratory research. After ensuring that the three individual scores were consistent—meaning each grader was dependably assessing students' written work within a similar range—the three marks from the three graders for each participant were then averaged to create one composite Essay Exam score for each individual. This composite score was then used in comparison to the Vocabulary Size Test scores.

VST & Essay Exam Are Not Normally Distributed

The Komogorov-Smirnov test and the Shapriro-Wilk test (Table 6.2) both show this sample is not normally distributed. This is most likely because the participants come from a highly selective group of students enrolled in a prestigious high school that requires outstanding entrance examination scores for admissions. In addition, it is because the sample size is small, and it is well known that fewer participants allows for more weight to be given to each data point permitting anomalies to take on greater significance than would be seen with a larger data set. Non-normal distributions require the data to be analyzed using different statistical tests from those of a normal distribution to ensure accuracy of results. Therefore, Levene's Test for Equal Variance was used with the T-test to determine if the numerical differences were statistically significant results or likely due to chance. The following sections highlight the results.

Age: Tigers, Rabbits, and Dragons

One way to understand age within the Chinese context is to use the Chinese system of zodiac signs because this is how students and parents think of themselves in relation to others in their world. Chinese zodiac signs are much more culturally important than the star signs used in the Western world. It is not uncommon for someone who has just been introduced to inquire into your zodiac animal (你属什么? Ni shu shenme?). Everyone from taxi drivers to colleagues will ask you about your sign when casually getting to know you. It is an easy, polite way to estimate a person's age and to make some generally believed predictions about personality and fortune.

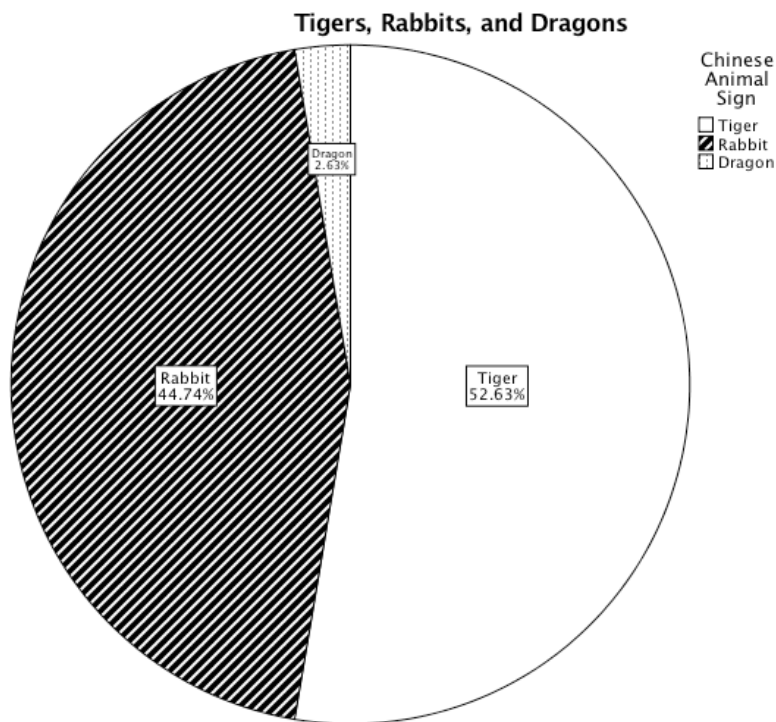


Figure 4. 3 Many of the students were 16 years old at the end of Grade 10. There were 20 Tigers (52.6%), 17 Rabbits (44.7%), and 1 Dragon (2.6%).

The Chinese belief in the zodiac is so strong that a significant number of parents plan to birth children in years with more auspicious fates; in 2012, the auspicious year of the Dragon, this led to a 5% increase in the number of children being born (Hsueh, 2016). The cultural beliefs around animal signs are now being empirically tested despite their relative mystical origins. For example, in an article entitled “Why Chinese Born in Years of the Dragon Are More Successful,” (2017) *The Economist* reported on a new study by Naci Mocan and Han Yu that assesses examination scores for 15,000 Chinese high schoolers; their data shows significant differences in performance among all the Chinese Zodiac signs, with the dragons outperforming all other signs. This new research shows how societal attitudes influences reality as parental expectations and investment in their young dragon’s education yield results. Unfortunately, my results did not confirm or deny any cultural beliefs about Chinese zodiac signs as the data could not statistically show that students born under one animal’s sign performed better than those born under another sign.

The Grade 10 English participants ($N = 38$) were divided by age into 15, 16, and 17 year olds. They were born under three different animal signs: the Tigers were born between January 28, 1998 and February 15, 1999; the Rabbits were born between February 16, 1999 and February 4, 2000; and, the one Dragon was born between February 5, 2000 and January 23, 2001. In total, there were 20 Tigers (52.6%), 17 Rabbits (44.7%), and 1 Dragon (2.6%). The majority were 16 years old at the end of Grade 10 when the examinations were administered.

Age and the Vocabulary Size Test

When I looked at the VST scores according to birth signs, I included the Tigers and Rabbits, but I excluded the single Dragon from the data because it was not statistically significant with only one data point. Despite a higher average score for Rabbits ($M = 50.18$) when compared to Tigers ($M = 47.95$), the discrepancy is likely an accidental phenomenon because it is in the range of standard error (Table 4.1). A larger test population size or a greater difference in mean scores would better test the hypothesis that Rabbits performed better than Tigers on the VST.

Table 4.1 On the Vocabulary Size Tests there is no significant difference between Tigers' and Rabbits' performance; the Rabbits' average higher scores may be attributed to chance (reference Table 6.3).

Chinese Zodiac	Number	Mean	Standard Deviation	Standard Error Mean
Tiger	20	47.95	10.102	2.259
Rabbit	17	50.18	11.159	2.707

Age and the Essay Exam

Table 4.2 On the Essay Exam there is no significant difference between Tigers' and Rabbits' performance; the Rabbits' average higher scores may be attributed to chance (reference Table 6.4).

Chinese Zodiac	Number	Mean	Standard Deviation	Standard Error Mean
Tiger	20	3.9333	0.88258	0.19735
Rabbit	17	4.1569	0.59064	0.14325

Then, I looked at the Essay Exam according to birth sign: Tigers and Rabbits. The single Dragon was excluded from the data. Despite a higher average score for Rabbits (M

= 4.16) when compared to Tigers ($M = 3.93$), the discrepancy is likely chance because it is in the range of standard error (Table 4.2). As stated previously, more data is needed to better test if zodiac signs (age) impact scores.

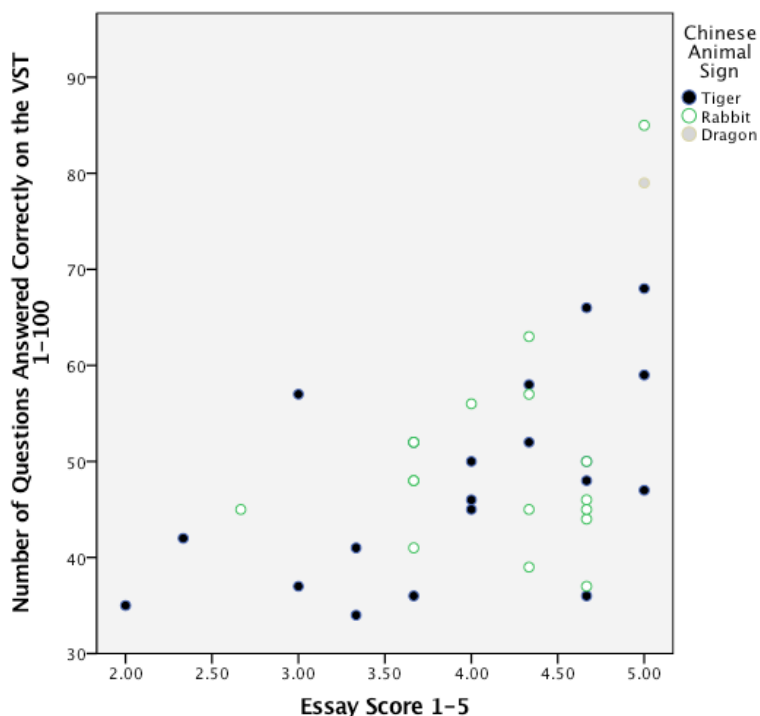


Figure 4. 4 This scatterplot presents the Vocabulary Size Test (VST) scores (vertical axis) and the Essay Exam scores (horizontal axis) for the Tigers (dark dots), the Rabbits (white spots), and the Dragon (gray).

Nonetheless, looking at the scatterplot graph (Figure 4.4) that shows age in relation to the Chinese Zodiac signs, a discernable pattern emerges at the bottom of the graph: the majority (6 out of 7) of the lowest performers are Tigers, who are older in age. This is unexpected because I believed that older individuals would be more mature and have more cognitive development enabling them to do better academically. However, this data does not support this claim. Instead, it hints to a more complicated picture since

Tigers have the largest range of Essay Exam scores spanning from highest to lowest.

Therefore, there is likely another external factor probably creating this result.

Gender

Gender differences in China are going through a transition from a traditional Confucian model where women are submissive to the male authority in the household to newer ideas of equality of everyone in a society of the proletariat (Hooper, 1975).

However, there is still a tension between older models of gender roles and contemporary beliefs about what it means to identify as male or female.

This population's majority-male demographic undeniably changed the dynamics of the course because many of the male students were more assertive when verbally asking and answering questions in class while the female students were much less outgoing. It is unfortunate that the stereotype of active males and shy females was so ingrained among the students when they arrived in my class. Therefore, I made it an imperative to pay close attention to balancing in-class participation and training students to share the floor in a more equitable manner.

