

September 2017

Measuring Productive Depth of Vocabulary Knowledge of the Most Frequent Words

Alireza Barouni Ebrahimi
The University of Western Ontario

Supervisor
Faez, Farahnaz
The University of Western Ontario

Joint Supervisor
Webb, Stuart
The University of Western Ontario

Graduate Program in Education

A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

© Alireza Barouni Ebrahimi 2017

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>

Recommended Citation

Barouni Ebrahimi, Alireza, "Measuring Productive Depth of Vocabulary Knowledge of the Most Frequent Words" (2017). *Electronic Thesis and Dissertation Repository*. 4894.
<https://ir.lib.uwo.ca/etd/4894>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact tadam@uwo.ca.

Abstract

Productive depth of vocabulary knowledge (PDVK) is associated with writing and speaking skills (Laufer & Goldstein, 2004). These skills are essential for English for Academic Purposes (EAP) students, who have difficulties with expressing themselves in oral presentations or written assignments (Evans & Green, 2007). As a result, diagnostic measurement of PDVK is of vital importance, especially in regard to the most frequent 1,000 word families because these word families cover 81% of written text and 85% of spoken text (Nation, 2006).

Depth of vocabulary knowledge has been investigated and measured in various studies (see Chen & Truscatt, 2010; Pigada & Schmitt, 2006; Schmitt & Meara, 1997; Schmitt, 1998, 1999; Webb, 2005, 2007a, 2007b, 2007c, 2009a, 2009b) leading to successful multi-dimensional batteries of tests for its measurement. However, no study, to date, has productively measured the depth (and strength) of knowledge of the most frequent words. Nation's (2013) conception of vocabulary knowledge—the proposition that vocabulary knowledge has three main aspects of Form, Meaning, and Use—structured the current study.

Considering that the development of a test battery to measure all aspects of vocabulary knowledge outlined by Nation (2013) was impractical (Ishii & Schmitt, 2009), the current Ph.D. project focused on four aspects of vocabulary knowledge: (a) word parts, (b) associations, (c) collocations, and (d) form and meaning. The study measured 46 Iranian university EAP students' productive vocabulary knowledge of the words at the 1,000 word frequency level. One productive test of word parts, two productive tests of semantic associations (synonym & antonym, and superordination & subordination tests), one productive test of collocation, and four corresponding productive tests of form-meaning connection for the aforementioned tests were developed for the present research.

The results showed that while the participants had a strong performance on form-meaning connection and superordination and subordination, their knowledge of collocations was considerably lower. The results also showed that the participants' performance on synonymy and antonymy, on association as a general term (synonym and antonym, superordination and subordination, and collocation altogether), and on word parts was not as strong as expected and was considerably lower than the maximum possible performance.

Together the findings indicate that while Iranian university students had the productive Meaning knowledge of the words at 1,000 level, they did not seem to have extensive Form knowledge of the same words, and their Use knowledge was limited. This assists in diagnosing areas of weakness and the degree to which instructional emphasis on high frequency words might improve their knowledge.

Keywords: Vocabulary Knowledge; Depth of Vocabulary Knowledge; Productive Vocabulary Knowledge; Productive Vocabulary Assessment; Measuring Productive Vocabulary Knowledge; Most Frequent Words

ACKNOWLEDGEMENTS

I would like to sincerely thank my supervisors, Dr. Farahnaz Faez and Dr. Stuart Webb, for all their keen insights and directions throughout this project. I also wish to express my gratitude to my advisory committee member, Dr. Steve Bird, for all his support and help.

I am very grateful to Western Ph.D. candidates in Education for all their help and support, and I dedicate the success of the study to my deceased mother and father.

To mom and dad, Nosrat and Abdolkarim.

*Alireza
July 2017*

Table of Contents

Abstract.....	i
Acknowledgements.....	ii
Table of Contents.....	v
List of Tables	viii
List of Figures	ix
CHAPTER ONE: INTRODUCTION.....	1
Background and Significance of the Study.....	1
Statement of the Problem and Research Purpose.....	9
Research Questions.....	11
Chapter Summary and Dissertation Organization	12
CHAPTER TWO: LITERATURE REVIEW.....	13
Overview.....	13
Vocabulary and Its Contribution to Language Skills.....	13
Importance of High Frequency Words.....	16
What Words to Measure?.....	18
Vocabulary Knowledge	19
Depth of Vocabulary Knowledge	24
Nation’s Aspects of Vocabulary Knowledge and How to Test Each Aspect	26
Importance of Nation’s Aspects of Vocabulary Knowledge	32
Why Measuring Depth is Necessary?	33
Why Measure Productive Depth of Vocabulary Knowledge?.....	34
Effects of Productive Depth of Vocabulary Knowledge on Students’ Performance	36
Approaches to Measuring Depth of Vocabulary Knowledge	37
Which Aspects of Vocabulary Knowledge to Measure and Why?.....	41
Word Parts	43
Associations	53
Collocations	59
Form and Meaning.....	66
Target Words	71
The Scoring of the Tests	74
Word Parts	74
Associations (Synonym & Antonym, Superordination & Subordination).....	76
Collocation.....	78

Chapter Summary and Conclusion	80
CHAPTER THREE: METHODOLOGY	82
Overview.....	82
Research Design	82
Word Parts	82
Associations	86
Collocations	89
Form and Meaning.....	92
Target Words	99
The Words at the 1,000 Word Frequency Level.....	99
Criteria for Selection of the Target Words.....	100
Target Words for the Word Parts Test.....	102
Target Words for the Synonym and Antonym Test.....	104
Target Words for the Superordination and Subordination Test.....	106
Target Words for the Collocation Test	108
Target Words for Form and Meaning Tests.....	109
Scoring the Tests.....	110
Scoring of the Word Parts Test.....	110
Scoring of the Synonym and Antonym Test.....	112
Scoring of the Superordination and Subordination Test.....	113
Scoring of the Collocation Test	115
Pilot Studies	117
Word Parts Test.....	118
Synonym and Antonym Test	119
Superordination and Subordination Test.....	120
Collocation Test.....	122
Pilot Study with 3 Iranian Graduate Students.....	124
Main Study.....	125
Participants.....	125
Research Instruments	127
The Vocabulary Levels Test, Version 1.....	127
Productive Levels Test.....	128
The Word Parts Test	128
The Synonym and Antonym Test	129
The Superordination and Subordination Test	129

The Collocation Test.....	129
The Form and Meaning Tests	129
Procedure	130
Part 1	130
Part 2	130
Chapter Summary	131
CHAPTER FOUR: RESULTS	132
Overview.....	132
Tests at a Glance	132
Form-Meaning Tests.....	133
Word Parts Test	136
Synonym and Antonym Test	141
Superordination and Subordination Test.....	147
Collocation Test.....	155
Performances on the Tests at a Glance	164
Chapter Summary and Conclusion	166
CHAPTER FIVE: DISCUSSION	168
Overview.....	168
Form-Meaning Tests: To what extent do Iranian EAP students have productive knowledge of form and meaning at the 1,000 word frequency level?.....	168
Word Parts Test: To what extent do Iranian EAP students have productive knowledge of word parts at the 1,000 word frequency level?	170
Association (Synonym & Antonym, and Superordination & Subordination) Tests: To what extent do Iranian EAP students have productive knowledge of associations at the 1,000 word frequency level?.....	173
Collocation Test: To what extent do Iranian EAP students have productive knowledge of collocations at the 1,000 word frequency level?.....	176
Summary of the Main Findings	179
Conclusion	180
Implications of the Study.....	181
Pedagogical Implications of the Study.....	183
Recommendations for Further Research.....	189
Limitations of the Study.....	191
References.....	194
APPENDICES	210
Appendix A: Word Parts Test.....	211

Appendix B: Synonym and Antonym Test.....	215
Appendix C: Superordination and Subordination Test.....	228
Appendix D: Collocation Test.....	231
Appendix E: Form and Meaning Tests.....	235
Appendix F: new-GSL 779 words.....	239
Appendix G: Ethics Approval.....	248

List of Tables

Table 2.1	What is Involved in Knowing a Word	27
Table 2.2	Aspects of Word Knowledge for Testing.....	31
Table 4.1	Information about the Designed Tests of the Study.....	132
Table 4.2	Descriptive Statistics of the Performance on 4 Form-Meaning Tests.....	134
Table 4.3	Descriptive Statistics of the Percentages on 4 Form-Meaning Tests.....	135
Table 4.4	Descriptive Statistics of the Performance on the Word Parts Test	136
Table 4.5	Repeated Measures ANOVA on the Mean Scores of the Word Classes	137
Table 4.6	Results of Bonferroni Post-hoc Comparisons.....	138
Table 4.7	Descriptive Statistics of the Number of Items Provided with Word Classes.....	139
Table 4.8	Descriptive Statistics of the Word Classes Provided for the Test by Each Participant..	140
Table 4.9	Descriptive Statistics of the Synonym and Antonym Test.....	142
Table 4.10	Descriptive Statistics of the Percentage on the Synonym and Antonym Test	142
Table 4.11	Results of Paired Samples Test on the Mean Scores of the Synonyms and Antonyms .	143
Table 4.12	Descriptive Statistics of the Items Provided with Synonyms & Antonyms.....	144
Table 4.13	Descriptive Statistics of the Average Number of Synonyms & Antonyms	146
Table 4.14	Descriptive Statistics of the Performance on the Superordination and Subordination Test	148
Table 4.15	Pearson Correlation between the Raters' Evaluations of the Superordination & Subordination Test.....	149
Table 4.16	Descriptive Statistics of the Superordinates and Subordinates	150
Table 4.17	Descriptive Statistics of the Percentages of the Provided Superordinates and Subordinates	151
Table 4.18	Results of Paired Samples Test on the Mean Percentages of the Superordinates and Subordinates.....	151
Table 4.19	Descriptive Statistics of the Items Provided with Reasonable Superordinates and Subordinates.....	152
Table 4.20	Descriptive Statistics of the Provided Superordinates and Subordinates.....	154
Table 4.21	Descriptive Statistics of the Performance on the Collocation Test.....	156
Table 4.22	Descriptive Statistics of the Percentage on the Collocation Test Evaluated with the Different Criteria.....	158
Table 4.23	Pearson Correlation between the Evaluations of the Collocation Test	160
Table 4.24	Descriptive Statistics of the Items Provided with 0 to 3 Collocates	161
Table 4.25	Descriptive Statistics of the Collocates Provided for the Test by Each Participant.....	163
Table 4.26	Descriptive Statistics of the Percentages of the Performance on the Tests.....	165
Table 5.1	A Range of Activities for Vocabulary Learning	184

List of Figures

Figure 1	The participants' mean percentages of the performance on form-meaning tests	135
Figure 2	The mean number of items provided with 4, 3, 2, 1, and 0 word class(es)	139
Figure 3	The participants' mean scores of the provided word classes for the items of the test.	141
Figure 4	The mean percentages of the performance on the synonym and antonym test.....	143
Figure 5	The mean scores of the items provided for synonym and antonym test.....	145
Figure 6	The participants' mean scores of the provided synonyms and antonyms for the items of the test.....	147
Figure 7	The mean scores of the number of items provided with 3, 2, 1, and 0 reasonable superordinate and subordinates	153
Figure 8	The participants' mean scores of the provided superordinates and subordinates for the items of the test.....	155
Figure 9	The participants' mean scores of the collocations	157
Figure 10	The participants' mean percentages of the collocations	159
Figure 11	The participants' mean scores of the number of items provided with 3, 2, 1, and 0 collocate(s).....	162
Figure 12	The participants' mean scores of the provided collocates for the items of the test.	164
Figure 13	The participants' mean percentages of the performance on 8 tests	166

CHAPTER ONE: INTRODUCTION

Background and Significance of the Study

When learning a foreign or second language, an individual's vocabulary knowledge is one of the most important components of language to develop. In fact, all language components including grammar, vocabulary, and pronunciation are important; however, it is far more difficult to communicate with no vocabulary than with no grammar. Wilkins (1972) states that "without grammar very little can be conveyed, without vocabulary nothing can be conveyed" (p. 111). This, in fact, shows the importance of lexical knowledge among all other language components. Lexical knowledge is related to success in reading, writing, general language proficiency and academic achievement (Laufer, Elder, Hill & Congdon, 2004; Milton, 2013; Milton, Wade, & Hopkins, 2010; Schoonen, 2010; Stæhr, 2008). Meara (1996) and Schmitt (2010) argue that vocabulary knowledge makes a significant contribution to almost all aspects of second language proficiency. In the same vein, Meara and Jones (1988) assert that vocabulary knowledge is heavily implicated in all language skills (listening, speaking, reading, and writing). Yet, the extent to which vocabulary knowledge contributes to language skills can be debated. The following, which discusses the contribution of lexical knowledge to reading, writing, listening, and speaking skills, supports the aforementioned argument.

Regarding the contribution of lexical knowledge to reading skill, Qian (1998; 1999; 2002) has found that in reading comprehension, both breadth (number of known words) and depth (quality of knowing the words) of vocabulary knowledge play important roles, and that

two aspects of depth of vocabulary knowledge – meaning and collocation¹ – are important variables. Qian (1999) explored the relationship between depth and breadth of vocabulary knowledge and reading comprehension in English as a Second Language (ESL) contexts. He found a significant correlation of .82 between the scores in a test of the depth of vocabulary knowledge and a reading comprehension test and .78 between the scores in a test of the breadth of vocabulary knowledge and the reading comprehension test. Qian (2002) also investigated the contribution of vocabulary breadth and some factors of vocabulary depth (synonymy, polysemy, and collocation) to academic reading comprehension. He found a significant correlation of .77 between the scores in a test of the depth of vocabulary knowledge and a TOEFL reading subtest and .74 between the scores in a test of the vocabulary breadth and the TOEFL reading subtest.

Regarding the contribution of lexical knowledge to reading and writing skills, Webb (2009a) investigated the effects of pre-learning vocabulary on reading comprehension and writing. The participants of the study demonstrated that they understood 80% of the sentences on the reading comprehension test for which 15 target words had been previously learned. The results show that L2 students who have studied target words may be able to understand sentences containing them and that vocabulary instruction improves reading comprehension. The participants of the study were also able to correctly use an average of 35% of the target words in sentences of a picture-description test. This result shows that L2 students who receive vocabulary instruction may be able to use some of the target words in sentences and may be able to successfully use recently taught words in their writing.

¹ A particular combination of words that are naturally used together such as “blond hair”

Regarding the contribution of lexical knowledge to listening skill, Stæhr (2009) conducted an empirical study to investigate the role of vocabulary knowledge in listening comprehension. He found that vocabulary breadth produced a significant correlation of .70 and depth of vocabulary knowledge displayed a slightly lower significant correlation of .65 with listening comprehension. Multiple regression analysis showed that breadth and depth of vocabulary knowledge together account for 51% of the variance in the listening scores. This result indicates a strong relationship between a learner's vocabulary knowledge and the quality of the learner's listening comprehension.

Regarding the contribution of lexical knowledge to speaking skill, Koizumi and In'nami (2013) examined the degree to which L2 speaking proficiency can be predicted by the breadth, depth, and speed of L2 vocabulary² among novice to intermediate Japanese learners of English. They found that vocabulary knowledge predicted 84% of speaking proficiency. Vocabulary breadth was found to predict 63% of speaking proficiency when it was first entered into the regression equation, and vocabulary depth predicted speaking similarly to breadth – 60% of speaking proficiency.

So far, the significant contribution of lexical knowledge to all four language skills has been discussed. However, it is a matter of real concern to know which vocabulary plays the most significant role in such a contribution. Among all words, the ones used frequently in a wide range of spoken and written texts are essential in effective comprehension and communication. For example, the most frequent 1,000 word families reoccur so often in spoken and written texts that they have a much greater value for comprehension and use than

² how fast learners can recognize and retrieve knowledge stored in the mental lexicon (Meara, 2005)

the second and the third most frequent ones (Webb & Chang, 2012). The following is the clarification of the significance of this group of words.

The significance of the most frequent 1,000 word families is revealed when their coverage of the British National Corpus (BNC) in comparison to other word families is considered. Nation (2013) illustrates that the most frequent 1,000 word families in the BNC account for 77.96% of the tokens in comparison to 8.10% and 4.36% coverage of the tokens in the second and third most frequent 1,000 word families. Nation (2006) also asserts that the greatest variation in vocabulary coverage occurs in the most frequent 1,000 word families and in the proper nouns which cover 78% to 81% of written text and approximately 85% of spoken text.

In addition to the frequency, the range—the extensiveness of the occurrence of the words in different texts and corpora—should be considered. Nation (2006) found that the most frequent 1,000 word families in the BNC accounted for just over 77.86% of the words in the Lancaster-Oslo/Bergen (LOB) corpus, while the second most frequent 1,000 word families made up 8.23%, and the third most frequent 1,000 word families accounted for 3.70% of the corpus. Moreover, Nation (2012) reports that when the most frequent 1,000 word families are run over the Wellington Written Corpus, they account for 75.22% of the tokens in comparison to 8.91% and 5.22% of the tokens that second and third 1,000 words respectively account for. Dang and Webb (2014) also found that the most frequent 1,000 word families accounted for 87.54% of academic spoken discourse (160 lectures and 39 seminars from four disciplinary sub-corpora of the British Academic Spoken English (BASE) corpus) while 5.4% and 1.76% of the academic spoken words were from the second and third most frequent 1,000 word families. Moreover, the most frequent 1,000 word families in the BNC were also found to make up 85.11% of the words in 88 television

programs (Webb & Rodgers, 2009a) and 86.52% of the words in 318 movies (Webb & Rodgers, 2009b). The second most frequent 1,000 word families constituted 4.42% of lexical items in the television programs and 4.15% of lexical items in the movies, and the third 1,000 word families represented 1.93% and 1.72% of lexical items in the television programs and movies respectively.

Last but not least, the significance of the most frequent 1,000 word families should be considered for the reason that vocabulary is largely acquired in the order of its frequency (Schmitt, Schmitt & Clapham, 2001; Nation, 2006), meaning that the first 1,000 words tend to be learned before the second, the second before the third, and so on. The following illustrates this finding which helps distinguish the most frequent 1,000 word families from other sets of 1,000 word families.

Read (1988) and Laufer et al. (2004) found that the learners' scores dropped on the Vocabulary Levels Test (VLT)³ as they moved from higher to lower frequency levels. Read (1988) administered the VLT at the beginning and end of a three-month English Proficiency Course and noticed a clear pattern of declining scores across frequency levels from highest to lowest. Laufer et al. (2004) found the same result in the process of the development and validation of a test of vocabulary size and strength of knowledge of meaning. Beglar (2010) also asserts that word frequency corpora make it clear that within large quantities of authentic data, the probability of meeting some words is far greater than that of meeting other words so that higher frequency words tend to be better known than lower frequency words. Beglar tested this hypothesis in his study to provide preliminary validity evidence for a 140-item

³ This test is considered as a diagnostic size test which can show in which levels of 2,000, 3,000, Academic, and 10,000 word levels the test takers may have problem

form of the Vocabulary Size Test (VST)⁴. The hypothesis was tested by computing the ensemble means for each of the fourteen 1000-word frequency levels that make up the VST. The results show that the mean ensemble difficulties of the 14 word frequency levels are generally consistent with theoretical expectations. The easiest group of words was the first 1000 word level items, which had a mean item difficulty estimate of 42.59, and the most difficult group was the fourteenth 1,000 word level items, which had a mean item difficulty estimate of 62.98.

Therefore, the aforementioned data shows that there is always a higher proportion of higher frequency words known than lower frequency words. This confirms that the first 1,000 word families are likely to be mastered before the second, and the second before the third 1,000 word families and so on. All in all, such extensive coverage and range of the words, in addition to the consideration of vocabulary learning order, demonstrate the value of the most frequent 1,000 word families. This issue may encourage researchers to investigate whether the most frequent 1,000 word families with the aforementioned significant frequency and range are known by language learners, and if they are, how well they know such words. In this regard, it is important to understand what it means to know a word. The following is a clarification of this issue.

Traditionally, knowledge of a word refers to the meaning and form of that word. Schmitt (2008) asserts that "many teachers and learners consider a word 'learned' if the spoken/written form and meaning are known" (p.333). However, it can be legitimately questioned whether language learners really know a word when they only know the meaning

⁴ This test is considered as a valid and reliable breadth of vocabulary knowledge test which measures the number of words the test takers know.

and form of that word. In other words, it is necessary to investigate what other aspects are considered in the knowledge of words.

Schmitt (2010) states that, in addition to a large vocabulary size (breadth), one must also know a great deal about each individual item in order to function well in a language. This is usually referred to as depth of vocabulary knowledge and is as important as vocabulary size or breadth. In other words, conventionally, while breadth refers to the quantity or number of words one knows, depth refers to the quality of knowledge or how well one knows those words (Anderson & Freebody, 1981; Mukarto, 2003; Nassaji, 2004). For example, a language student may know the meaning of 10,000 words (breadth), but may not be able to use them within a context, to associate them with other words, or to know how formal or informal those words are (depth). Schmitt (2014) states that such a distinction between breadth and depth of vocabulary knowledge is useful especially when one wishes to discuss the nature of the vocabulary with practitioners. That is, the breadth-depth distinction is useful when clarifying for the practitioners the need for rich, sustained instruction and input in order to develop knowledge beyond the simple form-meaning links. The following may facilitate a better understanding of vocabulary knowledge.

Nation's (2001, 2013) conception of vocabulary knowledge includes not just meaning but various other components including three main aspects, each with three subcategories, totaling nine aspects as follows:

- Form: pronunciation, spelling, word parts
- Meaning: form-meaning relationship, concept and referents, associations
- Use: grammatical functions, collocations, constraints on use (register, frequency ...)

Nation's conception of vocabulary knowledge is considered one of the most comprehensive descriptions. Schmitt (2010), in clarifying Nation's (2001, 2013) list of aspects of vocabulary knowledge, states that "he [Nation] presented a revised and expanded version in 2001, which is the best specification of the range of 'word knowledge' aspects to date" (p. 16). Milton (2009) also believes that Nation's (2001) table of what is involved in knowing a word is a "more complete and systematic summary of what the various types of word knowledge are" (p. 14). Read (2004) adds to this that "the most influential account in L2 vocabulary studies is Nation's analysis of what is involved in knowing a word" (p. 217). Webb (2013) also asserts that "Nation's (2001) description of what is involved in knowing a word is the most comprehensive account of depth" (p. 2).

The aforementioned nine aspects are introduced in two productive and receptive modes in Nation's (2001, 2013) table of what is involved in knowing a word. Receptive or passive knowledge is what is required to understand words when they are encountered in listening and reading while productive or active knowledge is what is required to use words in speech and writing (Schmitt, 2014; Webb, 2013). Therefore, productive depth of vocabulary knowledge is usually associated with productive language skills – writing and speaking (Laufer & Goldstein, 2004). These two language skills are essential for English for Academic Purposes (EAP) students who wish to study at English universities and colleges in which they would struggle with the language demands of university work. The following is a clarification of the problems they may have with these two language skills in academia.

Angelova and Riazantseva (1999) state that the growing tendency in universities toward establishing stricter standards of writing proficiency directly affects English non-native students (including EAP students) and especially English non-native graduate

students. They are held to the same stringent writing standards as their English first language counterparts, which places them at a severe disadvantage. Angelova and Riazantseva assert that writing in one's non-native language for academic purposes is an extremely complex cognitive and social task. Hyland (1997) surveyed 1,600 undergraduate students from eight disciplines at five Hong Kong English-medium universities to investigate the importance the students attribute to English, their main difficulties with English, and the value they place on EAP classes. They found that the students' language problems centered on the productive skills of writing and speaking and the acquisition of specialist vocabulary. Evans and Green (2007) also investigated the language problems experienced by 5,000 first year Cantonese-speaking students and found that a significant percentage of the participants experienced difficulty when studying content subjects through the medium of English. They suggested that their participants' problems centered on academic speaking and academic writing. Moreover, Liu and Jackson (2008) and Gan (2012) claim that lack of vocabulary knowledge is regarded as the main obstacle for spoken communication by their participants. Almost all of their participants indicated that the vocabulary problem was the major reason why they could not express themselves clearly and appropriately in English speaking universities.

Statement of the Problem and Research Purpose

Three main issues have been discussed so far: first, the importance of lexical knowledge generally and the knowledge of the most frequent 1,000 word families specifically; second, the lexical knowledge, including breadth and depth of vocabulary knowledge in which productive depth of vocabulary knowledge is related to productive language skills (speaking and writing); and third, the struggle of EAP students with academic speaking and writing. Based on all these discussions, it can be concluded that productive

depth of vocabulary knowledge of the most frequent 1,000 word families has a significant impact on EAP students' speaking and writing for academic success. Therefore, it is important to understand to what extent such students have productive depth of vocabulary knowledge of the most frequent 1,000 word families.

Depth of vocabulary knowledge has been investigated and measured in various studies (see Chen & Truscatt, 2010; Pigada & Schmitt, 2006; Schmitt & Meara, 1997; Schmitt, 1998, 1999; Webb, 2005, 2007a, 2007b, 2007c, 2009a, 2009b) leading to successful multidimensional batteries of tests for its measurement. Schmitt and Meara (1997) measured word associations and grammatical suffix knowledge receptively and productively. Schmitt (1998) measured knowledge of written form, associations, grammatical functions, and meaning while Schmitt (1999) measured meaning, associations, collocations, and grammatical word class. The participants' knowledge was extracted productively in interview sessions in both studies. Webb's (2005) study, which was the most inspiring for his series of studies, measured knowledge of orthography, association, syntax, grammatical functions, and form and meaning receptively and productively. Pigada and Schmitt (2006), in an interview, measured productively one participant's spelling, meaning, and grammatical characteristics knowledge. Chen and Truscatt (2010) measured knowledge of orthography, parts of speech, and associations receptively and productively, and form and meaning receptively.

None of these studies has provided an assessment of productive depth of vocabulary knowledge of the most frequent words because such studies that have measured productive vocabulary knowledge to some extent have focused on what might have been gained through completing an activity or activities, rather than looking at the degree to which words at a certain level are known productively. Moreover, while difficulties such as unfeasibility of

administration and participants' fatigue make it impractical to test all of Nation's (2013) aspects of vocabulary knowledge, many situations, including the current study, may not require such extensive testing (Webb, 2002). As a result, for the purpose of measuring depth of vocabulary knowledge, the current study measures knowledge of form and meaning, word parts, association, and collocations (the reasons for these choices will be discussed in detail in the next chapter). For this purpose, one productive test of word parts, two productive tests of semantic associations (synonym & antonym, and superordination & subordination tests), one productive test of collocation, and four corresponding productive tests of form-meaning connection were developed.

In summary, lexical knowledge generally and productive lexical knowledge of the most frequent words specifically play a significant role in EAP students' academic needs and success. Therefore, the necessity of knowing the extent to which these students have productive depth of vocabulary knowledge of high frequency words is a matter of real concern. Testing all nine aspects of vocabulary knowledge outlined by Nation (2013) seems to be impractical. Consequently, the objective of this dissertation is to measure productively word parts, associations, collocations, and form and meaning of the words at the 1,000 word frequency level to see to what extent these students have such knowledge.

Research Questions

To address the aforementioned problem, and considering the fact that the available participant pool for the current study was Iranian EAP students, the following research questions guided the design of the study:

- To what extent do Iranian EAP students have productive vocabulary knowledge of words at the 1,000 word frequency level?

1. To what extent do Iranian EAP students have productive knowledge of form and meaning at the 1,000 word frequency level?
2. To what extent do Iranian EAP students have productive knowledge of word parts at the 1,000 word frequency level?
3. To what extent do Iranian EAP students have productive knowledge of associations at the 1,000 word frequency level?
4. To what extent do Iranian EAP students have productive knowledge of collocations at the 1,000 word frequency level?

The extent to which Iranian EAP students have productive vocabulary knowledge of words at the 1,000 word frequency level will be reported and discussed based on their obtained scores of the aforementioned form-meaning, word parts, associations, and collocations tests.

Chapter Summary and Dissertation Organization

In this chapter, the background of the study is provided and its significance is explained, the statement of the problem and the research purpose are mentioned, and finally the research questions are raised. The remainder of this dissertation is organized as follows. Chapter 2 provides a comprehensive literature review of the study. Chapter 3 elaborates on the methodology and study design. Chapter 4 illustrates the results of the study. Chapter 5 reports the findings and draws the conclusion.

CHAPTER TWO: LITERATURE REVIEW

Overview

This chapter elaborates on the contribution of vocabulary to language skills, the importance of high-frequency words, and also discusses what high-frequency words to measure. The chapter explains the nature of vocabulary knowledge including depth of vocabulary knowledge, and argues why depth of vocabulary knowledge needs to be measured productively. The chapter also discusses approaches to measure depth of vocabulary knowledge, investigates which aspects to measure, and elaborates on the target words and scoring of the studies which measure this concept. Finally, the chapter ends with the chapter summary and conclusion.

Vocabulary and Its Contribution to Language Skills

Schmitt (2014) believes vocabulary has become mainstream and is now a major topic in language teaching research. Vocabulary is widely acknowledged as one of the key components necessary for second language proficiency (Schmitt, 1999). That is, vocabulary knowledge has a significant explicit role in academic success since it is highly correlated with language skills. This, in fact, is re-echoed clearly in the literature: Meara and Jones (1988) claim that vocabulary knowledge is significantly involved in all practical language skills; Laufer et al. (2004) argue that lexical knowledge has been shown to be related to success in reading and writing skills, general language proficiency, and academic achievement; Milton (2013) asserts that research studies (such as Milton, Wade, & Hopkins, 2010; Schoonen, 2010; Stæhr, 2008) show a moderate to strong relationship between vocabulary measures and the ability to read, write, listen, and also speak in the foreign language; and Schmitt (2010) also believes that typically high correlation is observed

between vocabulary and various measures of language proficiency. The following confirms the aforementioned claims.

Qian (1999) examined the role of both breadth and depth of vocabulary knowledge in measuring the performance of a group of adult ESL learners in carrying out academic reading comprehension tasks. His participants had a minimum vocabulary size of 3,000 word families. He mainly investigated how scores on vocabulary size, depth of vocabulary knowledge, and reading comprehension correlate with one another. He also investigated to what extent depth of vocabulary knowledge adds to the prediction of reading comprehension scores over and above the prediction provided by vocabulary size. The results of his study show significant correlations of .78, .82, and .64 between reading comprehension task and vocabulary size test, depth of vocabulary knowledge test, and morphological knowledge test respectively. The results also show that among three predictor variables (vocabulary size, depth of vocabulary knowledge, morphological knowledge), depth of vocabulary knowledge added a unique portion (11%) of explained variance in reading comprehension in addition to the 60% ($F = 110.15$, $p < .05$) variance accounted for by vocabulary size. Qian also carried out an additional procedure to determine the predictive power of depth of vocabulary knowledge in forecasting reading comprehension scores when depth of vocabulary knowledge was entered into the equation first. The results show that the variance was .68 ($F = 152.17$, $p < .05$) when depth of vocabulary knowledge alone was in the equation. Yet, after vocabulary size was added to the equation, the variance increased to .71 ($F \text{ change} = 8.03$, $p < .25$). In other words, vocabulary size added 3% explained variance in reading comprehension on top of the prediction afforded by depth of vocabulary knowledge.

Qian (2002) in a similar study also investigated the contribution of vocabulary breadth and some factors of vocabulary depth (synonymy, polysemy, and collocation) to

academic reading comprehension. He found a significant correlation of .77 between the scores in a test of the depth of vocabulary knowledge and a TOEFL reading subtest, and .74 between the scores in a test of the vocabulary breadth and the TOEFL reading subtest.

Regarding the contribution of lexical knowledge to reading and writing skills, Webb (2009a) investigated the contribution of pre-learning vocabulary to reading comprehension and writing. The study participants demonstrated that they understood 80% of the sentences on the reading comprehension test for which some target vocabulary had been previously learned. The results show that L2 students who have studied target vocabulary may be able to understand sentences containing them and that vocabulary instruction improves reading comprehension. The study participants also correctly used an average of 35% of the target words in sentences of a picture-description test. This result shows that L2 students who receive vocabulary instruction may be able to use some of the target words in sentences and may be able to successfully use recently taught words in their writing.

Regarding the contribution of lexical knowledge to listening skill, Stæhr (2009) conducted an empirical study to investigate the role of vocabulary knowledge in listening comprehension. The participants of the study were 115 advanced EFL students. The results of the study show significant correlations of .70 and .65 between listening comprehension test and vocabulary size test, and listening comprehension test and depth of vocabulary knowledge test respectively. Also, multiple regression analysis showed that breadth and depth of vocabulary knowledge together accounted for 51% of the variance in the listening scores. Stæhr (2009) believes that each of these two variables is a reliable predictor of the listening comprehension scores, and each can explain a significant portion of the variance in listening comprehension (.49 and .43 respectively).

Regarding the contribution of lexical knowledge to speaking skill, Koizumi and In'nami (2013) conducted two studies to examine the relationship between size, depth, speed, and speaking proficiency. They mainly investigated to what extent L2 speaking proficiency is predicted by L2 vocabulary knowledge (in terms of size, depth, and speed). They constructed a model in which the vocabulary knowledge factor was hypothesized to predict the speaking proficiency factor and adopted this model because of its good fit with their data. The model included size, depth, and speed variables in order to further examine the relationship between vocabulary knowledge and speaking proficiency. They found that vocabulary knowledge predicted 84% of speaking proficiency. Vocabulary breadth was found to predict 63% of speaking proficiency, and vocabulary depth predicted speaking similarly to breadth—60% of speaking proficiency. However, speed predicted speaking less than both size and depth—28% of speaking proficiency.

All in all, the sheer volume of the literature on the role of vocabulary knowledge in all language skills reveals that vocabulary knowledge makes a significant contribution to almost all aspects of L2 proficiency (Meara, 1996).

Importance of High Frequency Words

The significance of vocabulary as one of the key components of language skills is already clear; however, it is a matter of concern to know what vocabulary plays a significant role in language development. It may help to know that words can be broken into 1,000 frequency levels (word families) shown as K1, K2, K3, etc. Nation (2006, 2013) breaks British National Corpus (BNC) vocabulary into twenty 1,000 vocabulary frequency levels, proper nouns, marginal words, compounds, and one off-list group (not in the lists). The British National Corpus (BNC) is a 100 million word collection of samples of written and

spoken language from a wide range of sources, designed to represent a wide cross-section of British English, both spoken and written, from the late twentieth century (BNC official website, 2015).

We may consider in what frequency levels the words could be categorized to have a better understanding of the significance of the words. Generally, Nation (2013) categorizes words as high-frequency, mid-frequency, low-frequency and specialized words (including academic and technical words). The high-frequency words include function and content words which contain around 2,000 word families. Nation (2013) notes that this number has become a matter of debate as Schmitt and Schmitt (2012) argue for having a 3,000 word family high-frequency vocabulary list.

From these 2,000 or 3,000 high-frequency words, the first 1,000 word family level has the most significant role and importance. The most frequent 1,000 word families reoccur so often in spoken and written text that they have a much greater value for comprehension and use than the second most frequent 1,000 word families, and these words have much greater value than the third most frequent 1,000 word families (Webb & Chang, 2012). Nation (2013) illustrates that the most frequent 1,000 word families in the British National Corpus (BNC) account for 77.96% of the tokens in the BNC, the second most frequent 1,000 word families make up 8.10%, and the third most frequent 1,000 word families account for 4.36% of the corpus. The most frequent 1,000 word families in the BNC were also found to make up 85.11% of the words in 88 television programs (Webb & Rodgers, 2009a) and 86.52% of the words in 318 movies (Webb & Rodgers, 2009b). However, Webb and Rodgers' studies show that the knowledge of the second most frequent 1,000 word families includes 4.42% of the television programs and 4.15% of the movies, and the third 1,000 word families represent 1.93% and 1.72% of the television programs and movies respectively.

Such a large coverage of the discourse types demonstrates the value and the importance of the first 1,000 high frequency words in comparison to the second and the third ones.

What Words to Measure?

Brezina and Gablasova (2013) assert that although there are a number of different lists for English frequent lexical items, the most influential and widely used one is the General Service List ⁵(GSL, West, 1953), which has been adopted in pedagogical practice and vocabulary research (e.g. Cobb, 2012; Hirsh & Nation, 1992; McCarthy, 1990; Nation, 2004), and also has served as the non-academic baseline of the Academic Word List⁶ (AWL, Coxhead, 2000, 2011). In other words, West's GSL influences directly the way essential English vocabulary is conceptualized and also lies at the center of the distinction between general and academic vocabulary.

However, West's GSL has been criticized over the years mainly for being out-of-date (Carter, 2012; Nation, 1990; Richards, 1976). GSL was also criticized for the combination of objective and subjective criteria on which the wordlist was based (Gilner & Morales, 2008). In other words, Brezina and Gablasova (2013) argue that GSL's compilation involved a number of principles that brought subjectivity into the final product.

In response to the problems identified with the GSL, Brezina and Gablasova (2013) offered an objective approach to the development of the New General Service List (new-GSL) by means of examining frequent general words across a variety of language corpora. In addition, while the GSL is organized according to the word family principle (a headword, its

⁵ GSL is a list of approximately 2,000 words selected to represent the most frequent words of English, and taken from a corpus of written English.

⁶ AWL is a list of 570 words selected because they appear with great frequency in a broad range of academic texts.

inflected and closely related derivative forms, Nation, 2013), the new-GSL's lexical units are lemmas (a headword, its inflected and reduced forms, Nation, 2013).

Brezina and Gablasova (2013) argue that word family is a useful tool for the research and pedagogy concerned with receptive uses of vocabulary. This principle operates with the underlying assumption that “once the base word or even a derived word is known, the recognition of other members of the family requires little or no extra effort” (Bauer & Nation, 1993, p. 253). However, Brezina and Gablasova (2013) criticize this assumption for two reasons. First, the assumption is difficult to maintain for the semantic distance of the words that can be included under one headword in a word family, e.g. pairs of words such as *to train* and *trainers (shoes)*, *please* and *unpleasantly*, or *part* and *particle*. Second, the ability of using word families successfully depends on users' morphological skills which may not be necessarily at an adequate level.

Arguing the aforementioned criticism, they believe that the new-GSL is a list of the most frequent English vocabulary suitable for both receptive and productive use because in the creation of the new-GSL, the preference was given to lemmas rather than word families. Giving the preference to lemmas enabled them to limit the wordlist to the most frequent words with greater precision in comparison to following the word family principle. Considering all of these benefits, the extraction of the most frequent 1,000 words from the new-GSL seems reasonable.

Vocabulary Knowledge

Vocabulary knowledge has attracted much interest from linguists, psychologists, language learners and language teachers. Miller (1999) asserts that the question of what it means to know a word has fascinated many psychologists. To explain the reason, he adds

that knowing a word usually equals knowing its meaning, and meaning is believed to be one of the most important concepts of understanding the nature and limits of psychology. This important field of study (vocabulary knowledge) has been a matter of concern for quite a long time. However, one question may be raised when such knowledge needs to be investigated: what is vocabulary knowledge?

Vocabulary knowledge is mainly and fundamentally divided into two dimensions, breadth and depth (Anderson & Freebody, 1981; Aviad & Laufer, 2013; Daller, Milton & Treffers-Daller, 2007; Meara & Wolter, 2004; Milton, 2009). Depth of knowledge in contrast to breadth of knowledge was introduced as one dimension of vocabulary knowledge by Anderson and Freebody (1981) as follows:

The first [dimension] may be called ‘breadth’ of knowledge, by which we mean the number of words for which the person knows at least some of the significant aspects of meaning ... [there] is a second dimension of vocabulary, namely the quality or ‘depth’ of understanding. We shall assume that, for most purposes, a person has a sufficiently deep understanding of a word if it conveys to him or her all of the distinctions that would be understood by an ordinary adult under normal circumstances. (p. 92-93)

In other words, breadth refers to the number of words one knows and typically is measured by scores on vocabulary size tests such as the Eurocentres Vocabulary Size Test (Meara & Jones, 1990) and Vocabulary Size Test (Nation & Beglar, 2007). However, in addition to knowing a large number of words (breadth or size) to function well in a language, one must also know efficiently about each individual word in order to use it well (Schmitt,

2010). This is referred to as depth of vocabulary knowledge and is as important as vocabulary size. Shen (2008) asserts that in the most available vocabulary knowledge frameworks, these two dimensions, vocabulary breadth and depth, are at least included. The following illustrates this more.

Aviad and Laufer (2013) assert that vocabulary knowledge can be assessed qualitatively and quantitatively in terms of depth of knowledge, and breadth of knowledge and strength of knowledge of meaning. They quote from Richards (1976) that depth of knowledge refers to the degree of familiarity with the different form and meaning components of a given lexical entry, including its morphological structure, its grammatical or lexical patterns, and its relations with other lexical items, etc. They assert that breadth of knowledge refers to vocabulary size, the number of lexical entries stored in the mental lexicon. In measuring vocabulary size, a word is considered known when the correct meaning is associated with the correct word form. However, form-meaning association can take different forms. It can reflect different parameters according to which strength of knowledge of meaning is assessed (Laufer et al. 2004; Laufer & Goldstein, 2004).

The strength of knowledge of meaning is a combination of four aspects of knowledge of meaning: passive recognition, active recognition, passive recall, and active recall respectively (Laufer & Goldstein, 2004). Laufer and Goldstein (2004) assert that these four degrees of knowledge are mainly based on two dichotomous distinctions: (a) the provision of the form for a given meaning in comparison to the provision of meaning for a given form, (b) the ability to recall the form or meaning in comparison to the ability to recognize the form or meaning. The first distinction discusses that there is a difference in knowledge between people who can provide an L2 word (active knowledge) when meaning is provided, and those who can only provide the meaning when L2 word is presented to them (passive

knowledge). The second distinction is that there is a difference in knowledge between those people who can recall the form or the meaning and those who cannot recall but can recognize the form or meaning of a word in a set of possible options (Laufer & Goldsten, 2004).

All in all, the ability to link form to meaning is important because it increases the possibility of comprehending and using words. However, this ability does not guarantee that the words will be understood and used appropriately, and due to this fact, there is also a need to develop depth of vocabulary knowledge (Webb, 2013).

Milton (2009) believes that depth of knowledge should operate at least relatively independently from other qualities such as breadth. In this regard, Meara and Wolter (2004) assert that there are learners with similar vocabulary sizes (breadth), but different organization in their lexicons such as associations and collocations (depth). This fact turns our attention to learners with lots of words but poor lexicon organization in comparison to those with few words but a high degree of organization in their lexicon. This may explain why learners with the same amount of vocabulary knowledge can sometimes perform so differently from each other in academic examinations and in real communication (Milton, 2009).

The importance of depth in comparison to breadth of vocabulary knowledge has already been discussed; however, it should be argued that vocabulary knowledge is a multifaceted concept and has a variety of dimensions. In other words, lexical knowledge has been defined differently by different researchers. Richards (1976, p. 83) considers eight assumptions of knowing a word:

1. The native speaker of a language continues to expand his vocabulary in adulthood, whereas there is comparatively little development of syntax in adult life.
2. Knowing a word means knowing the degree of probability of encountering that word in speech or print. For many words we also know the sort of words most likely to be found associated with the word.
3. Knowing a word implies knowing the limitations imposed on the use of the word according to variations of function and situation.
4. Knowing a word means knowing the syntactic behavior associated with the word.
5. Knowing a word entails knowledge of the underlying form of a word and the derivations that can be made from it.
6. Knowing a word entails knowledge of the network of associations between that word and other words in the language.
7. Knowing a word means knowing the semantic value of a word.
8. Knowing a word means knowing many of the different meanings associated with a word.

Daller et al. (2007) propose a theoretical three-dimensional space that contrasts breadth and depth against fluency of word knowledge. In their model, fluency is the comfort and speed of accessing known words. The idea is that some learners can use their language knowledge easily, they communicate without hesitation and are highly communicative; however, some others may have difficulty accessing their knowledge, their communication has frequent pauses and hesitation, and they are not that communicative. Milton (2009)

believes that details are not elaborated in this model, but it can be assumed that breadth and depth are aspects of passive word knowledge while fluency is the aspect of productive word knowledge. Meara (1996) and Laufer and Nation (2001) also state that automaticity of access, or fluency, or the speed of performing some kind of operation on a word can be considered an additional component (dimension) of word knowledge.

Depth of Vocabulary Knowledge

Webb (2013) argues that depth of vocabulary knowledge is demonstrated by the extent to which Nation's (2001, 2013) aspects of vocabulary knowledge (18 receptive and productive aspects altogether which are explained in the upcoming section in detail) are present, and this, in fact, shows the extent to which words may or may not be used successfully. Nation's (2001, 2013) description of what is involved in knowing a word is considered the most comprehensive account of depth (Webb, 2013). However, there does not seem to be a general agreement on the definition of depth of vocabulary knowledge (Nation & Webb, 2011): Anderson and Freebody (1981) assert that "a person has a sufficient understanding of a word if it conveys to him or her all of the distinctions that would be understood by an ordinary adult under normal circumstances" (p. 93); Read (1993) considers depth of vocabulary knowledge as "the quality of the learners' vocabulary knowledge" (p. 357); and Wesche and Paribakht (1996) define depth "in terms of kinds of knowledge of specific words and in terms of degrees of such knowledge" (p. 13). In this regard, Henriksen (1999) and Read (2004) suggest that greater clarity is needed in defining depth of vocabulary knowledge and propose dimensions of vocabulary knowledge based on how vocabulary has been measured. Henriksen (1999) proposes three distinct dimensions of vocabulary knowledge as follows:

1. *Partial-precise knowledge*: this dimension refers to the varying degrees to which a word is known. It focuses on the extent that a word's meaning is understood and its form is known.
2. *Depth of knowledge*: this refers to the types of vocabulary knowledge proposed by Richards (1976) and Nation (1990, 2001, 2013). Since there are different aspects included in this dimension, it cannot be considered a single continuum. Instead, it is considered a process by which language learners develop a network between one word and other words in their mind.
3. *Receptive-productive*: the distinction here is between having the receptive knowledge of a word and being able to use it productively in speech and writing. Although it is difficult to define how and at what point words could be used productively, this dimension is often considered a continuum.

In comparison to other definitions of depth, Read (2000) argues that Henriksen's dimensions provide a better conceptualization of quality of vocabulary knowledge and a better recognition of what aspects of construct should be measured in particular research studies. Yet, "the fact remains that vocabulary knowledge is an inherently complex concept" (Read, 2000, p. 93). Read (2004) proposes three distinct lines of development in the application of depth to L2 vocabulary acquisition:

1. *Precision of meaning*: this is the difference between having an inadequate and vague idea of what a word means and having much more elaborated and precise knowledge of its meaning.

2. *Comprehensive word knowledge*: Knowledge of a word which includes not only its semantic features but also its orthographic, phonological, morphological, syntactic, collocational and pragmatic characteristics.
3. *Network knowledge*: The incorporation of the word into a mental lexical network, together with the ability to link it to, and to distinguish it from the related words.

Read (2004) believes that these three approaches overlap to a great extent. He believes that it can be argued that the comprehensive approach includes the other two conceptually, yet he states that it is useful to keep them separated for the sake of analysis since each one has been the basis for various scholars' accounts of what depth means and, also, somewhat different assessment procedures result from adopting one approach rather than the others.

Nation's Aspects of Vocabulary Knowledge and How to Test Each Aspect

Nation (2013) asserts that words are not isolated units of the language, but they fit into many related systems and levels. Because of this, there are many things to know about any particular word and there are many degrees of knowing. Nation (2013) identifies a wide range of aspects of word knowledge and suggests the following table of what is involved in knowing a word. The table divides each item into receptive and productive knowledge. Basically, he believes that receptive vocabulary use is the comprehension of the form of a word while listening or reading and retrieving its meaning. Productive vocabulary use is expressing a meaning through speaking or writing and retrieving and making the suitable spoken or written word form. He also asserts that these two terms apply to a variety of language knowledge and use. However, when they are applied to vocabulary, these two terms cover all of the 18 items mentioned in the following table.

Table 1

What is Involved in Knowing a Word

Form	Spoken	R	1. What does the word sound like?
		P	2. How is the word pronounced?
	Written	R	3. What does the word look like?
		P	4. How is the word written and spelled?
	Word parts	R	5. What parts are recognizable in this word?
		P	6. What word parts are needed to express the meaning?
Meaning	Form and meaning	R	7. What meaning does this word form signal?
		P	8. What form can be used to express this meaning?
	Concepts and referents	R	9. What is included in the concept?
		P	10. What items can the concept refer to?
	Associations	R	11. What other words does this make us think of?
		P	12. What other words could be used instead of this one?
Use	Grammatical functions	R	13. In what patterns does the word occur?
		P	14. In what patterns must we use this word?
	Collocations	R	15. What words or type of words occur with this one?
		P	16. What words or types of words must we use with this one?
	Constraints on use	R	17. Where, when and how often would we expect to meet this word?
		P	18. Where, when, and how often can we use this word?