Notably, the females had no difficulty keeping up with and surpassing their male counterparts in written communications and homework assignments. Lack of verbal participation did not mean lack of an ability to perform the English assignments, as one might assume. This only highlights the fact that all students must be encouraged to practice the different productive skills such as speaking and writing and receptive skills

like listening and reading in class to ensure higher and higher levels of fluency and with fewer and fewer errors in form and style.

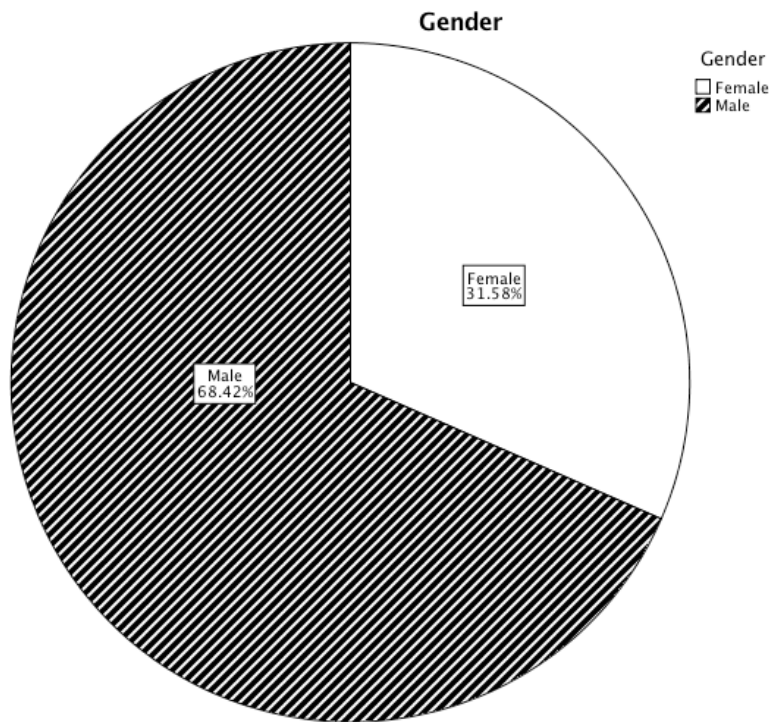


Figure 4. 5 The study population contained 12 females and 26 males.

Among the Grade 10 English participants ($N = 38$), there are 12 females (31.58%) and 26 males (68.42%); this is nearly a 1:2 ratio. Therefore, it was important to examine the VST scores along gender lines. Similarly as above, there are numerical differences between the mean male ($M = 52.17$) and female ($M = 48.65$) VST scores; however, the difference is not statistically significant (Table 4.3), so the discrepancy is likely due to chance.

Table 4.3 On the Vocabulary Size Test there is no significant difference between females' and males' performance; the females' average higher scores may be attributed to chance (reference Table 6.5).

Gender	Number	Mean	Standard Deviation	Standard Error Mean
Females	12	52.17	11.846	3.420
Males	26	48.65	11.331	2.222

Males and Females on Essay Exam

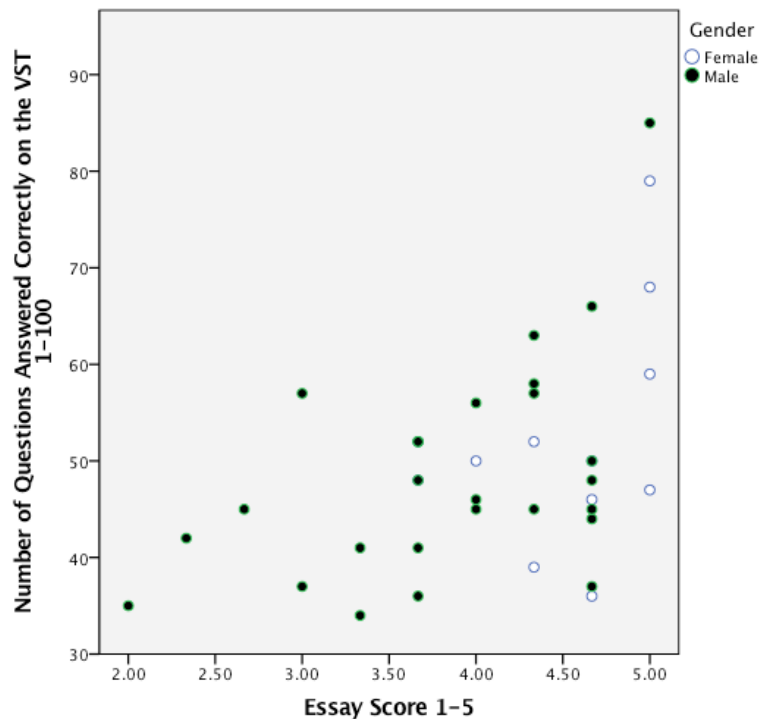


Figure 4. 6 This scatterplot presents the Vocabulary Size Test scores (vertical axis) and the Essay Exam scores (horizontal axis) for the males (dark dots) and the females (white dots).

The females earned scores ranging from 4-5 on the Essay Exam while the males range from the highest scores in the population to the lowest. This data confirms what I intuitively understood as a teacher: the females in the population, as a group, were

objectively good students, average to above average in the class, while the males had a wider range of abilities. The scatterplots graph (Figure 4.06) shows how the scores are dispersed across gender. When comparing the Essay Exam scores to gender it is apparent that the female participants' scores are clustered in a much tighter range closer to the top of the scale while the males' scores are spread out.

Subsequently, I evaluated the Essay Exam scores according to males and females, and found a numerical as well as statistical difference. The females ($M = 4.5$) had a higher average score than males ($M = 3.859$). In addition, this is a statistically significant difference between the two populations; that is, there is better than a 95% probability that the females performed better than the males on the Essay Exam (Table 4.4).

Table 4.4 On the Essay Exam, there is a numerical and statistical difference in scores between males and females, with females' performing better than males' performance (reference Table 6.6).

Gender	Number	Mean	Standard Deviation	Standard Error Mean
Females	12	4.5	0.50252	0.14506
Males	26	3.859	0.78979	0.15489

Relationships between Measures

After looking at the descriptive statistics for the scores, I found it necessary to see if there is a relationship between the examination results. To aid in determining whether, and if so to what extent, such a relationship exists, I looked at the Pearson Correlation Coefficient and performed a Regression Analysis.

Pearson's Correlation Coefficient

The data suggests that there is a moderate correlation ($r = 0.475$) between vocabulary knowledge and writing ability. When the scores from the VST were correlated with the scores from the Essay Exam using the Pearson's correlation coefficient, it was clear that there was a relationship between the number of words known and the ability to write a top-scoring composition.

Table 4.5 The Pearson's correlation coefficient is 0.475 (significance correlation shown above 0.01 level), which indicates that the two variables are positively correlated with 99% confidence (Sig of $0.003 < 0.01$). This is a moderate, statistically significant correlation between the two variables.

		Vocabulary Size Test 1-100	Essay Score 1-5
Vocabulary Size Test	Pearson Correlation	1	0.475**
	Sig. (2-tailed)		0.003
	N	38	38
Essay Exam	Pearson Correlation	0.475**	1
	Sig. (2-tailed)	0.003	
	N	38	38

As in other research conducted by Baba (2009) and Koda (1993), the correlation is not strong; instead it is moderate which suggests that vocabulary is important for composing prose but also more than words are needed to construct pieces of writing.

Regression Analysis

With the regression analysis, it was hoped to be able to make a predication about students' performance on the Essay Exam bases on their Vocabulary Size Test. However, the Among the Grade 10 English participants ($N = 38$), 33% of the Essay Exam score can

be accounted for by lexical knowledge as measured by the VST (20%) and by being female (13%); this leaves two-thirds of the Essay Exam score to be determined by other factors. Therefore, more research needs to be conducted to probe what additional factors influence writing proficiency on Essay Exams (Table 4.6).

Table 4.6 Regression Analysis

Model	R	R Squared	Adjusted R Squared	Std. Error of the Estimate
1 ^c	0.475 ^a	0.225	0.204	0.684
2 ^c	0.577 ^b	0.333	0.295	0.644

a. Predictors: (Constant), VST

b. Predictors: (Constant), Gender, VST

c. Dependent Variable: Essay Exam Score

Model 1 Analysis: R Square=0.225. From the data it can be speculated that 20% of the Essay Exam score can be ascribed to changes seen on the VST. Model 2 Analysis: R Square=0.333. From the data it can be estimated that 13% Essay Score can be attributed to gender. Therefore, 67% of the Essay Exam score is determined by other factors which have yet to be assessed and measured.

Summary

The Grade 10 participants ($N = 38$) include more than 2 males (26 males total: 68.42%) for every 1 female (12 females total: 31.58%), while the ages are split relatively evenly between the Tigers (20 Tigers: 52.6%) and Rabbits (17 Rabbits: 44.7%) with only a single Dragon (1 Dragon: 2.6%).

The results from the statistical analysis of the two examinations show that there is a moderate correlation between vocabulary knowledge as measured by the VST and writing ability as measured by the Essay Exam. To be more precise, 20% of participants'

Essay Exam scores can be attributed to their performance on the VST; furthermore, the female participants did statistically better than the males. However, 67% of the Essay Exam score is not influenced by vocabulary knowledge or gender, so ongoing research should address other possible factors.