Note: R = receptive knowledge, P = productive knowledge

Source: Learning vocabulary in another language, Nation (2013, p. 49)

Nation (2013) argues that at the most general level, knowing a word involves form, meaning, and use. He provides the word ‘*underdeveloped*’ as an example and asserts that knowing this word from the **receptive knowledge** point of view involves:

1. “being able to recognize the word form when it is heard” (p.48), that is, the way the word sounds like. (**What does the word sound like?**)

2. “being familiar with its written form so that it is recognized when it is met in reading” (p. 48), that is, the way the word looks like. (**What does the word look like?**)
3. “recognizing that it is made up of the parts *under-*, *-develop-*, and *-ed* and being able to relate these parts to its meaning” (p. 50), that is, recognition of the word parts (suffix, stem, prefix) which make the word. (**What parts are recognizable in this word?**)
4. “knowing that *underdeveloped* signals a particular meaning” (p. 50), that is, the core common meaning of the word or the most frequent homograph. (**What meaning does this word form signal?**)
5. “knowing what the word means in the particular context in which it has just occurred” (p. 50), that is, recognition of different meanings the word might have. (**What is included in the concept?**)
6. “knowing the concept behind the word which will allow understanding in a variety of contexts” (p. 50), that is, recognition of the components included in the concept which is represented by a word. (**What is included in the concept?**). Nation believes both 5 and 6 refer to receptive Concepts and Referents (personal communication, January 30, 2014).
7. “knowing that there are related words like *overdeveloped*, *backward* and *challenged*” (p. 50), that is, recognition of other words this word makes us think of. (**What other words does this make us think of?**)
8. “being able to recognize that *underdeveloped* has been used correctly in the sentence in which it occurs” (p. 50), that is, recognition of the grammatical patterns in which the word occurs. (**In what patterns does the word occur?**)

9. “being able to recognize that words such as *territories* and *areas* are typical collocations” (p. 50), that is, recognition of the words which are used with this word.
(What words or type of words occur with this one?)
10. “knowing that *underdeveloped* is not an uncommon word and is not a pejorative word” (p. 50), that is, where, when and how often this word is expected to be met.
(Where, when and how often would we expect to meet this word?)

On the other hand, knowing the word ‘*underdeveloped*’ from the point of view of **productive knowledge** is asserted by Nation (2013) as follows:

1. “being able to say it with correct pronunciation including stress” (p. 50), that is, the way the word is pronounced. **(How is the word pronounced?)**
2. “being able to write it with correct spelling” (p. 50), that is, the way the word is spelled and written. **(How is the word written and spelled?)**
3. “being able to construct it using the right word parts in their appropriate forms” (p. 50), that is, knowing and using necessary word parts (suffix & prefix) to express the desired meaning.
(What word parts are needed to express the meaning?)
4. “being able to produce the word to express the meaning *underdeveloped*” (p. 50), that is, knowing and using the appropriate word to express the desired meaning in the second language. **(What form can be used to express this meaning?)**
5. “being able to produce the word in different contexts to express the range of meanings of *underdeveloped*” (p. 50), that is, knowing and using the different meanings the word might have. **(What items can the concept refer to?)**

6. “being able to produce synonyms and opposites for *underdeveloped*” (p. 50), that is, production of other words this word makes us think of. (**What other words could be used instead of this one?**)
7. “being able to use the word correctly in an original sentence” (p. 50), that is, production of the grammatical patterns in which the word occurs. (**In what patterns must we use this word?**)
8. “being able to produce words that commonly occur with it” (p. 50), that is, production of the words that must be used with this word. (**What words or types of words must we use with this one?**)
9. “being able to decide to use or not use the word to suit the degree of formality of the situation” (p. 50), that is, where, when and how often the word is used. (**Where, when, and how often can we use this word?**)

Moreover, what Nation (2013) has considered aspects of word knowledge for testing are considerable in testing depth of vocabulary knowledge. He has brought up the following table which is an adapted version of the aforementioned table of what is involved in knowing a word.

Table 2.

Aspects of Word Knowledge for Testing

Form	Spoken	R	Can the learner recognize the spoken form of the word?
		P	Can the learner pronounce the word correctly?
	Written	R	Can the learner recognize the written form of the word?
		P	Can the learner spell and write the word?
	Word parts	R	Can the learner recognize known parts in the word?
		P	Can the learner produce appropriate inflected and derived forms of the word?
Meaning	Form and meaning	R	Can the learner recall the appropriate meaning for this word form?
		P	Can the learner produce the appropriate word form to express this meaning?
	Concepts and referents	R	Can the learner understand a range of uses of the word and its central concept?
		P	Can the learner use the word to refer to a range of items?
	Associations	R	Can the learner produce common associations for this word?
		P	Can the learner recall this word when presented with related ideas?
Use	Grammatical functions	R	Can the learner recognize correct uses of the word in context?
		P	Can the learner use this word in the correct grammatical patterns?
	Collocations	R	Can the learner recognize appropriate collocations?
		P	Can the learner produce the word with appropriate collocations?
	Constraints on use	R	Can the learner tell if the word is common, formal, infrequent, etc?
		P	Can the learner use the word at appropriate times?

Note: Source: *Learning vocabulary in another language*, Nation (2013, p. 538)

Table 2 clearly shows how to consider different aspects rather than just meaning of the word when learners' vocabulary knowledge is evaluated. Nation (2013) also argues how each of these aspects could be tested. The following test item types are based on Table 2.

- Spoken form
 1. Word or sentence dictation/ hear the word and choose the L1 translation
 2. Reading aloud/ cued oral recall
- Written form
 3. Say these written words/say these regularly spelled nonsense words.
 4. Word or sentence dictation

- Word parts
 5. Break the word into parts/choose or provide the meanings of the parts.
 6. Provide an affixed form of a known word.
- Form and meaning
 7. Translate these words into L1/choose the right picture.
 8. Translate these words into L2.
- Concept and referents
 9. Translate the underlined words into L1. ‘It was a hard frost.’
 10. Choose the words to translate this L1 word
- Associations
 11. Choose the words that you associate with this word
 12. Add to this list of associated words
- Grammatical functions
 13. Is this sentence correct?
 14. Use this word in a sentence
- Collocations
 15. Is this sentence correct?
 16. Produce collocations to go with this word
- Constraints
 17. Which of these words represent UK use?
 18. What is the formal word for X?

(pp. 551-552)

Importance of Nation’s Aspects of Vocabulary Knowledge

Nation’s (2001, 2013) aspects of what is involved in knowing a word seem the most comprehensive, practical and noteworthy. His conception of vocabulary knowledge includes not just meaning but various other components as well. In this regard, Read (2004) asserts that “several authors have outlined the scope of the area (Cronbach, 1942; Richards, 1976; Laufer, 1997; Nagy & Scott, 2000) but possibly the most influential account in L2 vocabulary studies is Nation’s analysis of what is involved in knowing a word” (p.217). In the same vein, Milton (2009) believes that Nation’s (2001) table of what is involved in knowing a word is a “more complete and systematic summary of what the various types of word knowledge are” (p. 14). Schmitt (2010) in clarifying Nation’s (2001) list of aspects of

word knowledge claims that “he [Nation] presented a revised and expanded version in 2001, which is the best specification of the range of ‘word knowledge’ aspects to date” (p. 16).

Webb (2013) also asserts that “Nation’s (2001) description of what is involved in knowing a word is the most comprehensive account of depth” (p. 2).

The value of Nation’s framework and its role in mastering second language words seems undeniable. Therefore, the current Ph.D. project considers Nation’s conception of vocabulary knowledge the theoretical framework of the study, and Nation’s proposed three main aspects of Form, Meaning, and Use inform the design of the current study.

Nation and Webb (2011) argue that in order to master a word, a person must gain knowledge of all nine aspects (receptively and productively). To assess this, researchers need to measure multiple aspects of knowledge and determine how well each of those aspects is known. “Measuring multiple aspects of knowledge is the most effective way to assess vocabulary depth, with the more aspects that are measured providing a greater indication of how well those words are known” (Nation & Webb, 2011, p. 227).

Why Measuring Depth is Necessary?

Breadth and depth of vocabulary knowledge have not received the same amount of attention despite the fact that their importance in second language acquisition is demonstrated by a considerable number of scholars (Jaen, 2007). Milton (2009) believes that some areas of vocabulary knowledge including vocabulary depth are less well researched and understood. In other words, since it has been easier to test lexical size than depth, measures of vocabulary size are further developed in comparison to depth (Read, 2000). However, as previously discussed (vocabulary and its contribution to language skills), recent studies show both depth and breadth play an important role in language skills. Jaen (2007) asserts that “more research

on the assessment of lexical depth is necessary and even urgent” (p. 128). In this regard, Read (2000) claims whatever the benefits of vocabulary size tests, they can only give a superficial understanding of how well any word is known.

Milton (2013) argues that on the one hand, it can be assumed that breadth and depth will be closely related so that scores on tests of breadth can be used to validate newly created tests of depth. On the other hand, it is still common to talk about vocabulary knowledge in terms of breadth and depth as two separate and contrasting dimensions, which should not be closely connected. Milton believes that it seems due to the fact that the term is ill-defined, there is an absence of well-established and standardized tests in this field.

Nation and Webb (2011) argue that measuring depth of vocabulary knowledge informs us of what learners know or do not know about words, and this helps to diagnose areas of weakness in vocabulary development which deserve more attention. For example, when learners combine words unusually, it often is a signal of limitation in their knowledge of collocation, while wrongly produced derivative forms of a word may indicate their lack of knowledge of word parts. Thus, Webb (2013) believes that measuring depth of vocabulary knowledge should be included in learners’ developing vocabulary knowledge because it demonstrates the importance of learning a variety of aspects as well as form and meaning, and also provides teachers with a better understanding of the learners’ knowledge.

Why Measure Productive Depth of Vocabulary Knowledge?

Schmitt (2010) asserts that productive vocabulary knowledge is more difficult and develops at higher levels of language learning than receptive knowledge. Nation (1990) states that the ability to use a word requires extended knowledge beyond what one needs just to understand it; as a result, production involves a higher level of knowledge than reception

does. Laufer et al. (2004) state that a learner's passive vocabulary is always larger than his or her active vocabulary. This indicates that many words are first acquired passively, and that active knowledge is a more advanced type of knowledge. For this reason, in Webb's studies (2005, 2009a, 2009b), the participants' scores on the receptive Vocabulary Levels Tests (Schmitt, Schmitt & Clapham, 2001) are higher than their scores on the productive Levels Test (Laufer & Nation, 1999). In the same vein, Webb (2005) believes that productive learning is superior to receptive learning not only in developing productive knowledge but also in producing larger gains in receptive knowledge. In other words, looking at individual aspects of vocabulary knowledge, productive knowledge covers receptive knowledge to a great extent, which means one cannot have productive knowledge of an aspect without receptive knowledge of that aspect (Webb, 2013).

However, productive vocabulary knowledge, contrary to its importance, has not received enough attention to date. Pearson, Hiebert, and Kamil (2007) claim that vocabulary assessment has almost always emphasized the receptive dimension of vocabulary, while the productive aspect of vocabulary is rarely examined. Schmitt (2010) asserts that examining the aspects of vocabulary knowledge productively can be one of the best ways to measure depth of vocabulary knowledge. In fact, face to face interviews, in which the interviewees are to provide the interviewer with productive performance on some aspects of vocabulary knowledge, although labour intensive, time-consuming and inconsistent from interview to interview, are considered the best ways of vocabulary knowledge measurement, and even are a method to validate newly designed vocabulary tests (Read, 2000; Schmitt, 2010). Therefore, it seems a productive procedure can be a valuable method to reveal knowledge of a word, and for this reason, in the current study, vocabulary knowledge is measured productively rather than receptively.

Effects of Productive Depth of Vocabulary Knowledge on Students' Performance

In contrast to the receptive or passive knowledge which is required to understand words when they are encountered in listening and reading, productive or active knowledge is required to use words in speech and writing (Schmitt, 2014; Webb, 2013). In the same vein, productive depth of vocabulary knowledge is usually associated with productive language skills – writing and speaking (Laufer & Goldstein, 2004). These two language skills are essential for English as a Second or Foreign Language students studying or wishing to study at English universities and colleges. Such students usually study English for academic purposes and are labeled “English for Academic Purposes (EAP)” students. The following is clarification of the problems they may have in academia.

Angelova and Riazantseva (1999) state that the growing tendency in universities toward establishing stricter standards of writing proficiency directly affects English non-native university students. Such students are held to the same writing standards as their English first language counterparts, which places them at a severe disadvantage. Hyland (1997) surveyed 1,619 Chinese first-year undergraduate students from eight disciplines at five Hong Kong English-medium universities and found that the students' language problems centered on the productive skills of writing and speaking and the acquisition of specialist vocabulary. A self-reporting questionnaire was used to investigate the students' perceptions of the role of English in academic success. The participants were asked to rank from 2 (the most difficult) to 5 (the least difficult) the difficulty level of six areas of their language proficiency (writing, speaking, specialist words, listening, reading, and assignments). The results suggest that the majority of these participants believe that productive skills are the

greatest problems. They ranked writing, speaking and specialist words between 3 and 3.5 while listening, reading and assignments were ranked between 4 and 5.

Evans and Green (2007) also investigated the language problems experienced by 5,000 first-year Cantonese-speaking undergraduate students at a large English-medium university. Their study, in fact, revisits the question posed by Hyland (1997) – is EAP necessary? The baseline data for the study were derived from surveys and interviews with the students and departmental program leaders. The findings of the study indicate that a significant percentage of the participants experienced difficulty when studying content subjects through the medium of English. The study shows that the participants' problems center on academic speaking and academic writing. The study also illuminates that the participants' receptive and productive vocabulary is generally inadequate. In this regard, Liu and Jackson (2008) and Gan (2012) claim that lack of vocabulary knowledge is regarded as the main obstacle for spoken communication by their participants. Almost all of their participants argued that the vocabulary problem was the major reason why they could not express themselves clearly and appropriately in English speaking universities.

Approaches to Measuring Depth of Vocabulary Knowledge

There are two main approaches to measuring depth of vocabulary knowledge, described by Schmitt (2010) as *developmental* and *dimensions (components)* approaches. The former uses scales to chart developing mastery of a lexical item (e.g. 0 = *no knowledge* to 5 = *full mastery*). The best-known test based on this approach is the Vocabulary Knowledge Scale (VKS) developed by Paribakht and Wesche (1993, 1996). Paribakht and Wesche developed the VKS in the context of research on the vocabulary development of ESL learners in a university setting in 1993. The VKS was designed to provide a means of

demonstrating certain changes in the receptive and initial productive knowledge of some target words. The changes were to result from instructional interventions such as vocabulary exercises or activities such as reading. Also, the VKS was supposed to show comparative gains resulting from different treatments. The VKS uses a written scale which combines self-report and performance items to elicit self-perceived and demonstrated knowledge of specific words. The scale ranges from *complete unfamiliarity* (score 1) to *the ability to use it with grammatical and semantic accuracy in a sentence* (score 5). The VKS uses the following scale to measure depth:

1. I don't remember having seen this word before.
2. I have seen this word before, but I don't know what it means.
3. I have seen this word before, and I think it means _____.
4. I know this word. It means _____.
5. I can use this word in the sentence: _____.

However, attempts, such as Wesche's and Paribakht's (1996) VKS, are not without their difficulties and in practice, as Wolter (2005) and Milton (2009) point out, they function as breadth rather than depth measures. Wolter (2005) argues that the test seems to be insensitive to many aspects of depth of knowledge. For example, it cannot measure different meanings of a word. It also seems that the full range of the scale (which scores from 1 to 5) is not used and scores gather at either ends which means scores of 3 and 4 are relatively rare. In addition, it is not very hard to achieve a score of 5 by writing a sentence containing the target word. It is possible to use a word correctly in a sentence without knowing its meaning.

Therefore, the VKS does not seem to function as a scale, yet it could be a binary test of *I know this word* and *I do not know this word* (Milton, 2009). Milton's own experience of

using the test to measure his own vocabulary knowledge showed him the limited use of the middle scales. In other words, he either did not recognize a word or he did and provided that with a translation and use in a sentence.

The problem of the developmental approach is the operationalization of the developmental process into a practical scale although, as Schmitt (2010) believes, it is undeniable that vocabulary is learned incrementally. In fact, Schmitt (2010) asserts that we have little idea about how vocabulary development advances. Therefore, creating a valid scale is rather hypothetical at the moment. In other words, for a scale to exist, there must be rational beginning, in between and ending points (stages). Although having absolutely no knowledge of a lexical item seems a straightforward beginning, the stages in between and ending remain problematic. For in between stages there is currently no principled way of knowing how many stages an acquisition scale should contain (Schmitt, 2010). In the case of ending stage especially, it must be something like ‘full knowledge of an item’ which is unjustifiable. There is no test that can verify if a word can be used accurately, appropriately, and fluently in every possible context. Therefore, any beginning and ending stage in addition to the stages in between will necessarily be approximations. In addition, while setting the number of steps in a scale may not seem to be a challenging problem, the equal intervals between scale steps can be a big problem (Schmitt, 2010).

The dimensions (components) approach identifies some aspects of vocabulary knowledge and measures test takers’ mastery of those aspects (Schmitt, 2010). The well-known test format based on the dimensions approach is the Word Associates Format/Test (WAF/WAT), developed by Read (1993, 1998). Following a suggestion from Meara, Read (1993) decided to develop a word association task by which the learners select the associated responses. The main idea was to provide the target word with a group of words (eight words

as in Read, 1998), some of which were related in meaning to the target word. The test taker would be required to select the related words (associates) as in the following example:

sudden

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

The words on the left side are adjectives, which are either synonyms of the target word or have a paradigmatic relationship with it (*quick* and *surprising* for *sudden*). On the right side, there are four nouns, among which a few are associations and can collocate with the target word and thus have a syntagmatic relationship with it (*change* and *noise* for *sudden*).

Read (2004) reports that some studies found a distinct role for depth of vocabulary knowledge measurement, and used the WAF as their test of vocabulary depth. However, this type of test represents just one way of operationalizing the concept of depth of knowledge. As a result, “a broader range of measures is needed before we can be more confident about the extent to which depth in some sense makes a contribution to the assessment of the lexical knowledge of L2 learners” (Read, 2004, p. 223). In addition, as Webb (2013) asserts, the WAT measures knowledge of three aspects – form and meaning, concept and referents, and collocation. However, it provides an overall score and does not provide a separate measure of each aspect. Therefore, it is not clear to what extent each aspect is known. Webb (2013) believes that “separate tests focusing on individual aspects may provide a more accurate measurement of depth” (p.4). This is what the dimensional approach can provide.

The dimensions approach breaks down the different aspects of word knowledge and then isolates and measures multiple aspects. The advantage of this approach is its comprehensiveness. While measuring several types of word knowledge can be time-

consuming and limits the number of lexical items in a study, it can produce a very rich description of vocabulary knowledge which is well worth the effort (Schmitt, 2010). Nation and Webb (2011) believe measuring multiple aspects of knowledge is the most effective way to assess vocabulary depth since measuring more aspects provides greater indication of how well the target words are known.

All in all, it seems very useful to develop a series of instruments to assess different aspects of vocabulary knowledge (depth of knowledge). Well-established instruments inform language teachers and learners of where greater attention is necessary. In addition, such instruments can also help examine vocabulary learning tasks and also help evaluate the degree to which a variety of tasks contribute to depth of vocabulary knowledge. Such tests not only show which tasks are more effective, but they also help show their strengths and weaknesses. This means that they show what each task contributes to vocabulary knowledge, and also which tasks may complement each other to improve vocabulary learning (Nation & Webb, 2011).

For all these reasons, it seems the dimensions approach and Nation's (2013) framework of vocabulary knowledge should be investigated as an avenue of research. The current dissertation mainly focuses on taking this approach into service; however, not all nine aspects of depth discussed by Nation (2013) will be measured for the purpose of the study. The following explains which aspects will be measured and why.

Which Aspects of Vocabulary Knowledge to Measure and Why?

The focus of the current dissertation will be on four (out of nine) aspects of vocabulary knowledge outlined by Nation (2013): (a) word parts, (b) associations, (c) collocations, and (d) form and meaning. Spoken form aspect is eliminated because

considering the variety of English accents, the evaluation of the appropriateness of pronouncing a word is very challenging. Written form and grammatical functions aspects are eliminated because a pilot study with 15 EAP students showed a ceiling effect on these two aspects. In the pilot study, 30 words were read by an English as a first language speaker to participants, and they were asked to write them down. Also, they were asked to use those 30 words in 30 sentences for the purpose of demonstrating grammatical knowledge. The results showed ceiling effects on both of these two aspects.

Concepts and referents aspect is eliminated because testing all possible meanings of the target words makes the form and meaning test so complicated that its administration would be problematic. In a pilot study with the aforementioned 15 participants, it was found that it would be challenging to instruct the participants to provide various meanings of a target word, and even when they were instructed properly, they mainly provided the core meaning. Nation (2013) asserts that the core meanings tend to be known and provided by language learners. Therefore, the current study just measures the core form and meaning of target words not their concepts and referents.

Constraints on use is eliminated because the many of the most frequent 1,000 words do not seem to have different constraints on use. Finally, developing a test battery to measure all aspects of vocabulary knowledge outlined by Nation (2013) is impractical (Ishii & Schmitt, 2009) due to the unfeasibility of administration and participants' fatigue. The following section reviews the literature on the four targeted aspects of vocabulary knowledge that will be measured in the present research: word parts, associations, collocations, and form and meaning.

Word Parts

The ability to use the appropriate word classes in a context is essential for developing grammatically appropriate language. For example, a language learner needs to produce *measurable* when an adjective is required while *measurement* may be required in a different context as a noun. Therefore, if language learners can use one word class appropriately, it is a matter of concern to assume that they know the other word classes of the same word as well. Generally, affix⁷ knowledge is an important aspect of vocabulary knowledge because it contributes to determining how well a language learner reads new words (Nagy et al. 1993) and can expand their vocabulary knowledge (Mochizuki & Aizawa, 2000).

Nation (2013) argues that knowing a word can involve knowing that it is made up of affixes and a stem that can occur in other words. He asserts that most of English content words can change their forms by adding prefixes and suffixes. These affixes are mainly divided into two types – inflectional and derivational. The English inflectional affixes are all suffixes including plural *-s*, *-ed*, *-ing*, third person singular *-s*, possessive *-s*, comparative *-er*, and superlative *-est*. Inflections, unlike most derivational suffixes, do not change the part of speech of the words and are added after a derivational suffix if the word has one.

Nation (2013) states that derivational affixes in English include prefixes and suffixes. He notes that most of the derivational suffixes and a few prefixes change the part of speech of the words, such as *happy* and *happiness*, or *able* and *enable*. He also notes that some of the affixes (especially prefixes) also change the meaning of the word, such as *judge* and *prejudge*, *happy* and *unhappy*, or *care* and *careless*. Also, he asserts that there are two kinds

⁷ Affix: a grammatical element that is combined with a word, stem, or phrase to produce derived and inflected forms.

of word stems: free forms and bound forms. Free forms can stand as a word such as *help* and *care*. Bound forms cannot stand as a word without an affix such as *-clude* for *exclude* and *include*.

Word formation type of vocabulary knowledge is strongly associated with writing ability. To investigate L2 learners' derivative skills (their ability to recognize and apply different affixes), in a part of her study, Nyssonen (2008) provided 327 Finnish students with a gap-fill task of 18 sentences. The sentences were provided with clues to the answers in Finnish. The gaps were to be filled in with the derivative forms of the word in Finnish. In addition, the students completed 4 structured writing tasks which were assessed by 4 raters. Performance on the affix test strongly correlated with writing ability at the level of .73 which indicated that affix knowledge was strongly associated with writing ability.

Knowledge of word parts in general, and the system of affixes in particular, is an essential part of overall word knowledge because morphological⁸ information is a valuable asset in expanding both receptive and productive vocabulary (Mantyla & Huhta, 2014). Word parts play an important role in vocabulary acquisition, and recognizing word parts and knowing what parts are needed to express a given meaning contribute considerably to the essence of knowing a word (Nation, 2013; Thornbury, 2002).

English has over 60 affixes (Jackson & Ze Amvela, 2000) so that this language is very rich with derivative forms. A study based on the American Heritage Corpus (Carroll, Davies, & Richman, 1971) shows that 21.9% (approximately one fifth) of the different word types in a written text are inflected, and 12.8% (approximately one eighth) have a derivational

⁸ Morphology: the patterns of word formation in a particular language, including inflection, derivation, and composition.

affix (Nagy & Anderson, 1984). Derivation is reported to be the most fruitful and frequent of all word formation methods in English (Yule, 2006). For each base word form, there are 1.5 to 4 derivations (Nation, 2013), in addition to the small number of inflections (Carstairs-McCarthy, 2002). Therefore, mastering affixes is considerably beneficial for language learners because it allows them to expand their vocabulary through morphological associations.

For the purpose of measuring word parts productively, the first question which needs to be addressed is the matter of measuring inflections, derivations, or both. Nation's (2013) framework for productive mode of word parts aspect indicates "what word parts are needed to express the meaning?" (p.49), and Nation (2013) interprets this as "being able to construct [the target word] using the right word parts in their appropriate forms" (p. 50). To test this aspect, Nation (2013) asks whether the learner can "produce the appropriate inflected and derived forms of the word" (p.538), and for the purpose of testing this aspect specifically, he asks whether the learner can "provide an affixed form of a known word" (p. 551). Therefore, Nation targets both inflections and derivations for this purpose, although the way he suggests testing this aspect does not seem to be efficient enough because an affixed form does not necessarily show both the inflected and derived forms of a target word. As a result, literature should be investigated to find how word parts have been measured productively by different researchers.

Generally, this area has received relatively little attention and there have been few attempts at developing instruments to measure word parts knowledge (Mantyla & Huhta, 2014). There are a limited number of research studies which have investigated word parts (e.g. Ishii & Schmitt, 2009; Mochizuki & Aizowa, 2000; Nagy et al. 1991, 1993; Schmitt & Meara, 1997; Schmitt, 1998, Schmitt, 1999, Schmitt & Zimmerman, 2002). The studies

which measure word parts productively are even fewer in number (e.g. Ishii & Schmitt, 2009; Schmitt & Meara, 1997; Schmitt, 1998, Schmitt, 1999, Schmitt & Zimmerman, 2002).

Schmitt and Meara (1997) designed their study to examine receptively and productively two individual vocabulary components, affix knowledge and word association knowledge, to investigate how they relate to each other, to overall vocabulary size, and to general language proficiency. For each target word (prompt verb) in the production task of the study, three classes of Japanese students with 36, 31 and 28 students respectively were asked to provide all allowable suffixes. In other words, a line space was provided in front of each prompt verb (target word) in the instrument of the study, and the participants were asked to write all allowable suffixes of the target prompt verbs. The study instrument was administered to the participants near the beginning of their school year (T1), and near the end of the school year (T2). This way of measuring word parts productively seemed appropriate for the purpose of the current study, as it could measure both inflections and derivations. Schmitt and Meara (1997) found that with the exception of the inflectional suffixes and “*ment*”, the participants did not perform very well on the suffixes and thus showed low proficiency in terms of suffix knowledge. Out of all possible native-like suffixes (the number of native-like suffixes for each item varied from 3 to 7 depending on the target word), 42% of them at T1 and 47% of them at T2 were provided productively by the participants.

Schmitt’s (1998) study describes the longitudinal acquisition of individual words by measuring the development of four types of word knowledge: written form, associations, grammatical behavior, and meaning. Schmitt conducted this study with 4 international students who had never previously resided in an English speaking country. Measuring knowledge of the grammatical behavior of this study informs the design of the current word parts study.

Schmitt (1999) argues that it is needed to have a better understanding of what vocabulary items, like those on the TOEFL test, actually measure. To do this, 30 ESL participants of his study are given a number of TOEFL items and are then interviewed to discover what they actually know about the target words' associations, grammatical properties, collocations and various meaning senses. Measuring knowledge of the grammatical properties of this study informs the design of the current word parts study.

Schmitt (1998, 1999) collected data through a one-on-one interview. In the first session of the interview, he explained to the participant about the format of the session. Next, he checked to see whether the participant was comfortable with the metalinguistic notions and terms such as *noun*, *verb*, *adjective*, and *adverb*. Schmitt measured grammatical knowledge by the following question: "What word class (part-of-speech) is _____?" Depending on the answer, Schmitt went on to ask "Is there a (noun, verb, adjective, adverb) form? If so, what is it?" for the remaining 3 word class forms. In fact, asking for different parts of speech can be a controlled way of measuring word parts which could be considered a benefit of such a way of measurement. Due to having a specific number of answers, such a version of test does not seem to take a very long time to complete. However, Schmitt originally developed the test to measure grammatical knowledge and not word parts. As a result, this measurement seems to have a disadvantage of overlapping grammatical functions with word parts aspects.

Schmitt (1998) found that even his advanced participants, who were capable of pursuing postgraduate studies in an English speaking university, did not have very good mastery of the word derivations. In the first session of the interview, his four participants' mean scores were 2.45, 2.09, 1.82, and 2.55 out of 4 (approximately 56% on average). However, the participants knew some individual word classes better than others—nouns and

verbs were the best mastered while adjectives and adverbs appeared to be learned at a later stage. Schmitt believes that this suggests that adjective and adverb forms are unlikely to be learned from general exposure, maybe due to their lower frequency of occurrence, and hence might be good candidates for explicit instruction.

Schmitt (1999) found that the participants who chose the correct TOEFL item could provide approximately two appropriate word classes for the target words while those who missed the TOEFL item could provide the word forms for 1.63 (41%) word classes. Thus, Schmitt found that only a very small number of the participants, even though advanced, gave all four word classes for the target words. He believes that this suggests that L2 learners do not seem to master all members of a word family until relatively late in the acquisition process, and even this may not happen at all.

Schmitt and Zimmerman (2002) investigated 106 ESL graduate and undergraduate students' ability to produce correct derivations in four major parts of speech (noun, verb, adjective, and adverb). They presented a series of four similar contextualized sentences for each target word. The participants were required to write the appropriate derivative form of the target word in each blank. They were to place an X in the blank if no derivative form existed and also the participants were informed that the target word could be the appropriate form without alteration. The sentences were mainly designed to limit the possible derivations for each sentence to one word class. Schmitt and Zimmerman also provided the word classes (parts of speech) for the missing word in each sentence to facilitate the task. The following is an example of an item:

stimulate		
stimulation	Noun	A massage is good _____.
stimulate	Verb	Massages can _____ tired muscles.
stimulating	Adjective	A massage has a _____ effect.
_____X_____	Adverb	He massaged _____.

(Schmitt & Zimmerman, 2002, p. 169)

This method has the advantage of measuring derivations in a controlled way (limited specific answers to each item). The test provides context for the derivative forms in contrast to the previous method which requires the participants to provide the word forms in isolation. In other words, using this method, the researchers did not rely on the participants' metalinguistic knowledge by framing the prompt in metalinguistic terms. Also, considering the fact that each derivative form is supposed to be used in a sentence, such a method may be a better way of measuring grammatical functions, as Nation's framework for productive mode of grammatical functions aspect necessitates using the word in a sentence.

On the contrary, such a test may take a long time to complete depending on the number of target words. The reason is that the participants need to read four sentences for each target word and try to find an appropriate form of the target word that can match the prompt sentence. Also, the production of the proper derivative form is dependent on the comprehension of the prompt sentences and this issue may pose some risks to the test. In other words, when designing such a test, the researcher needs to make sure that all prompt sentences are semantically similar, and easy to understand. Therefore, it may be claimed that such a test may measure more than just the production of derivative forms of the target words (it may also measure the participants' comprehension of the prompt sentences).

Schmitt and Zimmerman (2002) found that the participants' mean number of derivatives was 37.6 (58.8%) out of a possible 64 (16 target words x 4 derivatives per word).

This shows that the participants produced approximately two of the four possible derivatives. However, this number varied across two groups, with their undergraduate ESL students tending to provide 2 or 3 forms, and the graduate students tending to provide 3 or 4 forms. A relatively small percentage of the cases showed full mastery of the word family (all 4 forms) as even the advanced graduate participants provided fewer than 45% of the possible derivatives for words they believed they knew. In fact, only 18.2% of all four major word classes were known by their combined non-native participants productively. Verb derivatives were the best known with 67% produced. Nouns were next with 63% produced. Adjectives and adverbs were less known, 54 % and 52 % respectively.

Pacheco (2005) investigated the effect of direct instruction of affixes for promoting vocabulary acquisition to improve reading comprehension. In a part of her study, she administered two quizzes, before and after her treatment (instruction of affixes), on forming new words using affixes with the instructions that read “make all the new words and word families you can by adding prefixes and suffixes to the words given” (Pacheco, 2005, p. 175).

For 10 target words in the first quiz, her students wrote from 13 to 44 words while the number of correct answers ranged from 10 to 31 (with percentages of correct words ranging from 39% to 91%). For 15 target words in the second quiz, her students wrote from 15 to 43 words while the number of correct words were from 5 to 34 (with percentages of correct words ranging from 33% to 100%). Pacheco (2005) found that, in the first quiz, students had a high amount of guessing and most participants had added the inflectional suffixes to almost all target words (treated them as verbs). She also found that some students remembered *im*, *in*, *un* and *dis*, but could not always use them correctly. As a result, the student who wrote the largest number of words (44 words) had 31 of them correct for a 70%

of accuracy, and the student who wrote the smallest (13 words) had 10 of them correct for 76% of accuracy. The results of the second quiz show that 18 (out of 22) students, who took both quizzes, increased their scores while one student's score remained the same, and 3 students had lower scores. Increases ranged from 10% to 60%. Because the target words were different from the ones in the first quiz, the obtained results show an improvement in students' use of new words. However, the results of the second quiz did not seem to inform this study because the treatment (the instruction of affixes) could have an effect on students.

Ishii and Schmitt (2009), in a part of their study, measured the derivative word forms in an attempt to develop one size/depth vocabulary test battery. They propose a principled scoring scheme that combines both size and depth scores, and this can suggest a method of making the results of the scores accessible to the students. The following is the test format they prepared for measuring derivative word forms:

Targeted word	Noun	Verb	Adjective
Stimulate	stimulation	stimulate	stimulating
Educate			

(Ishii & Schmitt, 2009, p.10)

Five hundred and twenty-three university students in Japan were asked to write one word form under each part of speech. They piloted the test and found that the regularity in forming adverbs was problematic, and as deleting adverbs did not lower the reliability of their test, they decided to measure just three parts of speech. In the cases where more than one possible word form existed, the participants were required to write only one of the forms.

This method of measuring word forms is, in fact, a principled way of what Schmitt (1998, 1999) did. Therefore, it has the same aforementioned advantages and disadvantages.

Ishii and Schmitt found that the participants' mean number of derivatives was 17.61 out of a possible 45 (15 target words x 3 derivatives per word). Their learners scored 39.13% of derivatives correctly. Comparing the results of their four tests, they found a weakness in the derivative area. However, they attested that this test was the only productive test in their test battery. Interestingly, the participants with a larger vocabulary size performed better on the derivatives test than on the other two tests—the test of multiple meaning sense and the test of lexical choice between near-synonymy.

In summary, the literature shows that between derivative and inflective forms, derivations are prioritized. In other words, considering the fact that inflections and derivations impose different learning burdens (Schmitt & Zimmerman, 2002), as derivations are generally acquired after inflections (Berko, 1958; Ward & Chuenjundaeng, 2009), it can be a better option to measure derivations because the results may reveal deeper depth of vocabulary knowledge. This makes sense because inflectional affixes are often specifically taught, they are frequent, and there is usually a lot of input for them (Ward & Chuenjundaeng, 2009). Literature also shows that there is overlap between measuring grammatical functions and measuring word parts (see Ishii & Schmitt, 2009; Schmitt, 1998, 1999; Schmitt & Zimmerman, 2002). The discussed literature (Schmitt 1998, 1999 in particular) shows that even for the purpose of measuring grammatical functions, different derivative forms of the target words are required to be produced.

Associations

Word association research is consistent with lexical models which metaphorically are named 'network' or 'web' to explain the organization of the mental lexicon (Aitchison, 2003; Wilks, 2009). In other words, the associative links elicited in word association tasks are assumed to show the links in individual lexical and semantic networks (Alberchtsen, Haastrup, & Henriksen, 2008), and this seems to allow the researchers to identify similarities and differences in such networks between individuals. The properties of the associations can be categorized in a number of ways, but the most common categories, especially in earlier studies, are paradigmatic, syntagmatic, and clang (Fitzpatrick, 2006).

Meara (1983) defines syntagmatic, paradigmatic, and clang associations in his study. Syntagmatic association forms an obvious sequential link with the stimulus word. For example, words such as *bark*, *spotted*, and *bite* could be associated with *dog* syntagmatically. Paradigmatic association forms the same grammatical form class as the stimulus word. For example, for *dog*, words such as *cat*, *wolf* or *animal* could be associated paradigmatically. In addition to these two, Meara asserts that children produce clang associates as well. Clang associates are responses which are related to certain phonological features of the stimulus word which carry no clear semantic relationship to it. Assonance, rhyming responses, or responses with the same initial sounds as the stimulus are types of clang associates. For example, *save* and *cave* are clang associations or *auto*, *tomorrow*, *swallow*, *zoro*, and *borrow* could appear as clang associations.

Aitchison (2003) categorizes four main types of associations: coordination, collocation, superordination, and synonymy. She defines coordination as words which cluster together on the basis of the same level of detail, such as *salt* and *pepper*, and *sugar* and *salt*.

In addition, she believes words which go together in a group of only two members such as *day* and *night*, *left* and *right*, or common antonyms such as *hot* and *cold* are also coordinates. However, words of different classes that usually go together like *bright* and *red*, *cold* and *hands*, *salt* and *water* are categorized as collocations. Superordination refers to a term which serves as a cover or umbrella term for another such as *insect* for *butterfly*, or *animal* for *dog*. Synonymy is said to be the least common of the four associations by Aitchison as it is not common to find a near synonym of a word such as *starving* for *hungry*.

In defining syntagmatic, paradigmatic, and clang categories, Fitzpatrick and Munby (2014) refer back to Nation's (2013) aspects of word knowledge – form-based, meaning-based, and use-based knowledge. They point out that clang responses are form-based in that they have phonological similarities to the stimulus words, paradigmatic responses are meaning-based as they are from the same word class and have related meanings, and syntagmatic responses are use-based because they are usually found beside the stimulus word in a text. In studies such as Fitzpatrick (2006), the link between the classification method and Nation's framework is even more explicit with categories and sub-categories matching exactly those in Nation's framework.

Fitzpatrick (2006) used Nation's (2001) table of what is involved in knowing a word as a basis for identifying three main categories of association: meaning-based association, position-based association, and form-based association. Her meaning-based category is similar to the paradigmatic classification and is taking in most of Nation's 'meaning' category. Her position-based category is similar to syntagmatic and has the features in common with Nation's 'use' category. Her form-based category is similar to the clang category, includes orthographic as well as phonological associations, and is overlapping Nation's 'form' category. She also considered a category of 'erratic association' for

responses which were based on false cognates, or for which a link could not be identified even by the subject themselves. In other words, meaning-based associations are synonyms, context related associations and conceptual ones. Position-based associations are consecutive words and collocations. Form-based associations are words with different affix or similar form words with different meanings, and erratic ones are no link or blank replies. Fitzpatrick (2007, p. 324) rewords this as follows:

- meaning-based responses (i.e. those determined by semantic characteristics),
- position-based responses (determined by syntactic and collocational characteristics),
- form-based responses (determined by phonological, orthographical or morphological characteristics), and
- erratic responses (where no link between cue and responses was apparent, or where no response at all was given).

Finally, Nation (2013) asserts that the associations of a word are mainly the result of the various meaning systems that the word fits into. These include synonyms, antonyms, family members of the same general headword, words in a part-whole relationship, superordinate and subordinate words. Nation's (2013) framework for productive mode of associations aspect indicates "what other words could be used instead of this one?" (p.49), and Nation (2013) interprets this question as "being able to produce synonyms and opposites for [the target word]" (p. 50). To test this aspect, Nation (2013) asks whether the learner can "recall this word when presented with related ideas" (p.538), and for the purpose of testing

specifically this aspect, he asks whether the learner can “add to this list of associated words” (p. 551).

Nation’s instructions for measuring associations productively may not seem as straightforward as his instructions for measuring word parts productively. On the one hand, Nation is asking for the synonyms and antonyms of the target word; on the other hand, he is asking for all related ideas and possible associations of the same word. This can be interpreted as he probably is looking for more than synonyms and antonyms. Thus, a synonym and antonym test for a part of the purpose of the current study is necessary in the first place. In addition, to fully serve the purpose of measuring associations productively, literature needs to be investigated to see how associations were measured productively by different researchers.

For measuring associations, Schmitt’s and Meara’s (1997) instrument had 20 prompt words followed by three blanks to write three possible associations. A four-point Likert scale (0–3) was also attached to each prompt word to enable the participants to indicate how well they knew the verbs. Unfortunately, they have not provided the instructions to the test neither in the text nor in their sample instrument. The study instrument was administered to the participants near the beginning of their school year (T1). Also, near the end of the school year, using the same procedure, the same test was administered (T2).

They found that on the contrary to the increase of vocabulary size of the participants measured by the Vocabulary Levels Test, the association scores did not bear this out. In other words, Schmitt and Meara were surprised to see that their participants could supply approximately 50% of possible associations for target words they rated as known. This means that they could give an average of 1.5 nativelike associations out of 3 possible for

each target word. Therefore, they concluded that the participants did not have very high levels of associative knowledge (even for target words they thought they knew).

For measuring associations, Schmitt (1998) asked his 4 participants to give 3 responses to each target word. He elicited associations by using the following instructions, “Please give the first 3 words you think of when you hear the word _____.” Schmitt did not find very impressive results on associations. The four participants’ mean scores of associations were 1.09, 1.64, 1.09, and 1.18 (less than 50%) out of a possible 3 (11 target words \times 3 = 33/11 = 3) on the first session of the interview. Schmitt found that two participants’ associations became more native-like over time; however, the third one remained at about the same level, and one participant dropped out of program.

Schmitt (1999) tested association with an association elicitation instrument. The instrument consisted of 6 target words, each followed by 3 blanks. The participants were asked to provide the first 3 words they thought of when they saw the target words similar to the way Schmitt (1998) measured associations in his previous study. Schmitt (1999) found that of 136 correct associative responses to the TOEFL items, 45 (33%) matched with native-like associations. The results illustrate that if a participant answers a TOEFL vocabulary item correctly, it cannot necessarily be inferred that s/he associates that word with other words in his or her mental lexicon in a native-like way.

Wolter (2001) measured word associations productively in a part of his study to compare lexical knowledge of 9 English L1 speakers and 13 English L2 learners. For the purpose of his study, he used an aural-written test with the following instructions:

The following test is a word test. You will hear several words and you will be asked to respond with the first word that comes to mind upon hearing the word. There are no right or wrong answers, so try not to take a long time considering your response.

(Wolter, 2001, p.51)

The results showed that while English L1 participants produced a much bigger percentage of paradigmatic associations in comparison to English L2 participants (51.7% compared to 19.7% respectively), English L1's and L2's performance on syntagmatic association was almost the same (41% compared to 37.7% respectively). English L2 participants made 35.1% of clang-other association while English L1 participants made only 7.2%. Finally, while English L1 participants provided an association for all target words, English L2 participants had no response for 7.6% of the target words. All in all, while English L2 participants did not provide as many paradigmatic associations as English L1 participants, they provided almost the same number of syntagmatic associations. The provided associations were rated and assigned to the defined categories with two judges. The categories were defined in advance.

In a series of different studies, Webb (2005, 2007a, 2007b, 2007c, 2009b) measured associations, as a part of his studies, to investigate the effects of: (a) receptive and productive learning tasks on vocabulary knowledge (2005), (b) contextualized and decontextualized learning tasks on vocabulary knowledge (2007a), and (c) receptive and productive learning of word pairs on vocabulary knowledge (2009b). He also investigated the effects of: repetition (1, 3, 7, and 10 encounters) on vocabulary learning (2007b), and synonymy on vocabulary learning (2007c).

Webb, in his series of studies, had a test of productive knowledge of syntagmatic association in which the participants were asked to write an L2 syntagmatic associate beside each target word. He also had a test of productive knowledge of paradigmatic association in which the participants were required to write a paradigmatic associate beside each target word. Coordinates, superordinates, subordinates, antonyms, and synonyms were all scored as correct in his test. Syntagmatic associates were not scored as correct because the previous

test measured this aspect, and this issue was carefully explained in the instructions.

Following Webb's series of studies, one could conclude that in addition to synonyms and antonyms, coordination, superordination and subordination could also be considered when associations are measured. However, Webb's findings do not seem to inform the current study as he measured the participants' performance on disguised words after his treatments.

All of the aforementioned research studies have one factor in common. They all present target words and then ask for the words that come to mind as a response (either one word or three words). This method of measuring associations productively seems appropriate for the purpose of the current study.

Collocations

Interest in collocations as an aspect of vocabulary knowledge originates from the widespread view that language learners struggle with formulaic language in general and collocations in particular (Brown, 2014). Collocations are also viewed as a problem even for advanced learners of English (Nesselhauf, 2005). To address this issue, lists of collocations have been developed which mainly have focused on the needs of EAP students. For example, Simpson-Vlach and Ellis (2010) have made an academic formulas list, and Coxhead (2008) provides information on the recurrent phrases and collocates in her Academic Word List (Coxhead, 2000). Durrant (2009) made a list of collocations for academic purposes while Shin and Nation (2008) made a list of frequent collocations in spoken language.

Collocations attract a considerable number of researchers both in their own right and as a prominent part of formulaic language. However, collocations are one of the most complicated areas of investigation. The reason is that they can have a variety of definitions. Laufer and Waldman (2011) argue that collocations do not have one simple and precise

definition. In the same vein, Webb and Kagimoto (2011) state that a universal definition of collocation has yet to be reached. On the contrary, Wray (2009) asserts that there does not have to be universal agreement on one definition of collocation, but she believes there is value in reflecting on the implications of definitions used in different research studies.

Nesselhauf (2004, 2005), Granger and Paquot (2008), and Barfield and Gyllstad (2009) discuss two approaches for collocations: the phraseological approach and the frequency-based (statistical) approach. The phraseological approach defines multi-word units linguistically. In other words, this approach sets linguistic criteria to distinguish one type of phraseological unit from another (in particular phraseological units from free combinations). This approach has the idea of scales of opacity and fixedness. That is, collocations are assumed to occupy a certain space along the scales—less restricted than idioms which lie at one extreme, but more restricted than free combinations at the other extreme.

The frequency-based (statistical) approach considers collocations as words that co-occur within a certain distance of each other. This approach is linked with corpus linguistics – researchers use corpora and software to find collocations. In other words, they investigate the recurrent sequence of words or consider one word as a node and investigate the items appearing within a certain span, usually four orthographic words before or after the node. Statistical techniques are then used to determine the significance of the collocations found.

Collocations can be regarded as syntagmatic associations. Fitzpatrick (2013), in clarifying an approach to measuring associations response data, compares paradigmatic responses with syntagmatic responses and states that collocations can be included in syntagmatic responses which might be found in the same phrase as the cue. In the same vein, Read (2000) states that syntagmatic association is when two words often occur together in a phrase, that is, they collocate. Schmitt (2010) also defines syntagmatic associations as a

sequential relationship to the stimulus word usually having differing word classes. He states that syntagmatic relationships involve the contiguity (occurring in close proximity) of words – which is an aspect of collocations. In this regard, Webb and Kagimoto (2011) assert that the position of a collocate in relation to the node word is an area of interest that has been overlooked in classroom-based research. They argue that collocation is usually defined as a relationship between a collocate and the node in which the collocate may be placed before or after the node and not necessarily adjacent to it.

Nation (2013) regards collocations as a kind of *multi-word unit* and believes that there are four major kinds of multi-word units: (a) a multi-word unit can contain a group of words that commonly occur together such as *take a chance*; (b) a multi-word unit can contain a group of words where the meaning of the phrase is not obvious from the meaning of the parts such as *by and large* or *be taken in*; (c) a multi-word unit can simply refer to all the combinations of a particular word or type of a word and its accompanying words; and (d) a multi-word unit can refer to word groups that are naturally seen as being formulaic, that is, items stored as single choices. Nation argues that collocations differ greatly in size (the number of words involved in the sequence), in type (function words collocating with content words; content words collocating with content words), in closeness of collocates, and in the possible range of collocates (*commit* with *murder, crime, suicide, hara kiri, etc.*).

Boers et al. (2006) also regard collocations as a kind of multi-word expression, and regard other types as fillers such as *sort of*, functional expressions like *excuse me*, idioms such as *back to square one*, proverbs like *let's make hay while the sun shines*, and standardized phrases such as *there's a growing body of evidence that*.

Laufer and Waldman (2011) regard collocations as “habitually occurring lexical combinations that are characterized by restricted co-occurrence of elements and relative

transparency of meaning” (p. 648). They assert that restricted co-occurrence differentiates collocations from free combinations. On the other hand, relative semantic transparency of collocations distinguishes them from idioms whose meaning is much less transparent than that of collocations. Laufer and Waldman (2011) exemplify restricted co-occurrence as follows: *tea* collocates with *strong* but not with *powerful*, *discussion* collocates with *hold* or *have* but not with *deliver*, and *speech* collocates with *deliver* but not with *hold*. They also exemplify relative semantic transparency as follows: *face in face a problem* is not used with its original meaning, but the expression is clearer than *face the music*—an idiom that means *show courage*. Laufer and Waldman believe that many collocations are totally transparent if the learner is familiar with the individual words (e.g., *apply for a job*, *make a decision*, and *submit a proposal*). All in all, they conclude that restricted co-occurrence and semantic transparency place collocations on the continuum between free combinations and idioms. Thus, according to the definition used in their article, they regard *throw a disk* and *pay money* as free combinations, *throw a party* and *pay attention* as collocations, and *throw someone’s weight around* and *pay lip service* as idioms.

Webb, Newton and Chang (2013), Webb and Kagimoto (2011), and Nguyen and Webb (2016) define collocation from a statistical standpoint. This means that collocation refers to the regular co-occurrence of words within a given span which is demonstrating a statistical strength of co-occurrence. This definition has been widely accepted within the field of corpus linguistics (Halliday, 1966; McEnery & Wilson, 2001; Sinclair, 1991). In fact, measures such as mutual information (*MI*) scores, *t*-scores, and log-likelihood (explained in detail in the Scoring of the Tests) indicate the aforementioned statistical strength (Webb et al., 2013). Such measures indicate that two words occur more frequently together than would be expected by chance alone. The advantage of statistical approach for collocations is that

researchers can use these measures to quickly and objectively identify collocates for a word. Webb et al. (2013) exemplify the word *gain* and assert that a search for nouns that occur immediately after that in the British National Corpus shows that *access, control, entry, experience, confidence, support, power, and weight* all have acceptable mutual information scores, indicating that these words co-occur immediately after *gain* expectedly and not accidentally.

Webb et al. (2013) believe that the statistical approach has the advantage of removing a subjective component that can lead to different interpretations of what is and what is not a collocation. They also argue that this approach is more ecologically valid, “because in incidental learning contexts, learners are likely to encounter multiword items of varying degrees of semantic transparency” (Webb et al., 2013, p. 93).

Following Webb et al. (2013) and Webb and Kagimoto (2011), collocation will be defined from a statistical standpoint in the current study and will refer to the regular co-occurrence of words within a given span demonstrating a statistical strength of co-occurrence.

Nation’s (2013) framework for productive knowledge of collocations asks “what words or types of words must we use with this [target word]?” (p.49), and Nation (2013) interprets this as “being able to produce words that commonly occur with [the target word]” (p. 50). To test this aspect, Nation (2013) asks whether the learner can “produce the word with appropriate collocations?” (p.538), and for the purpose of testing this aspect specifically, he asks whether the learner can “produce collocations to go with this word” (p. 552). The framework is straightforward. Collocations (the words or types of words which can be used or can commonly occur with the target word) are the main target.

Schmitt (1999), in a part of his study as introduced in a previous section, measured the collocations productively in an innovative way. He asked his participants to verbally compose 3 sentences with the target words included. The participants were instructed that the words in these sentences needed to naturally occur together with the target word in the same way that words such as *blond* and *hair* could go together. They were instructed to give the most common sentence they could think of rather than an elaborative one. Schmitt gave the participants 3 situations or topics to guide them for this purpose because providing sentences with the target words seemed a difficult task. A sentence was without further comments if the participants included the target word in it even if it did not exactly relate to the given situation. This method of measuring collocations seems promising for a limited number of target words and a one-on-one procedure (interview). Using such a method, the researcher needs to keep prompting the participant to collect the intended data. Such a procedure would not be appropriate for classroom-based assessment with a large number of target items.

Schmitt (1999) found that for the items which were correctly answered in the TOEFL test, the participants gave collocates for 1.41 out of 3 sentences (47%) on average. For items which were missed on the TOEFL test, they composed 1.13 sentences (38%) which included a collocate. This shows that the participants were able to provide even less than half of possible collocates in sentences which is not an impressive finding on collocations.

Studies such as Schmitt and Meara (1997) and Schmitt (1998), which measured associations by asking the participants to provide the first words they might think of when they heard or read the target words, might have also measured collocations simultaneously. The reason is that the aforementioned instructions are not able to differentiate collocations from other associations—collocations can be considered syntagmatic associations

(Fitzpatrick, 2007, 2013; Read, 2000; Schmitt, 2010). As a result, the participants may think of collocations of the target words when they hear or see the words.

Webb (2005, 2007a, 2007b, 2007c, 2009b) addressed this problem in his series of research studies, as he had two separate tests to measure productive knowledge of syntagmatic association and paradigmatic association, and Webb asserts that this issue was carefully explained in the instructions.

Although a considerable number of studies on L2 vocabulary has been published in the last 15 years, they hardly include dedicated studies to L2 knowledge and development of collocations (Barfield & Gyllstad, 2009). Also, the number of studies on assessing collocations productively has been extremely limited.

In an attempt to view collocations as an independent construct and test them as a whole, Revier (2009) developed a new productive test format entitled CONTRIX. This test presents a modified form of a cloze test where participants are asked to select the combination of a verb, article, and noun that best complete a sentence. The following illustrates this better.

The quickest way to win a friend's trust is to show that you are able to	tell	a/an	Joke
	take	the	Secret
	keep	—	Truth

(Revier, 2009, p.129)

Revier asserts that although the test has the characteristic of receptive tests (selection), its format taps productive knowledge of participants. He argues that the test requires participants not only to combine the constituents of phrases but also to encode

grammatically the noun constituent for determination. Revier believes that this provides some insight into the participants' productive knowledge.

As Revier (2009) argues, the format of the test does not seem to tap only productive knowledge and receptive knowledge is also involved. A 45 item pilot version of the test with a group of Danish EFL learners met the criteria of validity and reliability and proved to distinguish among learners of different proficiency. The university student participants' (n=19) mean score was 28.8 out of 45 (64%). The mean scores for other participants (10th and 11th grade students) were 17.2 and 21.8 (38.22% and 48.44%) respectively.

The limited number of the studies to measure the production of collocations and to evaluate the strength of such production is proof of the necessity for conducting new studies in this realm.

Form and Meaning

It is believed that form and meaning link is the basic dimension of vocabulary knowledge and most probably the first to be acquired (Laufer & Goldstein, 2004) because words are units of meaning and they play the most important role in communication. Nation (2013) states that learners consider a word known as soon as they know its form (what the word sounds like or looks like) and its meaning. However, he believes that not only do language learners need to know the form and meaning of a word, but also they need to know how to connect these two. Nation exemplifies this with the word *brunch*. He states that a learner of English might know the form *brunch* and also might know that there is a concept for a single meal which takes the place of breakfast and lunch. The learner might know that *brunch* is the appropriate form for the concept of a meal combining breakfast and lunch. However, it is possible to know the form *brunch* but have no concept of its meaning, or it is

possible to know both but not to be able to connect the two. The strength of connection between these two will show how readily one can retrieve the meaning when encountering the form, and retrieve the form when intending to express the meaning (Nation, 2013).

Form and meaning association can take different forms and can reflect different parameters according to which strength of knowledge of meaning is assessed (Laufer et al. 2004; Laufer & Goldstein, 2004). The strength of knowledge of meaning is a combination of the four aspects of knowledge of meaning: passive recognition, active recognition, passive recall, and active recall (Laufer & Goldstein, 2004).

Nation's (2013) framework for productive knowledge of form and meaning states "what form can be used to express this meaning?" (p.49), and Nation interprets this as "being able to produce the word to express the meaning [of the target word]" (p. 50). To test this aspect, Nation asks whether the learner can "produce the appropriate word form to express this meaning" (p.538), and for the purpose of testing this aspect specifically, he asks whether the learner can "translate these words into L2" (p. 551).

The framework clearly is asking for the translation of the words into L2. However, Laufer and Goldstein (2004) assert that in multilingual classes, or in situations when the researcher is not familiar with the participants' L1, an L2 paraphrase may be asked for to demonstrate passive knowledge (receptive knowledge), and an L2 target word for a synonym or a paraphrase may be required to prove active knowledge (productive knowledge).

Laufer and Nation (1999) validated a vocabulary size test of controlled productive ability that is, in fact, a way of measuring form and meaning. For each item in their test, they presented a meaningful sentence context and the first letters of the target item. The first letters prevent the participants filling in another word which could be used semantically in the provided context. The following is an example of their test:

The book covers a series of isolated epis _____ from history.

(Laufer & Nation, 1999, p.37)

They decided on the number of first letters for each item by eliminating the possible alternatives to the tested word. They decided to provide the minimal number of letters because their test was a test of productive vocabulary ability. If two letters could start two possible words in the given sentence, they would add an additional letter. The test measures 18 items at 2,000, 3,000, 5,000, University Word List⁹, and 10,000 word levels. They finally conducted two studies, one to check the reliability and validity of the test, and one to check the equivalence of four parallel forms of the test.

They found that the higher the proficiency of their test takers, the better their performance on different levels of the test. The total mean scores of their high school 10th, 11th, 12th, and university English as a foreign language participants were 21.7, 33.4, 40.1, and 55.8 out of 90 (18 items x 5 word levels) respectively. The participants had studied English for 5, 6, and 7 years respectively, and the university students were majoring in English. Because English was their foreign not second language and class work was the main source of input, Laufer and Nation believed that the class grade was a fair indication of language proficiency. Their participants' mean scores for the 2,000 word level was 11.8 (65.55%), 15 (83.33%), 16.2 (90%), and 17 (94.44%) out of 18 (18 items).

Read (2000) believes that it is not clear what the blank-filling test (e.g. Productive Levels Test, Laufer & Nation, 1999) is measuring. In fact, it is assumed that the Productive Levels Test (PLT) measures controlled productive knowledge. However, Read (2000)

⁹ This list was first published by Xue and Nation (1984) and suggests words for learners moving beyond the core 2,000 high frequency words towards a more academic vocabulary.

believes that the PLT seems to assess the receptive knowledge rather than productive. In this regard, Webb (2008) argues that because it is a test of cued recall, the presence of the first few letters could be sufficient to recognize the target word receptively.

The PLT is also criticized for inferring general productive vocabulary size from a small vocabulary sample – 18 items only (Meara & Bell, 2001; Meara & Fitzpatrick, 2000; Meara, 2005). Also, the target words of the test are from West's (1953) General Service List which is now out of date (Nation, 2004). A study on the basis of the VLT (Schmitt et al. 2001) shows that 30 items (at least) are needed to get reliable results at the targeted frequency levels. Therefore, following Laufer's personal advice and taking into account Schmitt et al.'s findings, Llach and Espinosa (2014) combined the PLT Version A and Version C (equivalent versions) from the parallel Version 1 (Laufer & Nation, 1999) to have a 30 item test. Their 197 EFL Spanish participants' mean scores of the 2,000 most frequent words were 7.26, 8.78, and 12.30 out of 30 in three successive grades (1st, 2nd, and 3rd grades of secondary education).

Despite such criticisms, the PLT and the VLT are widely used in different research studies and are believed to be appropriate tests to measure form and meaning relationship because there is no perfect vocabulary test (Meara & Fitzpatrick, 2000), and the tests (PLT & VLT) have shown evidence of validity and reliability (see Laufer & Nation, 1999; Schmitt, Schmitt, & Clapham, 2001). For example, in a study to explore Chinese EFL vocabulary knowledge, Zheng (2009) administered the VLT and PLT to 88 Chinese university students. The mean score of her participants for the 2,000 word level was 15.61/18 (86.74%), indicating that her participants seemed to have mastery over the words at 2,000 level.

Laufer and Goldstein (2004) assert that the task in active recall (testing form and meaning productively) is to find the L2 target word for its provided L1 translation. They

believe that the first letter of the target word should also be provided to prevent the test takers from producing nontarget words that have the same meaning. They used the same task in a part of their study in an attempt: (a) to investigate the validity of the sequence of the four strength modalities (discussed in Vocabulary Knowledge section) in their bilingual computer adaptive test of size and strength (CATSS), and (b) to examine the connection between strength modalities and academic success. As a part of their study, they measured active recall of 30 target words at each word level of 2,000, 3,000, 5,000, and Academic Word List. The mean scores (active recall) of their participants' performance at the aforementioned levels were 6.51, 5.35, 2.50, and 3.18 out of 30 respectively which seems to be a poor result. The mean score of their high school participants' at the 2,000 word level was only 6.51 out of 30.

Inspired by Webb (2008), Nakata (2013) in his study to optimize second language vocabulary learning from flashcards challenges the idea that the first letter of the target word should be provided. In a part of his study, he administered active recall tests to control the position of initial productive retrieval during learning. Nakata asserts that in the productive retrieval test (active recall test), it is necessary to prevent test takers from providing synonyms with the same number of letters of the target words. He believes one letter in the target word and the number of letters in the word should be given in addition to the word's L1 translation. He states that the first or last letter of the word should not be given as the cue because learners tend to remember the beginning and ending of the words more than the middle (Barcroft & Rott, 2010), and providing the initial or final letter may have a serious effect on the productive retrieval performance.

Webb (2005, 2007a, 2007b, 2007c, 2009b) measured form and meaning aspects productively using translation tasks in his series of research studies as well. For this purpose,

the participants were given the L1 meanings of the target words and were asked to write the target disguised words. No letter of the target words as cues was provided though.

Target Words

Because the current study aims to measure depth of vocabulary knowledge with a battery of tests, it is important to know whether the aforementioned studies (discussed in detail in the previous 4 sections) with different batteries of tests had one or different groups of target words, and whether those words were selected randomly or purposefully.

Schmitt and Meara (1997) selected 50 verbs from the Brown frequency list (Francis & Kucera, 1982) because verbs take the widest range of suffixes. Also, verbs are neither the easiest nor the most difficult grammatical class of words (Morgan & Bonham, 1944). They used the following criteria as a guideline: each verb could have a variety of suffixes; the verbs and their related nouns did not have an identical form (i.e., *to smile, a smile*); some suffixes with high difficulty level (Bauer & Nation, 1993) were represented; verbs with different frequency were included; and the verbs required a one-letter change to take a suffix (*indicate; indication*). Twelve participants were asked to rate how well they knew these 50 verbs. This information was used to choose mainly verbs that were partially known, a few that were relatively well known, and a few that were not known at all. A final list of 20 target words (6 from K1, 7 from K2, 5 from K3, and 2 from K4) was finally selected with the aforementioned criteria.

Schmitt (1998) selected only a few target words because the study investigated them in detail. First, he mainly needed a few target words that the participants would be exposed to during their course of study. He chose the target words from the University Word List (UWL, Xue & Nation, 1984) because the participants were university students. Second, he

selected only polysemous UWL words that had 3 or more meaning senses to be able to examine the participants' growing awareness of the different senses. Third, because the participants might at least partly know the words, he needed a few low-frequency words in order to allow insights into the beginning stages of lexical acquisition. For this purpose, Schmitt selected 3 words from the 4,000 to 5,000 word level of the Brown word list (Francis & Kucera, 1982) with the same criterion of having at least 3 major meaning senses. He piloted 18 candidate words on 12 international students to check their familiarity with the words, and mainly did this to confirm that the eventual target words were likely to fall at various points on the acquisition continuum for the participants. Schmitt reduced the candidate list to 11 target words (one from K2, 7 from K3, 2 from K4, 1 from K6), which later piloting indicated was the maximum that could be addressed in a 2-hour long session.

Schmitt (1999) selected six TOEFL vocabulary items (2 words from K2, 3 words from K3, and one word from K1), which included target words with the greatest number of different meaning senses. Two passages were selected from *TOEFL Practice Tests* (1995), which had three vocabulary items each. The passages and vocabulary items in the *TOEFL Practice Tests* were taken from actual test forms and could be considered representative of TOEFL items.

Pigada and Schmitt (2006) selected 70 nouns and 63 verbs (133 French words in total). They intended to include all parts of speech in the study; however, the grammatical component of the test battery hindered that possibility. They limited their selection of target words to nouns and verbs—nouns which were used with an article, and verbs which were used with prepositions. Furthermore, they assert that it would be a logical start in this type of research, since nouns and verbs are the most common parts of speech found in natural text (Webb, 2005), and they hope that additional research will extend to other word classes.

Webb (2005, 2007b) selected ten words (one from K2, one from K5, three from K6, two from K7, one from K8, one from K10, and one from K13) from six graded readers. The words were then deleted from the text and replaced with nonsense words. The words included six nouns and four verbs. Such a ratio (6 to 4) was used because nouns and verbs are the most common parts of speech found in natural text, and the ratio approximates their proportional frequency of occurrence (Kucera & Francis, 1967).

Webb (2007a, 2007c) selected twenty low-frequency words as target words – ten with high-frequency synonyms (Set A), and the other ten without high-frequency synonyms (Set B). Both sets of target words were composed of six nouns and four verbs. All of the target words were taken from the fifth frequency band in the COBUILD dictionary—the 6,601st to 14,700th most frequent words. Disguised forms replaced the L2 forms of the 20 target words. Using disguised forms ensured that the participants would have no prior knowledge of the target words. Because the participants in the experiments were not aware that disguised forms were used, learning the disguised forms simulated for the learners the experience of learning actual words.

Webb (2009) selected nine nouns and six verbs from the fifth frequency band in the COBUILD dictionary—the 6,601st to 14,700th most frequent words. To ensure that participants had no prior knowledge of the target words, they were replaced with 15 nonsense words.

Schmitt's and Zimmerman's (2002) priority was to select target words and word families that their participants were expected to know and use. Because the participants were either university students or English as a second language speakers who intended to study in an English-medium university, they chose words from the Academic Word List (AWL, Coxhead, 2000). They selected 20 words from the range of 10 frequency-based sublists of the

AWL and tried to balance among the four word classes as much as possible. After piloting, they chose 16 target words—six words from K2, eight words from K3, and two words from K4. The AWL includes relatively few adverbs; as a result, they chose only one adverb as a target word. Based on frequency data from the British National Corpus (BNC, 1995), each word but one (an exception) was the most frequently occurring member of its word family. Frequency, rather than factors relating to morphological difficulty, was the main criterion in choosing the target words.

Ishii and Schmitt (2009) chose their items from the most frequent 2,000 lemmas in the British National Corpus. Words were randomly sampled from this list until 18 words which had three or more derivative forms of different parts of speech were found. After piloting a few versions, three were discarded, and they ended up with 15 words—nine from 1K and six from 2K.

Comparing the reviewed selection of target words with the research purposes of the studies (discussed in previous parts), it can be concluded that addressing the research purposes (research questions) was the main criterion for the selection of target words. Reviewing the selection of target words of the studies reveals that: (a) it is possible to choose purposefully or randomly different sets of target words to address specifically the current research questions, and (b) no study has already focused purposefully and systematically on the most frequent 1,000 words, and this justifies the significance of the current study.

The Scoring of the Tests

Word Parts

Three studies were innovative in the scoring of the word parts test: Schmitt (1998), Schmitt and Zimmerman (2002), and Ishii and Schmitt (2009). To measure grammatical

functions, Schmitt (1998) obtained the norming word class and derivational forms from 3 dictionaries. In his study, the participants scored 1 point if they knew the word class of the target word, and 1 point to know how to change it into each of the 3 other word classes. If a word class did not have a form, participants got credit for being able to state that fact. When 2 or more forms were possible for any word class, only one of them was required for credit. Thus, the possible scores ranged from 0 (knowledge for no word class) up to 4 (knowledge for all four word classes). However, during the development of this section, Schmitt found that the norming data from the dictionaries sometimes conflicted with the native English pilot participants' answers. In other words, the dictionaries sometimes were listed some forms (especially adverbials) that were strange for native English speakers. In such cases, Schmitt consulted the BNC to check those forms' frequency of occurrence. If the frequency was very low, he still accepted it as a possible word form, but he also considered acceptable an answer that no form existed. For example, the very rare adverb form of *circulate*, *circularly*, is so uncommon that he also accepted the answer "no form exists."

Schmitt and Zimmerman (2002) and Ishii and Schmitt (2009) produced a list of derivatives for a norming list as well. They accepted more than one alternative (if any) for some of the word classes. Ishii and Schmitt (2009) prepared the accepted answers based on the answers from 21 English native speakers when they piloted the test.

Schmitt and Zimmerman (2002), in order to have a principled norming list, drew on three sources of information. First, they consulted four learner dictionaries and extracted all listed derivatives. Second, they checked the frequency of these derivatives in the BNC and eliminated those that had very low frequency or those that did not exist in the corpus. Third, 36 English as a first language university students completed their instrument. Balancing the data from these three sources, they compiled a principled list of acceptable derivatives.

However, the native speakers' responses did not always agree with the norming list. Yet, since their responses were direct answers to the study instrument, Schmitt and Zimmerman weighted them more heavily than the dictionary and corpus evidence. In some other cases, a considerable number of English native speakers indicated a certain derivative form whereas another considerable number indicated that no typical derivative form existed. In such conditions, they accepted either a derivative form or X as the appropriate responses. They accepted misspellings as correct as long as the intended derivative could be discerned because they were more interested in knowledge of the derivations than spellings.

All in all, based on the aforementioned studies, it seemed the preparation of a norming list could be considered a reliable way to evaluate participants' performance on the word parts test. The preparation of the norming list in the present research will be explained in detail in Chapter 3.

Associations (Synonym & Antonym, Superordination & Subordination)

The association measurement procedure in Schmitt (1998) required participants to give 3 responses for each target word stimulus. Schmitt compared these responses to a native speaker norming list. He scored each of the participants' responses against his native speakers' norming list. The 3 scores, allocated to this part, were totaled and matched against a four-category profile of native-like behavior. In Category 0, none of the 3 responses matched any of those on the norming list, which meant no native-like association behavior was demonstrated. In Category 1, some responses matched infrequent ones on the norming list, which showed a minimal amount of native-like association knowledge. In Category 2, the responses were similar to those typical of the native speaker norming list, indicating native-like associations. Finally, in Category 3, the responses were similar to those in the top

half of the native speaker norming group, indicating a full native-like rating. Eighty-two percent of the association responses from the 3 native speakers belonged in Categories 2 or 3, while 12% fell into Category 1 and 6% into Category 0.

In Schmitt (1999), 50 native speakers gave three association responses each for every target word to form a norming list. The L2 participants were then given credit for matching associations on this list. A score of zero was given for matching no association with any of the native norming responses at all. It was found that about 10% of native speakers' responses were either unique to themselves or were given by only a few other respondents; as a result, they were not representative of the overall native norms. The participants matching only such responses were given a score of 1. The participants who provided associations which were relatively frequent native responses scored 2 and if the responses were among the most frequent ones they scored 3. Schmitt asserts that scores of 2 and 3 could be considered native like.

While it seemed that a norming list could be prepared for the synonym and antonym test (consulting thesaurus dictionaries and considering a number of English native speakers' responses), the results of the pilot study (to be discussed in Chapter 3) showed that a norming list could not be provided for the superordination and subordination test. The reason was that the native speakers' responses to the test were not consistent. Also, English non-native learners often respond with more variety and less uniformity than English native speakers on word association tests (Meara, 1983). Zareva and Wolter (2012) argue that English non-native speakers' shared language experiences are different from native English speakers, their culture and educational background may also be different, and even their conception of words may be different from English native speakers. Such issues may lead to different answers by English non-native speakers which may not seem consistent with what native

English speakers have provided. Zareva and Wolter (2012) argue that such answers show that non-native speakers maintain connectivity of their own lexicons in ways that are not identical to native speakers which by no means should suggest that it is of a lesser value.

Webb (2005) used a common-sense approach to evaluate the responses to his syntagmatic association test. He considered a comparison between English L1 and L2 responses. Unfortunately, the details of his approach are not mentioned in his study. However, a personal communication (March, 2015) revealed that he went through all responses and checked for responses that appeared to be common syntagmatic association. These were scored as correct. The syntagmatic associations that were questionable were scored by him and one other rater. When there was a difference of opinion then there was discussion of that response until agreement was achieved.

Employing raters to evaluate responses to the superordination and subordination test seemed reasonable and even necessary for the current study. The raters were to accept all reasonable responses (if any), and the calculation of inter-rater reliability showed how consistent they had done this. The details of such a common-sense approach will be discussed in Chapter 3.

Collocation

Providing a norming list for collocations could also be observed in some research studies such as Schmitt (1999). However, considering the fact that online collocational databases are currently available, using such databases seemed to be a more objective way of scoring collocations. The COCA/BNC Collocates Database (the COCA/BNC Collocates hereafter) and the Google Web 1T 5-Gram Database (the Web1T5 hereafter) are two examples in this regard.

The COCA/BNC Collocates is created by Mark Davies, a professor of linguistics at Brigham Young University, USA. The database contains millions of node/collocate pairs which are based on the 520 million word Corpus of Contemporary American English (COCA) and the 100 million word British National Corpus (BNC). The database is available at <http://corpus.byu.edu/>.

The Web1T5 consists of frequency counts from 1 trillion words of English Web text (Brants & Franz, 2006). While originally designed as a resource to build better language models for machine translation, its public release in 2006 was greeted by many researchers (e.g. Mitchell et al. 2008) in computational linguistics (Evert, 2010). Evert (2010) states that many corpus linguists and lexicographers are interested in using the Web1T5 database as a source of collocations, and he even investigates the validity of the linguistic usefulness of the Web1T5. However, Evert (2010, 2015) draws some legitimate criticism of the Web1T5 including the fact that its huge size has made it very noisy with results which even include pornography. His studies seem to show that the Web1T5 may not necessarily be the best database in computational linguistics. However, the pilot study for the present research (will be discussed in Chapter 3) involving scoring of English native speakers' performance on the collocation test showed that the Web1T5 tended to include more of their answers in comparison to the COCA/BNC Collocates. The database is available at http://corpora.linguistik.uni-erlangen.de/demos/cgi-bin/Web1T5/Web1T5_colloc.perl.

To use such collocational databases, collocations need to be defined from a statistical standpoint. Following Webb, Newton and Chang (2013) and Webb and Kagimoto (2011), the current study defined collocation from a statistical standpoint and referred to the regular co-occurrence of words within a given span demonstrating a statistical strength of co-occurrence. Measures such as mutual information (*MI*) scores, *t*-scores, and log-likelihood

indicate the aforementioned statistical strength (Webb et al., 2013). Such measures indicate that two words occur more frequently together than would be expected by chance alone. It has been suggested that a *t*-score of 2 or above and/or a *MI* score of 3 or above may be taken as indicative of collocation (Hunston, 2002; Stubbs, 1995). However, the two association measures tend to emphasize rather different sets of collocations. In particular, whereas rankings based on *t*-scores tend to highlight very frequent collocations, *MI* score tends to give prominence to word pairs which may be less common, but whose component words are not often found apart (Stubbs, 1995). Thus, pairs like *good example*, *long way*, and *hard work* attain high *t*-scores but low *MI* scores, while pairs like *ultimate arbiter*, *immortal souls* and *tectonic plates* attain the reverse (Durrant & Schmitt, 2009). As a result, the current study considered both *t*-score and *MI* score. On the contrary to the COCA/BNC Collocates, the Web 1T5 has *t*-score indicative option for collocations. Wolter and Gyllstad (2011) argue that the *t*-score is a measure of the strength of co-occurrence of words with a node word in a corpus, usually within a span of ± 4 words. It tests the null hypothesis by contrasting the observed frequency of co-occurrence in the corpus with the expected frequency of co-occurrence based on random distribution.

Chapter Summary and Conclusion

The literature review clarifies the contribution of vocabulary knowledge to language skills, the significance of the most frequent words, and the significance of Nation's aspects of vocabulary knowledge. It also justifies the reasons why such aspects need to be measured productively. Measuring aspects of vocabulary knowledge productively, as discussed, shows its significance when one considers the challenges EAP students usually meet with writing and speaking language demands of university work.

Productive vocabulary knowledge plays a significant role in writing and speaking skills; however, the current literature review shows that productive knowledge of the most frequent words has not been investigated to date. Therefore, the current dissertation mainly investigates EAP students' productive performance on four aspects of vocabulary knowledge of the most frequent words in a group of Iranian EAP university students. To do this, Nation's productive aspects of vocabulary knowledge are considered the main framework of the measurement, and a battery of depth of vocabulary tests are used whenever necessary to inform the current study. The next chapter elaborates on the methodology and study design.

CHAPTER THREE: METHODOLOGY

Overview

The following chapter clarifies: (a) the design of tests to measure four aspects of productive vocabulary knowledge—word parts, associations, collocations, and form and meaning, (b) the use of the appropriate criteria for selecting the target words and scoring the tests, and (c) the pilot of the designed tests to ensure that the instruments of the study were well-designed. The method including participants, instruments, and procedure are discussed in detail.

Research Design

The current study investigated the extent to which Iranian EAP students had productive (depth of) vocabulary knowledge of words at the 1,000 word frequency level based on their scores on the aforementioned tests. For this purpose, as discussed in Chapter 2, four aspects of vocabulary knowledge were the main focus: (a) Word Parts, (b) Associations, (c) Collocations, and (d) Form and Meaning. The following sections demonstrate how a test was designed to measure each aspect of vocabulary knowledge. The test forms were drafted, scrutinized, re-worded, and revised several times through discussion within the project team (my supervisors and I) and also through piloting with English L1 graduate students and international EAP students.

Word Parts

Nation's (2013) framework targets both inflections and derivations for measuring word parts. Therefore, inspired by Schmitt's and Meara's (1997) study, the following test was considered the first attempt to measure both inflections and derivations (word parts). The

target words were presented on the left side of the page followed by enough space to write their different inflectional and derivational forms. The following show the instructions and examples of the test:

Add any allowable letters to the beginning or end of the following words to change their meanings, uses or forms according to the provided examples. If no letter can be added to the beginning or end of the word, write NONE in front of the word.

For example: FLY: flying; flies; flyer; flyers; flier; fliers
 BOX: boxes; boxed; boxing
 FRESH: freshness; fresher; freshest; freshly; freshen; freshening;
 freshens; freshened
 ALREADY: NONE
 TOUCH: touched; touches; touching; untouched; touchable;
 untouchable; untouchables

Through discussion within the project team, it was noticed that the participants would find it easier to add *s*, *es*, *ed*, and *ing* to the stem and make inflectional forms of the target words than to recall the possible derivational forms. Therefore, separating inflections from derivations seemed reasonable. For the purpose of measuring inflectional forms, the following test was considered:

Add any allowable ***ed***, ***ing***, ***s***, or ***es*** to the end of the following words to change their meanings, uses or forms. If none of such letters can be added to the end of the word, write NONE in front of the word.

For example:

	ed, ing, s or es
BELIEVE	believed, believing, believes
MAKE	making, makes
BOX	boxed, boxing, boxes
FRESH	NONE
ALREADY	NONE

This method of measuring inflections, however, did not seem to be a truly productive test. In fact, it looked more like a receptive recall test (see Laufer & Goldstein, 2004) because the possible inflectional suffixes were provided, and the participants were asked to recall receptively which one matched the target words. Therefore, the test did not seem suitable to measure word parts productively.

For the purpose of measuring derivational forms, inspired by Pacheco's (2005) test instructions, the provided instructions of the aforementioned first attempt were rephrased and modified to the following:

Make all the new words by adding any letters (prefixes and suffixes) to the words given. Do **NOT** add *ed*, *ing*, *s* or *es*. If no letter can be added to the beginning or end of the word, write NONE in front of the word. Please write your answers as clearly as possible.

For example:

BELIEVE	believable; unbelievable; unbelievably; believer; unbeliever; disbelieve; disbeliever; disbelievingly
MAKE	maker; remake
BOX	NONE
FRESH	freshness; fresher; freshest; freshly; freshen
ALREADY	NONE
JUDGE	judgement; judgmental; misjudge; misjudgment; prejudge; prejudgment

The test seemed to have the potential for measuring derivative forms, but it was unable to measure both inflections and derivations as discussed in Nation's (2013) framework. There were a few other problems with the test as well. The test seemed complicated, time-consuming and tiring for the potential participants including the English L1 students. In fact, two English L1 graduate students provided feedback that they could only provide a limited number of derivative forms. The evaluation of the test was also challenging and subjective because some items had only a few possible answers while other items had a

wide range of answers. As a result, the form of the test did not seem to be appropriate for the purpose of this study. All in all, the aforementioned methods of measuring inflectional and derivational affixes did not seem promising.

The reviewed literature in Chapter 2 showed that between derivative and inflective forms, derivations should be prioritized, and it also seemed that measuring both derivations and inflections, as already discussed, was challenging and unfeasible. Therefore, although Nation's (2013) framework required the production of both derivations and inflections, for the purpose of the current study, it was decided to measure derivations only. Furthermore, considering the fact that inflections and derivations impose different learning burdens (Schmitt & Zimmerman, 2002), as derivations are generally acquired after inflections (Berko, 1958), it seemed it was a better option to measure derivations in the hope of having a better picture of the knowledge of the word parts aspect.

An adaptation of Schmitt (1998, 1999) and especially Ishii and Schmitt (2009) could be an appropriate method to serve the purpose of measuring derivative forms productively. Schmitt's and Zimmerman's (2002) method of measuring derivations also seemed promising. However, considering the fact that at least 30 target words would be tested, the time required to complete the test and consideration of participants' fatigue were the main reasons for not using an adaptation of Schmitt's and Zimmerman's method. In other words, adapting Schmitt's and Zimmerman's method would produce a test with at least 120 sentences to measure appropriate derivative forms of 30 target words. Such a test would be too lengthy and time-consuming. However, Ishii's and Schmitt's method could serve the same purpose as long as the participants in the current study were familiar with the metalinguistic concepts of *noun*, *adjective*, *verb*, and *adverb*. Therefore, inspired by Schmitt (1998, 1999), the word

parts test was administered after ensuring that participants of the study were comfortable with the metalinguistic notions and terms such as *noun*, *verb*, *adjective*, and *adverb*.

Appendix A includes the word parts test designed for the purpose of the study. Participants were required to provide four different word classes for each of the target words. If there was more than one possibility, they only needed to provide one, and if there was no form, they needed to write X in the blank box. The following example illustrates this.

	Noun	Verb	Adjective	Adverb
NATION	nation	nationalize	national	nationally
PAY	payment	pay	payable	X
FRESH	freshness	freshen	fresh	freshly

Associations

To measure associations, Schmitt's and Meara's (1997), Schmitt's (1998, 1999), Wolter's (2001), and Webb's (2005, 2007a, 2007b, 2007c, 2009b) research studies have one factor in common. They all present target words and then ask for the words that come to mind as a response (either one word or three words). This method of measuring associations productively seemed appropriate for the purpose of the current study; as a result, a test with the following instructions was considered:

Write down three words you think of when you read the following words according to the provided examples. There are no right or wrong answers, so try not to take a long time considering your response.

For example: *When you read the word "fly", you may think of "bird; plane; high"*
When you read the word "box", you may think of "container; square; wooden"
When you read the word "fresh", you may think of "new; clean; cool"

When you read the word “already”, you may think of “now; done; just”

- FLY: bird; plane; high
- BOX: container; square; wooden
- FRESH: new; clean; cool
- ALREADY: now; done; just

The form of the test seemed to have two problems though. First, the participants might provide some collocations of the target words, and this was not the intended response type. Instead, it was planned to measure collocations as a separate aspect of vocabulary knowledge with a separate test. Second, the participants would not necessarily provide synonyms and antonyms of the target words. This was a limitation of the test because, based on Nation’s (2013) framework, it was intended to measure synonymy and antonymy as a part of measuring the associations (a test was required to be adopted for this specific purpose though). Therefore, the idea of testing associations with this form of the test failed.

To address the aforementioned problems, measuring associations in a controlled way— control of producing associations for the test— seemed promising. It was noticed that coordinates, superordinates, subordinates, collocations, synonyms, and antonyms could be potential associations for measurement (see Webb, 2005). Collocations would be measured separately as one of the aspects of vocabulary knowledge mentioned by Nation (2013). Therefore, to have a controlled way of measuring other associations, the intention was to design one test to measure synonymy and antonymy, and another test to measure other forms of associations. As a result, two tests were adopted. Firstly, the following synonym and antonym test was considered:

For the following words, write down a synonym and an antonym.

For example:

	Synonym	Antonym
FRIEND	buddy	enemy
REAL	actual	unreal
HOLD	keep	release

For the purpose of testing other associations in a controlled way, three relationships were considered: (a) Coordination: bringing the different elements (words) into a relationship, (b) Superordination: representing a superior order or category within a system of classification, (c) Subordination: representing a lower rank or position within a system of classification. Coordination was, in fact, an umbrella term for all synonymy, antonymy, collocation, superordination, and subordination. Because synonymy, antonymy, and collocation were planned for separate measures, the focus needed to be on superordination and subordination. In an attempt to measure these two aspects productively, a test with the following instructions was adopted:

For the following main categories, write down three possible subcategories (examples).

For example:

MAIN CATEGORY	EXAMPLE	EXAMPLE	EXAMPLE
COLOR	red	blue	white
ART	dance	painting	drawing
VEHICLE	car	bus	truck

Although the test seemed practical, it was measuring superordination and subordination in a linear way – superordinates were provided and only subordination was

measured. Therefore, in an attempt to measure these two aspects in a mutual way, a new test with the following instructions was adopted:

For the following words, write down a possible main category and two other examples of the same category.

For example:

	POSSIBLE CATEGORY	EXAMPLE	EXAMPLE
RED	color	blue	white
PIZZA	food	burger	spaghetti
SHIRT	clothes	hat	skirt

This version seemed capable of measuring both superordination (as the possible category) and subordination (as examples of the same category). However, words such as *adjective*, *noun*, or *verb* as the possible category of target words could be provided by participants. As a result, the instructions of the test were revised as follows: “For the following words, write down a possible main category and two other examples of the same category. Do **NOT** write *noun*, *adjective*, *verb* or *adverb* for the possible categories”. Appendix B and C include the productive tests of associations (synonym/antonym and superordination/subordination tests) developed for the purpose of the study.

Collocations

Nation’s (2013) framework for measuring collocations seems to be straightforward; the words or types of words which can be used or can commonly occur with the target words are the main target. Thus, a test was needed, possibly with the same aforementioned format (explained in Collocations in Chapter 2) of presenting a target word and asking for the collocations the participants may think of, and with clear instructions to require collocations only. A part of Schmitt’s (1999) instructions which instructed that collocations were words

that “naturally occurred together with the target word” seemed informative for this purpose; however, “occur” seemed not to be a familiar word for some EFL learners. Therefore, Schmitt’s instructions were modified and a test with the following instructions was considered:

Write three words that can naturally be used together with the following words according to the provided examples.

For example:

The word “fly” can naturally be used with “air; plane; safe” as in “planes are safe to fly through the air”.

The word “box” can naturally be used with “gear; mail; lunch” as in “gear box; mail box; lunch box”.

The word “fresh” can naturally be used with “air; water; breath” as in “fresh air; fresh water; fresh breath”.

The word “already” can naturally be used with “mentioned; covered; exist” as in “already mentioned; already covered; already exist”.

- FLY: air; plane; safe
- BOX: gear; mail; lunch
- FRESH: air; water; breath
- ALREADY: mentioned; covered; exist

This version seemed to have two main problems though. First, the participants would provide some words such as “*I, my, the, that, etc.*” as collocates. Also, the prompt sentences for the examples did not provide any more help or information and just made the instructions lengthy. Therefore, the instructions of the test were revised as follows:

Write three words that can naturally be used together with the following words. Do **NOT** write pronouns (*I, you, he, she, me, him, her, ...*), possessive adjectives (*my, your, his, her, ...*), articles (*the, a, an*), or determiners (*this, that, these, those*).

For example:

- MAKE: sure; sense; mistake
- BOX: empty; chocolate; inside
- FRESH: air; water; breath
- ALREADY: mentioned; covered; exists

This version seemed more capable of measuring collocations, but it was not clearly instructing the participants that three desired words (collocates) could be used both before and after the target words. In order to guide the participants that collocates can be used both before and after the target words, the instructions of the test were revised as follows: “Write three words that can naturally be used together (before or after) with the following words. Do **NOT** write pronouns (*I, you, me, him, there, yourself...*), prepositions (*up, in, on, ...*), possessive adjectives (*my, your, his, her, ...*), articles (*the, a, an*), or determiners (*this, that, these, those*)”.

This version of the test still seemed to have a few problems. The facts that three collocates were required, and collocates could be used before as well as after the target words would be ignored. Furthermore, words such as “*up, in, and, can, ...*”, might still be provided as collocates. Therefore, the test instructions were revised. The number of needed collocates (three) was emphasized; it was emphasized that only content words were accepted; and in an attempt to make a more understandable version of the test, the target words were presented in a central column with possible collocates which could go before and after as follows:

Write **THREE** words that can naturally be used together (before or after) with the following words. **Content words (nouns, verbs, adjective and adverbs) ONLY are accepted.** Do **NOT** write pronouns (*I, you, me, him, there, yourself...*), prepositions (*up, out, for, in ...*), possessive adjectives (*my, your, his, her, ...*), articles (*the, a, an*), determiners (*this, that, these, those*), conjunctions (*and, or, but, ...*), auxiliary verbs (*can, could, will, would, am, is, are ...*) or negatives (*not, don't, ...*). Please write your answers as clearly as possible.

Example:

			FRESH	air	fruit	breath
		hot	COFFEE	beans	table	
	fully	gladly	ACCEPT	invitation		
live	left	completely	ALONE			

The last version seemed to have the potential to measure collocations for the purpose of the study and was adopted for use in the present research. Appendix D includes the collocation test designed for this study.

Form and Meaning

To measure form and meaning aspect productively, Nation's (2013) framework requires the translation of words into L2. However, Laufer and Goldstein (2004) assert that in multilingual classes, or in situations when the researcher is not familiar with the participants' L1, an L2 paraphrase can be required to demonstrate passive knowledge (receptive knowledge), and an L2 target word for a synonym or a paraphrase can be required to prove active knowledge (productive knowledge).

Providing synonyms of the target words was not a promising idea because a synonym and antonym test had already been designed. Therefore, it seemed the paraphrases (definitions) of the target words were to be provided. Inspired by Schmitt (1998), the online Longman Dictionary of Contemporary English, Cambridge Dictionary, Collins Cobuild Dictionary, Oxford Learners' Dictionary, and Merriam-Webster dictionary were consulted to determine the major meaning sense of each target word. If the dictionaries disagreed on a major meaning, decisions were made based on the corpus data – the first 100 BNC or COCA¹⁰ corpus sentences were consulted to choose the most frequent meaning sense of each target word. Comparing the major meaning definitions in different dictionaries, the definitions (of the major meaning senses) which looked the easiest and the most straightforward were chosen for the target words. Doing this, the definitions of all target words were provided and presented, followed by the same number of dashes as the number

¹⁰ COCA: Corpus of Contemporary American English

of letters of each target word to help guide the task. Furthermore, unaware of the legitimate arguments against providing the first letter of words in productive tasks (see Barcroft & Goldstein, 2010; Read, 2000; Webb, 2008) at that time, and inspired by Laufer and Goldstein (2004) and Laufer and Nation (1999), the first letter of each word was also provided to help guide the task. The participants were required to provide the English words for those definitions as follows:

Write down a suitable word for the following definitions. The first letter of each word is provided, and each provided line is one letter. For example, the word “*answer*” starts with “a” and has six lines as follows: a _ _ _ _ _

<i>For example:</i>	<i>to travel by plane or move through the air:</i>	<u>f</u> _ _	<i>answers:</i> <i>fly</i>
	<i>a container for putting things in:</i>	<u>b</u> _ _	<i>box</i>
	<i>recently picked, made or cooked:</i>	<u>f</u> _ _ _ _	<i>fresh</i>
	<i>before now or before a particular time:</i>	<u>a</u> _ _ _ _ _	<i>already</i>

The presence of the first letter of each target word, as discussed, seemed to change the task from a productive to a receptive recall though (see Laufer & Goldstein, 2004). Learners tend to remember the initial or ending letters more than the middle (Barcroft & Rott, 2010), and thus, providing the initial or final letter may have an effect on the productive nature of the task. In other words, the provided first letter could receptively remind the participants of the target word. Also, some of the definitions of the words seemed lengthy and time-consuming.

To address these two concerns, the provided first letter of each target word was removed, and the definitions were shortened. The definitions were revised to be as short but as clear as possible. However, for a few items, synonyms with the same number of letters could be provided. For example, “noise” could be provided as a response for “something that

you hear” while “sound” was the target word, or “weird” could be provided for “very strange, foolish or mentally ill” while “crazy” was the intended target word. To help guide the task and to prevent the participants from providing synonyms for the target words, inspired by Nakata (2013), one letter in the target word (neither the beginning nor the ending, but a letter in between) was given. The following criteria were used to provide that single letter of each target word:

- The most frequent vowels were to be used as cues. Five letters of “a, e, o, u, i” represent English vowel sounds. The decision-making process concerning the selection of the vowel as the cue was as follows: Using English letter frequency table, the most frequent vowel in each target word was identified and then it was provided as the cue (the provided letter). However, there were some exceptions. For example, for the word REMOVE, instead of selecting “e”, which was more frequent than “o”, “o” was selected as the cue, because REMOVE and DELETE shared the same definition and using “e” as the cue could elicit either word. In the same vein, “u” in QUICKLY was provided instead of “i”, which was more frequent, because the latter could elicit either QUICKLY or BRISKLY.
- For three-letter target words, the middle letter, irrespective of being a vowel or a consonant, was provided to keep the consistency and face validity of the test.
- If two vowels with the same letters were repeated in a word, such as “o” in DOCTOR, the second one was considered the cue.
- “Y” in the word TYPE was provided as the cue because it represented the vowel sound and was more frequent than “P”.

Addressing the aforementioned two concerns, the following instructions were provided:

Write down a suitable word for the following definitions. One letter of the desired word is provided and each line represents a letter.

For example:	To produce something:	__ <u>a</u> __ __	Answers:	make
	Recently picked, made or cooked:	__ __ <u>e</u> __ __		fresh
	Before now or before a particular time:	__ __ __ <u>e</u> __ __ __		already

At this point, although the definitions of the words and the task seemed as clear and simple as possible, it seemed the test was measuring more than the connection of the form and meaning. On the one hand, it was not intended to have another synonym or association test. On the other hand, it was noticed that the definitions potentially included synonyms, antonyms, or some related words, which could be associated with the target words. In fact, this format of measuring the form and meaning aspect was vulnerable to including synonyms, antonyms, or related words in provided definitions, and this could negatively affect the results of the association tests. For instance, in the aforementioned examples of the test instructions, it could be observed that “produce” and “make”, “recent” and “fresh”, and “before now” and “already” were synonyms, or at least they were related. Unfortunately, it was the same for a considerable number of other items. This could negatively affect the association tests if the form and meaning tests were to be administered first. In the case of administering the form and meaning tests after the association tests, the participants would have been exposed to the target words first, and this could devalue the results of the form and meaning tests. Therefore, that method was not necessarily the best way of measuring productively the connection of form and meaning.

Nation’s (2013) instructions for measuring the connection of form and meaning productively (word translation) seemed to remain as the best possible choice. Considering the potential participant pool for the current study, Persian to English word translation tests were

to be provided. In fact, four form and meaning tests were needed for the target words of the word parts, synonym and antonym, superordination and subordination, and collocation tests respectively.

To design the form and meaning tests, the Persian translations of the core meaning (usually the first entry in dictionaries) of each target word were provided. For this purpose, two English to Persian dictionaries, Farhang Moaser Hezareh (2003) and the online Aryanpour dictionary, were consulted. In other words, I compared what I knew as the translation of the core meaning of words with the translations provided by the aforementioned dictionaries. If there was an inconsistency between what I believed was the core meaning and what dictionaries showed as the core meanings, the COCA was consulted to judge which meaning was the core meaning. For this purpose, the first 100 examples of the corpus were consulted to select the most frequent meaning of the target word. For example, I believed that the core meaning of ASSOCIATION was “organization”. Farhang Moaser Hezareh (2003) showed “coalition” as the first entry for ASSOCIATION. Because ASSOCIATION had the meaning of “organization” in almost all of the first 100 COCA sentences, the translation of “organization” for this item was provided in the test.