Based on the TOEFL iBT® Integrated Writing Rubrics (Appendix C) test takers must have a strong command of English grammar and style, so ideas are not obscured by inaccuracies; they must understand coherent organization, so key points are well arranged; and the essay must include accurate and relevant information, so unnecessary material does not distract. Thus, robust measures to assess language learners' grammatical abilities, organizational proficiencies, and critical thinking skills are needed to gauge what portion of essay writing is dependent on the mastery of these components of the grading rubrics. In addition, many other factors which may impact writing ability have yet to be explored.

In conclusion, vocabulary knowledge as measured by the VST does influence writing ability as measured by the Essay Exam; however, there isn't a simple linear relationship that can predict results on either exam based on the results of the other. Therefore, further investigation is needed to determine if there are simple, efficient, and effective measures that can be regularly used as predictive assessments for writing ability as teachers do not always have time to assign and assess full essays when teaching a heavy load with large class sizes. Furthermore, researchers should continue to explore the Chinese passion for memorizing English vocabulary words and its usefulness as a study technique as an exceedingly large vocabulary may not proportionately influence writing

scores but it may be more predictive of results in other areas such as reading and listening assessments. The final chapter will further explore the implications for future research and practice.

Reference Tables

These tables are in Appendix D: Chapter 4 Supplemental Graphs and Tables.

Chapter 5: Chinglish to Standard Academic Prose—Chinese High Schoolers become Pros at Writing Essays

In mainland China, millions of students learn English as a Foreign Language (EFL) each year, and a small but growing number leave their homeland to attend high schools and universities around the world. In 2016, over a million Chinese students were studying in the US alone (Institute of International Education, 2016). With this kind of educational emigration, it is important to examine the Chinese EFL schooling system, pedagogical practices, and cultural beliefs to better understand the mainland students entering classrooms around the world; this would include answering basic questions such as how much English education do students receive during their formative education, which textbooks are used in the schools, and what teaching and study methods are recommended most frequently throughout China. With so many possible directions for inquiry, I chose to focus on how English is studied and taught in metropolitan China. More specifically, my research came from observing my students memorizing lists of vocabulary words in their free time and struggling to write essays during class, which highlighted the gap between their desires to learn English and their abilities to use English. My path to understanding the secondary education system in China and finding culturally suitable ways to teach English to high schoolers was a process of exploration: a circular activity of reading the literature, observing my students and colleagues, teaching my courses, and revising my thinking.

In the fall of 2012 when I arrived in China to teach English at a metropolitan public high school, I was unprepared for the realities of English teaching within the

Chinese education system. The vast differences in beliefs about the how and the what of the curriculum blindsided and humbled me while also leading me to question both the Chinese and Western ideologies about English-language education.

Having attended school in the US, my ideas about writing were strongly influenced by the historical circumstances of the US educational system. In the 1990s, I had been taught to write using process pedagogy with a strong dose of rhetoric and Expressivism (Murray, 2004), while at the beginning of the new millennium, my graduate studies had included coursework in multiliteracies and cultural studies (Cooper, 2010). It seems I am a product of my times as my educational experiences mirrored the trends of composition theory and teaching during the decades I was in school. Fulkerson (2005) describes the field's movement over the last half century through the influences of current-traditional thinking, Expressivism, process pedagogy, critical-cultural studies, and rhetoric. My experience with second language acquisition studies was through its overlap with composition studies in basic English and world Englishes and not through linguistics, which was the more traditional route described in Matsuda's (2006) history.

In China, the test-driven model (Huang & Brown, 2009) where students endlessly prepare for external examinations such as the Gaokao (Chinese university entrance exam), TOEFL (English proficiency exam), and SAT (US college readiness exam) made me wary. Another educational surprise was my students' insatiable desire to memorize vocabulary words as a way to learn English—called 背单词 (BeiDanCi) BDC. However, these were only the easily discernable differences; with historical research, it became clear that the field of English education in contemporary China was rooted in and

influenced by Russian communism more than Western language and composition theories (You, 2010). Translation and current-traditional methodologies were used in the twenty-first-century classroom alongside textbooks which supported the audio-linguistic method. Ultimately, all the theories of language teaching from the last century were used uncritically and in pieces to form a patchwork curriculum in Chinese classrooms.

Therefore, as a composition scholar educated in the West who is teaching EFL in China, I was required to sift through the school's course materials and pragmatically assess what worked with my students and what did not. In addition, I had many conversations with students and colleagues about their preferred language learning styles. Finally, this led to my desire to research English learning among Chinese high school students. I focused on the conflict zone where my students' desire to memorize vocabulary words met my desire to teach them essay writing skills.

This chapter reviews my doctoral research that statistically analyzes the Vocabulary Size Test (VST) scores and the Essay Exam results for Chinese Grade 10 students in a high school in metropolitan China. It summarizes the conclusions which show the degree of correlation between the number of English words known and one's ability to write an essay as well as other influences impacting the test outcomes. Then, it discusses the implications for the field of English language learning. Finally, the chapter concludes with recommendations for future research and future practice.

Summary of the Study

Once I determined the research population—metropolitan Chinese high school students learning EFL in mainland China—I needed to focus my research question. When looking around for a dissertation topic, I noticed the ubiquitous green and red TOEFL (Yu, 2011b) and SAT (Yu, 2011a) wordlist study guides published by New Oriental, which is a billion-dollar language-learning and test-prep empire. These books, which were in every locker and on every desk, flagged my students' desire to learn English vocabulary and signaled their belief that this was fundamental to their language learning process. However, when I reviewed their written assignments, it was also evident that they lacked the necessary essay writing skills needed to pass these same English language examinations essential for admissions to international universities. Therefore, I decided to focus on the space between what my students thought was essential in language learning—vocabulary—and what I thought was critical for their success—essay writing.

The first question I examined was the place of lexis in language learning. This seemingly simple concept is fraught with complexities of deciding what is a word and what does it mean to know one. For this doctoral research, I adopted Nation's ideas as they relate to his Vocabulary Size Test (VST) which I used to measure participants' vocabulary size or word knowledge (Nation, 2012); in this study, a word is a headword, or family of related words under one simple form, and knowing a word means that an individual can recognize the meaning when it is presented as a multiple-choice answer to a question on an exam. For example, one VST question asks test takers to identify the meaning of the word *soldier*. It is then assumed that if an examinee can recognize the

headword *soldier*, she or he would also be able to recognize *soldiers*—the related word.

A typical question on the examination would look like this:

- soldier: He is a soldier.
a. person in a business
b. student
c. person who uses metal
d. person in the army
(Nation, 2012)

My second line of inquiry concerned essay writing requirements for high school EFL students intending to go abroad for further study. These individuals usually take the TOEFL Exam, or a similar exam such as the IELTS, to prove their English proficiency level before being admitted to an international university. Therefore, my doctoral research used the TOEFL iBT® Integrated Writing Rubrics to look at compare-and-contrast essays written about the Greek myths unit in the Grade 10 curriculum. This external assessment tool was adapted for an internal essay prompt, enabling the essay graders to be more impartial when marking the written work that was completed as part of the coursework. According to the TOEFL Rubrics, a good essay demonstrates that the writer has read, understood, synthesized and summarized the content material and has written a piece of several paragraphs that clearly explains the ideas in fluid, grammatical sentences (TOEFL iBT® Rubrics: Appendix C).

Finally, this study statistically analyzes the results of the VST and Essay Exam, looking at participants' word knowledge and writing abilities as measured by the two different assessment tools. The results reinforce similar studies conducted among other populations who were studying Japanese in the US (Koda, 1993) or EFL in Japan (Baba,

2009). Koda (1993) found: “Vocabulary knowledge was correlated highly with the overall quality of ratings in both [descriptive and narrative writing] tasks (Descriptive, $r = .75$; Narrative, $r = .70$)” (p. 334), while Baba (2009) found: “Two lexical proficiency indices—vocabulary size and word definition ability—had moderate correlations of .40 and .51, respectively, with summary writing performance” (p. 199). Koda (1993) recognized the relationship between vocabulary knowledge and writing abilities in foreign language learning, while Baba (2009) determined correlation between definition-generation abilities and summary-writing abilities.

This doctoral research statistically shows that there is a relationship between vocabulary knowledge and essay writing ability, but the correlation is not enough to be predictive of either vocabulary knowledge or writing ability. More research needs to be conducted and reported to help better understand the place of lexical knowledge in language learning, in general, and composition writing, in particular.

Summary of Findings and Conclusions

The Grade 10 participants ($N = 38$) are majority male with more than 2 males (26 males total: 68.42%) for every 1 female (12 females total: 31.58%). When considering age, the participants are more evenly divided between the Tigers (20 Tigers: 52.6%), who were born between January 28, 1998 and February 15, 1999; and Rabbits (17 Rabbits: 44.7%), who were born between February 16, 1999 and February 4, 2000; with only a single Dragon (1 Dragon: 2.6%), who was born between February 5, 2000 and January 23, 2001.