After preparing the first draft, the test was checked with one of my supervisors. She is a Persian as a first language applied linguist and provided one high frequency Persian synonym for the translation of one item in the superordination and subordination form-meaning test, 6 for the synonym and antonym, 2 for the collocation, and 4 for the word parts form and meaning tests. These extra Persian words were close to the meaning of the provided translations, and seemed to be very helpful because, based on previous background knowledge, language learners may be more used to a specific translation. For example, in Persian, there are two translations for QUICKLY: “به سرعت” and “تند”. Both of them have

the same meaning; however, based on previous background knowledge, one may be more familiar for some language learners. In other words, students use different dictionaries or translators. When they are provided with a different translation, they may not be able to figure out the correct target word. Therefore, to help guide the task, those extra Persian translations were added to the items.

The test was given to a Persian as a first language (Persian L1) graduate student in Canada, and he was asked to provide the target words. The provided translation for item 8 in the word parts form-meaning test, TRAINING, was misleading as TEACHING had also the very same translation with the same number of letters and position of the cue. To help guide the task, one more translation was added (“کارآموزی” in addition to “آموزش”), and the cue changed—“i” for TRAINING was provided instead of the formerly provided “a” which could have prompted a test-taker to provide “teaching”.

In the same test, the cue of item 15, “a” for TRADITIONAL, needed to change because “conventional” had the same number of letters with the same position of the cue. Therefore, the first “a” was provided as the cue instead of the second one.

PRIVATE, FINAL, DISCOVER, REMOVE, USEFUL in the synonym and antonym form-meaning test, SITUATION, FORMER, PREVENT in the collocation, and ASSOCIATION, CONTINUE, CREATE, SIGNIFICANT, DETERMINE in the word parts were not provided by the Persian L1 graduate student. “Teaching” and “conventional” were also provided instead of TRAINING and TRADITIONAL.

The target words which were not provided were shown to the Persian L1 graduate student. He believed he knew them all. He was asked to provide his own translation of each word. His translations were compared with my translations and the translations of the aforementioned two dictionaries to ensure that the provided translations were appropriate.

One more Persian translation was added to the previously provided translations of PRIVATE, FINAL, REMOVE, SITUATION, PREVENT, ASSOCIATION, and DETERMINE to help guide the task.

The new versions of the tests were given to another Persian L1 graduate student in Canada. ASSOCIATION, TRAINING in the word parts, REMOVE, PREVIOUS in the synonym and antonym, and FORMER and PHONE in the collocation form-meaning tests were not provided. The student believed “60 minutes” needed to be added to the translation of HOUR because the provided translation may remind someone of “clock” or “watch”. The target words which were not provided were shown to her. She believed she was familiar with them all, and she asserted that at that time those words did not come to her mind. She was asked whether she agreed with the provided translations. She confirmed that those translations were appropriate and did not have any other better translations for those words to be replaced with the original ones.

The tests were shared with another Persian L1 graduate student. TRAINING, SIGNIFICANT, EFFORT in the word parts, ACTUALLY in the collocation, PRIVATE and WELL in the synonym and antonym were not provided. When he was provided with the words, he believed he knew those words but could not connect the meaning to the word forms at that time. He believed all items were well-translated.

The tests were finally given to the last Persian L1 graduate student with high proficiency in English, and he could provide all the target words. He confirmed that the words were translated appropriately. He pointed out that for the translation of the word REDUCE, “کم کردن”, the word DEDUCT could also be provided. It was not possible to change the cue for this item because REDUCE and DEDUCT had the same number of letters and also had the identical letters in between the first and last. Therefore, to help guide the

task, one more Persian translation, “کاهش دادن”, was added to the item as DEDUCT could not carry the latter meaning in Persian.

A few other synonymy varieties in the form-meaning tests were also observed later on. In the collocation form-meaning test, for item 10 (انتخاب: __ __ __ i __ __), both OPTION and CHOICE could be provided. For item 20 (فروختن: __ e __ __), both VEND and SELL could be provided, and for item 25 (به __ __ __ __ __ __ __ __ a __ __ __ __ به ویژه/بخصوص), both SPECIFICALLY and PARTICULARLY could be provided. Therefore, the cue letter “i” in CHOICE needed to change to “o”, the cue letter “e” in SELL needed to change to “l”, and the first “a” in PARTICULARLY was selected as the cue letter instead of the second one.

In the superordination and subordination form-meaning test, for item 10 (__ o __ __ پسر), both BOY and SON could be provided. Yet, because the target word had just 3 letters, no change was possible and both BOY and SON were counted as correct answers.

In the word parts form-meaning test, for item 1 (__ __ __ __ __ __ a __ __ __ __ سازمان/انجمن), both CORPORATION and ASSOCIATION could be provided. Therefore, the cue letter ‘a’ needed to change to “o”. See Appendix E for a copy of the final tests.

Target Words

The Words at the 1,000 Word Frequency Level

To extract the first 1,000 words from the new-GSL (Brezina & Gablasova, 2013), all words with ranks more than 1,000 were removed from the list. Function words needed to be removed as well because they have little lexical meaning or may have even ambiguous meaning. Nation’s (2013) function word list was removed first. This list of function words includes 161 word families and is made up of a list of auxiliary verbs, prepositions and

conjunctions, determiners and pronouns, and numbers. In addition to this list, the following function words and abbreviations, which were not available in Nation's (2013) function word list, were also removed for the same aforementioned reason:

1. A LOT/LOTS (OF) (x, 369)
2. ACCORDING TO (con, 410)
3. ACROSS (con, 372)
4. AGAINST (con, 162)
5. anyone (pron, 558)
6. away (avp, 908)
7. BECAUSE (con, 96)
8. BEHIND (con, 392)
9. despite (con, 679)
10. due (to) (adj, con, 524)
11. half (x, 746)
12. LIKE (con, 84)
13. MR (abr, 312)
14. Mrs (abr, 868)
15. onto (con, 970)
16. PER (con, 323)
17. SEVERAL (x, 356)
18. throughout (con, 705)
19. UPON (con, 429)
20. via (con, 899)
21. whatever (x, adv, 748)
22. WHETHER (con, 248)

In the new-GSL, “avp” stands for “adverbial particle in phrasal verbs”, “con” stands for “preposition or conjunction”, “pron” stands for “pronoun”, “x” stands for “determiner, quantifier or particle”, and “abr” stands for “abbreviation”. Following removal of the aforementioned words, the new-GSL came down to 779 words (Appendix F). This list served as the list of the most frequent words for the purpose of the current research study.

Criteria for Selection of the Target Words

The question posed at the beginning of the study was as follows: To what extent do EAP students have productive depth of vocabulary knowledge of the most frequent 1,000 words? To address this research question specifically, target words needed to be selected

randomly from the most frequent 1,000 words to represent that word level statistically. However, random selection of the words would result in a limited performance on word parts, synonym and antonym, and superordination and subordination tests because the target words selected randomly would have a limited number of derivative forms, synonyms and antonyms, and also possible superordinates and subordinates. As a result, the idea of random selection of the target words failed, and the research question was slightly revised to address this problem as follows: To what extent do EAP students have productive depth of vocabulary knowledge of words at the 1,000 word frequency level?

With regard to the target words, it was necessary to know what group of words would address the research questions (the extent of knowledge for four aspects), and could consequently work well for the study tests. As a result, it was necessary to find out whether one group of target words or more were needed, and also whether those words were to be selected randomly or purposefully.

Reviewing the selection of target words of the aforementioned studies in Chapter 2 seemed to support the idea that it was possible to choose purposefully or randomly different sets of target words to specifically address the different research questions. Word parts, associations, collocations, and form and meaning were measured in the study; as a result, it was determined that the following groups of words were needed: (a) a group of target words with various derivative forms (a variety of possible affixes) for the word parts test, (b) a group of target words with transparent synonyms and antonyms, and a group of target words to be categorized as superordinates and subordinates for the association tests; and (c) a group of words with common collocates for the collocation test. The translation of each of these groups of words into Persian was also needed to have one corresponding form and meaning

test for each of the aforementioned tests. In sum, four groups of target words altogether were needed, but the number of target words in each group was a matter of concern.

It was inferred, from Nation's (1983) and Schmitt's, Schmitt's and Clapham's (2001) experience with the number of items in the Vocabulary Levels Test, that 30 words could be statistically representative of a 1,000 word level. In fact, in Schmitt's, Schmitt's and Clapham's study, the Cronbach's alpha reliability figures (for dichotomously scored items) suggested that 10 clusters (30 items) per level would produce reliability figures above .90. However, 34 words were selected for each test because at least 30 would remain even if pilot studies resulted in removing a few items. In sum, it was concluded that four groups of 34 target words would be chosen. The following explains the criteria for choosing each group in detail.

Target Words for the Word Parts Test

To select the appropriate target words for the word parts test, two main criteria for each target word were considered: (a) the availability of four different word classes, and (b) the availability of high-frequency derivative forms. The reason for this selection was that the word parts test, as explained before, was to examine four different word classes; as a result, items with 4 word classes could provide better face validity for the test. Also, at that time, it was believed the derivative forms needed to be high-frequency to have the chance of being known and produced by the participants.

Words with 4 possible word classes were selected from the aforementioned 779 most frequent words first. To do this, Nation's 1K base word list, published on his website, and the aforementioned online dictionaries were consulted. If the presence of a possible derivative form was still in doubt, the BNC or COCA were consulted to see whether that form was

available. The list came down to 174 words. Then, words with the highest number of derivative forms were selected from this list because derivations were the main purpose of the word parts test. For this purpose, Nation's 1K base word list was mainly used and 57 words were selected from the aforementioned 174 words. Next, from the list of 57 words, the words for which I (as an English high-proficiency testee) could easily provide all 4 word classes were selected. In fact, I considered myself the first testee of the test. The list was reduced to 40 words. However, some of the remaining words had identical word classes. For example, "employ" had identical noun and verb word classes. This was not beneficial for the test because one single derivative form could occupy two different word class boxes in the word parts test. As a result, fewer derivative forms could be produced by the participants, and less depth of vocabulary knowledge of this aspect could be revealed. The second problem was that derivative forms are usually low in frequency. As a result, the first attempt to select appropriate target words for the word parts test was a failure. Thus, the criteria were revised as follows:

- a. The target words were to have no identical word classes. For example, "develop" is appropriate because it is a verb and does not have any identical word classes.
- b. The target words were to have at least three or four word classes. The participants would get credit for mentioning a missing word class (either noun, verb, adjective, or adverb form) in the appropriate box if the target word had just 3 word classes.

Such criteria did not allow balance among the four word classes – having the same number of nouns, verbs, adjectives, and adverbs. However, this did not seem to be a problem

because measuring derivatives (for the purpose of affix production) was intended, not measuring grammatical functions.

Putting the aforementioned criteria into effect, the 779-word list was reduced to 190 words. Therefore, there were 190 words with 3 or 4 possible derivative forms (word classes), which did not have items with identical word classes. Finally, these 190 words were randomized within an Excel spreadsheet and the first 34 words were selected as the target words. Appendix A shows the words included in the test.

Target Words for the Synonym and Antonym Test

For the synonym and antonym test, a group of target words with transparent synonyms and antonyms was needed. To serve this purpose, the following criteria were used:

- a. The target words were to have available synonyms and antonyms in online contemporary English Thesaurus dictionaries – online Thesaurus Dictionary, Merriam-Webster Thesaurus, and Collins Cobuild English Thesaurus.
- b. The possible synonyms and antonyms needed to be transparent, e.g. “large” and “small” as a synonym and an antonym for “big”, or “tall” and “short” as a synonym and an antonym for “long” seem transparent. However, “bill” as an antonym for “coin”, “office” as an antonym for “house”, or “decent” and “happy” as antonyms for “blue” do not seem to be transparent.
- c. The major meaning sense of each target word was to be considered for synonymy and antonymy. For this purpose, the online Longman Dictionary of Contemporary English, Cambridge Dictionary, Collins Cobuild Dictionary, Oxford Learners’ Dictionary, and Merriam-Webster dictionary were consulted to determine the

major meaning sense of each target word. If the dictionaries disagreed on a major meaning, decisions were made based on the corpus data – the first 100 BNC or COCA corpus sentences for the target word were consulted to choose the most frequent meaning sense of that specific target word.

To put the criteria into effect, first, by consulting the aforementioned online English Thesaurus dictionaries, words with possible transparent synonyms and antonyms were selected from the 779-word list. The list was reduced to 315 words. Then, these 315 words were randomized within an Excel spreadsheet and the first 34 words were selected.

A few words in these 34 randomly chosen words had equivalent high-frequency word classes—for example, “present”, as an adjective with a rank of 560, as a verb with a rank of 559, and as a noun with a rank of 1,789 in the new-GSL. Another example was “note” as a noun with a rank of 731 and as a verb with a rank of 730 in the new-GSL. Thus, it was highly possible that the participants might get confused about the word class of the target word for which they were required to provide synonyms or antonyms. Therefore, such items were replaced with the items with one high-frequency word class. For example, “present” was replaced with “necessary” as an adjective, and “note” was replaced with “announce” as a verb. Unlike the selection of the target words for the word parts test, the target words for this test were not selected based on having a single part of speech (not having equivalent word classes). The reason was that putting such a criterion into effect, words such as “big, short, small, good, bad, etc.”, with equivalent word classes, would have been removed from the potential selection list.

The target words might have other low-frequency word classes; however, those low-frequency entries did not seem to be problematic. For example, the verb form of “push” and

“pick” is much more frequent than their noun form, and it was unlikely that the participants would provide synonyms or antonyms for their noun forms (low-frequency word class).

However, if they did, those answers were accepted.

Last but not least, if the meaning of the items overlapped, such as “modern” and “recent” or “huge” and “big”, one of them was removed and another item with the same part of speech was selected. Appendix B includes the target words in the synonym and antonym test.

Target Words for the Superordination and Subordination Test

For the superordination and subordination test, 34 target words were required which could fall into a main category (superordinate) with examples (subordinates) within the same category. For this purpose, regardless of any criterion, the 779-word list was investigated to find all potential and possible relationships. An attempt was made to categorize all related words, and the following possible categories were made:

1. Bank account: amount; benefit; bill; money; number; percent; property; rate
2. Advice: agreement; council; discussion
3. Month: April; May
4. Area: land; location; region; space
5. Body part: arm; back; body; eye; face; foot; hair; hand; head; heart; leg
6. Education: book; course; research; result; student; study; teacher; term ; test; training; examine; school; learn; teach
7. Group: association; committee; community; meeting; member; team; society
8. Family: brother; wife; husband; daughter; father; mother; parent; son
9. **Furniture**: bed; seat; table
10. **Container**: box
11. Human: boy; child; guy; kid; man; people; person; population; user; baby
12. Job: leader; manager; officer; police; staff; worker; doctor
13. Relationship: couple; friend; partner
14. Career: work; business; duty; market, project; skill; task
15. **Entertainment**: game; film; music; party; play
16. Time: century; date; day; evening; hour; minute; moment; morning; night; period; summer; today; week; year

17. Place: church; city; class; club; country; department; field; home; hospital; school; station; town; village; court; company; office; building
18. Information: data; knowledge; news; technology
19. Cost: price; sale; money
20. Measure: degree; size; weight; scale
21. House: door; floor; room; wall; window
22. Energy: power; pressure; strength
23. Nature: environment; sea; tree; water; cloud
24. Subject: history
25. **Communication**: letter; contact; call; phone
26. Distance: mile; road; street; wide; narrow
27. Direction: north; south; west
28. Material: glass; stone; paper
29. Type: model; style; form; pattern; version
30. Feeling: happy; fear; love; sense
31. Size: big; small; short; little; huge
32. Technology: computer; website; internet; data
33. Art: music
34. Weight: heavy; light
35. Quality: poor; good; fine; bad
36. Activity: game
37. Department: education; history
38. War: fight; kill
39. Garden: tree; water
40. Medical: health; hospital; doctor
41. **Season**: summer
42. Country: city; town; village
43. Trade: buy; sell; money; business
44. Test: question; answer; example

All of the above words were in the 779-word list except five of them in bold for main categories. The reason was that some words were potentially related, but there was no suitable word in the 779-word list to name a category for them. For example, “game, film, music, party, play” were all in the 779-word list and seemed to be related, but there was no word in the list to name a category for them. Therefore, I named the category “entertainment” while the word (entertainment) was not in the list. However, the way the words were categorized was subjective; other researchers might categorize them differently. Making these categories, in fact, helped me recognize potential target words.

I consulted my supervisors about these categories, and 49 potential words became candidates for the test. The words (49 items) were listed in the superordination and subordination test, and the test was given to four English L1 graduate students. Their performance on each item was compared. Thirty-four items which had consistent replies to the main category section (superordination relationship) were chosen as the final target words of the test. Appendix C includes the target words for the superordination and subordination test.

Finally yet most importantly, it should be noted that each item could have several possible answers. For example, for TREE, in addition to “plant” (main category), and “grass” and “shrub” (examples of the same category), some other answers such as “garden, water, and flower” were also acceptable. This issue would remain valid for all other target words. That is, “April” could remind someone of “spring” as its possible category and “flower” and “rain” as examples in the same category in addition to answers such as “month, January, and February”. Therefore, participants were not expected to provide specific answers in the test, and a reasonable relationship (superordination and subordination) was considered correct (more detail has been provided in the scoring section).

Target Words for the Collocation Test

To select a group of target words for the collocation test, using the COCA/BNC Collocates and the Web1T5, it was noticed that any word could have a variety of frequent collocates. In other words, it seemed it was not needed to select the words purposefully for this test. Thus, the 779-word list was randomized within an Excel spreadsheet and the first 34 words were selected. Items were double-checked in the COCA/BNC Collocates and the Web1T5 to make sure there were frequent collocates for them.

In the first 34 words, there were some items (target words), such as “hope” and “visit”, which had high-frequency equivalent word classes. Such words could cause complications. In other words, “hope” as a noun with a rank of 920 and as a verb with a rank of 376, and “visit” as a noun with a rank of 950 and as a verb with a rank of 709 were available in the new-GSL. Such target words with equivalent high-frequency word classes could confuse the participants – whether they needed to provide collocates for the noun form or the verb form. These words were removed from the list and other items following (right after) the first 34 words with one single high-frequency word class were replaced.

There were a few other items in the randomized list such as “sell, water, scale” with low-frequency equivalent word classes. However, only their high-frequency word class, “sell” as a verb or “water” and “scale” as nouns, were in the new-GSL. Their other equivalent word classes, “sell” as a noun or “water” and “scale” as verbs, were so low frequency that it was unlikely that participants would produce collocates for them. However, even if the participants produced collocates for their low-frequency word classes, those collocates were accepted.

All in all, although the words were randomly chosen, the majority of them did not have equivalent word classes. There were a few with two possible equivalent word classes, but as mentioned before, their other possible equivalent word classes were so low frequency that they did not seem to cause a problem. Appendix D includes the target words in the collocation test.

Target Words for Form and Meaning Tests

The selected target words for the word parts, synonym and antonym, superordination and subordination, and collocation tests made up the target words for four form and meaning

tests. Appendix E includes the translation of these words into Persian in the form and meaning tests.

Scoring the Tests

Scoring of the Word Parts Test

The scoring approach for the word parts test is mainly inspired by three studies: Schmitt (1998), Schmitt and Zimmerman (2002), and Ishii and Schmitt (2009). Based on the aforementioned studies, a norming list for the purpose of scoring the word parts test was created. In order to compile the answer norms in a principled manner, four sources of information were drawn on. First, Nation's 1K base word list, published on his website, was consulted and all listed derivatives were extracted. Second, three online dictionaries (Longman, Cambridge, and Collins Cobuild dictionaries of English) were consulted and all other possible derivatives were added. Third, the presence of these derivatives in the BNC and COCA was double checked. If a form was not available in the corpora, it was removed. The frequency of these derivatives in the corpora was also checked. If a word class (especially adverbs) had low-frequency counts, both the derivative form and the answer "no form exists (X)" were accepted. For example, "manageably" as an adverb for MANAGER has a frequency count of 10 in the COCA and 5 in the BNC. Thus, for the adverb form of MANAGER, both "manageably" and the answer "no form exists", indicated as X, were accepted. Fourth, the test was given to 20 English L1 graduate students. If a new form, which was not in the norming list, was encountered, its availability and frequency in the BNC and COCA were checked. If it had a high frequency count, it was added to the norming list. If it had a low-frequency count, both that specific form and the answer "no form exists (X)" were accepted and added to the norming list. If the form was not in the corpus at all, it was not

accepted. For example, an English L1 graduate student stated that he could see himself using “managerially” as an adverb for MANAGER in a context. The frequency of “managerially” in the COCA and BNC was checked and it was found it had low-frequency counts of 10 and 5 respectively. Therefore, “managerially” or “manageably” as derivative forms of MANAGER or “no form exists (X)” were added to the norming list. If “managerially” had not been in the COCA or BNC, it would not have been accepted.

Based on such an approach, the following answers as four word classes for the word MANAGER were accepted:

	Noun	Verb	Adjective	Adverb
1. MANAGER	manager; management; mismanagement; manageress	manage; mismanage	manageable; unmanageable; managed; managerial; managing; unmanaged	manageably; X; managerially

However, a form such as “managery”, although provided by one English L1 graduate student, is not in the norming list because it does not exist neither in the BNC nor in the COCA.

Based on the norming list, the possible scores ranged from 0 (knowledge for no word class) up to 4 (knowledge for all four word classes) for each item. Misspellings were accepted (although they were not frequent) if the intended derivatives could be discerned, because the test measured knowledge of the derivations not spellings. In fact, the few encountered misspellings (e.g. ‘dicion’ instead of ‘decision’ or ‘managable’ instead of ‘manageable’) were clearly showing the intended derivatives; therefore, a second rater was not needed to be employed to rate misspellings. In addition, considering the fact that the original intent of the study was measuring word parts not grammatical functions, wrongly

positioned derivatives were also accepted. For example, “manageable” was mistakenly considered an adverb form by a participant, yet it was accepted as a correct answer.

Scoring of the Synonym and Antonym Test

The approach to scoring the word parts test inspired the study and a norming list for the purpose of scoring the synonym and antonym test was also prepared. In order to compile the answer norms in a principled manner, two sources of information were drawn on. First, three online thesaurus dictionaries (Merriam-Webster, Collins Cobuild, and Treasures.com) were consulted and all available synonyms and antonyms of the target words were extracted. Second, the test was given to 20 English L1 graduate students, and if two of them or more provided synonyms or antonyms which were not in the norming list, the presence of those words were checked in the BNC and COCA. On the condition that the provided words were in the corpora, they were added to the list. In addition, all possible multi-word synonyms and antonyms, which were either available in thesaurus dictionaries or provided by English L1 graduate students (on the condition that they were in the corpora), were accepted and added to the norming list. For example, in response to REMOVE, a multi-word synonym such as “take away” and a multi-word antonym such as “put back” were accepted and added to the norming list.

A variety of synonyms and antonyms were possible for any item, but only one synonym and one antonym were required for full credit. A score for synonyms, a score for antonyms and a total score for the whole test were considered – the score of each item was either 0 (knowledge of no synonymy or antonymy), 1 (knowledge of synonymy or antonymy), or 2 (knowledge of both synonymy and antonymy). Misspellings were accepted if the intended synonym or antonym could be discerned, because the test was to measure

knowledge of synonymy and antonymy, not spellings. Therefore, a second rater was not engaged to rate misspellings.

Based on this approach, as an example, the following table shows the possible answers for the word ABILITY:

	Synonym	Antonym
1. ABILITY	capability; competence; skill; talent; aptitude; capacity; capableness, competency, faculty; aptitude, expertise, proficiency, talent; intelligence; qualification; strength; understanding; ingenuity	inability; incapability; weakness; disability; incapableness, incapacity, incompetence, incompetency, ineptitude, ineptness, ignorance; impotence; incapacity; lack

Scoring of the Superordination and Subordination Test

A considerable number of studies on L2 word association (e.g. Politzer, 1978; Read, 1993; Soderman, 1993; Sokman, 1993; Schmitt, 1998; Wolter, 2002; Zivian, 1972) evaluated learners' responses using native speaker norms. The rationale, or maybe the assumption, for this is that as proficiency increases, word association performance becomes more like that of an adult native speaker. However, recent studies (e.g. Fitzpatrick, 2007; Zareva & Wolter, 2012) challenge the validity of such an assumption. Fitzpatrick's findings show that not only do native speakers vary in the actual words they produce, they also seem to vary in the types of association they make. The variety of different and inconsistent answers provided for the target words in the pilot study, which will be discussed in detail in the upcoming Pilot Studies section, confirm the findings of the recent studies. Therefore, instead of preparing a norming list, inspired by Webb (2005), a common-sense approach was used to score this test. However, two raters (in addition to myself) were employed to reduce the risk of subjective ratings. The raters were English L1 Ph.D. candidates in Education. They were trained for this

purpose. Inter-rater reliability was calculated to see to what extent the raters' judgment was consistent.

The raters (the employed raters and I) accepted all those answers which had a reasonable superordinate and subordinate relationships with the target words. For example, for HAPPY, any answer which could be categorized as "emotion, mood, feeling, etc.", such as "sad, angry, depressed, delighted, excited", were credited. Likewise, for TREE, any answer which could be categorized as "plants" such as "flower, shrubs, weed" and even fruits' name "mango, pear" were credited. In other words, credit was given to answers as long as a reasonable connection could be considered between the target word and the answers. Thus, "garden" in addition to "plant" was also accepted as a possible superordinate for TREE, and "water" and "flower", in addition to "shrubs" and "weed", were also accepted as TREE's subordinates. In other words, participants were not expected to provide necessarily some specific answers, and all reasonable connections between the words received full credit.

Based on this approach, the following answers for the word TREE, as an example, were accepted:

	ACCEPTABLE SUPERORDINATE	EXAMPLES OF ACCEPTABLE SUBORDINATES	EXAMPLES OF ACCEPTABLE SUBORDINATES
1. TREE	plant/garden	mango, flower, shrubs, pear, lotus, crop, grass	banana, apple, bush, weed, poison ivy, eggplant, water, rose

The possible scores for each item ranged from 0 (knowledge of no superordinate and subordinate) up to 3 (knowledge of both superordinate and subordinates). The few encountered misspellings were accepted because the intended superordinates or subordinates

could be discerned, and the study was measuring the knowledge of such relationships, not spellings.

Scoring of the Collocation Test

Preparing a norming list for the collocation test did not seem promising. I asked 4 English L1 graduate students to provide collocates for the items of the test. The results showed inconsistent responses to a considerable number of items. For example, for BLOOD, one English L1 graduate student provided “tainted, bath, line”, another one provided “red, thinner, brother”, while two others provided “money, moon, donation” and “type, orange, clot” respectively. For the same reason, making a norming list (generally) based on the collocation dictionaries (specifically) did not seem promising because it seemed the study would not have consistent answers by participants. Also, it was not clear what definition of collocation such dictionaries had provided and categorized collocates for words. Webb and Kagimoto (2011) state that a universal definition of collocation has yet to be reached, and they define collocation from a statistical standpoint for the purpose of their research study. Following Webb, Newton and Chang (2013) and Webb and Kagimoto (2011), collocation in the current study is defined from a statistical standpoint and refers to the regular co-occurrence of words within a given span demonstrating a statistical strength of co-occurrence.

The collocates provided by English L1 graduate students showed that the Web1T5 tends to include more of their answers in comparison to the COCA/BNC Collocates. Almost all of the provided English L1 graduate students’ collocates, which were available in the COCA/BNC Collocates, were also available in the Web1T5, yet there were some other collocates (provided by the same students) which were not recognized by the COCA/BNC

Collocates, but were available in the Web1T5. For example, while the collocate “donation” for BLOOD is both in the COCA/BNC Collocates and the Web1T5, the collocate “moon” just exists in the Web1T5. Also, while the collocate “peace” for BRING is available in both databases, the collocate “luck” is just available in the Web1T5. Three English L1 graduate students agreed that “moon” is a collocate for BLOOD and “luck” is a collocate for BRING.

As a result, the Web1T5 in addition to COCA/BNC Collocates was considered the norm, a span of ± 4 words in both was set, and also a display of the first 1,000 collocates (the maximum amount proved to be practical) with frequency of equal or more than 40 (the minimum amount provided by the corpus interface tool for the Web1T5) and 20 (the appropriate amount for the COCA and BNC Collocates, see Durrant & Schmitt, 2009; Laufer & Waldman, 2011) were set. All provided collocates of the target words with a *t*-score of 2 or above or a *MI* score of 3 or above were accepted.

In addition to the acceptance of all provided collocates with the aforementioned *t*-score and *MI* score, the collocations with frequency of co-occurrence equal to or greater than 37 (discussed in the upcoming Pilot Studies section) were also credited. The reason for this was that there were some collocations such as good YEAR, FORMER job, big HEART, or big EVENT that had lower *t*-scores and *MI* scores of 2 and 3. However, such collocations were provided by at least 3 English L1 graduate students in the pilot study. To include all potential and possible collocations, the frequency of co-occurrence in COCA was calculated as an additional scoring criterion. However, because frequency of co-occurrence was arbitrary, all inconsistent collocations with the first criterion (*t*-score and *MI* score), provided by the participants in the main study, were also given to two English L1 Ph.D. candidates in education to be evaluated whether they could be considered potential collocations or not. The raters were trained for this purpose.

Inter-rater reliability between the raters, and also between the results obtained by the raters and the results obtained by the frequency of co-occurrence were calculated to investigate the extent to which the ratings were consistent. Their evaluations were so highly correlated with each other and with the second criterion (frequency of co-occurrence of 37 or more) that it did not seem necessary to employ the third rater. In fact, there was a strong positive correlation between the first and the second rater, $r(46) = .991$, $p < .001$, the first rater and the frequency of co-occurrence evaluation, $r(46) = .943$, $p < .001$, and the second rater and the frequency of co-occurrence evaluation, $r(46) = .937$, $p < .001$. All in all, the possible scores for each item (target word) ranged from 0 (knowledge of no collocates) up to 3 (knowledge of all three collocates).

Pilot Studies

The study received ethics approval on April 29, 2016. Ten English L1 graduate students were invited to take part in the pilot study. The instructions of the tests were explained to them, and examples were reviewed. If they had any difficulty understanding the task, a clarification was provided. Care was taken to make sure they were comfortable with the metalinguistic notions and terms such as *noun*, *verb*, *adjective*, *adverb*, and *collocation*. They were advised that their comments on the tests or concerns about the tests were welcome. The following table shows how long each participant spent on each test:

	Word Parts	Sup & Sub	Syn & Ant	Collocation
PR	26	8	13	22
AR	19	15	20	22
JB	30	16	25	20
CW	27	30	18	25
RT	35	21	18	30

PM	34	9	20	25
MK	35	9	5	11
JR	20	15	15	15
JK	25	15	15	20
JN	20	10	15	25
MEAN TOTAL	27 min	15 min	16 min	21.5 min

Word Parts Test

The results of the pilot study showed that 81.69% of provided word classes were consistent with the derivatives extracted from Nation’s 1K base word list and the consulted dictionaries – Longman, Webster, Oxford, Cambridge, and Collins Cobuild. Surprisingly five participants had wrongly located some derivatives. That is, some nouns, verbs, adjectives, and adverbs were wrongly positioned, e.g. adverb was written for the adjective form or noun was written for the verb form, etc. They had wrongly written the word classes of 2 items, 11 items, 13 items, 8 items, and 2 items, respectively. Credit was given to such wrongly positioned derivatives because the test was measuring word parts not grammatical functions.

There were some derivatives provided such as “solved, imagined, or reflected” for SOLUTION, IMAGINE, and REFLECT. The words were identified as adjectives in the Babylon online dictionary, and they were used in the corpuses (COCA or BNC) after linking verbs as well, which was proof of them being adjectives. Therefore, credit was given to such answers in a second round of fine-tuning the scores. As a result, the consistency of the participants’ performance with the provided norming list increased from 81.69% to 87.35%. In addition, EXACTLY with 57.5% of consistency with the norming list, FACT with 70% of consistency, and EAT with 77.5% of consistency were removed from the test. The removal of the aforementioned words increased the consistency rate of the test from 87.35% to 89.19%.

Synonym and Antonym Test

Eighty-five point four four percent (85.44%) of provided synonyms and antonyms by 10 English L1 graduate students were available in the extracted synonyms and antonyms from online thesaurus dictionaries – Merriam-Webster, Collins Cobuild, and thesaurus.com. Adding one more criterion increased the percentage of the participants’ acceptable answers on the test from 85.44% to 90.14%. The criterion was as follows: if two or more English L1 graduate students provided an answer (either a synonym or an antonym), the word was double-checked in thesaurus dictionaries’ near synonymy or antonymy. If the word was a near synonym or antonym in one of the thesaurus dictionaries, it was added to the norming list, if not, the word was ignored. For example, from 10 English L1 graduate students, 7 of them provided “silly” as an antonym for “serious”. Webster thesaurus online dictionary considered “silly” a near antonym for “serious”. Therefore, “silly” was added to the antonyms of “serious”. “Periphery” and “tertiary” were provided by two English L1 graduate students as antonyms for MAIN. However, no thesaurus dictionary considered them either an antonym or a near antonym of MAIN. Therefore, they were ignored. The only exception to this recent criterion was the word “next” as an antonym for PREVIOUS which was provided by 7 English L1 graduate students, but it was not available as a possible antonym or near antonym for PREVIOUS in thesaurus dictionaries. Because many English L1 speakers consider “next” to be an antonym of PREVIOUS, ‘next’ was accepted in this study as a possible antonym for PREVIOUS.

The pilot study also showed that two items, STRENGTH and ACCEPT, needed to be removed from the test because their meanings seemed to overlap with the meaning of two other items, ABILITY and RECEIVE respectively. STRENGTH and ACCEPT were

removed because they were provided with a lower percentage of accepted synonyms and antonyms in comparison to ABILITY and RECEIVE, 95% and 75% in comparison to 100% and 95% respectively. In addition, INCLUDE was also removed from the test because only 55% of English L1 graduate students had provided acceptable synonyms and antonyms for it. After removing these 3 items, the test remained with 31 items and had 91.61% response rate in the pilot for accepted synonyms and antonyms for the target words.

The following are examples of the unaccepted synonyms and antonyms for USEFUL and ALONE with 75% of accepted answers:

1. USEFUL (75%)
 - a. unacceptable synonyms: important; needed; purposeful
 - b. unacceptable antonyms: unneeded
2. ALONE (75%)
 - a. unacceptable synonyms: ---
 - b. unacceptable antonyms: with; company; multiple; group; crowds

Superordination and Subordination Test

The results of the pilot study with the superordination and subordination test showed a high performance on the test. Ninety-eight point nine two percent (98.92%) of provided superordinates and subordinates were found to be acceptable. In fact, a reasonable triangle needed to be provided to have full credit for each item (a score of 3). That is, the provided possible category (superordinate, 1 score) and two examples of the same category (subordinates, 2 scores) needed to have a reasonable relationship with the target word. For example, APRIL was categorized as “month” and was provided with two subordinates, “January” and “February”. This was a reasonable relationship (triangle). Yet, one participant considered APRIL a name and provided “Mary” and “Huda” as two subordinates. This also seemed to be a reasonable and acceptable relationship. BOX could be categorized as “shape”

such as “triangle” and “square”, or it could be a “container” such as “bucket” and “bag”.

TREE could be categorized as both “plant” and “nature”, and “flower” and “bush” could be provided as examples of the same category. Such answers seemed to be reasonable and acceptable. The following table shows a number of unaccepted answers:

	Possible category	Example of the same category	Example of the same category
HOUR	Timing ✓	6 o'clock X	10 o'clock X
READ	Consumption X	See ✓	Hear ✓
CAR	Place X	Room X	House X
HISTORY	Literature X	Testimonial X	Fiction X
LETTER	Font X	Number ✓	Symbol ✓
BABY	Mother ✓	Child ✓	Caregiver X

The justification of rejected answers is as follows: (a) six and 10 o'clock did not seem to show any depth of knowledge for HOUR. If “timing” was considered the possible category, one would expect to see “seconds” or “minutes” as examples of the same category, (b) for the provided “consumption” for READ, the pilot participant asserted that when she reads, she is consuming entertainment. She believed seeing and hearing could also be considered the consumption of the entertainment. However, while it seemed that READ could be considered an activity like seeing and hearing, it was unlikely to be limited only to the consumption of the entertainment, (c) it did not seem reasonable to consider CAR a place, and it was very unlikely to assume that “room”, “house” and CAR belonged to the same category. Therefore, neither “place” nor “room or house” was credited as a reasonable answer for CAR, (d) “testimonial” and “fiction” seemed to be reasonable examples for “literature”, but they were unlikely to have a logical relationship with HISTORY. Therefore, there seemed to be a reasonable superordination and subordination relationship between the provided superordinate and subordinates, but this relationship did not seem to be suitable for HISTORY, (e) “font” was unlikely to be considered a superordinate (main category) for

LETTER, “number” and “symbol”. In fact, font (as set of letters) did not seem to be an appropriate superordinate for LETTER. Instead of font, “password” could be provided as a possible superordinate. Passwords could be LETTERS, numbers and symbols, and such a relationship seemed to work well, and finally (f) BABY and “caregiver” were unlikely to be subordinates of “mother”. In fact, BABY and caregiver did not seem to be in the same level and rank. “Caregiver” could replace “mother” as a possible superordinate for BABY. Therefore, it was not credited as a possible example (subordinate) for BABY.

No target word was provided with less than 90% of reasonable superordinates and subordinates. However, an attempt was made to keep the target words with the most consistent superordination (main category) as possible test items in the first place (discussed in the Target Words for the Superordination and Subordination Test). WALL and BABY, and LOOK and CITY were provided with 8 and 7 different superordination (main category) respectively by 14 English L1 graduate students (4 from the time the target words were to be selected plus 10 from the current pilot study). The rest of the items were provided with 5 or fewer than 5 varieties of superordination. In an attempt to have the target words with more consistent superordination, and to reduce the time of test administration, the aforementioned 4 items were removed from the test. The superordination and subordination test remained with 30 items and 98.88% of accepted answers provided by the pilot participants.

Collocation Test

The first criterion for the scoring of the collocation test was the calculation of a *MI* score of 3 or more and a *t*-score of 2 or more obtained by COCA/BNC Collocates and the Web1T5. In total, 78.13% of the provided collocates were consistent with this criterion. Out of 34 target words, the provided collocates for 4 of them (HUSBAND, OLD, TYPE,

USUALLY) had a rate of consistency lower than 70% with this criterion – 63.27%, 53.27%, 36.63%, 33.3% respectively. These 4 items were removed from the test; as a result, the rate of consistency with the criterion increased from 78.13% to 82.33%.

The assumption was that at least 90% of the provided collocates by English L1 graduate students should have been consistent with the scoring criterion. For this purpose, the inconsistent collocates (167 collocations altogether) with the first criterion were gathered together and their frequency of co-occurrence were checked in the COCA. They were ordered based on their frequency of co-occurrence from the highest to the lowest. The first 70 collocations (out of 167) with the highest frequency of co-occurrence increased the percentage of accepted collocations from 82.33% to 90%. The first 70 collocations had a frequency of co-occurrence between 37 (the lowest) and 6,349 (the highest). Therefore, 37 was chosen as the arbitrary cut-off point of the acceptable frequency of co-occurrence for the study and this criterion increased the rate of accepted collocations from 82.33% to 90.10%.

The aforementioned inconsistent collocations with the first criterion were given to three English L1 graduate students to be evaluated to see whether they agreed that those collocates could be used together naturally and habitually. From the aforementioned 70 collocations, 64 of them were rated as collocations by all three English L1 graduate students, 2 of them by two English L1 graduate students, 3 of them by one English L1 graduate student, and 1 collocation was evaluated as a non-collocate by all three, although it had a frequency of co-occurrence of 54 – BANK now (BANK as the node and ‘now’ as its collocate). Therefore, the selection of the minimum score of 37 as the arbitrary cut-off point of frequency of co-occurrence seemed to be appropriate for the purpose of the study. As a result, the collocation test was finalized with 30 items and two scoring criteria – having either

an acceptable *MI* or *t*-scores (3 and 2 or more respectively) or an acceptable frequency of co-occurrence (37 or more).

Pilot Study with 3 Iranian Graduate Students

Three Iranian graduate students in Canada were also invited to take part in the pilot study. They were born and raised in Iran and had high proficiency in English. They had been residing in Canada for approximately 6 years. The participants were told that they could take as long as needed to complete the tests. The form-meaning tests were administered first. They spent approximately 20 minutes on the tests. The collocation test was briefly explained and administered afterwards. Care was taken to make sure they had no difficulty understanding the instructions of the test. The test took approximately 30 minutes. The participants had a break for 15 minutes. Then, the superordination/subordination and synonym/antonym tests were administered in order. Care was taken to make sure they had no difficulty understanding the tasks. They completed the tests in approximately 45 minutes. They had another break for 15 minutes afterwards. The word parts test was administered last. Care was taken to make sure participants were familiar with metalinguistic terms such as adjective and adverb. They completed the word parts test in less than 40 minutes.

They believed the superordination and subordination was the easiest test while the collocation was the most challenging and the word parts the most time-consuming tests. The pilot participants were highly proficient in English and performed quite well. The following table shows their results.

	CFMT	SSFMT	SAFMT	WPFMT	CT	SST	SAT	WPT
BM	96.66% (29/30)	100% (30/30)	96.77% (30/31)	96.77% (30/31)	75.55% (68/90)	97.77% (88/90)	80.64% (50/62)	87.90% (109/124)
AH	96.66% (29/30)	100% (30/30)	93.54% (29/31)	93.58% (29/31)	76.66% (69/90)	96.66% (87/90)	88.70% (55/62)	77.41% (96/124)
RA	93.33% (28/30)	100% (30/30)	93.54% (29/31)	93.54% (29/31)	73.33% (66/90)	94.44% (85/90)	83.87% (53/62)	92.74% (115/124)

CFMT: Collocation Form-Meaning Test

SSFMT: Superordination and Subordination Form-Meaning Test

SAFMT: Synonym and Antonym Form-Meaning Test

WPFMT: Word Parts Form-Meaning Test

CT: Collocation test

SST: Superordination and Subordination Test

SAT: Synonym and Antonym Test

WPT: Word Parts Test

All in all, the results of the pilot with 3 Iranian graduate students showed that the participants had a strong knowledge of form-meaning connection and superordination and subordination at the 1,000 word frequency level. The results also showed that they knew just under three-fourths of all possible collocates, and well over three-fourths of synonyms and antonyms, and derivatives of the target words at the aforementioned word level. No glitch was observed in the study, and the tests were well-administered without encountering any problems. The tests, the scoring, and the procedures seemed appropriate, and the study was ready to be conducted.

Main Study

Participants

The study was conducted with 40 graduate and 6 undergraduate students majoring in Chemistry, Engineering, History, Biology, and Business from 2 Iranian universities. Their first language was Persian, and English was considered their foreign language. They were

enrolled in a private language center which was specialized in training students for academic purposes to prepare them for TOEFL and IELTS exams.

The participants had learned English continuously in private language schools for at least 3 years, to a maximum of 8 years, before their current enrollment in that specific program. Additionally, they had English courses during 7 years of junior and senior high school, 4 years of undergraduate, and 2 years of graduate studies. They aimed for an IELTS (academic module) score of 6.5 to 7.5 or more, and TOEFL score of 90 to 100 or more, because this score range usually satisfies the basic admission requirements of universities.

The participants had taken oral and written placement tests in that language center and their speaking and writing abilities had been evaluated by language instructors in the center. The instructors were MA graduates in TEFL with TOEFL scores around 112 and IELTS academic scores around 8.5.

Three types of courses were offered in the language center. The first included 80 hours of intensive TOEFL or IELTS training for those learners who were proficient enough in English to start proficiency test courses. The second included 40 hours of pre-TOEFL/pre-IELTS for those who were not proficient enough to start the proficiency test courses. The third types of courses were for low-proficiency learners registered in different levels of general English classes. Therefore, there were mainly 3 levels, general English (for low-proficiency students), pre-TOEFL or pre-IELTS for pre-intermediate or intermediate students, and TOEFL and IELTS intensive courses for advanced students.

All participants were registered for the third level and had already started the TOEFL or IELTS intensive courses. Their mean score on the Vocabulary Levels Test (VLT, Schmitt, Schmitt, & Clapham, 2001) at the 2nd 1,000 level was 29 (96.82%), indicating that they were proficient at that level, and that they had receptive form-meaning knowledge of almost all

2,000 most frequent words (Schmitt et al., 2001). Their mean score on a modified version (Llach & Espinosa, 2014) of the original Productive VLT (PLT, Laufer & Nation, 1999) at the 2nd 1,000 level was 26.43 (88.11%), indicating that they had mastered that level and had productive form-meaning knowledge of approximately 1,800 of the 2,000 most frequent words. All participants that remained in the data pool had answered correctly more than 80% of the items of the VLT and PLT at the 2,000 word level.

Research Instruments

The Vocabulary Levels Test, Version 1 (VLT, Schmitt, Schmitt & Clapham, 2001)

This test is considered a diagnostic test which can show in which levels of 2,000, 3,000, Academic, and 10,000 word levels the test takers may have problem and lack necessary form-meaning knowledge. Each level tests 30 words receptively. In each level there are ten groups of words, each consists of six words numbered from 1 to 6, and in front of each group, there are three word definitions or synonyms. Test takers should choose the right word to go with each meaning or synonym. Because the words are a representative sample of that specific word level, a learner's score at each level represents the proportion of all the words known at that level. For example, 15 out of 30 on the 2nd 1,000 level indicates that 50% or 500 out of 1000 words are known at that level. To remain in the data pool, the participants needed to answer correctly at least 80% of the items of 2,000 word level (at least 24 items out of 30 for each level). This percentage was accepted in Webb's series of studies.

The VLT was used as a pretest to choose those participants who had, at least, the form and meaning receptive knowledge of the high frequency words (see Nation, 2013). The reason is that it is unlikely to expect the participants to be able to provide derivative forms, associations or collocations for a group of target words when they do not even have the form

and meaning knowledge of those words. For example, the results of Schmitt's (1998) and Schmitt's and Meara's (1997) study lead to the conclusion that learners may not be able to produce associations unless they know at least one meaning sense of a word.

Productive Levels Test (PLT, Llach & Espinosa, 2014)

The original PLT (Laufer & Nation, 1999) is considered a productive diagnostic test which can show in which levels of 2,000, 3,000, Academic, and 10,000 word levels the test takers may have problem and lack necessary form-meaning knowledge. Each level consists of 18 sentences. However, Llach and Espinosa (2014) combined Version A and Version C (which are equivalent) from the Parallel version 1 of the original PLT (Laufer & Nation, 1995, 1999) to create a 30-item test that could be given in a single administration. They used the first band of the test that included a sample of the 2,000 high frequency words in English. Learners were presented with a set of 30 sentences including a blank and were required to write the missing target word of which a variable number of initial letters was provided.

The PLT was used as a pretest to ensure that the participants had the form and meaning productive knowledge of the high frequency words (at least) for the same reason mentioned in the VLT. To remain in the data pool, the participants needed to answer correctly at least 80% of the items of 2,000 word level (at least 24 items out of 30). This percentage was accepted in Webb's series of studies.

The Word Parts Test (Appendix A)

The word parts test measured productively the derivations of 31 target words. The participants were required to provide four different word classes (noun, verb, adjective and adverb) for each target word.

The Synonym and Antonym Test (Appendix B)

The synonym and antonym test measured productively the synonyms and antonyms of 31 target words. The participants were required to provide a synonym and antonym for each target word.

The Superordination and Subordination Test (Appendix C)

The superordination and subordination test measured productively the superordinates and subordinates of 30 target words. The participants were required to provide one superordinate (as a main category) and two subordinates (as two examples of the same category) for each target word.

The Collocation Test (Appendix D)

The collocation test measured the production of the collocates for 30 target words. The participants were required to write three words that could naturally be used together with the target words.

The Form and Meaning Tests (Appendix E)

The form and meaning tests measured productively the knowledge of the form and meaning connection of the target words. The translations (in Persian) of the aforementioned target words (target words of the word parts, synonym and antonym, superordination and subordination, and collocation tests—122 words altogether) were provided, and the participants were required to provide the English equivalence (L2) of each translation (L1).

Procedure

Part 1

The form-meaning tests, the PLT, and the VLT tests were administered respectively on a Friday morning in the language center to 46 participants. First, the form-meaning tests were administered for approximately 45 minutes. The results were collected and the participants had a break for 15 minutes. After the break, the PLT was administered for approximately 20 minutes. The results were collected and the VLT was administered for approximately 15 minutes. Before each test, I explained the test instructions to ensure that the participants were comfortable with the task. The suggested time for the tests were 30, 20, and 15 minutes; however, the participants were informed that they could spend as long as needed on the tests. The participants were served lunch and had a break for 45 minutes afterwards.

Part 2

The collocation test was administered in the afternoon and it took approximately half an hour. The data were collected, and the synonym and antonym test was administered afterwards which took approximately 20 minutes. The participants had a break for 15 minutes with refreshments after that. The superordination and subordination test was then administered for approximately 20 minutes followed by the administration of the word parts test for approximately 30 minutes. The suggested time for each test was 30 minutes; however, the participants were informed that they could spend as long as needed on the tests. I explained the instructions of each test before its administration and ensured that the participants were familiar with meta-linguistic words such as adjective, adverb or words such as collocations.

Chapter Summary

The chapter described in detail the design of tests to measure productively four aspects of vocabulary knowledge, the use of the appropriate criteria for selecting the target words and scoring the tests, and the pilot of the designed tests to ensure that the instruments of the study were well-designed. The method of the study including participants, instruments, and procedure were also discussed. The next chapter reports the results of the study.

CHAPTER FOUR: RESULTS

Overview

This chapter begins with a review of the instruments (tests) of the study, reports the results of the form-meaning, word parts, association, and collocation tests respectively, and then considers the results of the aforementioned tests all at a glance. The chapter ends with the chapter summary and conclusion.

Tests at a Glance

For the purpose of the current study, four productive depth of vocabulary knowledge tests, and four corresponding form-meaning tests, were developed (8 tests altogether). Table 4.1 shows specifications of each test.

Table 4.1

Information about the Designed Tests of the Study

Name of the Test	Task	Number of Items	Range of Scores for Each Item	Range of Scores for Each Test	TOTAL Score	TOTAL Percentage
Collocation Form-Meaning (CFM)	Provide the L2 target word for the word in L1.	30	0-1	0-30	30/30	100%
Superordination & Subordination Form-Meaning (SSFm)	Provide the L2 target word for the word in L1.	30	0-1	0-30	30/30	100%
Synonym & Antonym Form-Meaning (SAFM)	Provide the L2 target word for the word in L1.	31	0-1	0-31	31/31	100%
Word Parts Form-Meaning (WPFM)	Provide the L2 target word for the word in L1.	31	0-1	0-31	31/31	100%
Collocation Test (CT)	Provide 3 collocates for each target word.	30	0-3	0-90	90/90	100%
Superordination & Subordination Test (SST)	Provide one superordinate and two subordinates for each target word.	30	0-3	0-90	90/90	100%
Synonym & Antonym Test (SAT)	Provide one synonym and one antonym for each target word.	31	0-2	0-62	62/62	100%
Word Parts Test (WPT)	Provide 4 word classes of each target word.	31	0-4	0-124	124/124	100%

Table 4.1 includes test abbreviations, the tasks assigned to participants to complete each test, the number of target items for each test, the possible range of scores for each item and for the whole test, and the highest possible score. Based on the table, for example, the collocation form-meaning test was a translation test that required the participants to provide a translation equivalent in English for each Persian word. There were 30 items in the test, and each participant needed to provide all 30 items to score the highest (30/30). On the other hand, it can also be understood from the table that the collocation test was a test that required the participants to provide three collocates for each word. The test scored from 0 to 90, and the participants needed to provide all 90 collocates to score the highest (90/90).

Form-Meaning Tests

Participants were required to provide the English word for the same word in Persian for all four form-meaning tests. The collocation, and superordination and subordination form-meaning tests had 30 items while the synonym and antonym, and word parts form-meaning tests had 31 items. Each item was scored 0 if the target word was not provided and 1 if the word was provided. As a result, the mean scores for the tests ranged from 0 to 30 for the former two and 0 to 31 for the latter two. Table 4.2 shows the descriptive statistics of the participants' performance on the aforementioned form-meaning tests.

Table 4.2

Descriptive Statistics of the Performance on 4 Form-Meaning Tests

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
CFM	46	17.00	30.00	27.41	2.80	7.83
SSFM	46	28.00	30.00	29.74	.53	.29
SAFM	46	19.00	31.00	27.69	3.28	10.75
WPFM	46	17.00	31.00	27.37	3.00	9.00

Note. CFM: Collocation Form-Meaning; SSFM: Superordination & Subordination Form-Meaning; SAFM: Synonym & Antonym Form-Meaning; WPFM: Word Parts Form-Meaning. The mean scores are from a total number of 30, 30, 31, and 31 items respectively.

Table 4.2 (the descriptive statistics) shows that participants scored the highest in the superordination and subordination form-meaning test (29.74/30) while they scored lower in the other 3 form-meaning tests (27.41/30, 27.69/31, and 27.37/31). Based on Schmitt et al. (2001) criterion mastery, the participants had strong form-meaning knowledge of the target words. Schmitt et al. (2001) considered 86.66% (at least 26 items out of 30) to be the criterion for mastery of the word levels of the VLT.

The mean percentages of the performance may also provide a better understanding of the results. Table 4.3 shows the descriptive statistics of the participants' performance as percentages on the aforementioned form-meaning tests.

Table 4.3

Descriptive Statistics of the Percentages on 4 Form-Meaning Tests

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
CFM	46	56.66	100.00	91.36	9.33	86.97
SSFm	46	93.33	100.00	99.13	1.78	3.19
SAFM	46	61.29	100.00	89.33	10.58	111.89
WPFM	46	54.83	100.00	88.27	9.68	93.62

Note. CFM: Collocation Form-Meaning; SSFM: Superordination & Subordination Form-Meaning; SAFM: Synonym & Antonym Form-Meaning; WPFM: Word Parts Form-Meaning. The number of items are 30, 30, 31, and 31 respectively; however, the table shows the mean percentages of the performance.

Table 4.3 (descriptive statistics) shows that participants performed the best on the superordination and subordination form-meaning test (99.13%) while they performed with small variation on the other 3 form-meaning tests (91.36%, 89.33%, and 88.27% respectively). The following figure illustrates the results.

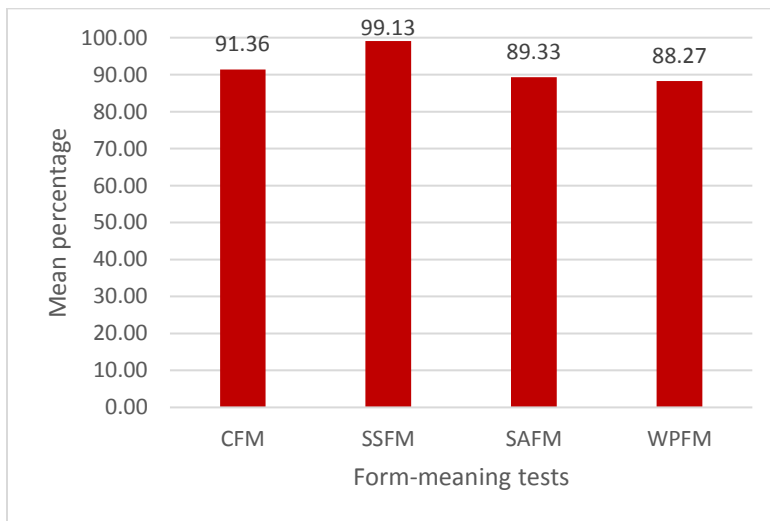


Figure 4.1. The participants' mean percentages of the performance on collocation, superordination & subordination, synonym & antonym, and word parts form-meaning tests respectively.

Word Parts Test

Participants were required to provide 4 word classes (noun, verb, adjective, and adverb) for 31 items. The score for each item ranged from 0 (no word class provided) to 4 (all 4 word classes provided). The total test score ranged from 0 (no knowledge) to 124 (the strongest performance). The answers were scored based on the provided norming list (discussed in detail in Scoring of the Word Parts Test, Chapter 3). Table 4.4 shows the descriptive statistics of the participants' performance on the word parts test.

Table 4.4

Descriptive Statistics of the Performance on the Word Parts Test

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Noun	46	13.00	31.00	24.83	3.55	12.59
Verb	46	20.00	31.00	26.30	2.62	6.88
Adjective	46	10.00	31.00	23.65	5.18	26.81
Adverb	46	3.00	30.00	19.80	6.74	45.45
Total mean	46	47.00	118.00	94.59	14.95	223.63

Note. The mean scores for noun, verb, adjective, and adverb are from a total of 31 while the total mean score is from a total of 124.

Table 4.4 shows the total mean score on the word parts test and the mean scores for each word class—noun, verb, adjective, and adverb. The descriptive statistics show that verbs had the highest mean score (26.30/31) followed by nouns (24.83/31), adjectives (23.65/31), and finally adverbs (19.80/31). The table also shows a total mean score of 94.59 out of 124, indicating that while approximately three-fourths (94.59/124) of all possible word

classes were provided, the total mean score was considerably lower than the maximum possible score (124/124).

A repeated measures ANOVA was run to examine whether the differences between the means of the word classes (noun, verb, adjective, and adverb) were statistically significant. Table 4.5 shows the results.

Table 4.5

Repeated Measures ANOVA on the Mean Scores of the Word Classes

: MEASURE_1		Type III Sum of Squares	df	Mean Square	F	Sig.
Wordparts	Sphericity Assumed	1068.016	3	356.005	29.810	.000
	Greenhouse-Geisser	1068.016	1.979	539.738	29.810	.000
	Huynh-Feldt	1068.016	2.069	516.128	29.810	.000
	Lower-bound	1068.016	1.000	1068.016	29.810	.000
Error(Wordparts)	Sphericity Assumed	1612.234	135	11.942		
	Greenhouse-Geisser	1612.234	89.044	18.106		
	Huynh-Feldt	1612.234	93.118	17.314		
	Lower-bound	1612.234	45.000	35.827		

Table 4.5 shows the differences between the means were statistically significant, $F(3,135)=29.81, p < .001$. Bonferroni Post-hoc comparisons were run to examine which mean difference was statistically significant. Table 4.6 shows the results.

Table 4.6

Results of Bonferroni Post-hoc Comparisons

: MEASURE_1

(I) wordparts		Mean Difference (I-J)	Std. Error	Sig. ^b	95% Confidence Interval for Difference ^b	
					Lower Bound	Upper Bound
Noun	Verb	-1.478*	.410	.005	-2.609	-.347
	Adjective	1.174	.512	.159	-.239	2.587
	Ad erb	5.022*	.918	.000	2.489	7.554
Verb	Noun	1.478*	.410	.005	.347	2.609
	Adjective	2.652*	.649	.001	.860	4.444
	Ad erb	6.500*	.845	.000	4.167	8.833
Adjective	Noun	-1.174	.512	.159	-2.587	.239
	Verb	-2.652*	.649	.001	-4.444	-.860
	Ad erb	3.848*	.841	.000	1.527	6.169
Adverb	Noun	-5.022*	.918	.000	-7.554	-2.489
	Verb	-6.500*	.845	.000	-8.833	-4.167
	Adjective	-3.848*	.841	.000	-6.169	-1.527

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Table 4.6 shows that while the performance on verbs was significantly better than the performance on other word classes, the performance on adverbs was significantly worse than the performance on other word classes. There was no significant difference between performance on nouns and adjectives.

Further analysis was conducted to investigate the number of items provided with 4, 3, 2, 1, and 0 correct word classes. Table 4.7 shows the descriptive statistics of the number of items provided by the participants with word classes consistent with the norming list.

Table 4.7

Descriptive Statistics of the Number of Items Provided with Word Classes Consistent with the Norming list

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
4 word classes	46	.00	25.00	12.02	6.66	44.38
3 word classes	46	1.00	26.00	10.83	4.20	17.61
2 word classes	46	.00	16.00	6.17	4.20	17.66
1 word class	46	.00	9.00	1.72	2.41	5.81
0 word class	46	.00	5.00	.26	.83	.69

Note. The mean scores are from a total number of 31 items.

Table 4.7 shows that on average 12.02 items (from a total of 31) were provided with 4 word classes consistent with the norming list while the numbers of items provided with 3, 2, 1, and 0 word class(es) were 10.83, 6.17, 1.72, and .26 respectively. This indicated that the majority of the items were not provided with all 4 word classes which was not very impressive for advanced students at university level. The following figure illustrates the results.

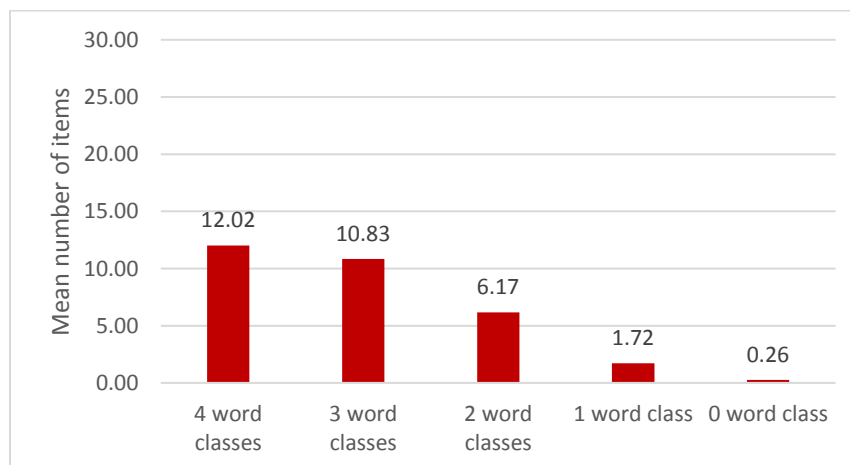


Figure 4.2. The mean number of items provided with 4, 3, 2, 1, and 0 word class(es) consistent with the norming list. The mean scores are from a total of 31 items.

More analysis was conducted to examine the extent to which the participants were able to provide the word classes of the items. Table 4.8 shows the descriptive statistics of the word classes provided for the test by each participant.

Table 4.8

Descriptive Statistics of the Word Classes Provided for the Test by Each Participant

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
P_1	31	.00	4.00	2.39	.88	.78
P_2	31	1.00	4.00	3.06	.77	.60
P_3	31	1.00	4.00	3.42	.85	.72
P_4	31	1.00	4.00	2.03	.80	.63
P_5	31	.00	4.00	1.52	.93	.86
P_6	31	1.00	4.00	2.90	.83	.69
P_7	31	1.00	4.00	3.10	.87	.76
P_8	31	1.00	4.00	3.42	.67	.45
P_9	31	.00	4.00	2.45	1.15	1.32
P_10	31	2.00	4.00	3.00	.86	.73
P_11	31	1.00	4.00	2.52	1.03	1.06
P_12	31	1.00	4.00	2.45	1.21	1.46
P_13	31	.00	4.00	2.71	.97	.95
P_14	31	.00	4.00	2.77	1.31	1.71
P_15	31	3.00	4.00	3.81	.40	.16
P_16	31	1.00	4.00	3.35	.80	.64
P_17	31	2.00	4.00	3.68	.65	.43
P_18	31	.00	4.00	3.13	.96	.92
P_19	31	1.00	4.00	2.74	.82	.66
P_20	31	1.00	4.00	2.45	.68	.46
P_21	31	2.00	4.00	3.23	.76	.58
P_22	31	1.00	4.00	2.84	.82	.67
P_23	31	2.00	4.00	3.39	.76	.58
P_24	31	1.00	4.00	3.39	.67	.45
P_25	31	2.00	4.00	3.16	.86	.74
P_26	31	.00	4.00	2.77	.99	.98
P_27	31	2.00	4.00	3.10	.83	.69
P_28	31	.00	4.00	2.84	.90	.81
P_29	31	1.00	3.00	2.77	.56	.31
P_30	31	.00	4.00	3.13	.99	.98
P_31	31	2.00	4.00	3.65	.61	.37
P_32	31	1.00	4.00	3.26	.82	.66
P_33	31	1.00	4.00	2.81	1.01	1.03
P_34	31	1.00	4.00	2.90	1.04	1.09
P_35	31	2.00	4.00	3.26	.77	.60
P_36	31	2.00	4.00	3.45	.68	.46
P_37	31	.00	4.00	2.97	1.02	1.03
P_38	31	2.00	4.00	3.61	.67	.45
P_39	31	3.00	4.00	3.74	.44	.20
P_40	31	2.00	4.00	3.74	.51	.26
P_41	31	1.00	4.00	2.45	.81	.66
P_42	31	2.00	4.00	3.00	.77	.60
P_43	31	2.00	4.00	3.58	.62	.38
P_44	31	2.00	4.00	3.52	.63	.39
P_45	31	1.00	4.00	3.19	.87	.76
P_46	31	3.00	4.00	3.55	.51	.26
Total mean	31	1.83	3.59	3.05	.33	.11

Note. “P” refers to the participants. The number of items is 31. The mean scores are out of 4.

Table 4.8 shows that while the total minimum mean score for provided word classes was 1.83, the maximum total mean score was 3.59. The table also shows that on average 3.05 word classes out of 4 were known and provided by the participants, indicating that the participants tended to know 3 word classes. If the results of the ANOVA test are taken into consideration, the participants seem to have more problem with providing adverbs. The following figure illustrates the results.

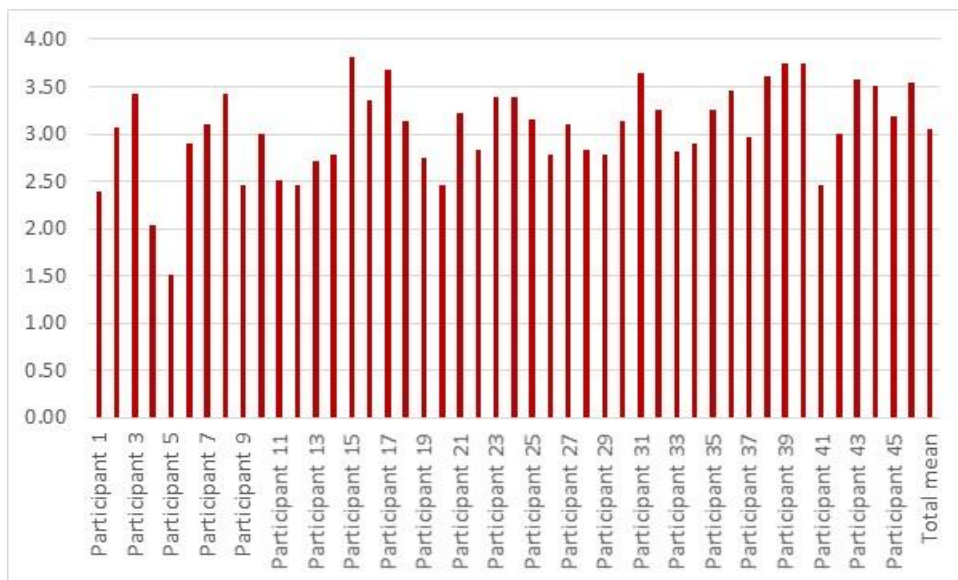


Figure 4.3. The participants' mean scores of the provided word classes for the items of the test. The mean scores are from a total of 4 (all 4 word classes)

Synonym and Antonym Test

Participants were required to provide one synonym and one antonym for 31 items (target words). The score of each item ranged from 0 (neither synonym nor antonym) to 2 (both synonym and antonym). Consequently, the test scores ranged from 0 to 62. The answers were scored based on the provided norming list (discussed in detail in Scoring of the

Synonym & Antonym Test, Chapter 3). Table 4.9 shows the descriptive statistics of the participants' performance on the synonym and antonym test.

Table 4.9

Descriptive Statistics of the Synonym and Antonym Test

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Synonym	46	12.00	30.00	23.33	4.61	21.25
Antonym	46	12.00	31.00	24.63	4.10	16.77
Total mean	46	28.00	61.00	47.96	7.90	62.40

Note. The test has 31 items so that the maximum scores are 31 for synonyms and antonyms, and 62 for the whole test.

Table 4.9 (the descriptive statistics) shows that the mean score of antonyms (24.63/31) was slightly higher than synonyms (23.33/31). The table also shows the mean score of the test was 47.96/62.

Table 4.10 shows the descriptive statistics of the percentage of the performance on the synonym and antonym test.

Table 4.10

Descriptive Statistics of the Percentage on the Synonym and Antonym Test

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Synonym percentage	46	19.35	48.39	37.62	7.44	55.28
Antonym percentage	46	19.35	50.00	39.73	6.61	43.63
Total mean percentage	46	45.16	98.39	77.35	12.74	162.33

Table 4.10 shows that on average 77.35% of the provided synonyms and antonyms were consistent with the norming list, from which synonyms had a share of 37.62%, and antonyms had a share of 39.73%. The total mean percentage (77.35%) was less than the

maximum possible percentage (100%), indicating that the participants did not seem to have full knowledge of synonymy and antonymy of the items. This did not seem very impressive for advanced students at university level. The following figure illustrates the results.

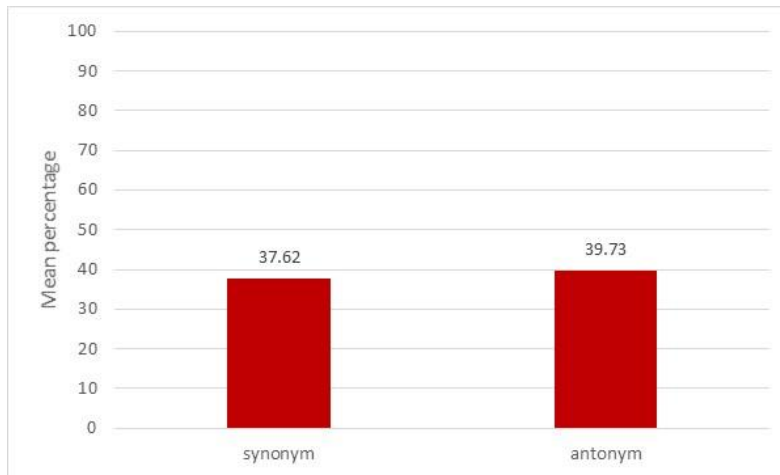


Figure 4.4. The mean percentages of the performance on the synonym and antonym test. The means are from 100% (the strongest performance).

A paired samples *t*-test was run to see whether the difference between the mean scores of the synonyms and antonyms was statistically significant. The following table shows the results.

Table 4.11

Results of Paired Samples Test on the Mean Scores of the Synonyms and Antonyms

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Synonym - Antonym	-1.30435	3.69305	.54451	-2.40105	-.20765	-2.395	45	.021

Table 4.11 shows that the difference between the mean scores (1.30) was statistically significant, $t(45) = 2.39$, $p = .02$, indicating that the participants had greater knowledge of antonyms of the target words than synonyms.

Further analysis was conducted to examine the mean number of items provided with both a synonym and an antonym (2), with a synonym or an antonym (1), and with neither a synonym nor an antonym (0). Table 4.12 shows the results.

Table 4.12

Descriptive Statistics of the Items Provided with Synonyms & Antonyms Consistent with the Norming List

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
2, Syn & Ant	46	7.00	30.00	19.61	5.58	31.18
1, Syn or Ant	46	1.00	16.00	8.78	3.86	14.89
no Syn & Ant	46	.00	11.00	2.61	2.60	6.78

Note. 2, Syn & Ant: items are provided with both a synonym and an antonym; 1, Syn or Ant: items are provided with either a synonym or an antonym; no Syn & Ant: items are provided with no answer consistent with the norming list. Maximum mean scores are 31.

Table 4.12 shows that the majority of the items (19.61, approximately 2/3) were provided with synonyms and antonyms consistent with the norming list, a lower number of items (8.78) were provided with either a synonym or antonym, and a considerably lower number of items (2.61) were provided with no correct answer consistent with the norming list. The following figure illustrates the results.

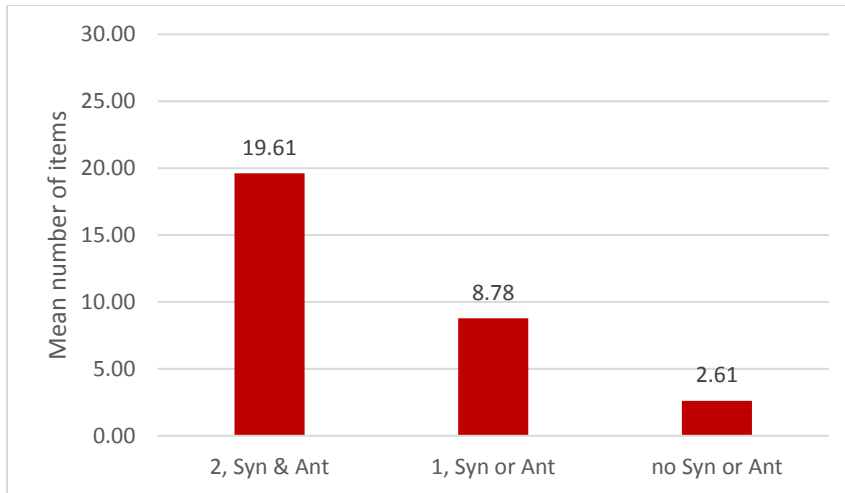


Figure 4.5. The mean scores of the items provided with both synonyms and antonyms, with either synonyms or antonyms, and with neither synonyms nor antonyms. The mean scores are from a total of 31 items.

More analysis was conducted to examine the extent to which the participants were able to correctly provide the synonyms and antonyms of the items (0, 1, and 2 for 31 items).

Table 4.13 shows the descriptive statistics of the average number of correct synonyms and antonyms that each participant provided for each item.

Table 4.13

Descriptive Statistics of the Average Number of Synonyms & Antonyms Provided for the Test by Each Participant

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
P_1	31	.00	2.00	1.16	.78	.61
P_2	31	.00	2.00	1.32	.65	.43
P_3	31	.00	2.00	1.68	.54	.29
P_4	31	.00	2.00	1.00	.77	.60
P_5	31	.00	2.00	.90	.75	.56
P_6	31	.00	2.00	1.58	.67	.45
P_7	31	1.00	2.00	1.65	.49	.24
P_8	31	.00	2.00	1.55	.68	.46
P_9	31	.00	2.00	1.29	.64	.41
P_10	31	.00	2.00	1.65	.61	.37
P_11	31	.00	2.00	1.32	.65	.43
P_12	31	.00	2.00	1.29	.82	.68
P_13	31	.00	2.00	1.32	.70	.49
P_14	31	.00	2.00	1.29	.69	.48
P_15	31	1.00	2.00	1.97	.18	.03
P_16	31	.00	2.00	1.65	.55	.30
P_17	31	.00	2.00	1.74	.58	.33
P_18	31	.00	2.00	1.45	.72	.52
P_19	31	.00	2.00	1.55	.68	.46
P_20	31	.00	2.00	1.48	.68	.46
P_21	31	.00	2.00	1.39	.72	.51
P_22	31	.00	2.00	1.35	.66	.44
P_23	31	.00	2.00	1.65	.61	.37
P_24	31	.00	2.00	1.61	.67	.45
P_25	31	1.00	2.00	1.74	.44	.20
P_26	31	.00	2.00	1.23	.67	.45
P_27	31	1.00	2.00	1.77	.43	.18
P_28	31	.00	2.00	1.77	.50	.25
P_29	31	.00	2.00	1.71	.64	.41
P_30	31	.00	2.00	.97	.84	.70
P_31	31	.00	2.00	1.77	.50	.25
P_32	31	.00	2.00	1.81	.48	.23
P_33	31	.00	2.00	1.32	.65	.43
P_34	31	.00	2.00	1.68	.54	.29
P_35	31	1.00	2.00	1.71	.46	.21
P_36	31	.00	2.00	1.68	.54	.29
P_37	31	.00	2.00	1.58	.67	.45
P_38	31	1.00	2.00	1.81	.40	.16
P_39	31	1.00	2.00	1.87	.34	.12
P_40	31	1.00	2.00	1.77	.43	.18
P_41	31	.00	2.00	1.35	.66	.44
P_42	31	.00	2.00	1.77	.50	.25
P_43	31	1.00	2.00	1.77	.43	.18
P_44	31	1.00	2.00	1.84	.37	.14
P_45	31	.00	2.00	1.68	.60	.36
P_46	31	.00	2.00	1.68	.65	.43
Total mean	31	.61	1.93	1.55	.28	.08

Note. “P” refers to the participants. The number of items is 31 for which both synonyms and antonyms could be provided (2), either a synonym or an antonym could be provided (1), or neither a synonym nor an antonym could be provided (0). The mean scores are from a total of 2 (both a synonym and an antonym).

Table 4.13 shows that on average the minimum total number of the provided synonyms and antonyms was less than 1 (.61) while the maximum was just under 2 (1.93). The table also shows that on average 1.55 synonyms and antonyms out of 2 were known and provided by the participants. The following figure illustrates the results.

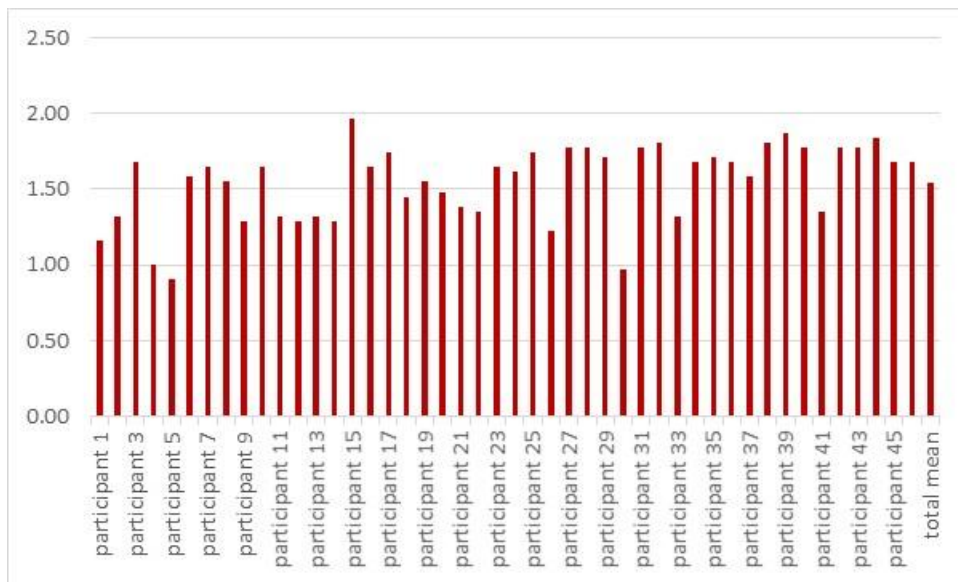


Figure 4.6. The participants' mean scores of the provided synonyms and antonyms for the items of the test. The mean scores are from a total of 2 (both synonyms and antonyms, the strongest performance).

Superordination and Subordination Test

Participants were required to provide one superordinate and two subordinates for 30 items. The range of scores for each item was from 0 (no answer) to 3 (1 correct superordinate and 2 correct subordinates). As a result, the score of the test could range from 0 to 90. Three raters evaluated the participants' performance on the superordination and subordination test. They were required to score all answers that had a reasonable

superordinate and subordinate relationship with the items. Table 4.14 shows the descriptive statistics of each rater's evaluation of the test.

Table 4.14

Descriptive Statistics of the Performance on the Superordination and Subordination Test Evaluated by 3 Raters

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Rater 1	46	63.00	90.00	83.56	6.85	46.91
Rater 2	46	63.00	90.00	84.52	6.89	47.49
Rater 3	46	63.00	90.00	84.02	6.77	45.89
Mean of raters	46	63.00	90.00	84.03	6.78	46.03

Note. The test has 30 items, and the mean scores are from a total of 90 (30 superordinates and 60 subordinates).

Table 4.14 shows the mean scores of 83.56/90, 84.52/90, and 84.02/90 from the evaluations of 3 raters. The three ratings were averaged to have one single score for the purpose of reporting the participants' performance on the superordination and subordination test. The mean score of 84.03 out of 90 was close to the maximum possible score (90/90).

Pearson product-moment correlation tests were run to ensure interrater reliability. Table 4.15 shows the results.

Table 4.15

Pearson Correlation between the Raters' Evaluations of the Superordination & Subordination Test

		Rater_1	Rater_2	Rater_3
Rater_1	Pearson Correlation	1	.973**	.982**
	Sig. (2-tailed)		.000	.000
	N	46	46	46
Rater_2	Pearson Correlation	.973**	1	.974**
	Sig. (2-tailed)	.000		.000
	N	46	46	46
Rater_3	Pearson Correlation	.982**	.974**	1
	Sig. (2-tailed)	.000	.000	
	N	46	46	46

** . Correlation is significant at the 0.01 level (2-tailed).

Table 4.15 shows that there were strong correlations between the raters' evaluations. In fact, there was a strong positive correlation between the first and the second rater, $r(46) = .973$, $p < .001$, the first and the third rater, $r(46) = .982$, $p < .001$, and the second and the third rater, $r(46) = .974$, $p < .001$.

Further analysis was conducted to examine the mean scores of the provided superordinates and subordinates. For this purpose, only the answers accepted as reasonable superordinates and subordinates for the items by all three raters were considered. This, in fact, helped calculate the most reliable results for the provided superordinates and subordinates, and also for the test. Table 4.16 shows the descriptive statistics of the performance on the test.

Table 4.16

Descriptive Statistics of the Superordinates and Subordinates

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Sup score	46	18.00	30.00	27.22	3.15	9.95
Sub score	46	43.00	60.00	56.35	4.15	17.21
Total	46	62.00	90.00	83.57	6.85	46.96

Note. Sup score: superordination mean score; Sub score: subordination mean score. The test has 30 items. For each item 1 superordinate and 2 subordinates are asked to be provided so that the superordination mean score is from a total of 30, the subordination mean score is from a total of 60, and the total mean score is from 90.

Table 4.16 shows a mean score of 27.22 from a total of 30 (30 items, one superordinate for each) for the superordination, a mean score of 56.35 from a total of 60 (30 items, two subordinates for each) for the subordination, and a total mean score of 83.57 from a total of 90 for the test. All three mean scores were so high that it could be claimed that there was a ceiling effect; however, their total mean score (83.57) was still slightly below the maximum possible score (90/90).

Further analysis was conducted to calculate the percentages of the provided superordination and subordination so that the results on these two could be compared with each other. Table 4.17 shows the descriptive statistics of the percentages of the provided superordinates and subordinates.

Table 4.17

Descriptive Statistics of the Percentages of the Provided Superordinates and Subordinates

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Sup percentage	46	60.00	100.00	90.73	10.52	110.59
Sub percentage	46	71.67	100.00	93.91	6.91	47.79

Note. Sup percentage: the mean percentage of the provided superordinates; Sub percentage: the mean percentage of the provided subordinates.

Table 4.17 shows a total mean percentage of 90.73 for the superordinates and 93.91 for the subordinates. It seems the participants had a better knowledge and performance on the subordination in comparison to the superordination; however, a paired sample *t*-test needs to be calculated to see whether the difference between the mean scores of the subordinates and superordinates was statistically significant. The following table shows the results of the *t*-test.

Table 4.18

Results of Paired Samples Test on the Mean Percentages of the Superordinates and Subordinates

		Paired Differences							Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	
					Lower	Upper			
Pair 1	Sup percentage - Sub percentage	-3.18804	6.95498	1.02546	-5.25342	-1.12267	-3.109	45	.003

Table 4.18 shows that the difference between the mean scores (3.18) was significant, $t(45) = 3.10, p = .003$, indicating that the participants had a better knowledge of the subordinates in comparison to the superordinates.

More analysis was conducted to examine the mean number of items provided with one reasonable superordinate and 2 reasonable subordinates accepted by all 3 raters. Table 4.19 shows the results.

Table 4.19

Descriptive Statistics of the Items Provided with Reasonable Superordinates and Subordinates Accepted by all 3 Raters

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
3 answers	46	16.00	30.00	25.96	3.60	12.93
2 answers	46	.00	10.00	2.61	2.37	5.62
1 answer	46	.00	5.00	.41	.86	.74
0 answer	46	.00	8.00	1.02	1.82	3.31

Note. 3 answers: 1 superordinate and 2 subordinates; 2 answers: 1 superordinate & 1 subordinate, or 2 subordinates; 1 answer: either 1 superordinate or 1 subordinate; 0 answer: neither a superordinate nor a subordinate. The mean scores are from a total of 30 items.

Table 4.19 shows that on average 25.96 items from a total of 30 were provided with reasonable superordinates and subordinates while this number was considerably lower for items provided with 2, 1, and 0 answer(s) – 2.61, .41, and 1.02 from 30 respectively. The results show that the majority of the items (approximately 26 out of 30) were provided with reasonable superordinates and subordinates, indicating that the participants had a strong performance on the test and that their performance was slightly lower than the maximum.

The following figure illustrates the results.

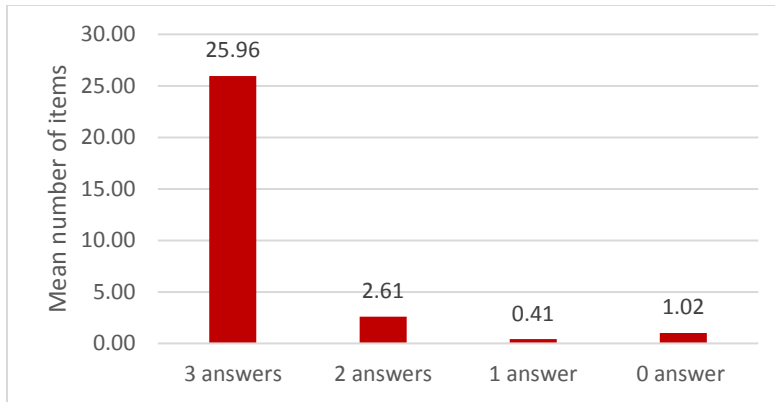


Figure 4.7. The mean scores of the number of items provided with 3, 2, 1, and 0 reasonable superordinate and subordinates. The mean scores are from a total of 30 items.

More analysis was conducted to examine the extent to which the participants were able to reasonably provide the superordinates and subordinates of the items (3, 2, 1, and 0 for 30 items). Table 4.20 shows the descriptive statistics of the superordinates and subordinates provided for the test by each participant.

Table 4.20

Descriptive Statistics of the Provided Superordinates and Subordinates for the Test by Each Participant

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
P_1	30	0	3	2.43	1.073	1.151
P_2	30	.00	3.00	2.07	1.31	1.720
P_3	30	2.00	3.00	2.93	.25	.064
P_4	30	.00	3.00	2.50	.73	.534
P_5	30	.00	3.00	2.40	1.07	1.145
P_6	30	.00	3.00	2.50	1.01	1.017
P_7	30	2.00	3.00	2.87	.35	.120
P_8	30	1.00	3.00	2.87	.43	.189
P_9	30	2.00	3.00	2.87	.35	.120
P_10	30	.00	3.00	2.80	.61	.372
P_11	30	.00	3.00	2.73	.78	.616
P_12	30	.00	3.00	2.53	.97	.947
P_13	30	.00	3.00	2.73	.64	.409
P_14	30	3.00	3.00	3.00	.00	.000
P_15	30	3.00	3.00	3.00	.00	.000
P_16	30	3.00	3.00	3.00	.00	.000
P_17	30	1.00	3.00	2.93	.37	.133
P_18	30	.00	3.00	2.73	.78	.616
P_19	30	.00	3.00	2.73	.69	.478
P_20	30	.00	3.00	2.70	.65	.424
P_21	30	1.00	3.00	2.83	.46	.213
P_22	30	.00	3.00	2.83	.59	.351
P_23	30	2.00	3.00	2.97	.18	.033
P_24	30	.00	3.00	2.87	.57	.326
P_25	30	3.00	3.00	3.00	.00	.000
P_26	30	.00	3.00	2.27	1.28	1.651
P_27	30	2.00	3.00	2.90	.31	.093
P_28	30	2.00	3.00	2.87	.35	.120
P_29	30	2.00	3.00	2.93	.25	.064
P_30	30	.00	3.00	2.57	.97	.944
P_31	30	.00	3.00	2.87	.57	.326
P_32	30	2.00	3.00	2.97	.18	.033
P_33	30	2.00	3.00	2.93	.25	.064
P_34	30	2.00	3.00	2.90	.31	.093
P_35	30	.00	3.00	2.37	.96	.930
P_36	30	3.00	3.00	3.00	.00	.000
P_37	30	2.00	3.00	2.97	.18	.033
P_38	30	2.00	3.00	2.97	.18	.033
P_39	30	2.00	3.00	2.97	.18	.033
P_40	30	2.00	3.00	2.83	.38	.144
P_41	30	1.00	3.00	2.37	.76	.585
P_42	30	1.00	3.00	2.83	.46	.213
P_43	30	2.00	3.00	2.80	.41	.166
P_44	30	2.00	3.00	2.97	.18	.033
P_45	30	3.00	3.00	3.00	.00	.000
P_46	30	2.00	3.00	2.97	.18	.033
Total mean	30	2.33	3.00	2.78	.18	.033

Note. “P” refers to the participants. The mean scores are from a total of 3 (the strongest possible performance).

Table 4.20 shows that the minimum total mean score of the provided answers was 2.33 out of 3 (1 superordinate and 2 subordinates) while the maximum total mean score was 3 out of 3. The table shows that on average 2.78 superordinates and subordinates (from a

total of 3) were known and provided by the participants. This indicated that the participants had a strong performance on the test and their performance was slightly lower than the maximum (2.78/3 compared to 3/3). The following figure illustrates the results.

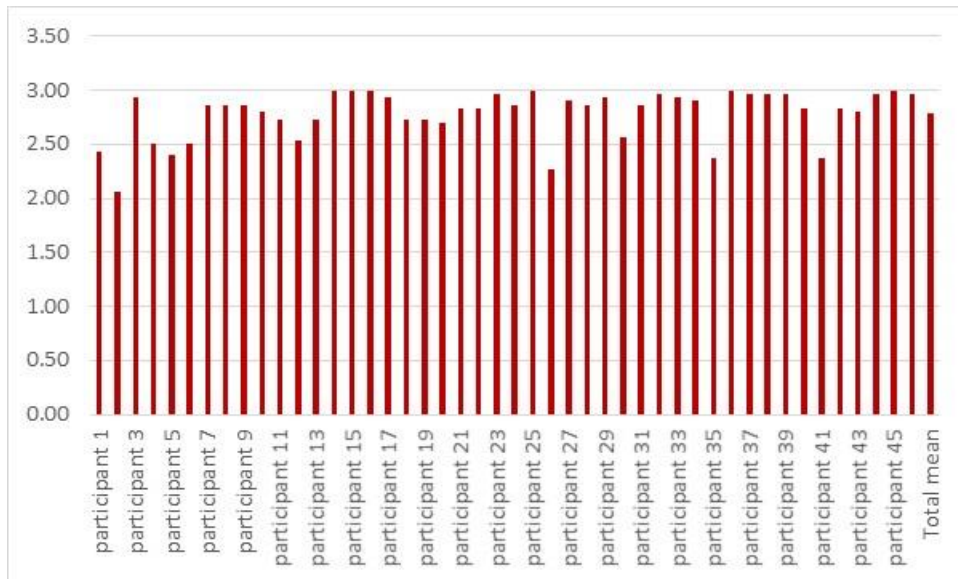


Figure 4.8. The participants' mean scores of the provided superordinates and subordinates for the items of the test. The mean scores are from a total of 3 (one superordinate and two subordinates).

Collocation Test

The collocation test included 30 items and participants were required to provide three collocates for each item. Therefore, the range of scores for each item was from 0 (no collocate) to 3 (all 3 collocates) and the minimum and maximum score for each participant ranged from 0 to 90. Four analyses were applied for scoring the responses provided by participants for the collocation test: (a) the calculation of *t*-scores and *MI* scores equal to or more than 2 and 3 respectively entitled “CT/raw”, (b) the calculation of *t*-scores and *MI* scores plus a frequency of co-occurrence of 37 or higher for items that did not meet the first criterion entitled “CT/FoC”, (c) the calculation of *t*-scores and *MI* scores plus two raters’

evaluations of items that did not meet the first criterion entitled “CT/R1 & CT/R2”, and (e) the calculation of the average of CT/FoC, CT/R1, and CT/R2 (entitled “Mean FoC & Rs”) to have one single score for the participants’ performance on the collocation test. Table 4.21 shows the descriptive statistics of the performance on the collocation test analyzed using the aforementioned criteria.

Table 4.21

Descriptive Statistics of the Performance on the Collocation Test Evaluated with Different Criteria

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
CT/raw	46	6.99	69.99	46.82	11.78	138.87
CT/FoC	46	7.99	76.00	53.28	12.74	162.26
CT/R1	46	7.99	78.99	58.32	13.78	189.90
CT/R2	46	7.99	79.99	61.27	13.84	191.48
mean FoC & Rs	46	7.99	77.99	57.62	13.26	175.85

Note. CT/raw: the calculation of *t*-scores and *MI* scores; CT/FoC: the calculation of *t*-scores and *MI* scores plus a frequency of co-occurrence (FoC) of 37 or higher for items that did not meet either of the first two criteria; CT/R1 & CT/R2: the calculation of *t*-scores and *MI* scores plus the first and second human raters’ evaluations of items that did not meet either of the first two criteria. Mean FoC & Rs: the mean of CT/FoC, CT/R1, and CT/R2. The test has 30 items, and for each item 3 collocates are asked to be provided. Therefore, the mean scores are from a total of 90 (the strongest performance).

Table 4.21 (the descriptive statistics) shows that a mean score of 46.82 out of 90 (CT/raw) was consistent with the *t*-score or *MI* score equal or more than 2 and 3 respectively. The table also shows that the highest mean scores were the participants’ mean scores of the collocations calculated by *t*-scores and *MI* scores plus the raters’ evaluations of items that did not meet either of the first two criteria (CT/R1 & CT/R2: 58.32/90 & 61.27/90). The calculation of the *t*-scores and *MI* scores plus a frequency of co-occurrence of 37 or higher for items that did not meet either of the first two criteria showed a lower mean score

(CT/Foc: 53.28/90) in comparison to the latter two mean scores. Generally, the participants' performance on the test, which was an average of the latter three mean scores (Mean FoC & Rs: 57.62/90), was considerably lower than the maximum possible performance (90/90). The following figure illustrates the results.

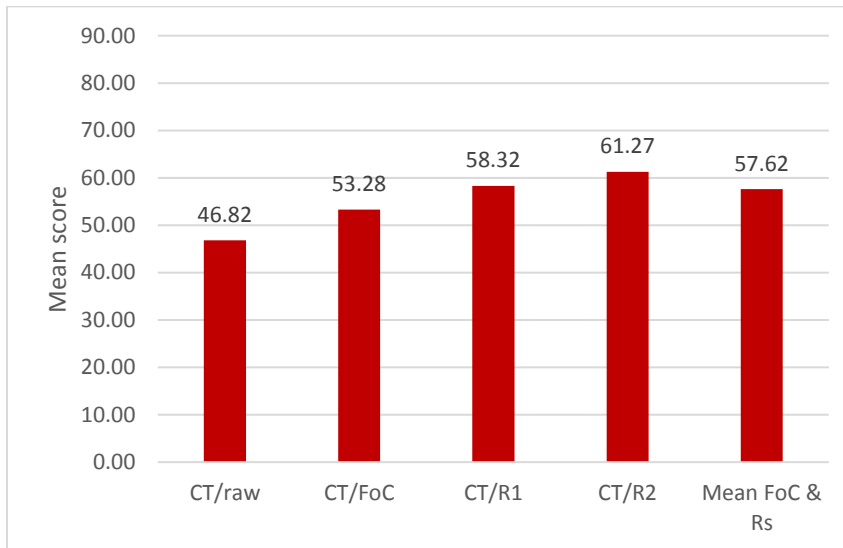


Figure 4.9. The participants' mean scores of the collocations resulted from the calculation of the t-scores and MI scores (CT/raw), the calculation of the aforementioned scores plus the frequency of co-occurrence (CT/FoC), the calculation of the aforementioned scores plus the human raters' evaluations (CT/R1 & CT/R2), and the calculation of the average of the latter two mean scores (Mean FoC & Rs). The test has 30 items for which 90 collocates can be provided (90/90, the strongest performance).

Table 4.22 shows the descriptive statistics of the percentage of the performance on the collocation test analyzed with the aforementioned criteria.

Table 4.22

Descriptive Statistics of the Percentage on the Collocation Test Evaluated with the Different Criteria

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
CT/raw	46	7.77	77.77	52.02	13.09	171.44
CT/FoC	46	8.88	84.44	59.20	14.15	200.32
CT/R1	46	8.88	87.77	64.80	15.31	234.45
CT/R2	46	8.88	88.88	68.07	15.38	236.40
mean FoC & Rs	46	8.88	86.66	64.02	14.73	217.10

Note. CT/raw: the calculation of *t*-scores and *MI* scores; CT/FoC: the calculation of *t*-scores and *MI* scores plus a frequency of co-occurrence (FoC) of 37 or higher for items that did not meet either of the first two criteria; CT/R1 & CT/R2: the calculation of *t*-scores and *MI* scores plus the first and second human raters' evaluations of items that did not meet either of the first two criteria. Mean FoC & Rs: the mean average of CT/FoC, CT/R1, and CT/R2. The test has 30 items for which 90 collocates should be provided (90/90) to show 100% of the strongest performance.