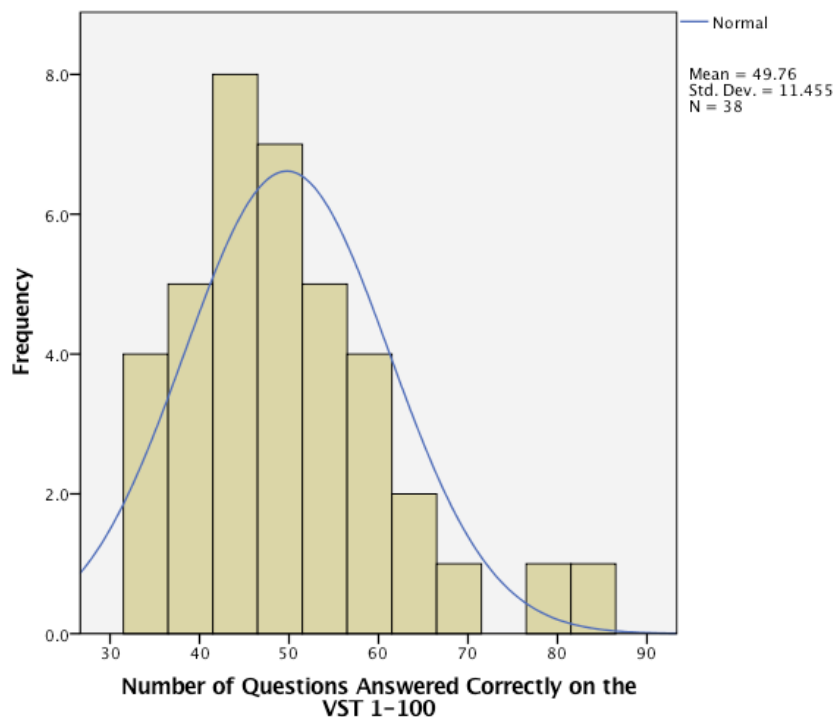


Figure 5. 1 The Vocabulary Size Test measures familiarity with the 20,000 most frequent words in the English language. The median score for the Chinese EFL participants is 48 questions correct out of 100. This score could be interpreted that on average participants recognize 9,600 English words and their meanings.

The participants' scores on the Vocabulary Size Test (VST) have a median score of 48 ($M = 49.76$, $SD = 11.46$) with a large range in the number of questions answered correctly between the lowest, who answered 34 correctly, to the highest, who answered 85 correctly (see Figure 5.01). According to Nation's (2012) matrix for estimating the number of words known in English, answering 48 questions correctly indicates that test takers know approximately 9,600 English word families. This is noteworthy since one study reported that an average 16-year-old native-English-speaking high schooler in New Zealand answered 65.46 questions correctly—approximately 13,092 English word families (Coxhead, Nation, & Sim, 2015). Another study estimated that an EFL student is

college ready when they know more than 10,000 words (Beglar & Nation, 2007).

Therefore, the participants in this study, on average, were nearly ready for university level coursework at the end of Grade 10. This was critical as many of them were preparing to study AP curriculum courses—university-level work—in their junior year.

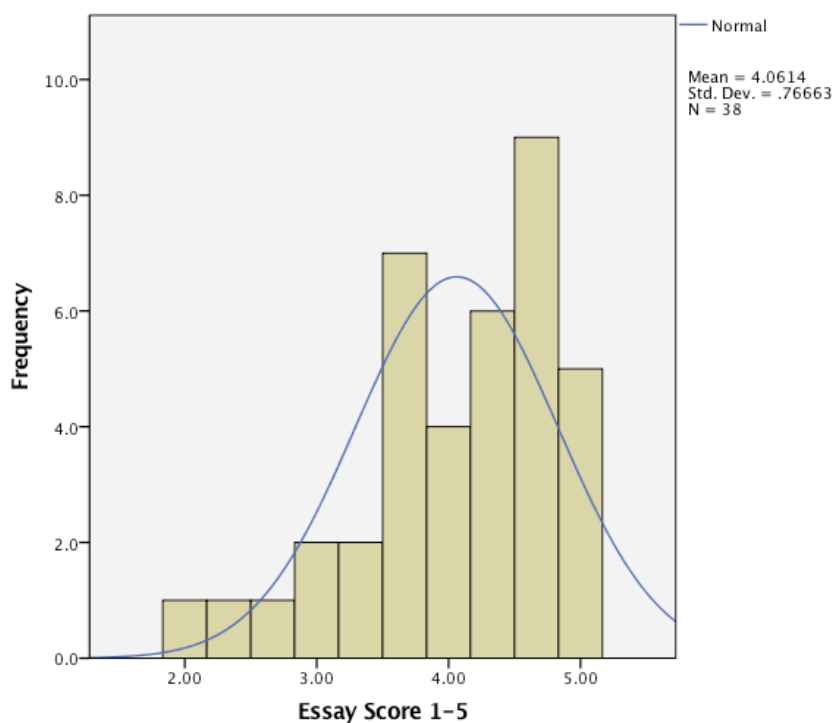


Figure 5. 2 The Essay Exam assessed writing ability among participants. It was graded on a scale of 1-5 using the TOEFL iBT® Integrated Writing Rubrics. The mean score for the Chinese EFL participants is 4.06 out of a total of 5 while the median score was 4.33 out of 5. This shows that the quality of the essays is generally good for high-school, foreign-language writers.

The participants' Essay Exam scores ranged from 2.00 to 5.00 with a median score of 4.33 ($M = 4.06$, $SD = 0.77$) (Figure 5.2). This average participant score indicates that they wrote strong enough essays to enable them to enter many well-respected tertiary institutions abroad—that is, if all other sections of the TOEFL (reading, listening, and speaking) were of comparable level. It should be noted that the participants scored much

closer to the high end on the grading rubrics for the Essay Exam. This is because it was designed as a measure for undergraduate college admissions; therefore, the Grade 10 participants, who were well-prepared for advanced high school material, were much closer to reaching the ideal score.

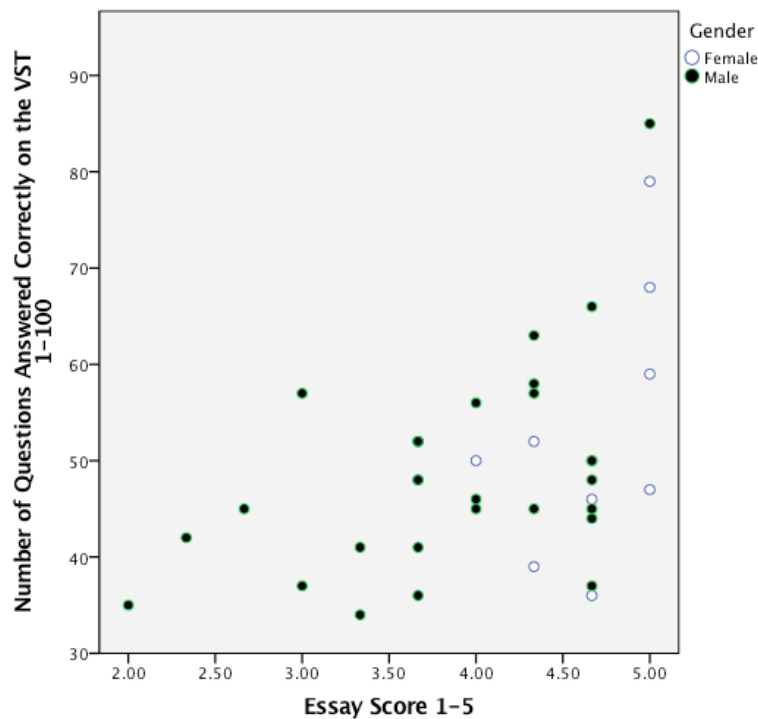


Figure 5. 3 This scatterplot presents the Vocabulary Size Test scores (vertical axis) and the Essay Exam scores (horizontal axis) for the males (dark dots) and the females (white dots).

Once the general descriptive statistics were analyzed, I looked in the data for possible trends in scores based on age or gender. The calculations did not show any statistically significant results to indicate that age impacts scores on either the VST or the Essay Exam in this particular population (see Chapter 4 for more details). Subsequently, I looked at male and female participants' performances on the VST and the Essay Exam; with the former exam, there were no statistically significant differences, but with the

latter, it became clear that the females' scores were statistically better. On the Essay Exam, the males' scores range from the highest to the lowest (range = 2-5; $M = 3.859$), while the females scores are tightly clustered at the top of the scale (range = 4-5; $M = 4.5$) (see Chapter 4 for more details). This data corroborates what I instinctively felt as their teacher: the female students were consistently getting As and Bs on their written classwork, while the males' grades ranged from As to Fs on class assignments. The scatterplot graph (Figure 5.03) illustrates the Essay Exam score dispersion.

When statistically analyzing the relationship between the scores on the VST and the Essay Exams, a moderate correlation was found between the VST, which measures vocabulary knowledge, and the Essay Exam, which measures writing ability. According to the data, 20% of participants' Essay Exam scores can be attributed to their performance on the VST; moreover, the female participants did statistically better than the males on the written assessment. However, 67% of the Essay Exam score is not influenced by vocabulary knowledge or gender (see Chapter 4 for more details). More research is needed to investigate what other factors, besides vocabulary knowledge and gender, affect writing ability among Chinese high school students and other EFL populations.

Implications

Vocabulary acquisition is not stressed among many second language teachers and theorists in the West where English is taught as a second, or additional language, among a large population of native English speakers. However, in China and other EFL

communities, word memorization is an important aspect of EFL learning, and this concept is emphasized. Lexes are the building blocks of language and are extremely important to meaning making, yet they are not enough. Organization is also critical to communication: words make phrases, phrases create sentences, sentences become paragraphs, and paragraphs build passages. The overemphasis on one part of the language learning process, such as vocabulary, grammar, or fluency, undermines the goal of recording or sharing ideas in academic and everyday situations. Therefore, both EFL and ESL communities need to reevaluate the assumptions delineating the place of lexes in language learning; the former may need to downplay the importance, while the latter may need to review vocabulary acquisition beliefs and practices and add more strategies for vocabulary acquisition.

One of the significant implications of my research was that I realized I must include more overt vocabulary instruction in my courses because my students wanted this, and it benefited them. First, my students were keen to keep vocabulary notebooks as they had been introduced to the idea of a “classic” vocabulary notebook “composed of lists of lexical words or phrases” (Ma, 2013, p. 241) by other successful Chinese students learning English. Wherever they looked it seemed, one person or another was encouraging them to keep a list of words they had yet to memorize. Second, it became increasingly clear that the more vocabulary a student knew the better they could read (Laufer & Ravenhorst-Kalovski, 2010) and write (see this chapter’s Summary of Findings and Conclusions).

Therefore, I initiated a word challenge in my Grade 10 English courses; each student was asked to find new words in their textbooks or course materials and keep a running list in their English folders (Appendix A). This was an ungraded assignment, but I encouraged participation each week by giving out candy to all students who successfully collected 30 new words and wrote definitions with example sentences or drawings. I pushed students to learn 500 new words in one semester by throwing a pizza party to all those who succeeded. The first time only a handful of students successfully completed 500 words in one semester, but the next competition required many more pizzas. By gamifying the acquisition of new vocabulary words instead of grading it as an assignment, I created a class culture where students motivated themselves and were rewarded socially and academically for their hard work.

Recommendations for Future Research

This dissertation research was an exploratory, pilot study conducted with an underrepresented population in the literature; therefore, there are many questions yet to be asked and researched. My list of suggested topics for future study has been compiled into general categories comprising the culture of English education, China's English education system, and English-language testing. English Education Culture

There is a lack of systematic research into the pedagogical practices of EFL teachers in different countries and education systems around the world. Much more research should be done to characterize and analyze textbooks, national curriculums, classroom practices, and the cultures of language learning that exist around the world.

The literature includes many individual accounts of what a single teacher or an individual school has done in a particular place, but there is little regular and methodical research into the EFL educational systems as they exist in countries as a whole. Often when this research does exist, it is not easily accessible outside of the country it concerns and it may not even be written in English—the shared language of English teachers and English scholars. Specifically, exploration into both formal education and informal tutoring programs from young children to adults needs to occur. We cannot continually focus on college students because it is convenient as university academics to study our immediate, proximate students; we must also look at formative educational practices that best prepare individuals to become successful and liberated life-long users of foreign languages not just published academics or professional linguists.

These kinds of questions would require more direct collaboration between the field of education and the EFL community to answer. English teachers and program directors must also be more consistently involved in inquiry as all would learn a great deal about best practices in language curriculum development, pedagogical methods, and classroom assignments when structures as well as components are scrutinized. For example, one large-scale project could be an international database compiling basic information about how many hours of English coursework students take in a year, what textbooks are used in the classroom and as extension assignments, how teachers are trained before starting to work and as they continue in their careers, and how efficient and effective are national programs for educating the population. This kind of research would encourage large programs and whole systems to assess how well they compare for time

and money spent on language education. Ideally this data would not be used to compel everyone to learn English but would encourage a diversity of language learning programs around the world so that more people can learn more languages that are beneficial or useful to their specific economic, social, cultural, political, and financial situations. In turn, such research could potentially increase language learners' competencies and improve international communication.

Chinese English Education

Within China, this kind of English education research is sorely missing—especially in English. From province to province and even city to city there are various arrangements in place for teaching and assessing English. For example, different provinces and cities use several state-approved textbooks to teach the national curriculum; in addition, students across the country take different versions of examinations such as the Zhongkao (high school entrance exam) and Gaokao (university entrance exam), which are rumored not to be equivalent in difficulty level. Therefore, even within China many argue about how to compare students' scores across the country to create a fair system that does not penalize or reward students based on their geographic circumstances and the required curriculum and examinations of their region.

This is a regular topic of interest for the national media: for example, in 2016, there were protests in Beijing because parents were concerned that the government was unfairly privileging minority students from the provinces in admission picks for prestigious Beijing universities when they were given permission to take the Gaokao (Chinese university entrance examination) in Beijing as opposed to their hometowns

(Xie, 2016). This is an example of unequal educational opportunities that occur across demographic boundaries and the social tensions that arise from these inequities. Because the provinces have different curriculums and different examinations to assess their students, it makes it hard to say if it is fair for the Gaokao immigrants who study in one place, then take their examinations elsewhere. What complicates this matter more is how top universities have admission quotas for enrolments from different provinces, so a mediocre student in a place with a difficult examination might become a top student in another area of the country with an easier test paper. To add more complexity, the minority students in the above-referenced article were selected from around the country to come and study at the High School Affiliated to Mingzu (Minorities) University in Beijing, so returning home would have been a double penalty because they would have taken an examination that they had not been prepared for, and they would have likely been negatively impacted by the university admission quotas. However, the Beijing parents just think of them as educational immigrants taking away seats at top universities from their children.

What's more, there are new competitors in China to the Chinese National Curriculum. Across the country, International Curriculum Centers that are attached to local high schools are teaching programs from around the world such as AP, A Level, and IB. That means that a small percentage of Chinese high school students are opting out of most of their national-curriculum courses to study another world-recognized curriculum. This new trend that started at the beginning of this millennium with the growing wealth of the middle classes has only accelerated in the last ten years. Each year

more and more ICC programs open across the country and the current programs increase enrollments. The exact data on how many students and schools are involved is not well documented (outside of the government agencies that monitor it closely), nor are the specifics of the exact curriculums and teaching methodologies well researched; although, some of the international curriculum bodies send their own observers to monitor the implementation of their programs in China.

Testing

There are many standardized English language tests available on the market for well financed schools and programs. However, a large percentage of EFL teachers do not have free and easy-to-use assessment tools for their classes, nor do administrators on a tight budget have widely accepted measures to gauge the effectiveness of their programs. This hinders English researchers and educators alike because neither can regularly and consistently run examinations that give meaningful results that can be compared across time and populations. Without data, it is difficult to impartially assess progress and to determine the ability of students and the efficacy of curricula and programs.

Therefore, it would be extremely useful for NGOs, professional organizations, or large university programs to create free tests for different components of English: reading comprehension, writing proficiency, listening level, speaking ability, vocabulary knowledge, and grammar skill. Granted, this work has already begun (Milton, 2009: Schmitt, 2010), but the variety of measures and the lack of stringent testing of those assessment tools make them difficult to adopt and reuse. In addition, not enough data is publically available to compare-and-contrast populations across time and space. Without

doubt, improvements in computer-aided learning tools will assist with accessibility to testing materials, collecting results from large numbers of diverse participants, and crunching the data. Furthermore, advances in the collection, use, and interpretation of big data should be harnessed toward this end.

Clearly there are pros and cons with the existing free assessment tools out there. However, more convergence between theory and practice—or thought and action—could improve English language education around the world and bring the field out of the domain of lore, opinions, and preferences by adding another layer of data analysis. Granted, we should not move so far into the world of big-data processing that we forget our roots in the humanities—statistical research should increase, not deter, our ability to teach English and communicate with each other effectively and efficiently.

Recommendations for Future Practice

The most important recommendations for future practice that emerge from this dissertation research directly relate to how vocabulary and writing should be taught in EFL environments. First, vocabulary should be consciously selected and consistently taught to minimize the burden of learning new words and to maximize the efficacy of time on task. Second, a wide variety of text types should be explicitly taught, explained and practiced; then, students can systematically learn different styles and forms of writing to address their academic and every-day needs.

Transparency of Vocabulary Used in Learning Materials

Vocabulary should be consistently taught. Language learners should follow a methodical process to learn the most frequent words first and the uncommon words as

needed. EFL textbooks and course materials should be much more explicit about lexical sophistication to allow better selection of resources and better preparation of students. Though many of the publishing houses have sophisticated processes for developing educational materials, which often include general information about reading level or the Common European Framework of Reference for Languages, textbooks' vocabulary levels are not always consistent nor is information about lexical sophistication available to those outside the company or organization publishing the materials.

Some publishers have developed leveled learning materials with specific lexical scores and even accompanying vocabulary materials, yet they are product-specific and not generalizable measures. For example, Penguin has a series of graded readers that use a specific set of words for each level with language outside this bandwidth being explained in a glossary. In addition, Nation has come up with a series of workbooks called *4000 Essential English Words* (Nation, 2009) to introduce the most frequent and common words in the English language. These kinds of learning materials, which focus on vocabulary, in particular, as opposed to a more general reading or language level score, are useful when educators want to add new resources or assign supplemental materials in a course.

Finally, more must be done to reassess best practices for learning vocabulary and language. With the plethora of technology available in the 21st Century, systematic testing of dictionaries, programs, apps and other learning tools must be conducted to establish the best assortment of methods that benefit diverse populations of learners at various stages in their cognitive development and language learning. At present, much of

this research seems to be in the hands of the free market where companies vie for economic success, not necessarily educational excellence. Furthermore, the research is unavoidably protected as trade secrets which run sophisticated algorithms in development.

Explicitly Teaching How to Write a Diversity of Text Types

Writing is a complicated process. There are many components to creating excellent written work; on the sentence level, vocabulary and grammar are critical, while on the macro level, the organization of ideas is paramount. The relationship between writers' ideas, sentences, paragraphs, and pages makes for endless arguments about what to teach when, why, and how. These discussions are complicated by the very different audiences of students who will experience the classes and their personal, professional, or academic needs.

Therefore, EFL instructors and composition theorists need to be cognizant of learners as they progress from simple writing tasks to more complex ones. More thought about the skills required to successfully complete different academic and real-life writing tasks should be considered. Whether to teach the essay or not is not the right question (Sztabnik, 2015; Sheppard, 2016; Caplan & de Oliveira, 2016). We need better pedagogy for teaching the multitude of text types that individuals will use in academic and daily life; letters, research papers, opinion pieces, and reports should have a place in the classroom alongside twenty-first-century communications such as e-mails, text messages, blogs, visual images, podcasts, and videos. Multiliteracies proponents (e.g. Zheng, Yim, & Warschauer, 2017) have been arguing this for years, but change comes slowly.