Table 4.22 shows 52.02% (CT/raw) of performance when the *t*-scores and *MI* scores were calculated. The table shows 59.20% (CT/FoC) of performance when the *t*-scores, *MI* scores, and frequency of co-occurrence of the collocations were calculated. This percentage rose to 64.80 and 68.07 (CT/ R1 & R2) when *t*-scores, *MI* scores, and the raters' evaluation of the collocations were considered. The mean percentage of 64.03 (the average of CT/FoC, CT/R1, and CT/R2) was considered the participants' performance on the collocation test to be able to report a single percentage as the participants' performance on the collocation test. Generally, the mean percentage of performance on the test (64.03%) seemed to be considerably below the maximum possible percentage (100%). The following figure illustrates the results.

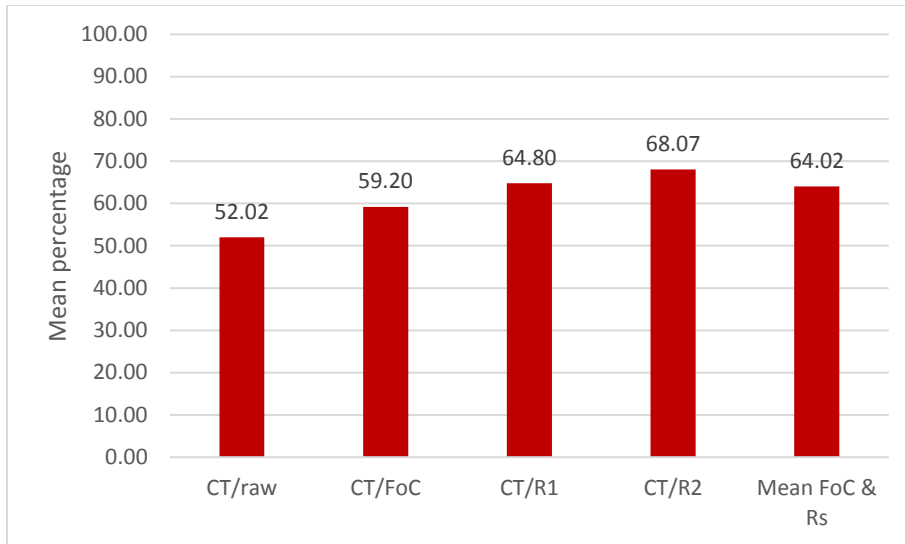


Figure 4.10. The participants' mean percentages of the collocations resulted from the calculation of the t-scores and MI scores (CT/raw), the calculation of the aforementioned scores plus the frequency of co-occurrence (CT/FoC), the calculation of the aforementioned scores plus the human raters' evaluations (CT/R1 & CT/R2), and the calculation of the average of the latter two mean percentages (Mean FoC & Rs). The test has 30 items for which 90 collocates should be provided (90/90) to show 100% of the strongest performance.

It should be noted that the participants provided 76 compound nouns (such as typewriter, smartphone, and heartbeat for items like TYPE, PHONE, and HEART) as collocations. This was 1.83% of the total collocates provided, which could increase all of the aforementioned percentages by 1.83%. However, because the test did not measure compound nouns, this percentage was not added to the participants' performance.

Pearson product-moment correlation tests were run to ensure reliability between the raters' evaluation (R1 & R2) and also between the raters' evaluation and the evaluation with the frequency of co-occurrence (FoC). Table 4.23 shows the results.

Table 4.23

Pearson Correlation between the Evaluations of the Collocation Test

		CT/FoC	CT/R1	CT/R2
CT/FoC	Pearson Correlation	1	.943**	.937**
	Sig. (2-tailed)		.000	.000
	N	46	46	46
CT/R1	Pearson Correlation	.943**	1	.991**
	Sig. (2-tailed)	.000		.000
	N	46	46	46
CT/R2	Pearson Correlation	.937**	.991**	1
	Sig. (2-tailed)	.000	.000	
	N	46	46	46

** . Correlation is significant at the 0.01 level (2-tailed).

Note. CT/FoC: the calculation of the *t*-scores and *MI* scores plus the frequency of co-occurrence; CT/R1 & R2: the calculation of the aforementioned scores plus the human raters' evaluations.

Table 4.23 shows that there were significant correlations between raters' evaluations and also between the evaluation with raters and the evaluation with the frequency of co-occurrence. In fact, there was a strong positive correlation between the first and the second rater, $r(46) = .991$, $p < .001$, the first rater and the frequency of co-occurrence evaluation, $r(46) = .943$, $p < .001$, and the second rater and the frequency of co-occurrence evaluation, $r(46) = .937$, $p < .001$. It should be noted that the inter-rater reliability between the first and second rater was strong enough so that a third rater was not required to evaluate the same collocates.

Further analysis was conducted to examine the number of collocates (0-3) the participants were able to provide for each item. For this purpose, the data was analyzed based on the results from the calculation of *t*-scores and *MI* scores plus a frequency of co-occurrence of 37 or higher for items that did not meet either of the first two criteria. The reason was that such criteria objectively judge whether the collocates were consistent with

the scoring criteria or not. This was not necessarily the same with human raters as one may have believed two words collocate, but another may not necessarily have had the same idea. Also, the strong inter-rater reliability between the human raters' evaluation and the evaluation with the frequency of co-occurrence of 37 or higher attested to the value of this objective criterion. Table 4.24 shows the descriptive statistics of the items provided with 3, 2, 1 and 0 collocate(s) consistent with the scoring criteria.

Table 4.24

Descriptive Statistics of the Items Provided with 0 to 3 Collocates Consistent with the Scoring Criteria

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
3 Collocates	46	.00	16.00	8.59	4.28	18.29
2 collocates	46	1.00	16.00	10.17	2.82	7.97
1 collocate	46	.00	15.00	7.20	3.26	10.61
0 collocate	46	.00	23.00	4.04	3.80	14.44

Note. The mean scores are from a total of 30 items. Three collocates are asked to be provided for each item.

Table 4.24 shows that while approximately one third of the items (10.17/30) were provided with 2 collocates consistent with the scoring criteria, a lower number of items was provided with 3 and 1 collocates (8.59/30 & 7.20/30 respectively), and a considerably lower number of items was provided with no collocate or collocates inconsistent with the criteria (4.04/30). The table shows that the majority of the items were not provided by 3 collocates consistent with the scoring criteria. The following figure illustrates the results.

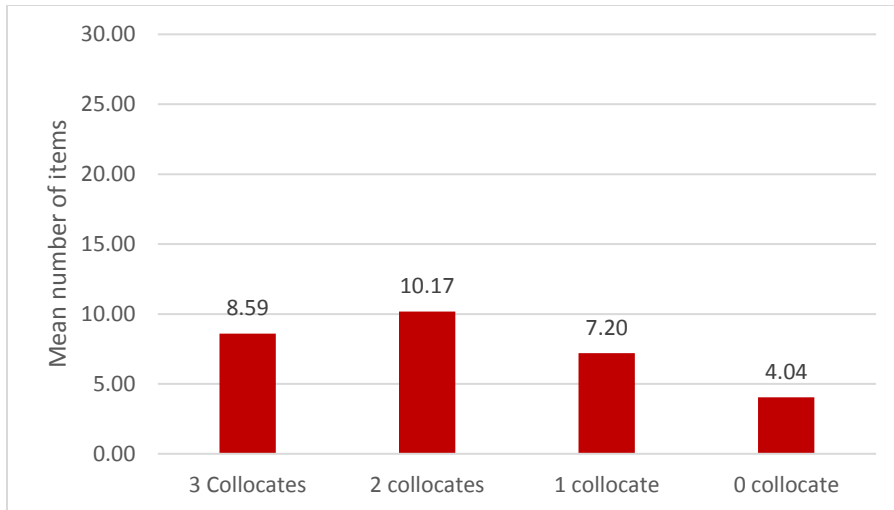


Figure 4.11. The participants' mean scores of the number of items provided with 3, 2, 1, and 0 collocate(s) consistent with the scoring criteria. The mean scores are from a total of 30 items.

Further analysis was conducted to examine the extent to which the participants were able to correctly provide collocates for the items (3 collocates for each item). Table 4.25 shows the descriptive statistics of the collocates provided by each participant.

Table 4.25

Descriptive Statistics of the Collocates Provided for the Test by Each Participant

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
P_1	30	.00	3.00	1.73	.98	.96
P_2	30	.00	3.00	1.20	.92	.86
P_3	30	.00	3.00	1.77	1.01	1.01
P_4	30	.00	3.00	1.20	.89	.79
P_5	30	.00	3.00	1.17	.99	.97
P_6	30	.00	3.00	1.83	1.12	1.25
P_7	30	.00	3.00	2.03	1.10	1.21
P_8	30	.00	3.00	1.53	.94	.88
P_9	30	.00	2.00	.27	.52	.27
P_10	30	.00	3.00	1.43	.77	.60
P_11	30	.00	3.00	1.17	.95	.90
P_12	30	.00	3.00	1.67	1.09	1.20
P_13	30	.00	3.00	1.83	.99	.97
P_14	30	.00	3.00	1.97	1.00	1.00
P_15	30	2.00	3.00	2.53	.51	.26
P_16	30	.00	3.00	1.13	.90	.81
P_17	30	.00	3.00	1.90	.88	.78
P_18	30	.00	3.00	1.80	1.19	1.41
P_19	30	.00	3.00	1.93	.94	.89
P_20	30	.00	3.00	1.70	.88	.77
P_21	30	.00	3.00	1.97	.93	.86
P_22	30	.00	3.00	1.90	.96	.92
P_23	30	1.00	3.00	2.23	.86	.74
P_24	30	.00	3.00	2.20	.96	.92
P_25	30	1.00	3.00	2.17	.79	.63
P_26	30	.00	3.00	1.50	1.04	1.09
P_27	30	.00	3.00	2.17	.87	.76
P_28	30	.00	3.00	1.87	.94	.88
P_29	30	.00	3.00	2.17	.91	.83
P_30	30	.00	3.00	1.63	1.10	1.21
P_31	30	.00	3.00	2.03	.96	.93
P_32	30	.00	3.00	1.93	.87	.75
P_33	30	.00	3.00	1.97	.81	.65
P_34	30	.00	3.00	1.97	.76	.59
P_35	30	.00	3.00	1.17	1.02	1.04
P_36	30	.00	3.00	2.20	.92	.86
P_37	30	.00	3.00	1.30	.88	.77
P_38	30	.00	3.00	2.03	.93	.86
P_39	30	.00	3.00	1.90	.88	.78
P_40	30	.00	3.00	2.37	.85	.72
P_41	30	.00	3.00	1.67	1.09	1.20
P_42	30	.00	3.00	2.33	.80	.64
P_43	30	.00	3.00	1.80	1.00	.99
P_44	30	.00	3.00	1.93	.83	.69
P_45	30	.00	3.00	1.33	.99	.99
P_46	30	.00	3.00	2.20	.92	.86
Total mean	30	.74	2.39	1.78	.44	.19

Note. “P” refers to the participants. The test has 30 items, and 3 collocates need to be provided for each item. The mean scores, therefore, are from a total of 3.

Table 4.25 shows that on average the minimum total number of the collocates provided for the items was less than 1 (.74) while the maximum was more than 2 (2.39). The table shows that on average 1.78 collocates out of 3 were known and provided by the participants. This was not a very impressive result as it reflected that the average mean score of the provided collocates was considerably less than the maximum possible score (3/3). The following figure illustrates the results.

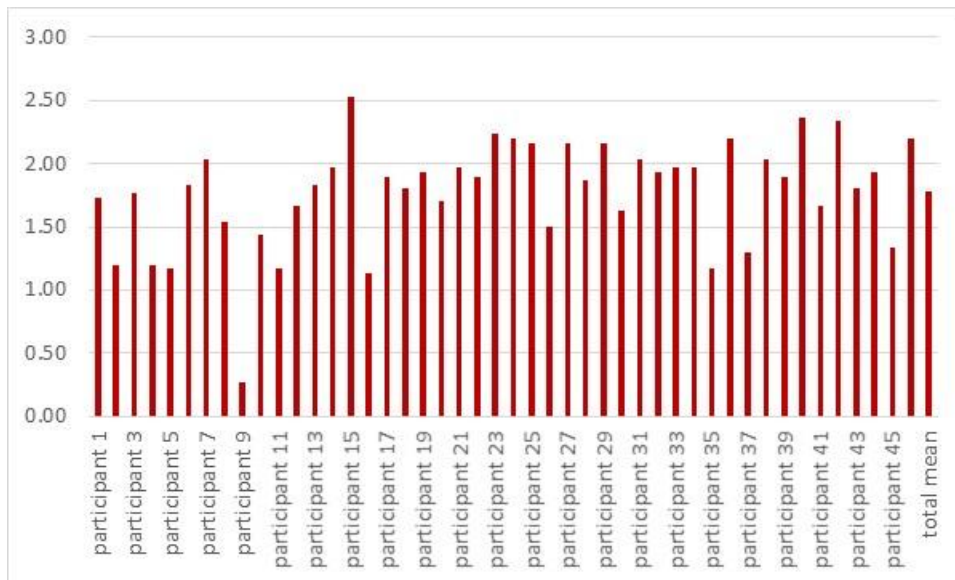


Figure 4.12. The participants' mean scores of the provided collocates for the items of the test. The mean scores are from a total of 3.

Performances on the Tests at a Glance

Due to the fact that the tests had different number of items and different range of scores and mean scores, the percentage of the performance seemed to be the only common scale of evaluation of the tests. Table 4.26 shows the descriptive statistics of the percentages of the performance on 8 tests (4 depth of vocabulary knowledge and their corresponding form-meaning tests)—collocation form-meaning, superordination and subordination form-

meaning, synonym and antonym form-meaning, word parts form-meaning, collocation, superordination and subordination, synonym and antonym, and word parts tests respectively.

Table 4.26

Descriptive Statistics of the Percentages of the Performance on the Tests

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
CFM	46	56.66	100.00	91.36	9.33	86.97
SSFm	46	93.33	100.00	99.13	1.78	3.19
SAFM	46	61.29	100.00	89.33	10.58	111.89
WPFM	46	54.83	100.00	88.27	9.68	93.62
CT	46	8.88	86.66	64.02	14.73	217.10
SST	46	70.00	100.00	93.37	7.54	56.82
SAT	46	45.16	98.38	77.33	12.73	162.06
WPT	46	37.90	95.16	76.27	12.06	145.40

Note. CFM: Collocation Form-Meaning; SSFM: Superordination & Subordination Form-Meaning; SAFM: Synonym & Antonym Form-Meaning; WPFM: Word Parts Form-Meaning; CT: Collocation Test; SST: Superordination & Subordination Test; SAT: Synonym & Antonym Test; WPT: Word Parts Test.

Regarding the form-meaning tests, Table 4.26 shows that while the superordination and subordination form-meaning test was performed the best (99.13%), the word parts form-meaning test was performed the least well (88.27%). Generally, while the participants' performance showed a peak of 99.13% on the superordination and subordination form-meaning test, their performance showed lower percentages of 91.36, 89.33, and 88.27 on the collocation, synonym and antonym, and word parts form-meaning tests respectively. The results showed a strong performance on the form-meaning tests as all four of them had a mean percentage higher than 86.66, reaching the criterion for mastery of the word levels (Schmitt et al. 2001).

Regarding three aspects of depth knowledge, while the superordination and subordination test was performed the best (93.37%), the collocation test was performed the worst (64.02%). The synonym and antonym, and the word parts tests also showed 77.33% and 76.27% of performance respectively. The results showed that the participants' performance on the collocation, synonym and antonym, and word parts tests were lower than the maximum possible performance (100/100). Results are summarized in the following figure.

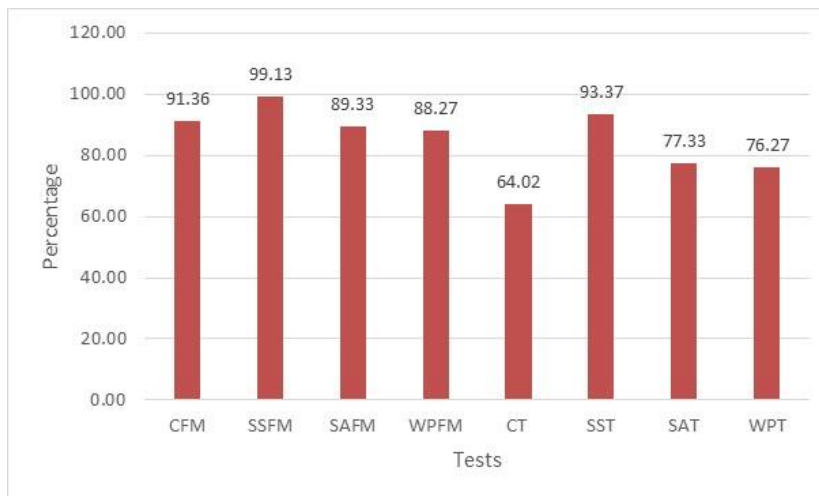


Figure 4.13. The participants' mean percentages of the performance on 8 tests. The acronyms stand for collocation form-meaning (CFM), superordination and subordination form-meaning (SSFM), synonym and antonym form-meaning (SAFM), word parts form-meaning (WPFM), collocation test (CT), superordination and subordination test (SST), synonym and antonym test (SAT), and word parts test (WPT).

Chapter Summary and Conclusion

This chapter reported the obtained results of the study. All in all, the findings show a high level of achievement on the form-meaning tests based on Schmitt et al.'s (2001) criterion mastery, and a ceiling effect on the superordination and subordination test.

Performance on the collocation, synonym and antonym, and word parts tests was considerably lower than the maximum possible score, indicating that advanced EAP students at university level with years of studying English both at public and private educational systems did not have complete mastery of the aforementioned depth aspects of vocabulary knowledge. This is an important issue as it applies to the most frequent words whose range and coverage are considerable in English. The next chapter interprets the results and findings in detail.

CHAPTER FIVE: DISCUSSION

Overview

This chapter discusses the interpretation of the scores and the results obtained from the form-meaning, word parts, synonym and antonym, superordination and subordination, and collocation tests respectively. The results are compared with previous studies where applicable. Summary of the main findings, conclusion, implications of the study, and recommendations for further studies are also discussed. The chapter ends with the limitations of the current study.

Form-Meaning Tests: To what extent do Iranian EAP students have productive knowledge of form and meaning at the 1,000 word frequency level?

The results show that the participants' performance on the form-meaning tests reached Schmitt et al.'s (2001) criterion of mastery as the mean percentage on the superordination and subordination form-meaning test was 99.13% (29.74/30) while the percentages for the other 3 form-meaning tests, the word parts, synonym and antonym, and collocation form-meaning tests were 88.27% (27.37/31), 89.33 % (27.69/31), and 91.36% (27.41/30) respectively. Laufer and Nation (1999) assert that 85% to 90% performance is satisfactory for the 2,000 word level of the productive VLT. The same percentage considered by Laufer and Nation would likely be satisfactory for the first 1,000 word level as well. The main reason could be the fact that the same approximate percentage is also considered the criterion mastery of the receptive Levels Test. Schmitt et al. (2001) considered 86.66% (26 items out of 30) to be the criterion mastery of the word levels of the VLT. However, Webb,

Sasao, and Balance (2017) adopted a stricter cutting point for mastery of the 1,000 level in the updated VLT (Webb, Sasao, & Ballance, 2017) arguing for a score of 97% or higher, indicating that one knows that level very well before moving on to learning the next level.

Considering the fact that the study had four productive form-meaning tests with target words at the 1,000 word level, it seems a single mean percentage of the aforementioned percentages could be representative of the participants' productive knowledge of the form-meaning aspect. The mean percentage is 92% $((91.36+99.13+89.33+88.27)/4)$, indicating that the participants have a strong productive form-meaning mastery of the words at the first 1,000 level ($>86.66\%$, Schmitt et al.'s (2001) criterion for mastery). This percentage may represent the participants' productive form-meaning knowledge because it, in fact, is the result of testing 122 items (30 + 30 + 31 + 31 items of each form-meaning test respectively) of the most frequent 1,000 words. Schmitt et al. (2001) claim that it is important to have as high a sampling rate as possible for each word level to be able to test each level reliably, leading them to suggest using at least 30 items for testing each word level. The current finding (92% resulted from testing 122 items) is backed by 4 times as high a sampling rate as Schmitt et al. considered necessary for testing each level.

The purposeful selection of the target words for the synonym and antonym, superordination and subordination, and word parts form-meaning tests (see Chapter 3) may endanger the generalizability of the obtained result (92%) as the statistical representative of the first 1,000 words. Therefore, the mean percentage of 91.36 on the collocation form-meaning test seems to be a more reliable result for this purpose because the target words were semi-randomly chosen for this test (see Chapter 3). This high mean percentage (91.36%) indicates a level of productive knowledge of the form-meaning of the most

frequent 1,000 words by the participants that exceeds 86.66%, Schmitt et al.'s (2001) criterion for mastery.

To the best of my knowledge, no study has specifically measured productive form-meaning knowledge of the most frequent 1,000 words to date; however, the current results of the form-meaning tests are consistent with what Laufer and Nation (1999) and Zheng (2009) found for the 2,000 word level. Zheng (2009) used Laufer's and Nation's (1999) Productive Levels Test (PLT) for measuring the word levels. The total mean scores of their university student participants were 17 out of 18 (94.44%) and 15.61 out of 18 (86.72%) respectively at the 2,000 word level. The findings of the current study, Laufer's and Nation's study, and Zheng's study show that the form-meaning of the most frequent words (either the first or the second 1,000 words or both) may often be known at university level. It should also be noticed that a modified version of the PLT, as discussed in Methodology (Chapter 3), was used in the current study as a diagnostic test to ensure the participants had at least the productive form-meaning knowledge of the words at 2,000 level. The mean percentage of the PLT in the current study was 88.11, which reconfirms the fact that the productive form-meaning of the most frequent words, as discussed before, seems to be known by the university-level participants in this research.

Word Parts Test: To what extent do Iranian EAP students have productive knowledge of word parts at the 1,000 word frequency level?

The study shows 76.27% of responses on the word parts test were consistent with the norming list (see Chapter 3). The participants produced just over three-quarters of possible word classes for the target words (76.27/100). In other words, the participants did not provide approximately one-quarter of the word classes. This finding shows that participants

may not necessarily be able to produce derivatives of all the words at the 1,000 level that have 3 or 4 possible derivative forms. Considering the great coverage and range of the first 1,000 words, this is not a very impressive finding for advanced students at a university level.

Twelve out of 31 items (an approximate mean number of items) were provided with all word classes consistent with the norming list. Thus, maximum performance occurred for just over one-third of the target words (12/31). This reveals that advanced participants at the university level do not seem to be able to produce all word classes for the items at 1,000 level with 3 to 4 derivatives. Schmitt (1999) believes that this suggests L2 learners may be unable to master all members of a word family until relatively late in the acquisition process, and even this may not happen at all.

Approximately 11, 6, and 2 items out of 31 (an approximate mean number of items) were provided with 3, 2, and 1 derivative form(s) consistent with the norming list. Yet, the total mean number of derivatives provided by the participants is 3 out of 4, indicating that the participants can produce approximately three of the four possible derivatives. Adverbs have a lower mean score compared to the other 3 word classes, indicating that the participants tend to have more difficulty with providing adverbs.

The results also show that the verb derivatives were the best known with 84.84% produced. Nouns were next with 80% produced. Adjectives and adverbs were less known, 76% and 63.88% respectively. This finding is consistent with Schmitt's (1998), and Schmitt's and Zimmerman's (2002) findings. They also found that the participants knew some individual word classes better than others—nouns and verbs were the best known while adjectives and adverbs appeared to be less known. Schmitt (1998) believes that this suggests that adjective and adverb forms are unlikely to be learned from general exposure, maybe due to their lower frequency of occurrence, and hence might be good candidates for explicit

instruction. In other words, nouns and verbs are the most common parts of speech found in natural text (Kucera & Francis, 1967; Webb, 2005, 2007b), and that may cause more exposure to them which may result in learning them better.

Knowledge of word parts was also found to be relatively weak in previous studies. In a hierarchical order: (a) Schmitt and Meara (1997) found that their participants provided only 42% and 47% of all possible native-like suffixes in the beginning and near the end of the academic year respectively, (b) Schmitt (1998) found that his advanced participants had a performance of approximately 56% on possible word classes, (c) Schmitt (1999) found that the participants who chose the correct TOEFL item could provide approximately two appropriate word classes (50%) for the target words while those who missed the TOEFL item could provide the word forms for 1.63 (41%) word classes, (d) Schmitt and Zimmerman (2002) found that the participants' mean number of derivatives was 37.6 (58.8%) from a possible 64 (16 target words * 4 derivatives per word), indicating that the participants produced approximately two of the four possible derivatives. In fact, only 18.2% of all four major word classes were known by their combined non-native participants productively, and (e) Ishii and Schmitt (2009) found that the participants' mean number of derivatives was 17.61 from a possible 45 (15 target words * 3 derivatives per word). Their learners scored 39.13% of derivatives correctly.

Performance on the test of word parts in the current study was higher than in the earlier studies; however, it should be noted that the aforementioned studies on word parts evaluated target words from lower levels of frequency than 1K (even words from 4K and 6K). In other words, the lower performance on the word parts of the reviewed studies may have been the result of performing on lower frequency target words. In addition, lack of knowledge or limited knowledge of words in lower levels of frequency than 1K may not

necessarily make a big difference for participants because the coverage and range of such word levels are considerably less than the 1K level. Therefore, the findings of the current study, with the consideration of the great coverage and range of 1K, seem to be consistent with the findings of the previous studies.

Association (Synonym & Antonym, and Superordination & Subordination) Tests: To what extent do Iranian EAP students have productive knowledge of associations at the 1,000 word frequency level?

The results showed that the percentages of correct responses were 77.33% and 93.37% on the synonym and antonym, and the superordination and subordination tests. The participants produced over three-quarters (77.33%) of possible synonyms and antonyms for the target words with a similar percentage of synonyms and antonyms (37.62% compared to 39.73% respectively). In other words, the participants did not provide approximately one-quarter of all possible synonyms and antonyms. This may show that advanced students at the university level may not necessarily be able to produce synonyms and antonyms for all the words at 1,000 level. It should be noted that care was taken to select target words with transparent synonyms and antonyms for the purpose of the study. In contrast, the study revealed a ceiling effect (93.37%) for knowledge of superordinates and subordinates (with similar percentages as 90.73% of superordinates and 93.31% of subordinates were correct) for the words at 1,000 level with reasonable and transparent superordination and subordination.

Approximately 26 out of 30 items were provided with reasonable superordinates and subordinates, indicating that the participants gave the maximum possible performance on the majority of the items. Also, the total mean number of the superordinates and subordinates produced by the participants is 2.78 out of 3 with superordinates and subordinates having

approximately the same share (comparing the mean scores of 27.22 from 30 for superordinates with 56.35 from 60 for subordinates). The mean score of 84.03 out of 90 was close to the maximum score (90/90) for superordination and subordination. Knowledge of synonyms and antonyms was clearly lower than the knowledge of superordinates and subordinates.

On average, 20 out of 31 items (an approximate mean number of items) were provided with both synonyms and antonyms consistent with the norming list in which the synonyms and antonyms seemed to have the same share (comparing the mean scores of 23.33 from 31 for synonyms with 24.63 from 31 for antonyms). This is the maximum possible performance for approximately two-thirds of the words, indicating mastery over the majority of the items.

On average, 3 out of 31 items (an approximate mean number of items) were provided with no synonym or antonym. That is, the study shows no synonymy or antonymy mastery for over 10% of the items. In other words, advanced participants at university level may have not necessarily had the synonymy or antonymy knowledge of the words at 1,000 level with transparent synonyms and antonyms. Further analysis also shows that, on average, 1.55 (out of 2) synonyms and antonyms were provided by the participants for the test.

Purposeful selection of the target words with reasonable transparent superordinates and subordinates (as some of them are illustrated in Appendix C) would probably have helped the production of reasonable and transparent answers for the test. However, even purposeful selection of the target words with possible transparent synonyms and antonyms (illustrated in Appendix B) did not seem to guarantee the production of the synonyms and/or antonyms for all items at the 1,000 word level.

Previous studies on associations did not show impressive results. Schmitt and Meara (1997) report that their participants could only supply around 50% of possible associations. Schmitt (1998) also did not find very impressive results on associations, as his participants performed even less than 50% on the test in the first session of his interview. Schmitt (1999) found that only 33% of his participants' associative responses matched with his norming list. However, studies on associations (e.g. Schmitt & Meara, 1997; Schmitt, 1998; Wolter, 2001) had a holistic view on associations and probably tested all possible associations (collocates, coordinates, superordinates and subordinates, synonyms and antonyms) together. The main reason for this is that the aforementioned studies required participants to provide the first 3 words they could think of when they heard the target words. These studies compared the responses with the responses provided by English L1 participants (norming list), and it was not clear what type of associations were provided both by native and non-native participants. The current study, however, tested associations in a controlled way (controlled way of evaluating superordination and subordination, and synonymy and antonymy), and the findings may not necessarily be comparable with findings of previous studies.

The fact that the aforementioned studies on associations evaluated target words from lower levels of frequency than 1K (even words from 4K and 6K) should also be noted. In other words, lower performance on the associations of the earlier studies, although not necessarily comparable with the findings of the current study, may have been the result of performing on lower frequency target words. In addition, lack of knowledge or limited knowledge of words in lower levels than 1K may not necessarily make a big difference for participants because the coverage and range of such word levels are considerably less than the 1K level.

For the purpose of discussing the participants' associational knowledge, it seems reasonable to consider the results of the performance on the collocation test in addition to these two superordination and subordination, and synonym and antonym tests. Collocations are, in fact, syntagmatic associations (Fitzpatrick, 2013; Read, 2000; Schmitt, 2010), and the findings on these three tests may show a better picture of the participants' productive knowledge of associations.

The participants performed the lowest on the collocation test (64.02%), the highest on the superordination and subordination test (93.37%), and in between on the synonym and antonym test (77.33%). If the results of these three tests are considered together, it can be concluded that the participants seem to have more than basic or fundamental associational knowledge of words at 1,000 level. However, when it comes to the considerable coverage and range of the most frequent 1,000 words, it may not necessarily be the best idea to rely on this general conclusion. The study shows that advanced students may not have complete mastery of collocations, and synonyms and antonyms for words at the 1,000 level. This may be unexpected at the university level for EAP students who have studied English for years both at public and private educational systems.

Collocation Test: To what extent do Iranian EAP students have productive knowledge of collocations at the 1,000 word frequency level?

The study shows that an average of 64% performance on the collocation test (the average of the calculation of *t*-scores and *MI* scores, frequency of co-occurrence, and two raters' evaluations of items that did not meet the first criterion) was consistent with the scoring criteria, meaning that over one-third of the possible collocates were not produced by the participants. This percentage of performance went down to 59.2 if the objective criteria of

scoring (*t*-score or *MI* score equal or more than 2 and 3 respectively plus the frequency of co-occurrence equal or more than 37) was considered. Also, the results show that 52% of the produced collocates were consistent with the first criterion (*t*-score or *MI* score equal or more than 2 and 3 respectively). This percentage was lower than the percentage of produced collocates (consistent with the first criterion) by the English L1 pilot participants (52% compared to 82.23%). Considering the variety of possible collocates for the target words, the performance on the collocation test demonstrates that advanced participants at the university level may not have complete mastery over the production of the collocates for the words at the 1,000 level.

The 30 target words (items) of the collocation test were chosen semi-randomly, and it seems they can be considered a major indicator of the words at the 1,000 level. The results show that 4 items (an approximate mean number of items) out of 30 were not provided with any collocate consistent with the scoring criteria. This indicates that these learners may have difficulty using language because the most frequent 1,000 words cover approximately 81% of written text and 85% of spoken text (Nation, 2006). If the results of the current collocation test are considered the major indicator of the first 1,000 words, it can be inferred that the participants lack productive collocational knowledge of a considerable amount of written and spoken text. It is challenging to calculate the exact amount though because the target words were not selected randomly, and they may not be statistically representative of the words at 1,000 level. Such a lack of knowledge can hinder students' oral and written presentations. In this regard, Chui's (2006) findings show a very high correlation between the collocational knowledge and the sentence production task. In her study, she found that limited collocational knowledge of her participants at the university level hindered their productive use of the words in sentences.

The results also show that approximately 9 items (an approximate mean number of items) out of 30 were provided with all 3 collocates. That is, if the productive maximum performance on the test was criterion for a correct response, slightly less than one-third of the items would be scored as correct.

Approximately 10 items (an approximate mean number of items) were provided with 2 collocates and 7 items with one collocate consistent with the scoring criteria. However, the participants' average collocational knowledge of the items is 1.78 out of 3.

Schmitt (1999) found that his participants, on average, gave collocates for 1.41 (47%) sentences (out of 3 possible sentences) for the items which were correctly answered in the TOEFL test. His participants composed 1.13 (38%) sentences which included a collocate for items that were missed on the TOEFL test. This shows that his participants were able to provide fewer than half of the possible collocates in sentences, indicating relatively little knowledge of collocations.

Schmitt (1999) tested six TOEFL vocabulary items—2 words from K2, 3 words from K3, and one word from K1 (based on vocabprofilers, Cobb, 2012). Top 60,000 lemmas from the COCA (Davies, 2015) shows the rank of Schmitt's target words as follows: 2,006, 2,379, 1,882, 3,458, 1,040, and 21,185 (2 words from K2, 2 words from K3, 1 word from K4, and 1 from K22 respectively). As a result, the lower performance on the production of the collocates in Schmitt's study may have been due to testing lower frequency words. The current study tests the most frequent words (lemmas from 1K), and if the frequency of the target words is considered, it seems the performance on collocations in both studies are consistent with each other. The results of the current study also seem to be similar to Revier's (2009), as his university student participants' (n=19) mean score was 28.8 out of 45 (64%). However, his participants were supposed to select the combination of 45 collocational

phrases that best completed 45 sentences, and such a task does not necessarily tap productive knowledge and might be considered receptive knowledge.

There are similarities between the participants in Schmitt's study and the participants in the current study. Schmitt's participants were 30 L2 learners of English. The majority of them (27) were international students attending a summer course designed to improve their academic English skills in preparation for their admission entrance into British universities. They were all students who had either taken the TOEFL test before or would take the TOEFL test if they had chosen to study in the United States instead of Britain. The findings of Schmitt's study and the current study show that even advanced students at a the university level may not necessarily have mastery over the production of collocations for high frequency words.

Collocations, as a kind of association, seem to be tested in studies of associations (e.g. Schmitt & Meara, 1997; Schmitt, 1998) as well; however, such studies have not reported what percentage of the produced associations have been collocations. Wolter's (2001) study on associations differentiates between the produced syntagmatic associations (collocations) by his English L2 participants and he reported that 37.7% of their produced associations were syntagmatic. However, the instructions of his test do not necessarily ask for collocations (as discussed in Literature Review), and consequently, his findings may not be compared with the current findings.

Summary of the Main Findings

Iranian advanced EAP university students demonstrated a level of mastery in productive form-meaning knowledge for the words at the 1,000 level. However, they did not have complete mastery over productive derivational, associational, and collocational

knowledge of the words at the 1,000 level. This should be a matter of concern given the extensive coverage and broad range of the most frequent 1,000 words.

Conclusion

The most frequent 1,000 words comprise at least 80 percent of whatever an English learner may say, write, listen, or read. For this reason, when it comes to evaluating learners' vocabulary knowledge of the most frequent 1,000 words, every identified weakness merits attention. The most frequent words are so basic and fundamental components of English that they are usually assumed to be very well known by advanced students who have learned English for years both at the public educational system and private language schools. However, it is a real matter of concern that advanced students do not show strong depth knowledge of such words.

The findings of the current study show that Iranian advanced learners at the university level from an EFL context have mastery of form-meaning knowledge of the first 1,000 words (breadth), but do not have mastery over derivations, associations, and collocations (depth) of the most frequent 1,000 words. EAP students usually study a large variety of vocabulary books published for academic purposes, e.g. vocabulary for IELTS, TOEFL, GRE, etc., and appear to gain the form-meaning knowledge of a considerable number of words in this way. However, as the current study shows, they may not necessarily achieve mastery over other aspects of the same words. One of the outcomes of such a difference may be limited performance in oral and written English. While there are many possible reasons for this limited knowledge, a lack of awareness of the aspects involved in vocabulary knowledge and possibly greater focus on receptive input rather than productive use of the language likely play a large role.

The research suggests that it is still necessary to highlight terms such as depth of vocabulary knowledge, especially for advanced learners. Although scholars such as Read (2004) and Milton (2009) suggest dispensing with the term and concentrating on more specific measures of the quality of vocabulary knowledge, when it comes to the discussion of the nature of vocabulary in general, particularly with practitioners, Schmitt (2014) asserts that the distinction of breadth and depth seems to be still useful. Schmitt (2014) believes the breadth-depth distinction is useful to inform the practitioners that rich and sustained instruction and input are needed to develop the knowledge beyond the simple form-meaning links. This seems necessary to be clarified for high frequency words and particularly for advanced students because for elementary students with small vocabulary sizes there is often little difference between size and a variety of depth measures (Schmitt, 2014).

Furthermore, instructional emphasis on productive outcome of language in general and productive aspects of vocabulary knowledge (productive mode) in particular need to be highlighted. It seems it is necessary to highlight that the language received from listening and readings needs to be produced in speech and writing, and for this purpose, productive aspects of vocabulary knowledge need to be emphasized.

Implications of the Study

The study provides a clear picture of the Iranian EAP students' areas of vocabulary weakness because the instruments measure the three components of productive vocabulary knowledge outlined by Nation (2013): Form, Meaning, and Use. That is, the word parts test taps into Form, the form-meaning and association tests tap into Meaning, and the collocation test taps into Use. Nation and Webb (2011) assert that, to assess depth of vocabulary knowledge, researchers need to measure multiple aspects and determine how well each of

those aspects is known. They believe measuring the extent to which an aspect is known provides a measure of strength of vocabulary knowledge. The current study seems to have both of the aforementioned characteristics—it measures multiple aspects and determines how well each of those aspects is known. For this reason, the method used in the study may help to evaluate productive depth of vocabulary knowledge of language learners in different contexts, e.g. EAP, ESL, EFL, EAL¹¹, or ESOL¹² students. Different teaching methods such as focusing on reading versus focusing on oral skills also may have influence on the findings of the study.

Traditionally, knowledge of a word refers to the meaning and form of that word (Schmitt, 2008). The findings of the present research indicate that language learners, even at advanced levels of proficiency, still may primarily focus on gaining knowledge of the form and meaning aspect. However, the current study can raise awareness of some important aspects involved in vocabulary knowledge other than the link between form and meaning.

It is worthwhile to note that higher scores in knowledge of form and meaning and lower scores in other aspects suggest that depth of vocabulary knowledge do not seem to develop equally and in tandem. In fact, the form and meaning is baseline knowledge needed in learning vocabulary and other aspects may develop incrementally (Schmitt, 2010) with repeated exposures to words (see Webb, 2007b) in varied context of use. However, the findings of the current study highlight the fact that learning of the latter aspects may require more focused instruction.

The study targeted productive (rather than receptive) knowledge of the most frequent words. To the best of my knowledge, no study, to date, has measured the depth and strength

¹¹ English as an Additional Language

¹² English to Speakers of Other Languages

of productive knowledge of the most frequent words with the same detailed and purposeful methodology. In contrast to the many studies investigating receptive vocabulary knowledge, relatively few studies that investigate productive vocabulary knowledge can be found (see Pearson, Hiebert, & Kamil, 2007). The current study not only enriches the available literature but also sheds light on the significance of measuring productive vocabulary knowledge.

Last but not least, the instruments developed for this study (the designed depth tests) are currently available in a paper-based format. This helps teachers administer the tests without the need for any special equipment in classrooms. The tests can be administered on different occasions and can be scored in the classrooms. This could provide an opportunity for learners to have a better understanding of the purpose of the tests.

Pedagogical Implications of the Study

This study presents a clearer picture of Iranian advanced EAP students' derivational, associational, and collocational knowledge of the most frequent words. The study informs language teachers and learners of the need to pay greater attention to different productive aspects of vocabulary knowledge. Teachers may help learners through the use of different teaching methods or learning activities to overcome their weaknesses in their knowledge of word parts, associations, or collocations. Among all available activities and procedures that improve students' derivational, associational, and collocational knowledge, Nation's (2013) suggestions are particularly helpful and practical.

Nation (2013) classified activities for vocabulary learning according to the different aspects of knowing a word (see Chapter 2). Rich instruction involves using several of activities to enrich students' vocabulary knowledge. The following adaptation of Nation's (2013) table of a range of activities for vocabulary learning shows suggestions of activities

that may help to develop derivational, associational, and collocational vocabulary knowledge.

Table 5.1

A Range of Activities for Vocabulary Learning

Form	word parts	Filling word part tables Cutting up complex words Building complex words Choosing a correct form Finding etymologies
Meaning	associations	Finding substitutes Explaining connections Making word maps Classifying words Finding opposites Suggesting causes or effects Suggesting associations Finding examples
Use	collocates	Matching collocates Finding collocates Analyzing and classifying collocates

Note. An adaptation of Nation's (2013) table of a range of activities for vocabulary learning

The following is an explanation of activities that may be useful to further develop knowledge of word parts, associations, and collocations.

Word parts

Knowledge of word parts enables learners to use different members of the word families they know, and also helps them remember new complex words. Nation (2013) discusses five activities to teach and improve word parts knowledge.

Filling word part tables. This activity is similar to the way the word parts aspect is measured in the current study. Learners are required to work in pairs to complete a table like the following. All spaces may not necessarily be filled. Learners check their answers with one another before the teacher provides the answer.

noun	verb	adjective	adverb
argument			
	evaluate		
		distinct	
			normally

Cutting up complex words. Learners are provided with a list of words that they divide into parts. They can also be required to give the meaning of some of the parts. For example, the teacher may ask learners to divide *effortless, disappointed, misuse, disrespectful* into parts. Learners can also be asked to give the meaning of word parts such as *-less, dis-, mis-, -ful*.

Building complex words. Learners are provided with word stems and are required to make words (using e.g., *mis-, dis-, un-*) from them. For example, words such as *use, respect, employed* are provided and learners are asked to make their negatives such as *misuse, disrespect, unemployed*.

Choosing the correct form. Learners are provided with sentences containing a blank and a word stem in brackets. They are required to change the stem to the appropriate inflectional and derivational form to fill in the blank. The following is an example.

I went to the doctor for a _____ (consult).

Finding etymologies. Learners are required to find etymologies in dictionaries to investigate how new words are connected to what they already know. For example, Online Etymology Dictionary shows *disorganize* is from French *désorganiser*, from *dés-* (not) plus *organiser* (organize).

Associations

Associations help students to understand the full meaning of words and recall the word forms or meanings in contexts. Nation (2013) believes associations are to a great extent the result of various meaning systems the words fit into. These include synonyms, antonyms, family members of the same headword, words in part-whole relationship, and superordinate and subordinate words. The following explains his suggested activities to develop associational vocabulary knowledge.

Finding substitutes. Learners are required to choose words from a list to replace the target underlined words in a text or sentences. The following is an example of this activity.

Use the words in the box to find synonyms for the italicized words:

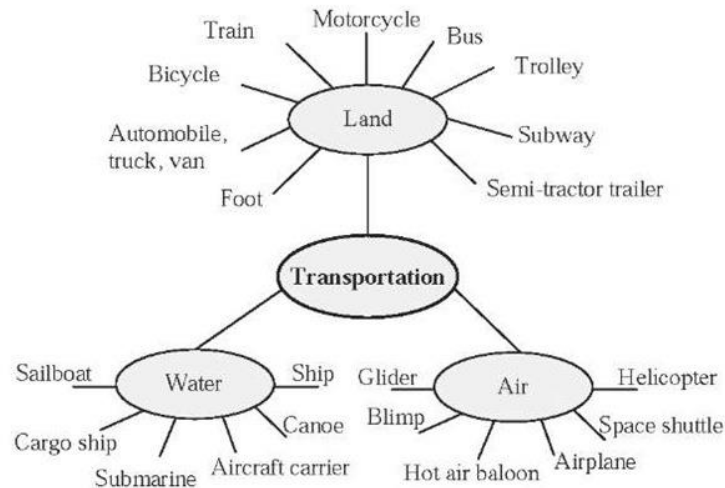
allow	wet	stop
angry	expensive	buy
thin	go down	fix

- My clothes are still *damp* so I'll have to wait for them to dry.
- They are starting to *descend* the mountain.
- He was *furious* when he saw the damage to the car.

Explaining connections. Learners are required to work in pairs to explain connections between a group of related words like the following:

analyse criteria exclude justify classify

Making word maps. Learners are required to work in pairs or groups to make a semantic map for a target word. The following is an example:



Classifying words. Learners are provided with lists of words to classify them into groups based on certain criteria—classifying words based on the availability of their positive or negative connotations, or whether the words are living or non-living. For example, words such as *birds, animals, insects, plants* can be classified as living while *furniture, cars, buildings* can be classified as non-living. Also, *professional* may remind someone of *skill* or *excellence* while *poverty* may remind someone of *misery* or *famine*.

Finding opposites. Learners are provided with lists of words from a text or sentences and are required to provide opposites for the words. For example, *elementary, presence,* and *deny* can be provided as opposites of *advanced, absence,* and *admit*.

Suggesting causes or effects. Learners are required to provide causes to go with effects and effects to go with causes. For example, for a phrase like *medical consultation* learners may think of some causes such as *illness, pain, tiredness* and some effects such as *medicine, hospital, reassurance*.

Suggesting associations. Learners are provided with four or five words. They are required to work in groups to provide associates for those words. Then, they scramble the words and give them to another group who are required to classify the associates under the

aforementioned four or five words. Groups compare and discuss their classifications. For example, in a list of four words, *poultry* is provided. Learners work together to provide associates for that such as *chicken, turkey, duck, meats*. These words will be scrambled and will be given to another group of learners to classify the provided associates under *poultry* again.

Finding examples. Learners are provided with a list of categories such as *food, household objects, numbers, jobs*, and so on. Each learner needs to choose one category. Then, the learner is required to write as many words as possible under the chosen category. For example, *food* can contain items like *bread, meat, and vegetable*. The learner then passes the list of provided words to the next learner who tries to provide words not already provided by the first learner. The list of words is passed on until all learners have their original lists back. The provided lists of words become a class dictionary.

Collocation

Knowing which words can naturally and habitually occur with other words helps learners to use language fluently. The following are some suggestion for improving collocations.

Matching collocates. Learners are provided with lists of nodes (target words) and collocates to match. For example, learners will be asked to match *sugar, crime, mistake*, and *hair* with *commit, blood, blond*, and *make*.

Finding collocates. Learners are required to look in dictionaries, rely on their experience, use concordancers, and use parallels with their first language to find collocates for given nodes (target words). For example, words such as *crime, account, risk*, and *bed* are provided for learners. Using the different sources, they are asked to find appropriate collocations such as *commit a crime, open an account, take a risk*, and *make the bed*.

Analyzing and classifying collocates. Learners are required to group collocates to see whether they share the same general meanings. For example, collocations such as *open an account*, *forgive a debt*, *land a deal*, and *receive a discount* can be grouped together because they can share some general meanings.

All these activities could help learners increase and improve their derivational, associational, and collocational vocabulary knowledge. Including intentional and systematic instruction of synonyms, antonyms, word classes, and collocations may help to reduce the chance that students' knowledge of these aspects lags behind form and meaning. There are also books designed to help students develop different aspects of vocabulary knowledge, e.g. *Focus on Vocabulary* by Schmitt, Schmitt, and Mann (2011). Such instructional emphasis on high frequency words may help to improve the speaking and writing skills of EAP students.

Recommendations for Further Research

The current study shows that depth aspects were not known to the same extent as the form-meaning aspect (breadth). One of the outcomes of such a difference (between breadth and depth knowledge) could be the limited performance in oral and written English. Further studies are needed to investigate why, for such high frequency words and such advanced students, there are still differences between size (the knowledge of form-meaning) and a variety of depth measures. The potential reasons for this issue need to be investigated because, as discussed before, vocabulary knowledge has a significant impact on learners' speaking and writing. Further studies are needed to investigate what aspects language learners consider involved in vocabulary knowledge and whether they are informed of the various aspects involved in vocabulary knowledge such as word parts, associations, or collocations.

Research is also needed to investigate the productive depth of vocabulary knowledge of the AWL (Coxhead, 2000) or the high frequency 2,000 words. For such studies, the selection of participants seems to be very important. The potential participants need to be proficient enough because learners with a small vocabulary size cannot be expected to perform reasonably on depth aspects. Furthermore, a considerable number of target words are needed from one specific group or level of frequency to better generalize the findings to that specific group.