Finally, my doctoral research results suggest that vocabulary knowledge impacts writing exam scores, though it typically accounts for only one-fifth of a given mark (see Chapter 4). This means that learning thousands of words is important but not necessarily as important as learning how to use those words effectively in a piece of writing. The interplay between lexical sophistication and composition must continue to be parsed to help EFL students efficiently learn the language and effectively communicate their ideas.

Summary

Clearly, a great deal more research is needed among China's vast and diverse population of ELLs, especially as a growing number of Chinese students are emigrating around the world for secondary and tertiary educational opportunities. This dissertation is a small addition to the literature on vocabulary acquisition and writing skills among Chinese ELLs within the Chinese education system in mainland China. This doctoral research, which statistically analyzed the correlations between the Chinese Grade 10 participants' VST and Essay Exam scores, can only illuminate part of the gap between word knowledge and composing texts, so more scholarship is required. Future research is needed to better understand the culture of English education around the world, China's English education system throughout the country, and improved methods for English-language testing; in addition, we must improve future practice by encouraging transparency in the variety of vocabulary used in widely available learning materials, and we must find better ways to explicitly teach people how to write the diversity of text types needed in the 21st Century. Without doubt, much can be done to improve the way the world learns languages to facilitate communication for generations to come.

References

- Adams, J. H., & Sargent, T. C. (2012). *Curriculum transformation in China: Trends in student perceptions of classroom practice and engagement* Retrieved from http://repository.upenn.edu/gansu_papers/34
- Adamson, B., & Morris, P. (1997). The English Curriculum in the People's Republic of China. *Comparative Education Review*, 41(1), 3-26.
- Ash, A. (2016, October 12). Is China's gaokao the world's toughest school exam?. *The Guardian*. Retrieved from <https://www.theguardian.com/world/2016/oct/12/gaokao-china-toughest-school-exam-in-world>
- Baba, K. (2009). Aspects of lexical proficiency in writing summaries in a foreign language. *Journal of Second Language Writing*, 18, 191-208.
- Barton, L., and Dupaquier Sardinas, C. (2008). *NorthStar 3 Reading and Writing 3rd Ed.* New York: Pearson Longman.
- Bauer, L. and Nation, I.S.P. (1993) Word families. *International Journal of Lexicography* 6, 4: 253-279.
- Beglar, D. (2010). A Rasch-based validation of the Vocabulary Size Test. *Language Testing*, 27(1), 101-118.
- Beglar, D., & Nation, I. S. P. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9-13.
- Berlin, J. A. (1987). *Rhetoric and reality: writing instruction in American colleges, 1900-1985*. Carbondale: Southern Illinois University Press.
- Breitman, R. (1990). Hitler and Genghis Khan. *Journal of Contemporary History*, 25(2/3), 337-351. Retrieved from <http://www.jstor.org/stable/260736>
- Canagarajah, A. (2006). Toward a Writing Pedagogy of Shuttling between Languages: Learning from Multilingual Writers. *College English*, 68(6), 589-604. doi:10.2307/25472177
- Caplan, N. A., de Oliveira, L. C. (2016). Why we still won't teach the 5-paragraph essay. *TESOL Blog*. Retrieved from <http://blog.tesol.org/why-we-still-wont-teach-the-5-paragraph-essay/>
- Celce-Murcia, M. (2013a). Chapter 1: An overview of language teaching methods and approaches. In M. Celce-Murcia, D.M. Brinton, & M.A. Snow (Eds.), *Teaching*

- English as a Second or Foreign Language* (4th ed.) (pp. 2-14). Boston, MA: Cengage Learning.
- Celce-Murcia, M. (2013b). Chapter 5: Teaching English in the context of world Englishes. In M. Celce-Murcia, D.M. Brinton, & M.A. Snow (Eds.), *Teaching English as a Second or Foreign Language* (4th ed.) (pp. 63-70). Boston, MA: Cengage Learning.
- Chen, J. F., Warden, C. A., & Chang, H.-T. (2005). Motivators that do not motivate: The case of Chinese EFL learners and the influence of culture on motivation. *TESOL Quarterly*, 39(4), 609-633.
- Cobb, T. (2017). Frequency lists for teachers/researchers: GSL 1000 families. Retrieved from http://www.lexutor.ca/freq/lists_download/1000_families.txt
- Collins. (2012). Skills for the TOEFL iBT® Test: Reading and writing. London: HarperCollins Publishers.
- Compulsory Education Law of the People's Republic of China of 1986. (2009). Retrieved from http://en.moe.gov.cn/Resources/Laws_and_Policies/201506/t20150626_191391.html
- Cooper, M. M. (2010). Sustainable writing. In B. Horner, M. Z. Lu, & P. K. Matsuda (Eds.), *Cross-Language Relations in Composition* (pp. 236-243). Carbondale: Southern Illinois University Press.
- Cowan, R., Light, R. L., Mathews, B. E., & Tucker, G. R. (1979). English Teaching in China: A Recent Survey. *TESOL Quarterly*, 13(4), 465-482.
- Coxhead, A., Nation, P., & Sim, D. (2014). Creating and trialing six forms of the Vocabulary Size Test. *TESOLANZ Journal*, 22, 13-26. Retrieved from http://www.tesolanz.org.nz/Site/Publications/TESOLANZ_Journal/
- Coxhead, A., Nation, I. S. P., & Sim, D. (2015). Measuring the vocabulary size of native speakers of English in New Zealand secondary schools. *New Zealand Journal of Educational Studies*, 50(1), 121-135.
- Cumming, A., Kantor, R., Baba, K., Erdosy, U., Eouanzoui, K., & James, M. (2005). Differences in written discourse in independent and integrated prototype tasks for next generation TOEFL. *Assessing Writing*, 10, 5-43.
- D'Aulaire, I., & D'Aulaire, E. P. (1962). *D'Aulaires' book of Greek myths* (1st ed.). Garden City, N.Y.: Doubleday.
- Dan, Y., & Liu, Y.. (2014). *SAT 写作高分范文*. 北京语言大学出版社.

- Doyle, A. C. (1992). *Six great Sherlock Holmes stories* (unabridged. ed.). New York: Dover; London: Constable.
- Eisner, C., & Vicinus, M. (2008). *Originality, imitation, and plagiarism: Teaching writing in the digital age*. Ann Arbor: University of Michigan Press.
- Ellis, R. (2006). Current issues in the teaching of grammar: An SLA perspective. *TESOL Quarterly*, 40(1), 83-107.
- ETS. (2014). *TOEFL iBT® Integrated Writing Rubrics*. Retrieved from https://www.ets.org/s/toefl/pdf/toefl_writing_rubrics.pdf
- Fang, X., & Warschauer, M. (2004). Technology and curricular reform in China: A case study. *TESOL Quarterly*, 38(2), 301-323.
- Flowerdew, J., Levis, J., Zhang, Q., & Watkins, D. (2007). Conceptions of a good tertiary EFL Teacher in China *TESOL Quarterly*, 41(4), 781-790.
- Fox, H. (1994). *Listening to the world: Cultural issues in academic writing*. Urbana, IL: National Council of Teachers of English.
- Fulkerson, R. (2005). Composition at the turn of the Twenty-First Century. *College Composition and Communication*, 56(4), 654-687. Retrieved from <http://www.jstor.org/stable/30037890>
- Fulkerson, R. (1979). Four Philosophies of Composition. *College Composition and Communication*, 30(4), 343-348. doi:10.2307/356707
- Hazenbergh, S., & Hulstijn, J. H. (1996). Defining a minimal receptive second-language vocabulary for non-native university students: An empirical investigation. *Applied Linguistics*, 17(2), 145-163.
- High School Affiliated Renmin University of China Joint Program, The (2010). International section. Retrieved from http://www.rdfz.cn/gjb/gjbjj/jj1/201105/t20110524_22524.html
- Hinkel, E. (2006). Current perspectives on teaching the four skills. *TESOL Quarterly*, 40(1), 109-131.
- Hooper, B. (1975). Women in China: Mao "v." Confucius. *Labour History*, (29), 132-145. doi:10.2307/27508202
- Hsueh, S. L. (2016). The Chinese zodiac, explained [Video file]. Retrieved from http://www.ted.com/talks/shaolan_the_chinese_zodiac_explained
- Huang, J., & Brown, K. (2009). Cultural factors affecting Chinese ESL students' academic learning. *Education*, 129(4), 643-653.

- Institute of International Education. (2016). "International Student Enrollment Trends, 1948/49-2015/16." *Open Doors Report on International Educational Exchange*. Retrieved from <http://www.iie.org/opendoors>
- Jeffery, J. V., Kieffer, M. J., & Matsuda, P. K. (2013). Examining conceptions of writing in TESOL and English Education journals: Toward a more integrated framework for research addressing multilingual classrooms. *Learning and individual differences*, 28, 181-192.
- Jones, F. H., Jones, P., & Jones, J. L. T. (2007). *Tools for teaching: Discipline, instruction, motivation* (2nd ed.). Santa Cruz, CA: F.H. Jones & Associates.
- Koda, K. (1993). Task-induced variability in FL composition: Language-specific perspectives. *Foreign Language Annals*, 26(3), 332-346.
- Laufer, B., & Ravenhorst-Kalovski, G.C. (2010). Lexical threshold revisited: Lexical test coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1), 15-30.
- Li, C.-L. (2011). Chinese High School Graduates' Beliefs About English Learning. *Studies in Literature and Language*, 3(2), 11-18.
- Lu, M. (2006). Living-English Work. *College English*, 68(6), 605-618. doi:10.2307/25472178
- McCall Smith, A. (2002). *The No. 1 Ladies' Detective Agency* (1st Anchor Books ed.). New York: Anchor Books.
- Ma, Q. (2013). Matching vocabulary learning process with learning outcomes in L2 academic writing: An exploratory case study *Linguistics and Education*, 24, 237-246.
- Macrorie, K. (1988). *The I-search paper*. Portsmouth, NH: Boynton/Cook Publishers, Heinemann.
- Matsuda, P. K. (2006). Second-language writing in the Twentieth Century: A situated historical perspective. In P. K. Matsuda, M. Cox, J. Jordan, & C. Ortmeier-Hooper (Eds.), *Second language writing in the composition classroom: A critical sourcebook* (pp. 14-30). New York, NY: Bedford/St. Martins.
- Miller, S. (1991). *Textual carnivals: The politics of composition*. Carbondale: Southern Illinois University Press.
- Milton, J. (2009). *Measuring Second Language Vocabulary Acquisition*. Toronto: Multilingual Matters.