It seems also worthwhile to investigate to what extent learners other than EAP students, such as EFL, ESL, ESOL, or EAL students, have knowledge of the most frequent 1,000 words. For all these learners, parallel versions of the instruments of the study can be developed. One version of the tests can be used to diagnose learners' weakness in depth of vocabulary knowledge prior to the purposeful instruction of word parts, associations, and collocations of the most frequent words. After instruction, another parallel form (equivalent version) may be administered to see whether the learners have mastered depth of knowledge of the most frequent 1,000 words.

It is also worthwhile to develop electronic versions of the tests. The instruments of the study (the designed depth tests) are currently available only in the paper-based format. However, the tests could be written in a web-based format so that a large variety of language learners can take them at their own preference and ease and also receive prompt feedback on their results right after completing the tests. Such online tests can help researchers have a deeper understanding of the learners' quality and strength of knowing the most frequent 1,000 words because a much larger number of participants would take the tests.

Last but not least, the current study only measured the extent to which the learners have knowledge of four aspects. It is also useful to examine the relationship between the

knowledge of each aspect and overall language proficiency. In that case, some specific aspects can probably be recognized as indicators of language proficiency.

Limitations of the Study

The following issues seem to be the main limitations of the current study:

- a. Measuring depth of vocabulary knowledge, especially when it is measured productively, is time-consuming. This, in fact, is the nature of depth of vocabulary knowledge which takes time to be measured. As a result, any test that intends to measure it may fail to test a large number of target words, and this study is not an exception. In addition, measuring all of the aforementioned nine aspects of vocabulary knowledge, although well worth the effort, seems to be unfeasible. For this reason, the study only focused on four aspects of vocabulary knowledge. However, any effort to look at other aspects of vocabulary knowledge in future studies is very useful. In fact, the most comprehensive way of measuring vocabulary knowledge is measuring all nine aspects.
- b. The designed productive tests do not seem suitable for low proficiency English learners because measuring some of the contextual types of word knowledge such as collocations or associations is possible only when the words are more established at higher levels of proficiency (Schmitt, 2010). That is the only way to determine the degree of higher-level mastery of lexical items.
- c. The scoring of the collocation test seems to be challenging for teachers who are not familiar with corpora (either COCA or BNC). Teachers need

some practice and guidelines to be able to look up the value of *t* score, *MI* score, and the frequency of co-occurrence of the provided collocates. Such a procedure is time-consuming and probably is not suitable to be carried out in a classroom setting.

- d. There does not seem to be a clear criterion of mastery to enable us to interpret the obtained scores from the depth tests yet. For example, 76.27% (94.59 items out of 124) performance on word parts test by the current participants shows that they do not have mastery over the word parts of the most frequent words, but it does not show whether their performance is strong or less than satisfactory. A criterion for mastery like the one available for the VLT (Schmitt et al., 2001), 86.66/100 or 26/30, can considerably help the interpretation of such scores.
- e. Considering the quantitative nature of the study, the number of participants plays an important role in finding more reliable results and in generalizing them to other contexts. No doubt the larger the pool of participants for such studies, the more trustworthy the results would be. Unfortunately, it was very challenging to find a larger sample size for the current research study, as the potential participants needed to be from a homogenous group with the same L1 at their advanced level of English proficiency. The participants needed to be willing to take 10 laborious vocabulary tests and interested enough in the findings of the study to provide all their potential knowledge and bring all their efforts to the data collection sessions. For these reasons, the study only recruited EFL participants from Iran, but it would be useful to examine knowledge of

learners from other contexts as well. In case of using the tests on a larger scale, however, the tests need further development and validation. The instruments of the study (the productive tests) were designed for the purpose of the current study, and they seemed to be appropriate to address the research questions of the study. Evidence of reliability and validity needs to be searched for if the tests are to examine knowledge of learners from other contexts.

References

- Aitchison, J. (2003). *Words in the mind*. Oxford: Blackwell Publishing.
- Alberchtsen, D., Haastrup, K., & Henriksen, B. (2008). *Vocabulary and writing in a first and second language: Procedures and development*. Basingstoke: Palgrave Macmillan.
- Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77-117). Newark, DE: International Reading Association.
- Angelova, M., & Riazantseva, A. (1999). If you don't tell me, how can I know?: A case study of four international students learning to write the US way. *Writing Communication, 16*(4), 491-525.
- Aviad, T. L., & Laufer, B. (2013). Lexical properties in the writing of foreign language learners over eight years of study: single words and collocations. In C. Bardel, C. Lindquist, & L. Laufer (Eds.), *L2 vocabulary acquisition, knowledge and use new perspectives on assessment and corpus analysis* (pp. 127-148). European Second Language Association.
- Barcroft, J., & Rott, S. (2010). Partial word form learning in the written mode in L2 German and Spanish. *Applied Linguistics, 31*, 623–650.
- Barfield, A., & Gyllstad, H. (2009). *Researching collocations in another language*. New York: Palgrave Macmillan.
- Bauer, L., & Nation, P. (1993). Word families. *International Journal of Lexicography, 6*(4), 253-79.
- Berko, J. (1958). The child's learning of English morphology. *Word, 14*, 150–177.
- Beglar, D. (2010). A rasch-based validation of the vocabulary size test. *Language Testing*,

27(1), 101-118.

Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequence and perceived oral proficiency: putting a lexical approach to the test.

Language Teaching Research, 10(3), 245–261

Brants, T., & Alex, F. (2006). Web 1T 5-gram Version 1. Linguistic Data Consortium, Philadelphia, PA. Retrieved from

<http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006T13>

Brezina, V., & Gablasova, D. (2013). Is there a core general vocabulary? Introducing the New General List. *Applied Linguistics*, 1-23.

Brown, D. (2014). Knowledge of collocations. In J. Milton, & T. Fitzpatrick (Eds.),

Dimensions of vocabulary knowledge (pp 123-139). UK: Palgrave Macmillan.

Carroll, J. B., Davies, P., & Richman, B. (1971). *The American heritage word frequency book*. New York: Houghton Mifflin, Boston American Heritage.

Carstairs-McCarthy, A. (2002). *An introduction to English morphology*. Edinburgh: Edinburgh University Press.

Carter, R. (2012). *Vocabulary: Applied Linguistic Perspectives*. New York: Routledge.

Chen, C., & Truscott, J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics*, 31(5), 693–713.

Chui, A. S. Y. (2006). A study of the English vocabulary knowledge of university students in Hong Kong. *Asian Journal of English Language Teaching*, 16, 1-23.

Cobb, T. (2012). Compleat Lexical Tutor. Retrieved from <http://www.lextutor.ca/>

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34 (2), 213-238.

Coxhead, A. (2008). Phraseology and English for academic purposes: challenges and

- opportunities. In F. Meunier, & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 149-161). Amsterdam: John Benjamins Publishing Company.
- Coxhead, A. (2011). The Academic Word List 10 years on: Research and teaching implications. *TESOL Quarterly*, 45(2), 355–62.
- Daller, H., Milton, J., & Treffers-Daller, J. (2007). Editors' introduction: Conventions, terminology and an overview of the book. In H. Daller, J. Milton, & J. Treffers-Daller (Eds.), *Modelling and assessing vocabulary knowledge* (pp. 1-32). Cambridge: Cambridge University Press.
- Dang, T. N. Y., & Webb, S. (2014). The lexical profile of academic spoken English. *English for Specific Purposes*, 33, 66-76.
- Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *IRAL*, 47, 157–177.
- Evans, S., & Green, C. (2007). Why EAP is necessary: A survey of Hong Kong tertiary students. *Journal of English for Academic Purposes*, 6, 3-17.
- Evert, S. (2010). Google Web 1T5 n-grams made easy (but not for the computer). *Proceedings of the 6th Web as Corpus Workshop (WAC-6)*, 32-40, Los Angeles, CA.
- Evert, S. (2015). An NLP approach to the evaluation of Web Corpora.
Retrieved from http://www.stefan-evert.de/PUB/Evert2015_WebEval_handout.pdf
- Francis, W. N., & Kucˇera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.
- Fitzpatrick, T. (2006). Habits and rabbits: word associations and the L2 lexicon. *EUROSLA Yearbook*, 6, 121- 46.
- Fitzpatrick, T. (2007). Word association patterns: Unpacking the assumptions. *International*

Journal of Applied Linguistics, 17(3), 319–31.

Fitzpatrick, T. (2013). Word associations. *The Encyclopedia of Applied Linguistics*.

Retrieved from

<http://onlinelibrary.wiley.com/doi/10.1002/9781405198431.wbeal1283/abstract?userIsAuthenticated=false&deniedAccessCustomisedMessage=>

Fitzpatrick, T., & Munby, I. (2014). Knowledge of associations. In J. Milton, & T.

Fitzpatrick (Eds.), *Dimensions of vocabulary knowledge* (pp 92-105). UK: Palgrave Macmillan.

Gan, Z. (2012). Understanding L2 Speaking Problems: Implications for ESL Curriculum

Development in a Teacher Training Institution in Hong Kong. *Australian Journal of Teacher Education*, 37(1), 43-59.

Gilner, L., & Morales, F. (2008). Corpus-based frequency profiling: Migration to a word list

based on the British National Corpus. *The Buckingham Journal of Language and Linguistics*, 1, 41–58.

Granger, S., & Paquot, M. (2008). Disentangling the phraseological web. In S. Granger, & F.

Meunier (Eds.), *Phraseology: an interdisciplinary perspective* (PP. 27-49). Amsterdam and Philadelphia: John Benjamins.

Gyllstad, H. (2013). Looking at L2 vocabulary knowledge dimensions from an assessment

perspective - challenges and potential solutions. In C. Bardel, C. Lindquist, & L.

Laufer (Eds.), *L2 vocabulary acquisition, knowledge and use new perspectives on assessment and corpus analysis* (pp. 11-28). European Second Language Association.

Halliday, M. A. K. (1966). Lexis as a linguistic level. In C. E. Bazell et al (Eds.), *In Memory*

of J.R. Firth (pp.150-161). London: Longman.

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second*

Language Acquisition, 21, 303-317.

Hirsh, D., & Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, 8, 689-96

Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

Hyland, K. (1997). Is EAP necessary? A survey of Hong Kong undergraduates. *Asian Journal of English Language Teaching*, 7, 77-99.

Ishii, T., & Schmitt, N. (2009). Developing an integrated diagnostic test of vocabulary size and depth. *RELC*, 40(1), 5-22.

Jackson, H., & Ze Amvela, E. (2000). *Words, meaning and vocabulary*. London: Continuum.

Jaen, M. M. (2007). A corpus-driven design of a test for assessing the ESL collocational competence of university students. *IJES*, 7 (2), 127-147.

Koizumi, R. & In'nami, Y. (2013). Vocabulary knowledge and speaking proficiency among second language learners from novice to intermediate levels. *Journal of Language Teaching and Research*, 4(5), 900-913.

Laufer, B., & Nation, P. (1999). A Vocabulary Size Test of Controlled Productive Ability. *Language Testing*, 16(1), 36-55.

Laufer, B., & Nation, P. (2001). Passive vocabulary size and speed of meaning recognition: Are they related? In S. Foster-Cohen, & A. Nizegorodcew (Eds.), *EUROSLA yearbook* (pp. 7-28). Amsterdam and Philadelphia: John Benjamins.

Laufer, B., Elder, C., Hill, K., & Congdon, P. (2004). Size and strength: do we need both to measure vocabulary knowledge? *Language Testing*, 21(2), 202-226.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: size, strength, and computer adaptiveness. *Language Learning*, 54(3), 399-436.

- Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, 61(2), 647- 672.
- Lea, M.R., & Street, B.V. (1999). Writing as academic illiteracies: Understanding textual practices in higher education. In C. Candlin, & K. Hyland (Eds.), *Writing: Texts, processes and practices* (pp. 62 -81). London: Longman.
- Liu, L. (2007). *Acquiring English for academic purposes: Challenges ESL students experience in an English-Speaking Canadian university* (Unpublished doctoral dissertation). University of Regina: Regina, Saskatchewan.
- Liu, M., & Jackson, J. (2008). An exploration of Chinese EFL learners' unwillingness to communicate and foreign language anxiety. *Modern Language Journal*, 92, 71-86.
- Llach, M. P. A., & Espinosa, S. M. (2014). Knowledge of form and meaning. In J. Milton, & T. Fitzpatrick (Eds.), *Dimensions of vocabulary knowledge* (pp. 60-72). UK: Palgrave Macmillan.
- McCarthy, M. (1990). *Vocabulary*. Oxford: Oxford University Press.
- McEney, T., & Wilson, A. (2001). *Corpus linguistics: An introduction* (2nd ed.). Edinburgh, UK: Edinburgh University Press.
- Mantyla, K., & Huhta, A. (2014). Knowledge of word parts. In J. Milton, & T. Fitzpatrick (Eds.), *Dimensions of vocabulary knowledge* (pp. 45-59). UK : Palgrave Macmillan.
- Meara, P. (1983). Word associations in a foreign language. *Nottingham Linguistics Circular*, 11, 29-38.
- Meara, P. (1984). The study of lexis in interlanguage. In A. Davies, C. Criper, & A. P. R. Howatt (Eds.), *Interlanguage* (pp. 225-235). Edinburg: University of Edinburg Press.
- Meara, P., & Jones, G. (1988). Vocabulary size as a placement indicator. In P. Grunwell (Ed.), *Applied linguistics in society* (pp. 80-87). London: CILT.

- Meara, P., & Jones, G. (1990). *Eurocentres Vocabulary Size Test 10KA*. Zurich, Switzerland: Eurocentres.
- Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K., Malmkjaer, & J., Williams (Eds.), *Performance and competence in second language acquisition* (pp. 35-53). Cambridge: Cambridge University Press.
- Meara, P., & Fitzpatrick, T. (2000). Lex30: an improved method of assessing productive vocabulary in an L2. *System*, 28, 19-30.
- Meara, P., & Bell, H. (2001). P-Lex: a simple and effective way of describing the lexical characteristics of short L2 texts. *Prospect*, 16(3), 323-37.
- Meara, P., & Wolter, B. (2004). V_Links: Beyond vocabulary depth. In D. Albrechtsen, K. Haastrup, & B. Henriksen (Eds.), *Angles on the English-speaking world* (pp. 85-96). Copenhagen: Museum Tusulanum.
- Meara, P. (2005). Lexical frequency profiles: A monte carlo analysis. *Applied Linguistics*, 26(1), 32-47.
- Miller, G.A. (1999). On knowing a word. *Annual Review of Psychology*, 50, 1-19.
- Milton, J. (2009). *Measuring second language vocabulary acquisition*. UK: Short Run Press.
- Milton J., Wade, J., & Hopkins, N. (2010). Aural word recognition and oral competence in a foreign language. In R. Chacón-Beltrán, C. Abello-Contesse, & M. Torreblanca-López (Eds.), *Further insights into non-native vocabulary teaching and learning* (pp. 83-98). Bristol: Multilingual Matters.
- Milton, J. (2013). Measuring the contribution of vocabulary knowledge to proficiency in the four skills. In C. Bardel, C. Lindquist, & L. Laufer (Eds.), *L2 vocabulary acquisition, knowledge and use new perspectives on assessment and corpus analysis* (pp. 57-78). European Second Language Association.

- Mitchell, T. M., Shinkareva, S. V., Carlson, A., Chang, K. M., Malave, V. L., Mason, R. A., & Just, M. A. (2008). Predicting human brain activity associated with the meanings of nouns. *Science*, *320*, 1191–1195.
- Mochizuki, M., & Aizawa, K. (2000). An affix acquisition order for EFL learners: An exploratory study. *System*, *28*, 291-304.
- Morgan, C. L., & Bonham, D. N. (1944). Difficulty of vocabulary learning as affected by parts of speech. *Journal of Educational Psychology*, *35*, 369–377.
- Mukarto, F. X. (2003). *Assessing the depth of second language vocabulary knowledge*. Paper presented at the 38th RELC International Seminar, Singapore.
- Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly*, *19*(3), 304-330.
- Nagy, W. E., Diakidoy, I. N., & Anderson, R. C. (1991). The development of knowledge of derivational suffixes. Retrieved from https://www.ideals.illinois.edu/bitstream/handle/2142/17820/ctrstreadtechrepv01991i00536_opt.pdf?sequence=1
- Nagy, W. E., Diakidoy, I. N., & Anderson, R. C. (1993). The acquisition of morphology: learning the contribution of suffixes to the meanings of derivatives. *Journal of Reading behaviour*, *25*(2), 155-170.
- Nguyen, T. M. H., & Webb, S. (2016). Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research*, 1-23.
- Nakata, T. (2013). *Optimising second language vocabulary learning from flashcards* (Unpublished doctoral thesis). Victoria University of Wellington, Wellington, New Zealand.
- Nassaji, H. (2004). The relationship between depth of vocabulary knowledge and L2

- learners' lexical inferencing strategy use and success. *The Canadian Modern Language Review*, 61(1), 107-134.
- Nation, I. S. P. (1983). Testing and teaching vocabulary. *Guidelines*, 5, 12–25.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. UK: Cambridge University Press.
- Nation, I. S. P. (2004). A study of the most frequent word families in the British National Corpus. In P. Bogaards, & B. Laufer (Eds.), *Vocabulary in a Second Language* (pp. 3-14). Amsterdam: John Benjamins.
- 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., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9-13.
- Nation, I. S. P., & Webb, S. (2011). *Researching and analyzing vocabulary*. USA: Heinle.
- Nation, I. S. P. (2012). The BNC/COCA word family lists. Retrieved from <http://www.victoria.ac.nz/lals/about/staff/stuart-webb>
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd edition). UK: Cambridge University Press.
- Nesselhauf, N. (2004). Learner corpora and their potential for language teaching. In J. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 125-152). Amsterdam: John Benjamins.
- Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam: John Benjamins.
- Nyssonen, L. (2008). *Understanding English word formation: a study among 6th-grade*

- pupils in Finish comprehensive school* (Unpublished MA Thesis). University of Jyvaskyla, Finland.
- Pacheco, E. C. R. (2005). *Affixes as a strategy for vocabulary acquisition in a first year ESL college reading course* (Unpublished doctoral dissertation). University of Puerto Rico, Puerto Rico.
- Paribakht, T. S., & Wesche, M. (1993). Reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada Journal*, 11(1), 9-29.
- Pearson, P. D., Hiebert, E. H., & Kamil, M. L. (2007). Vocabulary assessment: What we know and what we need to learn. *Reading Research Quarterly*, 42(2), 282-296.
- Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, 18(1), 1-28.
- Politzer, R. L. (1978). Paradigmatic and syntagmatic associations of first year French students. In V. Hausa (Ed.), *Papers in linguistics and child language* (pp. 203-210). The Hague, Netherlands: Mouton.
- Qian, D. D. (1998). *Depth of vocabulary knowledge: assessing its role in adults' reading comprehension in English as a second language* (Unpublished doctoral dissertation). University of Toronto, Toronto, Canada.
- Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review*, 56, 282-308.
- Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, 52(3), 513-536.
- Qian, D. D., & Schedl, M. (2004). Evaluation of an in-depth vocabulary knowledge measure

- for assessing reading performance. *Language Testing*, 21(1), 28-52.
- Richards, J. C. (1974). The role of vocabulary teaching. *TESOL Quarterly*, 10(1), 77-89.
- Read, J. (1988). Measuring the vocabulary knowledge of second language learners. *RELC Journal*, 19(2), 12-25.
- Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing*, 10(3), 355-371.
- Read, J. (1995). Refining the word associates format as a measure of depth of vocabulary knowledge. *New Zealand Studies in Applied Linguistics*, 1, 1-17.
- Read, J. (1998). Validating a test to measure depth of vocabulary knowledge. In A. Kunnan (Ed.), *Validation in language assessment* (pp. 41-60). Mahwah, NJ: Lawrence Erlbaum.
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Read, J. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined? In P. Bogaards, & B. Laufer (Eds.), *Vocabulary in a second language. Selection, acquisition, and testing* (pp. 209-227). Amsterdam/Philadelphia, PA: John Benjamins.
- Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework: Word associations and verbal suffixes. *SSLA*, 20, 17-36.
- Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language Learning*, 48(2), 281-317.
- Schmitt, N. (1999). The relationship between TOEFL vocabulary items and meaning, association, collocation and word class knowledge. *Language Testing*, 16, 189-216.
- Schmitt, N. (2000). *Vocabulary in language teaching*. New York: Cambridge University Press.

- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, 18 (1), 55–88.
- Schmitt, N., & Zimmerman, C. B. (2002). Derivative word forms: What do learners know? *TESOL Quarterly*, 36(2), 145-171.
- Schmitt, N. (2008). Review article, Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329-363.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. UK: Palgrave Macmillan.
- Schmitt, N., Schmitt, D., & Mann, D. (2011). *Focus on Vocabulary*. USA: Pearson Education.
- Schmitt, N., & Schmitt, D. (2012). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 47(4), 484-503.
- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 1-39.
- Schoonen, R., & Verhallen, M. (2008). The assessment of deep word knowledge in young first and second language learners. *Language Testing*, 25(2), 211-236.
- Schoonen, R. (2010). *The development of lexical proficiency knowledge and skill*. Paper presented at the Copenhagen Symposium on Approaches to the Lexicon, Copenhagen.
- Shen, Z. (2008). The roles of depth and breadth of vocabulary knowledge in EFL reading performance. *Asian Social Science*, 4(12), 135-137.
- Shin, D., & Nation, P. (2008). Beyond single words: The most frequent collocations in spoken English. *ELT Journal*, 62(4), 339-348.
- Simpson-Vlach, R., & Ellis, N.C. (2010). An academic formulas list: New methods in

- phraseology research. *Applied Linguistics*, 31(4), 487-512.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Soderman, T. (1993). Word associations of foreign language learners and native speakers: Different response types and their relevance to lexical development. In B. Hammerberg (Ed.), *Problems, process and product in language learning* (pp. 157-169). Abo Finland: Abo Academic.
- Sokman, A. (1993). Word associations results: A window to the Lexicons of ESL students. *JALT Journal*, 15(2), 135-150.
- Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, 36(2), 139-152.
- Stæhr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *SSLA*, 31, 577-607.
- Stubbs, M. (1995). Collocations and semantic profiles: On the cause of the trouble with quantitative methods. *Functions of language*, 2 (1), 1–33.
- Tang, R. (2012). *Academic writing in a second or foreign language*. UK: Continuum International Publishing Group.
- Thornbury, S. (2002). *How to teach vocabulary*. Harlow: Longman.
- Ward, J., & Chuenjundaeng, J. (2009). Suffix knowledge: Acquisition and applications. *System*, 37(3), 461-469.
- Webb, S. (2002). *Investigating the effects of learning tasks on vocabulary knowledge* (Unpublished doctoral dissertation). Victoria University of Wellington, Wellington, New Zealand.
- Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition*, 27(1), 33-52.

- Webb, S. (2007a). Learning word pairs and glossed sentences: The effects of a single context on vocabulary knowledge. *Language Teaching Research*, 11(1), 63-81.
- Webb, S. (2007b). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28(1), 46-65.
- Webb, S. (2007c). The effects of synonymy on vocabulary learning. *Reading in a Foreign Language*, 19, 120-136.
- Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition*, 30, 79-95.
- Webb, S. (2009a). The effects of pre-learning vocabulary on reading comprehension and writing. *Canadian Modern language Review*, 65, 441-470.
- Webb, S. (2009b). The effects of receptive and productive learning of word pairs on vocabulary knowledge. *RELC Journal*, 30(3), 360-376.
- Webb, S., & Rodgers, M. P. H. (2009a). The vocabulary demands of television programs. *Language Learning*, 59, 335-66.
- Webb, S., & Rodgers, M. P. H. (2009b). The lexical coverage of movies. *Applied Linguistics*, 30, 407-427.
- Webb, S., & Kagimoto, E. (2011). Learning collocations: Do the number of collocations, position of the node word, and synonymy affect learning? *Applied Linguistics*, 32(3), 259-276.
- Webb, S., & Chang, A. C. (2012). Second language vocabulary growth. *RELC Journal*, 43(1), 113 - 126.
- Webb, S. (2013). Depth of vocabulary knowledge. *The Encyclopedia of Applied Linguistics*. Retrieved from

<http://onlinelibrary.wiley.com/doi/10.1002/9781405198431.wbeal1325/abstract?userIsAuthenticated=false&deniedAccessCustomisedMessage=>

- Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning*, 63(1), 91–120.
- Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL - International Journal of Applied Linguistics*, 168(1), 34-70.
- Wesche, M., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. *The Canadian Modern Language Review*, 53(1), 13-40.
- West, M. (1953). *A general service list of English words: with semantic frequencies and a supplementary word-list for the writing of popular science and technology*. London: Longman.
- Wilkins, D. A. (1972). *Linguistics in Language Teaching*. London: Arnold.
- Wilks, C. (2009). ‘Tangled webs ... complications in the exploration of L2 lexical networks’. In T. Fitzpatrick, & A. Barfield (Eds.), *Lexical processing in second language learners* (pp. 25-37). Bristol: Multilingual Matters.
- Wolter, B. (2001). Comparing the L1 and L2 mental lexicon. *SSLA*, 23, 41 – 69.
- Wolter, B. (2002). Assessing proficiency through word associations: Is there still hope? *System*, 30(3), 315-329.
- Wolter, B. (2005). *V_Links: A new approach to assessing depth of word knowledge* (Unpublished Ph.D. dissertation). University of Wales, Swansea.
- Wolter, B., & Gyllstad, H. (2011). Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Applied Linguistics*, 32(4), 430-449.
- Wray, A. (2009). Conclusion: Navigating L2 collocation research. In A. Barfield, & H.

- Gyllstad (Eds.), *Researching collocations in another language: multiple interpretations* (pp. 232-244). Basingstoke: Palgrave Macmillan.
- Xue, G., & Nation, I.S.P. (1984). A university word list. *Language Learning and Communication*, 3(2), 215-229.
- Yule, G. (2006). *The study of language* (3rd ed.). Cambridge: Cambridge University Press.
- Zareva, A., & Wolter, B. (2012). The promise of three methods of word association analysis to L2 lexical research. *Second Language Research*, 28(1), 41-46.
- Zheng, Y. (2009). Exploring Chinese EFL learners' receptive and productive vocabulary knowledge: Implications for EFL vocabulary teaching. *The Journal of Asia TEFL*, 6(1), 163-188.

APPENDICES

Appendix A: Word Parts Test

Write different word classes (noun, verb, adjective, and adverb) of the following words. If there is more than one possibility (e.g., more than one adjective form), you only need to write one. **If there is no form, put an “X” in the box.** Please write your answers as clearly as possible.

For example:

	Noun	Verb	Adjective	Adverb
NATION	nation	nationalize	national	nationally
PAY	payment	pay	payable	X
FRESH	freshness	freshen	fresh	freshly

	Noun	Verb	Adjective	Adverb
1. ASSOCIATION				
2. BELIEVE				
3. SUPPOSE				
4. CONTINUE				
5. CREATE				
6. SPEAK				
7. MANAGER				
8. TRAINING				
9. REFLECT				
10. FORGET				
11. SOLUTION				
12. PRODUCT				
13. ARGUE				
14. UNDERSTAND				
15. TRADITIONAL				
16. RECOGNIZE				
17. SIGNIFICANT				
18. ENVIRONMENT				
19. SIMILAR				
20. IMAGINE				
21. PERSON				
22. DETERMINE				
23. SECURITY				
24. VARIOUS				
25. INFORMATION				
26. EFFORT				
27. DIRECTION				
28. MOVEMENT				
29. FINANCIAL				
30. ECONOMIC				

31. GOVERNMENT

Word Parts Test

Write different word classes (noun, verb, adjective, and adverb) of the following words. If there is more than one possibility (e.g., more than one adjective form), you only need to write one. **If there is no form, put an “X” in the box.** Please write your answers as clearly as possible.

For example:

	Noun	Verb	Adjective	Adverb
NATION	nation	nationalize	national	nationally
PAY	payment	pay	payable	X
FRESH	freshness	freshen	fresh	freshly

	Noun	Verb	Adjective	Adverb
1. ASSOCIATION	Association; associate	Associate; disassociate	Associative; disassociated; associated; associational; associate; associable	Associatively; X
2. BELIEVE	Belief; disbelief; believer; unbeliever; disbeliever; unbelief; believability; believing	Believe; disbelieve	Believable; unbelievable; disbelieving; unbelieving	Unbelievably; disbelievingly; unbelievably; believably
3. SUPPOSE	supposition	suppose	supposed	supposedly
4. CONTINUE	Continuance; continuity; continuation; continuum; discontinuation; discontinuance; continuer	Continue; discontinue	Continual; continuous; continued; continuing	Continually; continuously; continuingly
5. CREATE	Creation; creativity; creator; creationist; creationism; creature	Create; recreate	Creative; creationist, creational	creatively
6. SPEAK	Speaker; speech	speak	Spoken; unspoken; unspeakable; speaking; speakable, speechless	Unspeakably; X

7. MANAGER	Manager; management; mismanagement; manageress	Manage; mismanage	Manageable; unmanageable; managed; managerial; managing; unmanaged	Manageably; X; managerially
8. TRAINING	Training; trainee; trainer	train	Trained; untrained; trainable	X; trainably
9. REFLECT	Reflection; reflectiveness; reflector	reflect	Reflective; reflected; reflecting	reflectively
10. FORGET	Forgetfulness; forgetter	forget	Forgetful; forgotten; unforgotten; forgettable; unforgettable	Forgetfully; forgettably
11. SOLUTION	Solution; solvency; solvent	solve	Solvable; solvent; soluble; solved	X
12. PRODUCT	Product; productivity; production; produce; producer	Produce	Productive; unproductive	productively
13. ARGUE	Argument; arguer	argue	Arguable; unarguable; argumentative	Arguably; unarguably; argumentatively
14. UNDERSTAND	Understanding; misunderstanding	Understand; misunderstand	Understandable; understood; misunderstood; understanding; misunderstanding	Understandably; understandingly
15. TRADITIONAL	Tradition; traditionalism; traditionalist	X; traditionalize	Traditional; non- traditional; traditionalist; untraditional	traditionally
16. RECOGNIZE	recognition	recognize	Recognizable; unrecognizable	recognizably
17. SIGNIFICANT	Significance; signifier; signified	Signify	Significant; signifiable	significantly
18. ENVIRONMENT	Environment; environmentalist; environmentalism	X	environmental	environmentally
19. SIMILAR	Similarity; dissimilarity	X	Similar; dissimilar	Similarly
20. IMAGINE	Imagination; imaginary; image; imaging	imagine	Imaginative; unimaginative; imaginable; unimaginable; imaginal; imagined	Imaginatively; unimaginatively; imaginably; unimaginably

21. PERSON	Person; personality; personification	Personalize; impersonate; personify	Personal; personable	Personally; personably
22. DETERMINE	Determination; determiner; determinant	determine	Determined; indeterminable; determinable; determinative; determine; indeterminate	X; indeterminably; determinedly; determinately
23. SECURITY	Security; insecurity;	Secure; securitize	Secure; insecure; unsecured; securable; secured; securable	securely
24. VARIOUS	Variety; variability; variance; variate; variousness; variation	vary	Various; variable; varied	Variously; variably
25. INFORMATION	Information; informant; informer	inform	Informational; informed; informative; informable	X ; informatively
26. EFFORT	effort	X	Effortless; effortful	Effortlessly; effortfully
27. DIRECTION	Direction; directive; directness; director; directionality	direct	Directional; directionless; directive; direct; directed; directable; indirect;	Directly; direct; directionally; indirectly;
28. MOVEMENT	Movement; mover; movability; move	move	Moved; moving; unmoving; unmoved; moveable; unmovable; immovable; immovable; movable	Movably; movingly; immovably
29. FINANCIAL	Finance; financier; financing; financials	finance	Financial; financeable; financed	financially
30. ECONOMIC	Economy; economist; economics; economist	Economize	Economical; economic; uneconomic; uneconomical; economy	Economically; uneconomically
31. GOVERNMENT	Government; governor; governance; governorship	govern	Governmental; intergovernmental; governing; governable	governmentally

Appendix B: Synonym and Antonym Test

For the following words, write down a synonym and an antonym. Please write your answers as clearly as possible.

For example:

	Synonym	Antonym
FRIEND	buddy	enemy
REAL	actual	unreal
HOLD	keep	release

	Synonym	Antonym
1. ABILITY		
2. PRIVATE		
3. TRUE		
4. FINAL		
5. HUGE		
6. POPULAR		
7. NECESSARY		
8. ADVANTAGE		
9. DISCOVER		
10. DIFFERENT		
11. RECEIVE		
12. COME		
13. LEGAL		
14. CLEAR		
15. FULL		
16. IMPROVE		
17. REMOVE		
18. MODERN		
19. SERIOUS		
20. PUSH		
21. USEFUL		
22. HARD		
23. NORMAL		
24. QUICKLY		
25. WELL		
26. ALONE		

27. AWARE		
28. REDUCE		
29. PREVIOUS		
30. SURE		
31. MAIN		

Synonym and Antonym Test

For the following words, write down a synonym and an antonym. Please write your answers as clearly as possible.

For example:

	Synonym	Antonym
FRIEND	buddy	enemy
REAL	actual	unreal
HOLD	keep	release

	synonym	antonym
2. ability	calibre; capability; competence; skill; talent; aptitude; capacity; capability, capableness, capacity, competence, competency, faculty; skill, aptitude, capability, competence, expertise, proficiency, talent; capability; capacity; competence; intelligence; qualification; skill; strength; talent; understanding; expertise; ingenuity; proficiency	inability; incapability; weakness; disability; disability; inability; disability, inability, incapability, incapableness, incapacity, incompetence, incompetency, ineptitude, ineptness, ignorance; impotence; inability; incapability; incapacity; incompetence; ineptness; lack; weakness
3. private	personal; secret; hidden; intimate; non-public; nonpublic; discrete; behind-the-scenes, behind the scenes; confidential, esoteric, hushed, hush-hush, inside, intimate, nonpublic, privy, secret; backstairs, clandestine, covert, furtive, hole-and-corner, hugger-mugger, sneak, sneaking, sneaky, stealth, stealthy, surreptitious, undercover, underground, underhand, underhanded; idiomatic, individualized, particular, patented, peculiar, personal, personalized, individual, separate, singular, subjective, unique; exclusive, individual, intimate, own, personal, reserved, special; secret, clandestine, confidential, covert, hush-hush, off the record, unofficial; secluded, concealed, isolated, secret, separate, sequestered, solitary; personal; intimate; hidden; isolated; confidential; exclusive; independent; individual; secret; separate; special; quiet; secluded; soldier	public; open; common, open, public, overt; general, generic, popular, shared, universal; common; known; ordinary; public; open;
4. true	correct; real; accurate; valid; genuine; authentic; right, bona fide, certifiable, certified, echt, genuine, honest, pukka, real, right, sure-enough, authentic;	false; incorrect; fake, bogus, counterfeit, fake, false, mock, phony, pseudo, sham, spurious, suppositious, supposititious, unauthentic,

	<p>accurate, , dead-on, exact, good, on-target, precise, proper, right, so, spot-on, correct, veracious; concrete, de facto, effective, existent, factual, genuine, real, sure-enough, actual, very; constant, dedicated, devoted, devout, down-the-line, fast, good, loyal, pious, staunch, steadfast, steady, faithful, true-blue; accurate, authentic, exact, precise, right, strict, faithful, veracious; artless, genuine, honest, ingenuous, innocent, naive, natural, real, simple, sincere, guileless, unaffected, unpretending, unpretentious; documentary, hard, historical, literal, matter-of-fact, nonfictional, objective, factual; calculable, good, reliable, responsible, safe, secure, solid, steady, sure, tried, tried-and-true, dependable, trustable, trustworthy, trusty; archetypal, average, characteristic, normal, regular, representative, standard, typical; all right, decent, ethical, honest, honorable, just, moral, nice, right, righteous, right-minded, straight, good, upright, virtuous; correct, accurate, authentic, factual, genuine, precise, real, right, truthful, veracious faithful, dedicated, devoted, dutiful, loyal, reliable, staunch, steady, trustworthy exact, accurate, on target, perfect, precise, spot-on, unerring; real; valid; loyal; accurate; appropriate; authentic; bona fide; correct; genuine; honest; legitimate; natural; normal; perfect; proper; pure; sincere; truthful; typical; dedicated; devoted; faithful; honest; honorable; pure; reliable; sincere; steadfast; sure; worthy</p>	<p>unreal; false, improper, inaccurate, incorrect, inexact, off, untrue, wrong; conjectural, hypothetical, ideal, inexistent, nonexistent, platonic, possible, potential, suppositional, theoretical; disloyal, faithless, false, fickle, inconstant, perfidious, recreant, traitorous, treacherous, unfaithful, untrue; corrupt, corrupted, false, imprecise, inaccurate, inauthentic, inexact, loose, unfaithful; affected, artful, artificial, assuming, dishonest, dissembling, dissimulating, fake, false, guileful, insincere, phony, pretentious; fictional, fictionalized, fictitious, nondocumentary, nonfactual, nonhistorical, unhistorical; dodgy, uncertain, undependable, unreliable, unsafe, untrustworthy; aberrant, abnormal, anomalous, atypical, deviant, irregular, nonrepresentative, nontypical, untypical; bad, black, dishonest, dishonorable, evil, evil-minded, immoral, indecent, sinful, unethical, unrighteous, wicked, wrong; abnormal; affected; counterfeit; deceptive; different; dishonest; false; flawed; illegitimate; imprecise; improper; inaccurate; incorrect; insincere; lying; misleading; uncommon; unconventional; unfitting; unguine; unsuitable; untrustworthy; untruthful; unusual; wrong; apathetic; deceptive; dishonest; disloyal; false; inconstant; indefinite; treacherous; uncertain; unfaithful; unreliable; untrustworthy</p>
5. final	<p>last; end; ultimate; finished, bottommost, closing, concluding, last, hindmost, lag, latest, latter, rearmost, terminal, terminating, ultimate; certain, determinate, fixed, firm, flat, frozen, hard, hard-and-fast, inexpugnable, set, settled, stable; last, closing, concluding, latest, terminal, ultimate definitive, absolute, conclusive, decided, definite, incontrovertible, irrevocable, settled; ending; last; conclusive; definitive; closing; concluding; eventual; last-minute; ultimate; decisive; definite; finished</p>	<p>beginning; first; beginning, earliest, first, foremost, headmost, inaugural, initial, leadoff, maiden, opening, original, pioneer, primary, starting; indefinite</p>
6. huge	<p>enormous; gigantic; big; large; colossal; astronomical, brobdingnagian, bumper, colossal, cosmic, cyclopean, elephantine, enormous, galactic, gargantuan, giant, gigantesque, gigantic, grand, herculean, heroic, himalayan, humongous, immense, jumbo, king-size, king-sized, leviathan, mammoth, massive, mega, mighty, monster, monstrous, monumental, mountainous, oceanic, pharaonic,</p>	<p>small; tiny; miniature; bantam, bitty, diminutive, infinitesimal, lilliputian, little bitty, micro, microminiature, microscopic, midget, miniature, minuscule, minute, pocket, pygmy, teensy, teensy-weensy, teeny, teeny-weeny, tiny, wee; dwarfed; insignificant; limited; little; miniature; miniscule; minor;</p>

	planetary, prodigious, super, super-duper, supersize, supersized, titanic, tremendous, vast, vasty, walloping, whacking, whopping; large, colossal, enormous, gigantic, immense, mammoth, massive, monumental, supersize, tremendous, vast; extremely large; colossal; enormous; extensive; gargantuan; giant; great; humongous; immerse; magnificent; mammoth; massive; monstrous; monumental; towering; tremendous; vast	minute; narrow; poor; short; small; teeny; tiny; unimportant
7. popular	famous; favored; trendy; liked; in fashion; common; preferred; reputed, widespread, big, crowd-pleasing, du jour, faddish, faddy, fashionable, favorite, happening, hot, in, large, modish, pop, popularized, red-hot, vogue, voguish; conventional, customary, going, current, prevailing, prevalent, standard, stock, usual; common, majority, overall, general, prevailing, public, received, ruling, vulgar; democratic, republican, self-governing, self-ruling; affordable, accessible; affordable, bargain-basement, budget, cheapie, cheapo, chintzy, cut-price, cut-rate, dime-store, dirt cheap, el cheapo, inexpensive, low, low-end, cheap, reasonable; well-liked, accepted, approved, cool, fashionable, favorite, in, in demand, in favor, liked, phat, sought-after common, conventional, current, general, prevailing, prevalent, universal; well-known; favorite; common; standard; attractive; beloved; famous; fashionable; favored; prominent; suitable; trendy; accessible; familiar; prevalent; public; rampant; ubiquitous; universal	unknown; unpopular; disliked; out of fashion; out-of-fashion; infamous; out, unfashionable, unpopular; conventional, customary, going, current, prevailing, prevalent, standard, stock, usual; uncommon, unpopular; nondemocratic, undemocratic; costly, dear, deluxe, expensive, high, high-ticket, precious, premium, pricey, valuable; inconspicuous; obscure; old-fashioned; unfashionable; unknown; unpopular; limited; particular; uncommon; unknown
8. necessary	vital; crucial; essential; needed; mandatory; compulsory, forced, imperative, incumbent, involuntary, mandatory, nonelective, obligatory, peremptory, required; certain, ineluctable, ineludible, inescapable, inevitable, sure, unavoidable, unescapable; all-important, critical, imperative, indispensable, integral, must-have, essential, necessitous, needed, needful, required, requisite, vital; needed, compulsory, essential, imperative, indispensable, mandatory, obligatory, required, requisite, vital certain, fated, inescapable, inevitable, inexorable, unavoidable; essential; inevitable; basic; crucial; decisive; fundamental; imperative; mandatory; needed; paramount; required; significant; unavoidable; urgent; vital; certain; imminent; unavoidable	inessential; optional; unnecessary; frivolous; unneeded, extraneous, elective, optional, voluntary; avoidable, evadable, uncertain, unsure, dispensable, inessential, needless, nonessential, unessential, unnecessary, unneeded; additional; extra; inessential; minor; nonessential; optional; secondary; trivial; unimportant; unnecessary; voluntary; doubtful
9. advantage	benefit; asset; profit; upper hand; edge; lead; privilege; strength; better, bulge, catbird seat, drop, edge, high ground, inside track, jump, pull, stead, upper hand, vantage, whip hand; help, aid, asset,	disadvantage; weakness, penalty; loss; harm; disfavor; hurt; waste; disadvantage, drawback, handicap, liability, minus, penalty, strike; disadvantage, drawback, encumbrance,

	benefit, boon; benefit, ascendancy, dominance, good, help, lead, precedence, profit, superiority, sway; benefit; asset; choice; convenience; dominance; edge; favor; gain; improvement; influence; interest; lead; leverage; position; power; preference; profit; protection; recognition; return; superiority; support; upper hand; wealth	hindrance, impediment, minus; block; disadvantage; disapproval; disfavor; harm; hurt; impotence; incapacity; inferiority; injury; loss; subservience; waste; weakness
10. discover	find; detect; uncover; ascertain, catch on to, catch on, find out, get on to, get on, hear, learn, realize, see, wise up; wise, ascertain, descry, detect, determine, dig out, dig up, find, dredge up, dredge, ferret out, ferret, find out, get, hit on, hit upon, hit, hunt down, hunt up, hunt, learn, locate, nose out, root out, root, rout out, rout, rummage, run down, scare up, scout up, scout, track down, track, turn up; bare, disclose, reveal, divulge, expose, let on, let about, spill, tell, unbosom, uncloak, uncover, unmask, unveil; find, come across, come upon, dig up, locate, turn up, uncover, unearth find out, ascertain, detect, learn, notice, perceive, realize, recognize, uncover; find; uncover; catch; come upon; design; detect; determine; devise; disclose; explore; hear; identify; invent; learn; locate; notice; observe; realize; recognize; reveal; see; spot; unearth	lose; hide; conceal; miss, overlook, pass over; cloak, conceal, cover up, cover, enshroud, hide, mask, shroud, veil; conceal; hide; ignore; lose; miss; misunderstand; neglect; overlook power
11. different	dissimilar; unlike; odd, strange; contrasting; uncommon; unique, diverse, disparate, dissimilar, distant, distinct, distinctive, distinguishable, diverse, nonidentical, other, unlike, unlike; separate, individual, respective, unlike, altered, changed, contrasting, disparate, dissimilar, divergent, inconsistent, opposed, various, assorted, diverse, miscellaneous, sundry, varied unusual, atypical, distinctive, extraordinary, peculiar, singular, special, strange, uncommon, dissimilar, unlike, contrasting, disparate, distant, distinct, distinctive, divergent, diverse, offbeat; other; particular; peculiar; various; separate; distinct; bizarre; diverse; extraordinary; original; other; particular; peculiar; rare; special; specific; strange; unconventional; unique; unusual; various; miscellaneous; various; disparate; divergent; diverse; varied	same; similar; alike; ordinary; alike, identical, indistinguishable, kin, kindred, like, parallel, same, similar, same, alike; common; like; normal; ordinary; resembling; same; similar; standard; uniform; usual; common; commonplace; familiar; general; indefinite; like; normal; ordinary; regular; same; similar; standard; uniform; unimportant; usual; alike; like; same; similar
12. receive	get; accept; take; enter, admit, take; get, accept, acquire, be given, collect, obtain, pick up, take experience, bear, encounter, suffer, sustain, undergo, greet, accommodate, admit, entertain, meet, welcome; accept; collect; draw; earn; gain; get; hear; hold; make; obtain; pick up; pocket; reap; secure; seize; take; take in; win; endure; sustain; suffer;	send; give; deny; donate; ban, bar; deny; disallow; discourage; dissuade; drop; fail; forfeit; forsake; free; give; halt; let go; lose; misconceive; miss; misunderstand; offer; pass; refuse; reject; release; renounce; repulse; spend; stop; throw away; reject

	undergo; accept; bring in; introduce; meet; take in; welcome	
13. come	arrive; appear; approach; advance, approach, near, nigh, appear, arrive, get in, get through, land, show up, turn up, become, get, go, grow, run, turn, wax, be, befall, betide, chance, happen, come about, come down, come off, cook, do, go down, go on, hap, occur, pass, transpire, advance, go, come along, do, fare, forge, get along, get on, go along, go off, march, pace, proceed, progress, move towards, advance, approach, draw near, near, arrive, appear, enter, materialize, reach, show up, turn up, happen, fall, occur, take place result, arise, emanate, emerge, flow, issue, originate, reach, extend, be available, be made, be offered, be on offer, be produced, advance, approach, appear, arrive, become, enter, get, happen, hit, materialize, mover, occur, reach, show up, happen, break, develop, fall, occur, take place, extend, reach, become, develop, get, go, grow, join, run, turn,	go; leave; go away, go, recede from, recede, retreat, withdraw, go, leave, remain, stand, stay, stop, conceal, depart, disappear, discourage, dissuade, fail, go, hide, leave, lose, miss, stop, cease, decrease, halt, leave, lessen, stop, surrender
14. legal	lawful; legitimate; proper; sanctioned; right, allowed, just, lawful, legit, legitimate, licit, clean, fair, sportsmanlike, sportsmanly, legitimate, allowed, authorized, constitutional, lawful, licit, permissible, sanctioned, valid, judicial, forensic, juridical, allowable, permissible, constitutional, contractual, fair, juridical, lawful, legitimate, proper, statutory, valid,	illegal; unlawful; illegal, illegitimate, illicit, lawless, unlawful, wrongful, dirty, foul, nasty, unfair, unsportsmanlike, illegal, illegitimate, invalid, unacceptable, unlawful,
15. clear	understandable; transparent; unobstructed; free; luminous, clean, obvious, crystal, crystal clear, crystalline, limpid, liquid, lucent, pellucid, see-through, transparent, apparent, bald, bald-faced, barefaced, bright-line, broad, clear-cut, crystal clear, decided, distinct, evident, lucid, lustrous, luminous, manifest, nonambiguous, obvious, open-and-shut, palpable, patent, pellucid, perspicuous, plain, ringing, straightforward, transparent, unambiguous, unambivalent, unequivocal, unmistakable, assured, certain, cocksure, confident, doubtless, implicit, positive, sanguine, sure, bright, fair, cloudless, sunny, sunshiny, unclouded, absolute, conclusive, deciding, decisive, definitive, last, open, cleared, free, unclogged, unclosed, unobstructed, unstopped, blameless, clean handed, innocent, faultless, guiltless, impeccable, inculpable, irreproachable, lily-white, beaming, bedazzling, brilliant, candescent, bright, dazzling, effulgent, fulgent, glowing, incandescent, lambent, lucent, lucid, luminous, lustrous, radiant, refulgent, sheeny, shining, shiny, splendid, certain,	blurry, ambiguous; unclear; opaque; obstructed; foggy; cloudy, , cloudy, opaque, ambiguous, clouded, cryptic, dark, enigmatic, equivocal, indistinct, mysterious, nonobvious, obfuscated, obscure, unapparent, unclarified, unclear, unclouded, doubtful, dubious, uncertain, unsure, bleak, cloudy, dirty, foul, nasty, overcast, rainy, raw, rough, squally, stormy, sunless, tempestuous, turbulent, inconclusive, indecisive, unclear, blocked, clogged, closed, jammed, obstructed, plugged, shut, stopped, stuffed, uncleared, guilty, dim, dull, lackluster, unbright, unbrilliant, ambiguous, complicated, difficult, equivocal, fuzzy, imprecise, indefinite, indistinct, inexact, mistakable, obscure, questionable, uncertain, unclear, undefined, unintelligible, unsure, vague, doubtful, indefinite, uncertain, unsure, blame, charge, condemn, convict, hold, incriminate, sentence, discourage, dissuade, fail, forfeit, lose