- Ministry of Education of the People's Republic of China (2015a). Education law of the People's Republic of China of 1995. Retrieved from http://en.moe.gov.cn/Resources/Laws_and_Policies/201506/t20150626_191385.html
- Ministry of Education of the People's Republic of China (2015b). Number of schools, educational personnel and full-time teachers by type and level. Retrieved from http://en.moe.gov.cn/Resources/Statistics/edu_stat_2015/2015_en01/201610/t20161012_284511.html
- Ministry of Education of the People's Republic of China (2015c). Number of students of formal education by type and level. Retrieved from http://en.moe.gov.cn/Resources/Statistics/edu_stat_2015/2015_en01/201610/t20161012_284510.html
- Ministry of Education of the People's Republic of China (2015d). Teachers Law of the People's Republic of China. Retrieved from http://en.moe.gov.cn/Resources/Laws_and_Policies/201506/t20150626_191389.html
- Morrill Act of 1862, 7 U.S.C. Ch. 130, § 1, 12 Stat. 503 (1862).
- Murray, D. M. (2004) *Write to Learn* (8th ed.). Boston: Cengage Learning.
- 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. (2009). *4000 Essential English Words*, Books 1-6. Seoul, Korea: Compass Publishing.
- Nation, I. S. P. (2012). *Vocabulary Size Test instructions and description* Victoria University. School of Linguistics and Applied Language Studies. Retrieved from <http://www.victoria.ac.nz/lals/about/staff/publications/paul-nation/Vocabulary-Size-Test-information-and-specifications.pdf>
- Nation, I.S.P. (2013). *What should every EFL teacher know?* Seoul: Compass Publishing.
- Needham, J., & Wang, L. (1954). *Science and civilisation in China*. Cambridge Eng.: University Press.
- New Oriental (2015). Who we are: Overview. Retrieved from <http://www.neworiental.org/english/who/201507/8213540.html>
- Nguyen, L., & Nation, I. S. P. (2011). A bilingual vocabulary size test of English for Vietnamese learners. *RELC Journal*, 42(1), 86-99.
- Plakans, L. (2010). Independent vs. integrated writing tasks: A comparison of task representation. *TESOL Quarterly*, 44(1), 185-194. Retrieved from <http://www.jstor.org/stable/27785076>

- Reagan, T. G. (1996). *Non-Western educational traditions: Alternative approaches to educational thought and practice*. Mahwah, N.J.: L. Erlbaum Associates.
- Reid, J. (2006). 'Eye' learners and 'Ear' learners: Identifying the language needs of international students and U.S. resident writers. In P. K. Matsuda, M. Cox, J. Jordan, & C. Ortmeier-Hooper (Eds.), *Second language writing in the composition classroom: A critical sourcebook* (pp. 76-88). New York, NY: Bedford/St. Martins.
- Riazi, A. M. (2016). Comparing writing performance in TOEFL-iBT and academic assignments: An exploration of textual features. *Assessing Writing*, 28, 15-27.
- Roberts, M. (1979). *Chinese fairy tales and fantasies* (1st ed.). New York: Pantheon Books.
- Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. New York: Palgrave Macmillan.
- Schoonen, R., van Gelderen, A., Stoel, R. D., Hulstijn, J., & de Glopper, K. (2011). Modeling the development of L1 EFL writing proficiency. *Language Learning*, 61(1), 31-79.
- Sheppard, R. (2016). In defense of the 5-paragraph essay. *TESOL Blog*. Retrieved from <http://blog.tesol.org/in-defense-of-the-5-paragraph-essay/>
- Solorzano, H.S., and Schmidt, J.P.L. (2008). *NorthStar 3 Listening and Speaking 3rd Ed.* New York: Pearson Longman.
- Spiegelman, A. (1986). *Maus: A survivor's tale*. New York: Pantheon Books.
- Stewart, J. (2014). Do multiple-choice options inflate estimates of vocabulary size on the VST? *Language Assessment Quarterly*, 11, 271-282.
- Sullivan, P., Zhang, Y., & Zheng, F. (2012). College writing in China and America: A modest and humble conversation, with writing samples. *College Composition and Communication*, 64(2), 306-331.
- Sztabnik, B. (2015). Let's bury the 5-paragraph essay: Long live authentic writing. *Talks with Teachers*. Retrieved from <http://talkswithteachers.com/5paragraphessayvsauthenticwriting/>
- Taggart, C. (2009). *A classical education: The stuff you wish you'd been taught at school*. London: Michael O'Mara.

- Tomba, L. (2015). Chinese Families going global? In Barmé G., Jaivin L., & Goldkorn J. (Eds.), *Shared Destiny* (pp. 230-239). ANU Press. Retrieved from <http://www.jstor.org/stable/j.ctt19893k8.28>
- Why Chinese born in years of the dragon are more successful. (2017, September, 4). *The Economist*. <https://www.economist.com/blogs/graphicdetail/2017/09/dayly-chart-0>
- Wiesel, E. (1960). *Night* (1st American ed.). New York: Hill and Wang.
- Wong, H. K., & Wong, R. T. (2005). *The first days of school: How to be an effective teacher* (3 ed.). Mountain View, CA: Harry K. Wong Publications.
- Xie, W.T. (2016, June 6). Beijing parents protest against “favorable treatment” of minority students in college entrance exams. *Global Times*. <http://www.globaltimes.cn/content/987217.shtml>
- You, X. (2010). *Writing in the devil's tongue: A history of English composition in China*. Carbondale: Southern Illinois University Press.
- Yu, M. H. (2011a). *SAT vocabulary root + associate memory method*. Hangzhou: Zhejiang Education Publishing House.
- Yu, M. H. (2011b). *TOEFL vocabulary root + associate memory method*. Hangzhou: Zhejiang Education Publishing House.
- Zhao, D., Zhu, F., & Liu, J. (2009). Practicing participatory teaching methodology in western China's rural areas: Classroom teaching observations and analysis. *Chinese Education and Society*, 41(5), 58-69.
- Zhao, N. (2009). Metacognitive strategy training and vocabulary learning of Chinese college students. *English Language Teaching*, 2(4), 123-129.
- Zheng, B., Yim, S. and Warschauer, M. 2017. Social Media in the Writing Classroom and Beyond. *The TESOL Encyclopedia of English Language Teaching*. 1–5.
- Zhou, S. (2010). Comparing receptive and productive academic vocabulary knowledge of Chinese EFL Learners. *Asian Social Science*, 6(10), 14-17.
- Zimmerman, C. B. (1997). Historical trends in second language vocabulary instruction. *Second Language Vocabulary Acquisition: A Rationale for Pedagogy*, 5-19.
- Zimmerman, C. B. (2013). Chapter 19: Teaching and learning vocabulary for second language learners. In M. Celce-Murcia, D.M. Brinton, & M.A. Snow (Eds.), *Teaching English as a Second or Foreign Language* (4th ed.) (pp. 288-302). Boston, MA: Cengage Learning.

Appendix A: Class Handouts and Teaching Materials

Full Name

Class Name

Instructor Name

Month 00, YEAR

Compare-and-Contrast Exam Essay

The introductory paragraph starts here. Remember you need a hook to catch the readers' attention; this can either be a good quote or question. Next, you will need to write a clear thesis statement that tells your reader what your essay is about. For example: this template will help you write a good compare-and-contrast essay for your English course.

This paragraph is the first paragraph of the body section of your paper. You must start the paragraph with a clear topic sentence and then follow up with details and examples. This first paragraph compares, shows the similarities between, whatever you are comparing. Finish the paragraph with a transition sentence so you can smoothly move to your next idea.

This is the second paragraph in the body section of your paper. You must start with a clear topic sentence and then add examples with illustrative details. This paragraph must contrast, show the differences between, whatever you are contrasting. End this paragraph with a clear closing statement or transition.

Finally, you must write a conclusion to end your essay well. Start this paragraph with a summary that is similar to, but not the same as, your thesis statement in the introductory paragraph. Last but not least, add a clincher, which releases your reader from the essay in much the same way that you say: "Good bye." You can either end with a quote or a personal comment. Remember, never start and end with the same trick (quote/quote or question/question) try to use two different ways to begin and end your essay (quote/comment or question/quote). Now, go and write your compare-and-contrast essay following this advice.

Figure 6. 1 This is the example Compare-and-Contrast Essay Template. It was given to students to educate them about how to organize their essay writing.

Grade 10 Compare-and-Contrast Essay: Greek Myths

Prepare a compare-and-contrast essay about some aspects of Greek Mythology. The final essay will be handwritten during the final exam.

Here are three questions. You will choose one to answer:

- A. compare and contrast two Greek gods
(e.g. Dionysus and Hermes// Zeus and Cronus)
- B. compare and contrast two Greek goddesses
(e.g. Athena and Artemis// Hera and Demeter)
- C. compare and contrast two Greek heroes
(e.g. Hercules and Jason// Theseus and Perseus)

If you want a harder option, you can try one of these with your teacher's permission:

- D. compare and contrast the Greek gods with the Greek goddesses and explain some similarities and differences between gender roles
(e.g. Zeus and Hera//Ares and Athena)
- E. compare and contrast the Greek gods with Greek heroes and explain some similarities and differences between immortals and mortals
(e.g. Prometheus and Heracles//Aphrodite and Helen)

Exam Preparation: In order to prepare for the final exam, you will write two paragraphs to describe the Greek gods, goddesses, or heroes in Greek mythology, and you will write two paragraphs to narrate an important story about those Greek gods, goddesses, or heroes. These paragraphs will help you think about and write about the important details and examples needed for your final exam essay.

	Choice 1	Choice 2	Choice 3
1. write a paragraph DESCRIBING	a Greek god	a Greek goddess	a Greek hero
2. write a paragraph NARRATING	a story about the above god	a story about the above goddess	a story about the above hero
3. write a paragraph DESCRIBING	another Greek god	another Greek goddess	another Greek hero
4. write a paragraph NARRATING	a story about the above god	a story about the above goddess	a story about the above hero
5. write an essay COMPARING & CONTRASTING	two Greek gods	two Greek goddesses	two Greek heroes

Homework: Think about which gods, goddesses, or heroes you would like to research and write about for this final project. Be prepared to discuss your choice with your classmates and your teacher.

Figure 6. 2 This is an example Greek Myth Compare-and-Contrast Essay Assignment Sheet. It was given to students a month prior to the exam to assist with brainstorming.

<p>word part of speech 翻译</p> <p>Sentence...</p> <p>+ other forms of the word + + picture/drawing +</p>	<p>app n. 应用, 程序</p> <p>I just downloaded a cool new app on my smartphone.</p> <p>application</p>
<p>blog n., v. 博客</p> <p>That movie star has a blog to posts pictures from her concerts and writes about her performances.</p> <p>blogging, blogger</p>	<p>comic n. 漫画</p> <p>I love that comic strip, so I read it each week in the newspaper.</p> <p>comical, comically, comedic</p>
<p>digital adj. 数字的, 数码</p> <p>Now, we all take digital photos on our cellphones instead of photographs with film in a camera.</p>	<p>download v., n. 下载到, 下载</p> <p>I downloaded a new app on my tablet to help me learn to write Chinese characters.</p>
<p>image n. 图片</p> <p>It is hard to know if a digital image is real or not.</p>	<p>puzzle n., v. 难题, 使困惑</p> <p>I love puzzles because they make me think hard.</p> <p>puzzling, puzzlement</p>
<p>tablet (computer) n. 平板电脑</p> <p>She got a new tablet for her 15th birthday.</p>	<p>trainers n., v. 运动鞋, 教练员, 训练</p> <p>Each August, my mother would buy me a new pair of trainers for gym class.</p>

Figure 6. 3 This is an example Vocabulary Word List. Students were encouraged to create their own word lists according to this format, with a semester goal of 500 words.

Appendix B: Essay Exam Writing Prompt

Grade 10 Compare-and-Contrast Essay Exam: Greek Myths

Write a 4-paragraph compare-and-contrast essay about the Greek Myths we have studied in class. The essay **MUST** compare-and-contrast some aspects of Greek myth that we have studied in class.

Here are the five topics. Choose one (1) to answer:

- A. compare and contrast two Greek gods
- B. compare and contrast two Greek goddesses
- C. compare and contrast two Greek heroes
- D. compare and contrast the Greek gods with the Greek goddesses and explain some similarities and differences between gender roles
- E. compare and contrast the Greek gods with Greek heroes and explain some similarities and differences between immortals and mortals

Brainstorm: Complete a mind map or an outline, but you do not need both.

Mind map (draw here):

Outline:

Introduction

with thesis:

Compare

- i. example
- ii. example

Contrast

- i. example
- ii. example

Conclusion

with summary:

Figure 6. 4 This is an example Greek Myth Compare-and-Contrast Essay Exam Question. It was given to students under exam conditions as a final assessment.

Appendix C: TOEFL iBT® Integrated Writing Rubrics

TOEFL iBT® Test

Integrated WRITING Rubrics

SCORE	TASK DESCRIPTION
5	A response at this level successfully selects the important information from the lecture and coherently and accurately presents this information in relation to the relevant information presented in the reading. The response is well organized, and occasional language errors that are present do not result in inaccurate or imprecise presentation of content or connections.
4	A response at this level is generally good in selecting the important information from the lecture and in coherently and accurately presenting this information in relation to the relevant information in the reading, but it may have minor omission, inaccuracy, vagueness, or imprecision of some content from the lecture or in connection to points made in the reading. A response is also scored at this level if it has more frequent or noticeable minor language errors, as long as such usage and grammatical structures do not result in anything more than an occasional lapse of clarity or in the connection of ideas.
3	A response at this level contains some important information from the lecture and conveys some relevant connection to the reading, but it is marked by one or more of the following: <ul style="list-style-type: none"> ■ Although the overall response is definitely oriented to the task, it conveys only vague, global, unclear, or somewhat imprecise connection of the points made in the lecture to points made in the reading. ■ The response may omit one major key point made in the lecture. ■ Some key points made in the lecture or the reading, or connections between the two, may be incomplete, inaccurate, or imprecise. ■ Errors of usage and/or grammar may be more frequent or may result in noticeably vague expressions or obscured meanings in conveying ideas and connections.
2	A response at this level contains some relevant information from the lecture, but is marked by significant language difficulties or by significant omission or inaccuracy of important ideas from the lecture or in the connections between the lecture and the reading; a response at this level is marked by one or more of the following: <ul style="list-style-type: none"> ■ The response significantly misrepresents or completely omits the overall connection between the lecture and the reading. ■ The response significantly omits or significantly misrepresents important points made in the lecture. ■ The response contains language errors or expressions that largely obscure connections or meaning at key junctures or that would likely obscure understanding of key ideas for a reader not already familiar with the reading and the lecture.
1	A response at this level is marked by one or more of the following: <ul style="list-style-type: none"> ■ The response provides little or no meaningful or relevant coherent content from the lecture. ■ The language level of the response is so low that it is difficult to derive meaning.
0	A response at this level merely copies sentences from the reading, rejects the topic or is otherwise not connected to the topic, is written in a foreign language, consists of keystroke characters, or is blank.

Copyright © 2014 by Educational Testing Service. All rights reserved. ETS, the ETS logo, TOEFL and TOEFL iBT are registered trademarks of Educational Testing Service (ETS) in the United States and other countries. 27121



Quality Beyond Measure.

Figure 6. 5 The TOEFL iBT® Integrated Writing Rubrics from ETS (2014) are included with permission.

Appendix D: Chapter 4 Reference Tables

Table 6.1 The interrater reliability for the three Essay Exam scores is satisfactory as the reliability coefficient of Cronbach Alpha is 0.736.

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	Number of Items
0.736	0.726	3

Table 6.2 This is the Test of Normality for Vocabulary Size Test and Essay Exam.

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Stat	df	Sig.	Stat	df	Sig
Vocabulary Size Test	0.159	38	0.016	0.912	38	0.006
Essay Exam	0.165	38	0.011	0.911	38	0.005

a. Lilliefors Significance Correction

VST Test of Normality: Sig of Kolmogorov-Smirnov is $0.016 < 0.05$, indicating that the data tested is not a normal distribution. Essay Exam Tests of Normality: Sig of Kolmogorov-Smirnov is $0.011 < 0.05$, indicating that the data tested is not a normal distribution.

Table 6.3 T-test results: given the results of Levene's test and t-test results, F statistic value of Sig is $0.248 > 0.05$, and accepting the original hypothesis of equal variances, the first line of the t-test Sig value of $0.501 > 0.05$, so there is no significant difference among the Tigers' and Rabbits' Vocabulary Size Test.

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variance assumed	1.380	0.248	-0.680	35	0.501	-.0214	0.314	-0.851	0.424
Equal variance not assumed			-0.693	34.884	0.493	-.0214	0.308	-0.839	0.412

Table 6.4 T-test results: given the results of Levene's test and t-test results, F statistic value of Sig is $0.109 > 0.05$, and accepting the original hypothesis of equal variances, the first line of the t-test Sig value of $0.564 > 0.05$, so there is no significant difference among the Tigers' and Rabbits' Essay Exams.

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variance assumed	2.707	0.109	-582	35	0.564	-0.178	0.305	-0.797	0.442
Equal variance not assumed			-	33.597	0.553	-0.178	0.296	-0.780	0.424

Table 6.5 T-test results: given the results of Levene’s test and t-test results, F statistic value of Sig is $0.455 > 0.05$, and accepting the original hypothesis of equal variances, the first line of the t-test Sig value of $0.278 > 0.05$, so there is no significant difference between males and females on the Vocabulary Size Test.

	Levene’s Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variance assumed	0.570	0.455	1.102	36	0.278	0.374	0.339	-.314	1.062
Equal variance not assumed			1.147	23.740	0.263	0.374	0.326	-.299	1.047

Table 6.6 T-test results: given the results of Levene's test and t-test results, F statistic value of Sig is $0.711 > 0.05$, and accepting the original hypothesis of equal variances, the first line of the t-test Sig value of $0.007 < 0.05$, so there are significant differences between males and females on the Essay Exams.

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variance assumed	0.139	0.711	2.839	36	0.007	0.851	0.230	0.243	1.460
Equal variance not assumed			2.947	23.581	0.007	0.851	0.289	0.255	1.448