	<p>convinced, decided, definite, positive, resolved, satisfied, sure, obvious, apparent, blatant, comprehensible, conspicuous, distinct, evident, manifest, palpable, plain, pronounced, recognizable, unmistakable, transparent, crystalline, glassy, limpid, pellucid, see-through, translucent bright, cloudless, fair, fine, light, luminous, shining, sunny, unclouded, unobstructed, empty, free, open, smooth, unhindered, unimpeded, unblemished, clean, immaculate, innocent, pure, untarnished, unblock, disentangle, extricate, free, loosen, open, rid, unload, pass over, jump, leap, miss, vault, brighten, break up, lighten, clean, cleanse, erase, purify, refine, sweep away, tidy, tidy up, wipe absolve, acquit, excuse, exonerate, justify, acquire, earn, make, reap, secure, cloudless, bright, fair, sunny, understandable, apparent, clear-cut, coherent, definite, distinct, evident, explicit, obvious, precise, sharp, simple, straightforward, transparent, unambiguous, unequivocal, unmistakable, open, unhindered, free, stark, transparent, apparent, clean, definite, positive, sure, clean, clear away, clarify, eliminate, erase, free, open, liberate, absolve, acquit, exonerate, release, pass over, negotiate, profit, get, make, receive, secure, win,</p>	
16. full	<p>packed; complete; filled; stuffed; whole; crammed; abounding; loaded, all, all of, all over, altogether, clean, completely, dead, enough, entire, entirely, even, exactly, fast, flat, fully, heartily, out, perfectly, plumb, quite, soundly, thoroughly, through and through, totally, utterly, well, wholly, wide, achingly, almighty, archly, awful, awfully, badly, beastly, blisteringly, bone, colossally, corking, cracking, damn, damned, dang, deadly, desperately, eminently, enormously, especially, ever, exceedingly, extra, extremely, fabulously, fantastically, far, fiercely, filthy, frightfully, very, greatly, heavily, highly, hugely, immensely, incredibly, intensely, jolly, majorly, mightily, mighty, monstrous, mortally, most, much, particularly, passing, rattling, real, really, right, roaring, roaringly, seriously, severely, so, sore, sorely, spanking, specially, stinking, such, super, supremely, surpassingly, terribly, that, thumping, too, unco, uncommonly, vastly, vitally, way, whacking, wicked, wildly, due, exactly, just, precisely, right, sharp, smack-dab, squarely, saturated, brimming, complete, filled, loaded, replete, satiated, stocked plentiful, abundant, adequate, ample, comprehensive, exhaustive, extensive, generous rich, clear, deep,</p>	<p>empty; half, halfway, incompletely, part, partially, partly, little, negligibly, nominally, slightly, somewhat, defective, deficient, empty, inadequate, incomplete, insufficient, lacking, wanting, broken, deficient, flawed, imperfect, incomplete, inferior, insufficient, limited, narrow, part, partial, restricted, short, unfinished, wanting</p>

	distinct, loud, resonant, rounded, plump, buxom, curvaceous, rounded, voluptuous, loose, baggy, capacious, large, puffy, voluminous, brimming, filled, adequate, big, chock-full, complete, crowded, entire, intact, packed, stocked, sufficient, thorough, absolute, broad, complete, comprehensive, detailed, entire, exhaustive, extensive, generous, maximum, perfect, whole, clear, rich	
17. improve	enhance; update; better; ameliorate, correct, ameliorate, amend, better, enhance, enrich, help, meliorate, perfect, refine, upgrade, enhance, advance, better, correct, help, rectify, touch up, upgrade progress, develop, make strides, pick up, rally, rise, advance, better, boost, correct, develop, enhance, help, increase, lift, progress, promote, raise, recover, reform, revamp, revise, rise, upgrade	worsen; ruin; deteriorate; decline; degrade; worsen, blow, decline, decrease, demote, depress, descend, deteriorate, diminish, drop, halt, hinder, hurt, lessen, lose, lower, recede, repress, retreat, retrogress, stop, worsen,
18. remove	take, extract, discard; take away; eliminate; erase; subtract; displace, lead, length, distance, spacing, spread, stretch, way, doff, douse, peel off, peel, put off, shrug off, take off, clear, draw, take out, withdraw, budge, dislocate, displace, disturb, relocate, move, reposition, shift, transfer, transpose, ax, axe, bounce, can, cashier, discharge, fire, muster out, pink-slip, release, dismiss, retire, sack, terminate, turn off, take away, take off, take out, abolish, delete, detach, displace, eject, eliminate, erase, excise, extract, get rid of, wipe from the face of the earth, withdraw, dismiss, depose, dethrone, discharge, expel, oust, throw out, move, depart, flit, relocate, take off, away, abolish, clear away, cut out, delete, discard, discharge, dismiss, eliminate, erase, evacuate, expel, extract, get rid of, oust, pull out, raise, separate, ship, take out, transfer, transport, withdraw, kill, clear away, dispose of, eliminate, eradicate, erase, exclude, get rid of, take out,	add; keep; put back; replace, fill, don, put on, slip into, slip, throw on, throw, place, position, put, employ, engage, hire, retain, sign up, sign on, sign, take on, accept, add, allow, continue, employ, engage, hire, hold, include, insert, keep, load, permit, put in, ratify, remain, retain, welcome, accept, add, include, keep, welcome
19. modern	new; contemporary; current; hip; recent, present, contemporary, current, designer, hot, mod, modernistic, new, new age, newfangled, new-fashioned, new fashioned, present-day, present day, red-hot, red hot, space-age, space age, state-of-the-art, state of the art, ultramodern, up-to-date, up to date, up-to-the-minute, up to the minute, modernist, ultramodernist, current, contemporary, fresh, new, newfangled, novel, present-day, recent, up-to-date, new, up to date, contemporary, current, modernized, present day, state of the art, stylish	old; old-fashioned; old fashioned; ancient; outdated; historic; antique, traditional, antiquated, archaic, dated, fusty, musty, oldfangled, old-fashioned, old-time, old time, out-of-date, out of date, passé, antediluvian, dodo, fogy, fossil, fuddy-duddy, reactionary, stick-in-the-mud, old, old fashioned, past
20. serious	important; solemn; critical; grave; sincere; somber, genuine, earnest, grave, humorless, no-nonsense, po-faced, sedate, severe, sober, sobersided, solemn,	unimportant; trivial; deceptive; frivolous; silly; playful, goofy, facetious, flip, flippant, humorous, jesting, jocular, joking, kittenish,

	<p>staid, uncomic, unsmiling, weighty, grave, heavy, weighty, grave, grievous, hazardous, jeopardizing, menacing, parlous, perilous, risky, dangerous, threatening, unhealthy, unsafe, venturesome, severe, acute, critical, dangerous important, crucial, fateful, grim, momentous, no laughing matter, pressing, significant, urgent, worrying solemn, grave, humorless, sober, unsmiling sincere, earnest, genuine, honest, in earnest, somber, humorless, deliberate, genuine, honest, severe, sincere, thoughtful, crucial, weighty, dangerous, deep, difficult, far-reaching, grievous, important, major, meaningful, severe, significant, tough, urgent</p>	<p>ludic, playful, light, unserious, harmless, innocent, innocuous, nonhazardous, nonthreatening, safe, unthreatening, deceptive, dishonest, false, insincere, misleading, untrustworthy, calm, easy, facile, friendly, insignificant, little, meaningless, minor, small, trivial, unimportant, worthless, funny</p>
21. push	<p>press; shove; force; urge; effort, drive, bandwagon, blitz, cause, crusade, drive, juggernaut, movement, campaign, drive, propel, shove, thrust, bore, bull, bulldoze, crash, elbow, jam, jostle, muscle, shoulder, squeeze, shove, depress, drive, press, propel, ram, thrust, make one's way, force one's way, elbow, jostle, move, shoulder, shove, squeeze, thrust, urge, encourage, hurry, impel, incite, persuade, press, spur, shove, butt, nudge, thrust, drive, ambition, dynamism, energy, enterprise, go, initiative, vigor, vitality, advance, assault, attack, drive, effort, offensive, thrust, drive, initiative, thrust, accelerate, bump, depress, drive, force, launch, move, nudge, pressure, propel, shift, shove, incite, urge, encourage, pressure, speed, spur, advertise, promote, advance, boost</p>	<p>pull; hesitation, discourage, dissuade, hinder, hold, keep, remain, retard, stop, discourage, dissuade, leave alone, decrease, halt</p>
22. useful	<p>helpful; effective; handy; usable; useable; actionable, applicable, applicative, applied, functional, practicable, serviceable, ultrapractical, usable, practical, workable, working, available, employable, exploitable, fit, functional, operable, practicable, serviceable, usable, conducive, facilitative, helpful, helpful, advantageous, beneficial, effective, fruitful, practical, profitable, serviceable, valuable, win-win, worthwhile, beneficial, valuable, advantageous, appropriate, convenient, effective, favorable, fruitful, good, handy, helpful, practical, pragmatic, profitable, proper, suitable,</p>	<p>useless; unuseful, unnecessary, impracticable, impractical, inapplicable, nonpractical, unusable, unworkable, useless, impracticable, inoperable, nonfunctional, unavailable, unemployable, unusable, unhelpful, useless, awkward, bad, disadvantageous, disagreeable, harmful, hurtful, impotent, improper, inappropriate, incapabale, inconvenient, injurious, misbehaving, unacceptable, unfitting, unhandy, unhelpful, unproductive, unskilled, unsuitable, unsuited, useless, worthless</p>
23. hard	<p>solid; tough; difficult; firm; strong, affectless, callous, case-hardened, cold-blooded, compassionless, desensitized, hard-boiled, hard-hearted, heartless, indurate, inhuman, inhumane, insensate, insensitive, ironhearted, merciless, obdurate, pachydermatous, pitiless, remorseless, ruthless, slash-and-burn, soulless, stony, stoney, stonyhearted, take-no-prisoners, thick-skinned,</p>	<p>soft; easy; weak, charitable, compassionate, humane, kind hearted, kindly, merciful, sensitive, soft hearted, sympathetic, tender, tender hearted, warm, warm hearted, cheap, easy, effortless, facile, light, mindless, simple, soft, undemanding, delicate, nonhardy, soft, tender, weak, groundless, illogical, invalid, irrational, nonrational, nonsensical, nonvalid,</p>

	<p> uncharitable, unfeeling, unmerciful, unsparing, unsympathetic, arduous, augean, backbreaking, challenging, demanding, difficult, effortful, exacting, formidable, grueling, gruelling, heavy, hellacious, herculean, killer, laborious, moiling, murderous, pick-and-shovel, rigorous, rough, rugged, severe, stiff, strenuous, sweaty, tall, testing, toilsome, tough, uphill, cast-iron, hardy, hard-bitten, hardened, inured, rugged, stout, strong, sturdy, tough, toughened, vigorous, commonsense, commonsensible, commonsensical, firm, good, informed, just, justified, levelheaded, logical, rational, reasonable, reasoned, sensible, sober, solid, valid, well-founded, bitter, brutal, burdensome, cruel, excruciating, grievous, grim, harsh, hardhanded, heavy, inhuman, murderous, onerous, oppressive, rough, rugged, searing, severe, stiff, tough, trying, acute, almighty, blistering, deep, dreadful, excruciating, explosive, exquisite, fearful, fearsome, ferocious, fierce, frightful, furious, ghastly, intense, heavy, heavy-duty, hellacious, intensive, keen, profound, terrible, vehement, vicious, violent, austere, authoritarian, flinty, severe, harsh, heavy-handed, ramrod, rigid, rigorous, stern, strict, tough, compact, firm, rigid, solid, stiff, unyielding, certain, determinate, final, firm, flat, frozen, fixed, hard-and-fast, inexpugnable, set, settled, stable, acrid, acrimonious, embittered, bitter, rancorous, resentful, sore, adamant, adamantine, bullheaded, dogged, obstinate, hardened, hardheaded, hard-nosed, headstrong, immovable, implacable, inconvincible, inflexible, intransigent, mulish, obdurate, opinionated, ossified, pat, pertinacious, perverse, pigheaded, self-opinionated, self-willed, stiff-necked, stubborn, unbending, uncompromising, unrelenting, unyielding, willful, wilful, documentary, factual, historical, literal, matter-of-fact, nonfictional, objective, true, solid, firm, inflexible, rigid, rocklike, stiff, strong, tough, unyielding, strenuous, arduous, backbreaking, exacting, exhausting, laborious, rigorous, tough, difficult, complicated, intricate, involved, knotty, perplexing, puzzling, thorny, unfeeling, callous, cold, cruel, hardhearted, pitiless, stern, unkind, unsympathetic, painful, disagreeable, distressing, grievous, intolerable, unpleasant, energetically, fiercely, forcefully, forcibly, heavily, intensely, powerfully, severely, sharply, strongly, vigorously, violently, with all one's might, with might and main, diligently, doggedly, industriously, persistently, </p>	<p> unfounded, uninformed, unjustified, unreasonable, unreasoned, unsound, easy, light, soft, light, moderate, soft, clement, forbearing, gentle, indulgent, lax, lenient, tolerant, flabby, soft, spongy, squashy, squishy, unbitter, acquiescent, agreeable, amenable, compliant, complying, flexible, pliable, pliant, relenting, yielding, fictional, fictionalized, fictitious, nondocumentary, nonfactual, nonhistorical, unhistorical, delicate, soft, vulnerable, weak, easy, facile, happy, light, nice, trivial, unimportant, bright, cheerful;, easy, gentle, kind, nice, pleasant, sunny, doubtful, indefinite, uncertain, unsure, lightly, </p>
--	---	--

	steadily, untiringly, rocklike, solid, strong, tough, difficult, exhausting, arduous, complicated, heavy, rough, serious, terrible, tough, troublesome, cruel, ruthless, bleak, grim, harsh, painful, tough, true, indisputable, positive, sure, heavily, seriously, strongly, vigorously, closely, badly, close, fast, tight	
24. normal	sane, usual; ordinary; regular; standard; typical; average; average, common, commonplace, cut-and-dried, cut-and-dry, everyday, garden-variety, ordinary, prosaic, routine, run-of-the-mill, standard, standard-issue, unexceptional, unremarkable, usual, workaday, balanced, clearheaded, compos mentis, lucid, sane, right, stable, archetypal, archetypical, average, characteristic, typical, regular, representative, standard, true, usual, average, common, conventional, natural, ordinary, regular, routine, standard, typical sane, rational, reasonable, well-adjusted, sane, rational, cool, healthy, reasonable, well-adjusted, whole, common, usual, natural, orderly, regular, routine, traditional, typical	abnormal; different; irregular; unusual; extraordinary; weird, unique, abnormal, exceptional, extraordinary, odd, out-of-the-way, strange, unusual, brainsick, crazed, crazy, demented, deranged, insane, lunatic, mad, maniacal, maniac, mental, unbalanced, unsound, aberrant, abnormal, anomalous, atypical, deviant, irregular, nonrepresentative, nontypical, untypical, abnormal, different, disorderly, extraordinary, irregular, rare, uncommon, untraditional, unusual, poor, sick, unhealthy
25. quickly	fast; rapidly; speedily; apace, briskly, chop-chop, double-quick, fleetly, full tilt, hastily, hell-for-leather, hot, lickety-split, posthaste, presto, pronto, quick, fast, rapidly, snappily, soon, speedily, swift, swiftly, swiftly, abruptly, apace, briskly, fast, hastily, hurriedly, promptly, pronto, rapidly, soon, speedily, fast, expeditiously, hastily, hurriedly, immediately, instantaneously, instantly, promptly, rapidly, speedily, swiftly,	slowly; slow, slow, slowly, eventually, later, slowly
26. well	healthy, fine; hearty; good; satisfactory, able-bodied, bouncing, fit, hale, hearty, robust, sound, healthy, well-conditioned, whole, wholesome, acceptably, adequately, all right, alright, creditably, decently, fine, good, middlingly, nicely, ok, okay, passably, respectably, satisfactorily, serviceably, so-so, sufficiently, tolerably, amply, bounteously, bountifully, freehandedly, freeheartedly, generously, handsomely, lavishly, liberally, munificently, openhandedly, unstintingly, ably, adeptly, adroitly, artfully, capably, competently, consummately, deftly, expertly, masterfully, masterly, proficiently, skillfully, considerately, courteously, graciously, kindly, nicely, reasonably, thoughtfully, agreeably, charmingly, delectably, deliciously, delightfully, dreamily, enchantingly, enjoyably, favorably, felicitously, fetchingly, gloriously, gratifyingly, great, nicely, palatably, pleasantly, pleasingly, pleurably, prettily, satisfyingly, splendidly, sweetly, swimmingly, welcomingly, winningly, all, all of, all	poor; bad; ill; unwell; sick, ailing, diseased, ill, sick, unfit, unhealthy, unsound, unwell, bad, badly, deficiently, ill, inadequately, insufficiently, intolerably, poorly, unacceptably, unsatisfactorily, parsimoniously, stingily, ungenerously, amateurishly, artlessly, incapably, incompetently, inefficiently, ineptly, inexpertly, poorly, unskillfully, discourteously, inconsiderately, thoughtlessly, badly, disagreeably, ill, unpleasantly, half, halfway, incompletely, part, partially, partly, arduously, hardly, laboriously, strenuously, inappropriate, incomplete, infirm, little, poor, rough, sick, small, unfit, unhealthy, unsuitable, unsuited, weak, failing, inappropriate, languishing, unfitting, unsuitable, inadequately, incompletely, incorrectly, insufficiently, mistakenly, partially, wrongly, difficultly, inadequately,

	<p>over, altogether, clean, completely, dead, enough, entire, entirely, even, exactly, fast, flat, full, heartily, out, perfectly, plumb, quite, soundly, thoroughly, through and through, totally, utterly, fully, wholly, wide, easy, effortlessly, facilely, fluently, freely, handily, hands down, lightly, painlessly, readily, smoothly, easily, ah, aha, come on, fie, indeed, my word, pshaw, no, what, why, cradle, font, fountain, fountainhead, origin, root, seedbed, spring, source, wellspring, satisfactorily, agreeably, nicely, pleasantly, smoothly, splendidly, successfully, skillfully, ably, adeptly, adequately, admirably, correctly, efficiently, expertly, proficiently, properly, prosperously, comfortably, suitably, fairly, fittingly, justly, properly, rightly, intimately, deeply, fully, profoundly, thoroughly favorably, approvingly, glowingly, highly, kindly, warmly, considerably, abundantly, amply, fully, greatly, heartily, highly, substantially, thoroughly, very much healthy, fit, in fine fettle, sound satisfactory, agreeable, fine, pleasing, proper, right, thriving, hole, bore, pit, shaft flow, gush, jet, pour, spout, spring, spurt, surge, healthy, strong, together, lucky, fortunate, good, accurately, adequately, carefully, closely, completely, correctly, effectively, efficiently, fully, nicely, properly, smoothly, strongly, successfully, thoroughly, sufficiently, adequately, appropriately, completely, easily, far, freely, fully, properly, quite, right, smoothly, thoroughly, hole, mine, pit, pool, reservoir, source</p>	<p>inappropriately, incompletely, insufficiently, partially, unsuitably,</p>
27. alone	<p>lonely; singular, solitary; solo; single; isolated, lone, lonely, lonesome, single, solitary, solo, unaccompanied, only, lone, one, one-off, singular, sole, solitary, special, sui generis, unique, independently, single-handed, single-handedly, singly, solely, unaided, unassisted, solely, exclusively, just, only, purely, simply, by oneself, apart, detached, isolated, lonely, only, on one's tod, separate, single, solitary, unaccompanied, separate, apart, only, unattended, unique, solely</p>	<p>together; accompanied, coupled, popular, accompanied</p>
28. aware	<p>alert; mindful; cognizant; conscious; sentient; attentive, knowledgeable, alive, apprehensive, conscious, cognizant, mindful, sensible, sentient, ware, witting, informed, enlightened, in the picture, knowledgeable, knowledgeable, alive, appreciative, apprehensive, attentive, cognizant, conscious, familiar, informed, mindful, receptive, wise</p>	<p>unaware; oblivious; clueless, ignorant, insensible, oblivious, unaware, unconscious, unmindful, unwitting, heedless, ignorant, inattentive, neglectful, unaware, unconscious, unfamiliar, unfriendly, uninformed, unknowledgeable</p>
29. reduce	<p>decrease; lower; lessen; minimize, lower, degrade, break, bust, degrade, disrate, downgrade, demote,</p>	<p>increase; advance, elevate, promote, raise, aggrandize, amplify, augment, boost, enlarge,</p>

	abate, de-escalate, dent, deplete, diminish, downscale, downsize, drop, dwindle, ease, knock down, lessen, lower, decrease, attenuate, break, cheapen, depress, devalue, devalue, downgrade, lower, mark down, depreciate, sink, write down, write off, lessen, abate, curtail, cut down, decrease, diminish, lower, moderate, shorten, weaken, degrade, break, bring low, downgrade, humble, make less, decrease, curtail, cut down, diminish, dwindle, lessen, lower, pare, scale down, shorten, slash, trim, weaken, defeat, bring, force, undermine, weaken, humble, humiliate, lower	escalate, expand, increase, raise, appreciate, enhance, mark up, upgrade, compliment, develop, enlarge, expand, extend, grow, increase, lengthen, praise, prolong, raise, strengthen
30. previous	former; prior; earlier; past; before; earlier; preceding, antecedent, anterior, foregoing, former, precedent, preceding, prior, earlier, erstwhile, foregoing, former, past, preceding, prior, former, prior, earlier, past, preceding, premature	latter; after; subsequent; last; next , after, ensuing, following, later, posterior, subsequent, succeeding, current, future, later, on time, present, timely, well-timed
31. sure	certain; convinced; positive; confident, assured, clear, cocksure, confident, doubtless, implicit, positive, sanguine, certain, certain, ineluctable, ineludible, inescapable, necessary, inevitable, unavoidable, unescapable, can't-miss, certain, fail-safe, infallible, sure fire, unfailing, calculable, good, reliable, responsible, safe, secure, solid, steady, dependable, tried, tried-and-true, true, trustable, trustworthy, trusty, accomplished, certain, inarguable, incontestable, incontrovertible, indisputable, indubitable, irrefragable, positive, irrefutable, unanswerable, unarguable, unchallengeable, undeniable, unquestionable, all right, alright, assuredly, certainly, clearly, definitely, doubtless, easily, forsooth, hands down, inarguably, incontestably, incontrovertibly, indisputably, plainly, really, so, indeed, surely, truly, unarguably, undeniably, undoubtedly, unquestionably, certain, assured, confident, convinced, decided, definite, positive, reliable, accurate, dependable, foolproof, infallible, undeniable, undoubted, unerring, unfailing, inevitable, assured, bound, guaranteed, inescapable, certain, definite, assured, clear, confident, convinced, positive, physically stable, fast, safe, strong, inevitable, assured, certain, self-confident, assured, certain, confident, decisive, positive	unsure; uncertain; tenuous; unconfident; unsure; unclear, doubtful, dubious, uncertain, unsure, avoidable, evadable, uncertain, unsure, fallible, dodgy, uncertain, undependable, unreliable, unsafe, untrustworthy, answerable, arguable, contradictable, controvertible, debatable, disputable, doubtful, moot, negotiable, problematic, problematical, questionable, refutable, doubtful, uncertain, unsure, doubting, hesitant, indefinite
32. main	principal; central; major; first, primary; prime, arch, big, capital, cardinal, central, chief, dominant, first, grand, great, greatest, highest, key, leading, foremost, master, number one, numero uno, overbearing, overmastering, overriding, paramount, predominant, preeminent, premier, primal, primary, principal,	secondary; minor; peripheral, extra, last, least, auxiliary, unimportant, common, inessential, inferior, minor, ordinary, secondary, subordinate, trivial, unimportant, usual,

	<p>prior, sovereign, sovran, supreme, beef, brawn, muscle, thew, continent, landmass, mainland, brunt, bulk, chief, core, generality, heft, body, mass, staple, weight, chief, central, essential, foremost, head, leading, pre-eminent, primary, principal, conduit, cable, channel, duct, line, pipe, principal, predominant, central, essential, fundamental, leading, major, particular, preeminent, primary, prime, special, vital, absolute, utter, entire</p>	
--	--	--

Appendix C: Superordination and Subordination Test

For the following words, write down a possible main category and two other examples of the same category. Do **NOT** write *noun*, *adjective*, *verb* or *adverb* for the possible categories. Please write your answers as clearly as possible.

For example:

	POSSIBLE CATEGORY	EXAMPLE	EXAMPLE
APPLE	fruit	banana	strawberry
SHIRT	clothes	hat	skirt
CAT	animal	dog	cow

	POSSIBLE CATEGORY	EXAMPLE	EXAMPLE
1. TREE			
2. HAPPY			
3. APRIL			
4. BIG			
5. ARM			
6. COMPUTER			
7. BROTHER			
8. BED			
9. RED			
10. BOY			
11. DOCTOR			
12. BOX			
13. HOUR			
14. HOUSE			
15. YOUNG			
16. SUMMER			
17. HOT			
18. CAR			
19. HISTORY			
20. LETTER			

21. NORTH			
22. DAY			
23. SHORT			
24. ADD			
25. CENTURY			
26. LOVE			
27. MORNING			
28. READ			
29. ROOM			
30. GLASS			

Superordination and Subordination Test

For the following words, write down a possible main category and two other examples of the same category. Do **NOT** write *noun*, *adjective*, *verb* or *adverb* for the possible categories. Please write your answers as clearly as possible.

For example:

	POSSIBLE CATEGORY	EXAMPLE	EXAMPLE
PIZZA	food	burger	spaghetti
SHIRT	clothes	hat	skirt
TORONTO	name	John	London

	POSSIBLE CATEGORY	EXAMPLE	EXAMPLE
1. TREE	Plant/nature	Lotus/Mango/flower/shrubs/flower	Poison ivy/Apple/bush/flowers/weed
2. HAPPY	Emotion/mood	Happy/depressed	Angry/sad
3. APRIL	Month	May/Aug/Jan/Jun	June/Dec/Feb/Aug
4. BIG	Size	Little/small/medium	Huge/medium/small/large
5. ARM	Body/appendage/body part	Leg/head	Hands/finger/head
6. COMPUTER	Technology/device/ /electronics/machine	Whiteboard/phone/smart phone/radio	One cloud/Ipad/tablet/TV
7. BROTHER	Family /relative/family member	Sister/dad	Mother/sister/father
8. BED	Furniture	Dresser/table/couch	Table/couch/dresser/chair
9. RED	Color	Purple/blue/yellow/	Blue/green
10. BOY	Human/gender/sex	Girl	Man/transsexual/woman

11. DOCTOR	Profession/career/job	Teacher/pilot	Nurse/lawyer/chemist
12. BOX	Container/object	Bag/basket	Barrel/bag/bucket/pot
13. HOUR	Time	45 minutes/minute/day	Half an hour/second/minute
14. HOUSE	Building/place/residence type/dwelling/asset	School/apartment/barn	Police Station/hospital/condo/church
15. YOUNG	Age	Youth/old	Old/middle-aged/older
16. SUMMER	Season	Winter/spring	Fall/spring
17. HOT	Temperature	Cold/cool	Lukewarm/humid/comfortable/cold
18. CAR	Vehicle	Truck/boat/bus	Minivan/airplane/van/truck
19. HISTORY	Subject	Math	Physics/geology/English/English
20. LETTER	Notice/communication /communication method /communication form	Email/phone	Text/voicemail/phone call/email
21. NORTH	Direction	East/south/east	West
22. DAY	Time/period	Monday/week/night/month	Tuesday/month/evening/year
23. SHORT	Stature/height/length description/size/dimension/distance	Tall/long	Average/medium/shorter
24. ADD	Operations/math/ operation in math/math operation	Subtract/subtract	Divide/multiply
25. CENTURY	One hundred/time/time period	Years/millennium/day/millennium	Scores/decade/month
26. LOVE	Affection/emotion/feelings	Hug/hate	Kiss/joy/rage/like
27. MORNING	Time /time of day	Afternoon	evening
28. READ	Education/action/learn/school activity/activity/learning/consumption	Write/speak/listen/write	Think/listen/watch
29. ROOM	Apartment/place/part of house/building parts/place/space	Bedroom/kitchen/garage/lobby	Bathroom/gym/basement/corridor
30. GLASS	Material/advanced material	Wood/fiber optic cable/metal	Stone/metal/IR lenses/wood

Appendix D: Collocation Test

Write **THREE** words that can naturally be used together (before or after) with the following words on next page. **Content words** (nouns, verbs, adjective and adverbs) **ONLY are accepted**. Do **NOT** write:

- pronouns (*I, you, it, me, him, there, yourself...*),
- prepositions or particles (up, out, for, in, off, down, away ...),
- possessive adjectives (*my, your, his, her, ...*),
- articles (*the, a, an*),
- determiners (*this, that, these, those*),
- conjunctions (*and, or, but, ...*),
- auxiliary verbs (*can, could, will, would, am, is, are ...*), or
- negatives (*not, don't, ...*)

Please write your answers as clearly as possible.

Example:

			FRESH	air	fruit	breath
		hot	COFFEE	beans	table	
	fully	gladly	ACCEPT	invitation		
live	left	completely	ALONE			

			BLOOD ¹			
			SPEND ²			
			BRING ³			
			PRICE ⁴			
			COMPLETELY ⁵			
			POSSIBLE ⁶			
			OPPORTUNITY ⁷			
			YEAR ⁸			
			WORLD ⁹			
			CHOICE ¹⁰			
			BANK ¹¹			
			SITUATION ¹²			
			FORMER ¹³			
			ENCOURAGE ¹⁴			
			USER ¹⁵			
			ACTUALLY ¹⁶			
			POPULATION ¹⁷			
			PREVENT ¹⁸			
			SIT ¹⁹			
			SELL ²⁰			
			DEGREE ²¹			
			PHONE ²²			
			HEART ²³			
			SEND ²⁴			
			PARTICULARLY ²⁵			
			WATER ²⁶			
			TEACH ²⁷			
			WEAR ²⁸			
			ONLINE ²⁹			
			EVENT ³⁰			

Collocations Test

Write **THREE** words that can naturally be used together (before or after) with the following words on next page. **Content words** (nouns, verbs, adjective and adverbs) **ONLY are accepted**. Do **NOT** write:

- pronouns (*I, you, it, me, him, there, yourself...*),
- prepositions or particles (up, out, for, in, off, down, away ...),
- possessive adjectives (*my, your, his, her, ...*),
- articles (*the, a, an*),
- determiners (*this, that, these, those*),
- conjunctions (*and, or, but, ...*),
- auxiliary verbs (*can, could, will, would, am, is, are ...*), or
- negatives (*not, don't, ...*)

Please write your answers as clearly as possible.

Example:

			FRESH	air	fruit	breath
		hot	COFFEE	beans	table	
	fully	gladly	ACCEPT	invitation		
live	left	completely	ALONE			

				BLOOD	pressure	sugar	tests	samples
				SPEND	time	money	night	weekend
				BRING	help	joy	peace	tears
gas	retail	low	pay	PRICE	tag	increases	declines	controls
				COMPLETELY	agree	lost	different	free
made	best	entirely	quite	POSSIBLE	explanation	solution	reasons	consequences
excellent	equal	great	unique	OPPORTUNITY				
past	last	next	each	YEAR				
modern	whole	real	third	WORLD	war	trade	cup	bank
excellent	rational	obvious	multiple	CHOICE				
				USUALLY	means	involves	occurs	requires
data	food	national	central	BANK	account	robbery	loans	teller
bad	difficult	political	current	SITUATION				
				FORMER	president	boss	member	employees
		actively	strongly		students	people	parents	investment
				OLD	years	man	woman	enough
internet	individual	computer	drug	USER	friendly	name	account	fee
				ACTUALLY	increased	happened	exist	declined
large	entire	local	total	POPULATION	growth	size	control	density
				PREVENT	cancer	injury	disease	infection
beloved	dead	second	former	HUSBAND				
				SIT	comfortably	quietly	together	alone
				SELL	tickets	product	drug	books
each	certain	particular	different	TYPE				
graduate	high	college	certain	DEGREE				
	mobile	pay	cell	PHONE	calls	number	rings	booth
				HEART	attack	surgery	rate	failure
				SEND	email	profile	message	money
				PARTICULARLY	important	useful	relevant	effective
boiling	cold	clean	hot	WATER	quality	resources	sport	supply
				TEACH	students	English	music	science
				WEAR	gloves	glasses	shoes	makeup
			available	ONLINE	courses	shopping	dating	movies
upcoming	historic	special	annual	EVENT				

Synonym and Antonym Form-Meaning Test

Translate the following words into English. One letter of the desired English word is provided and each line represents a letter. Please write your answers as clearly as possible.

For example:

make	: پاسخ	ساختن: <u> a </u> _ _
fresh		تازه: _ _ <u> e </u> _ _
sun		آفتاب: _ _ <u> u </u> _ _

- | | | | |
|---------------|------------------|---------------|------------------------|
| _____ i _____ | 1. توانایی: | _____ o _____ | 16. بهتر کردن: |
| _____ a _____ | 2. شخصی/خصوصی: | _____ o _____ | 17. برداشتن/زدودن: |
| _____ u _____ | 3. درست/صحیح: | _____ e _____ | 18. امروزی/جدید: |
| _____ a _____ | 4. پایانی/نهایی: | _____ o _____ | 19. جدی: |
| _____ u _____ | 5. عظیم: | _____ u _____ | 20. هل دادن: |
| _____ a _____ | 6. مشهور/محبوب: | _____ e _____ | 21. مفید: |
| _____ e _____ | 7. لازم: | _____ a _____ | 22. سخت: |
| _____ a _____ | 8. مزیت/فایده: | _____ a _____ | 23. معمولی/طبیعی: |
| _____ e _____ | 9. کشف کردن: | _____ u _____ | 24. به سرعت/تند: |
| _____ e _____ | 10. متفاوت: | _____ e _____ | 25. خوب/سالم: |
| _____ e _____ | 11. دریافت کردن: | _____ o _____ | 26. تنها: |
| _____ o _____ | 12. آمدن: | _____ a _____ | 27. آگاه: |
| _____ e _____ | 13. قانونی: | _____ u _____ | 28. کم کردن/کاهش دادن: |
| _____ e _____ | 14. واضح: | _____ e _____ | 29. پیشین/قبل: |
| _____ u _____ | 15. پر/لبریز: | _____ u _____ | 30. مطمئن: |
| | | _____ a _____ | 31. مهم/اصلی: |

Superordination and Subordination Form-Meaning Test

Translate the following words into English. One letter of the desired English word is provided and each line represents a letter. Please write your answers as clearly as possible.

For example:

make	پاسخ :	ساختن: <u> a </u>
fresh		تازه: <u> e </u>
sun		آفتاب: <u> u </u>

- | | |
|---------------------------------------|------------------------------|
| 1. درخت: <u> e </u> | 16. تابستان: <u> e </u> |
| 2. خوشحال: <u> a </u> | 17. داغ: <u> o </u> |
| 3. یکی از ماههای میلادی: <u> i </u> | 18. اتومبیل: <u> a </u> |
| 4. بزرگ: <u> i </u> | 19. تاریخ: <u> o </u> |
| 5. دست و بازو: <u> r </u> | 20. نامه: <u> e </u> |
| 6. رایانه: <u> e </u> | 21. شمال: <u> o </u> |
| 7. اخوی/داداش: <u> e </u> | 22. روز: <u> a </u> |
| 8. تختخواب: <u> e </u> | 23. کوتاه: <u> o </u> |
| 9. قرمز: <u> e </u> | 24. اضافه کردن: <u> d </u> |
| 10. پسر: <u> o </u> | 25. قرن: <u> e </u> |
| 11. پزشک: <u> o </u> | 26. عشق: <u> o </u> |
| 12. جعبه: <u> o </u> | 27. صبح: <u> o </u> |
| 13. ساعت (۶۰ دقیقه): <u> o </u> | 28. خواندن: <u> e </u> |
| 14. خانه: <u> o </u> | 29. اتاق: <u> o </u> |
| 15. جوان: <u> o </u> | 30. شیشه: <u> a </u> |

Collocation Form-Meaning Test

Translate the following words into English. One letter of the desired English word is provided and each line represents a letter. Please write your answers as clearly as possible.

For example:

make : پاسخ
fresh
sun

ساختن: _ a _ _ _
تازه: _ e _ _ _
آفتاب: _ u _ _ _

27. درس دادن: _ e _ _ _
28. پوشیدن: _ e _ _ _
29. متصل به اینترنت: _ i _ _ _
30. رویداد/ اتفاق: _ e _ _ _
1. خون: _ _ _ _ o _ _
2. خرج کردن/صرف کردن: _ _ _ e _ _ _
3. آوردن: _ _ i _ _ _
4. قیمت: _ _ _ i _ _ _
5. کاملاً: _ _ _ _ _ e _ _ _ _ _
6. ممکن: _ o _ _ _ _ _ _ _ _ _
7. فرصت: _ _ _ _ _ o _ _ _ _ _ _ _
8. سال: _ e _ _ _
9. جهان: _ o _ _ _ _
10. انتخاب: _ _ _ o _ _ _ _
11. محلی برای نگهداری پول: _ a _ _ _ _
12. موقعیت/ وضعیت: _ _ _ _ _ a _ _ _ _ _
13. پیشین/ سابق: _ _ _ _ _ e _ _ _
14. تشویق کردن: _ _ _ _ _ a _ _ _ _ _
15. کاربر/ استفاده کننده: _ e _ _ _
16. در واقع/ در حقیقت: _ _ _ _ _ a _ _ _ _ _
17. جمعیت: _ _ _ _ _ a _ _ _ _ _
18. جلوگیری کردن/پیشگیری کردن:
_ _ _ _ _ e _ _ _ _ _
19. نشستن: _ i _ _ _
20. فروختن: _ _ _ l _ _ _
21. درجه/ رتبه: _ _ _ _ _ e _ _ _
22. وسیله‌ای برای برقراری تماس: _ _ _ o _ _ _ _
23. قلب: _ e _ _ _ _
24. فرستادن: _ e _ _ _ _
25. به ویژه/بخصوص:
_ _ a _ _ _ _ _ _ _ _ _ _
26. آب: _ _ _ _ e _ _ _ _

Appendix F: new-GSL 779 words

NEW GSL – UP TO 1000 RANK – NO NATION’S (2013) FUNCTION WORDS & NO OTHER FUNCTION & ABBREVIATION WORDS

- | | |
|---------------------------------|---------------------------------|
| 1. ability (n, 800) | 41. ASK (v, 140) |
| 2. ABLE (adj, 252) | 42. aspect (n, 787) |
| 3. accept (v, 541) | 43. associate (v, 880) |
| 4. access (n, 871) | 44. association (n, 844) |
| 5. account (n, 508) | 45. assume (v, 767) |
| 6. achieve (v, 569) | 46. attempt (n, 715) |
| 7. act (n, 676) | 47. attend (v, 815) |
| 8. act (v, 762) | 48. attention (n, 555) |
| 9. ACTION (n, 406) | 49. AUTHORITY (n, 467) |
| 10. ACTIVITY (n, 449) | 50. AVAILABLE (adj, 405) |
| 11. ACTUALLY (adv, 391) | 51. avoid (v, 707) |
| 12. ADD (v, 276) | 52. aware (adj, 925) |
| 13. addition (n, 888) | 53. AWAY (adv, 208) |
| 14. admit (v, 795) | 54. baby (n, 964) |
| 15. advantage (n, 851) | 55. BAD (adj, 304) |
| 16. advice (n, 935) | 56. bank (n, 740) |
| 17. affect (v, 661) | 57. BASE (v, 389) |
| 18. AGE (n, 321) | 58. basic (adj, 958) |
| 19. agency (n, 913) | 59. basis (n, 782) |
| 20. AGREE (v, 436) | 60. BEAR (v, 426) |
| 21. agreement (n, 1000) | 61. BECOME (v, 116) |
| 22. air (n, 523) | 62. bed (n, 644) |
| 23. ALLOW (v, 212) | 63. BEGIN (v, 199) |
| 24. alone (adv, 663) | 64. behaviour (n, 903) |
| 25. amount (n, 535) | 65. BELIEVE (v, 237) |
| 26. announce (v, 975) | 66. benefit (n, 651) |
| 27. answer (n, 659) | 67. BIG (adj, 309) |
| 28. answer (v, 874) | 68. bill (n, 941) |
| 29. APPEAR (v, 279) | 69. BIT (n, 473) |
| 30. application (n, 961) | 70. black (adj, 589) |
| 31. APPLY (v, 497) | 71. blood (n, 896) |
| 32. approach (n, 606) | 72. board (n, 781) |
| 33. April (n, 955) | 73. BODY (n, 290) |
| 34. AREA (n, 187) | 74. BOOK (n, 288) |
| 35. argue (v, 780) | 75. box (n, 916) |
| 36. argument (n, 824) | 76. BOY (n, 435) |
| 37. arm (n, 544) | 77. BREAK (v, 388) |
| 38. arrive (v, 624) | 78. BRING (v, 167) |
| 39. art (n, 672) | 79. brother (n, 914) |
| 40. article (n, 821) | 80. BUILD (v, 361) |

81. **building** (n, 518)
 82. **BUSINESS** (n, 266)
 83. **BUY** (v, 417)
 84. **call** (n, 895)
 85. **CALL** (v, 143)
 86. **CAR** (n, 367)
 87. **card** (n, 933)
 88. **care** (n, 525)
 89. **career** (n, 963)
 90. **CARRY** (v, 272)
 91. **CASE** (n, 156)
 92. **catch** (v, 577)
 93. **cause** (n, 902)
 94. **CAUSE** (v, 443)
 95. **CENTRE** (n, 492)
 96. **century** (n, 540)
 97. **CERTAIN** (adj, 358)
 98. **CERTAINLY** (adv, 444)
 99. **CHANCE** (n, 494)
 100. **CHANGE** (n, 253)
 101. **CHANGE** (v, 265)
 102. **character** (n, 809)
 103. **charge** (n, 714)
 104. **check** (v, 873)
 105. **CHILD** (n, 169)
 106. **choice** (n, 534)
 107. **CHOOSE** (v, 448)
 108. **church** (n, 850)
 109. **circumstance** (n, 947)
 110. **CITY** (n, 414)
 111. **claim** (v, 634)
 112. **CLASS** (n, 462)
 113. **CLEAR** (adj, 415)
 114. **clearly** (adv, 626)
 115. **close** (adj, 628)
 116. **close** (v, 641)
 117. **club** (n, 804)
 118. **collect** (v, 996)
 119. **colour** (n, 724)
 120. **COME** (v, 67)
 121. **committee** (n, 785)
 122. **COMMON** (adj, 434)
 123. **COMMUNITY** (n, 421)
 124. **COMPANY** (n, 274)
 125. **compare** (v, 597)
 126. **complete** (adj, 728)
 127. **complete** (v, 829)
 128. **completely** (adv, 773)
 129. **computer** (n, 861)
 130. **concern** (n, 743)
 131. **concern** (v, 666)
 132. **CONDITION** (n, 387)
 133. **CONSIDER** (v, 241)
 134. **contact** (n, 884)
 135. **contain** (v, 519)
 136. **CONTINUE** (v, 298)
 137. **CONTROL** (n, 404)
 138. **control** (v, 837)
 139. **COST** (n, 453)
 140. **council** (n, 689)
 141. **COUNTRY** (n, 194)
 142. **couple** (n, 573)
 143. **COURSE** (n, 189)
 144. **court** (n, 680)
 145. **COVER** (v, 432)
 146. **CREATE** (v, 371)
 147. **current** (adj, 605)
 148. **currently** (adv, 799)
 149. **cut** (v, 532)
 150. **data [PL]** (n, 828)
 151. **date** (n, 642)
 152. **daughter** (n, 934)
 153. **DAY** (n, 95)
 154. **dead** (adj, 936)
 155. **deal** (n, 721)
 156. **deal** (v, 646)
 157. **DEATH** (n, 425)
 158. **DECIDE** (v, 319)
 159. **decision** (n, 521)
 160. **deep** (adj, 741)
 161. **degree** (n, 717)
 162. **demand** (n, 823)
 163. **department** (n, 869)
 164. **depend** (v, 702)
 165. **DESCRIBE** (v, 398)
 166. **design** (n, 793)
 167. **design** (v, 796)
 168. **DETAIL** (n, 498)
 169. **determine** (v, 593)
 170. **DEVELOP** (v, 338)
 171. **DEVELOPMENT** (n, 327)
 172. **DIE** (v, 464)
 173. **DIFFERENCE** (n, 476)
 174. **DIFFERENT** (adj, 170)
 175. **DIFFICULT** (adj, 466)
 176. **difficulty** (n, 814)

177. **direct** (adj, 898)
 178. **direction** (n, 806)
 179. **directly** (adv, 997)
 180. **discover** (v, 691)
 181. **discuss** (v, 665)
 182. **discussion** (n, 865)
 183. **distance** (n, 998)
 184. **doctor** (n, 855)
 185. **DOOR** (n, 360)
 186. **DRAW** (v, 375)
 187. **drive** (v, 507)
 188. **drop** (v, 939)
 189. **duty** (n, 972)
 190. **earlier** (adv, 986)
 191. **EARLY** (adj, 280)
 192. **easily** (adv, 727)
 193. **EASY** (adj, 455)
 194. **eat** (v, 686)
 195. **economic** (adj, 710)
 196. **edge** (n, 981)
 197. **education** (n, 554)
 198. **EFFECT** (n, 281)
 199. **effective** (adj, 993)
 200. **effort** (n, 585)
 201. **element** (n, 832)
 202. **enable** (v, 897)
 203. **encourage** (v, 768)
 204. **END** (n, 182)
 205. **end** (v, 516)
 206. **energy** (n, 971)
 207. **enjoy** (v, 591)
 208. **ensure** (v, 640)
 209. **enter** (v, 578)
 210. **environment** (n, 751)
 211. **ESPECIALLY** (adv, 479)
 212. **essential** (adj, 906)
 213. **establish** (v, 608)
 214. **evening** (n, 692)
 215. **EVENT** (n, 401)
 216. **evidence** (n, 501)
 217. **exactly** (adv, 755)
 218. **examine** (v, 956)
 219. **EXAMPLE** (n, 315)
 220. **exist** (v, 708)
 221. **EXPECT** (v, 278)
 222. **EXPERIENCE** (n, 359)
 223. **EXPLAIN** (v, 446)
 224. **explore** (v, 945)
 225. **express** (v, 967)
 226. **extra** (adj, 952)
 227. **EYE** (n, 244)
 228. **FACE** (n, 314)
 229. **FACE** (v, 458)
 230. **FACT** (n, 181)
 231. **factor** (n, 769)
 232. **fail** (v, 556)
 233. **FALL** (v, 296)
 234. **FAMILY** (n, 216)
 235. **FATHER** (n, 477)
 236. **fear** (n, 894)
 237. **feature** (n, 685)
 238. **FEEL** (v, 141)
 239. **feeling** (n, 742)
 240. **field** (n, 514)
 241. **fight** (v, 891)
 242. **FIGURE** (n, 337)
 243. **fill** (v, 667)
 244. **film** (n, 807)
 245. **final** (adj, 625)
 246. **finally** (adv, 547)
 247. **financial** (adj, 690)
 248. **FIND** (v, 83)
 249. **fine** (adj, 765)
 250. **finish** (v, 798)
 251. **fire** (n, 778)
 252. **firm** (n, 803)
 253. **fit** (v, 922)
 254. **floor** (n, 772)
 255. **focus** (v, 599)
 256. **FOLLOW** (v, 159)
 257. **FOOD** (n, 468)
 258. **FOOT** (n, 439)
 259. **FORCE** (n, 416)
 260. **force** (v, 849)
 261. **foreign** (adj, 978)
 262. **forget** (v, 750)
 263. **FORM** (n, 254)
 264. **form** (v, 627)
 265. **former** (adj, 766)
 266. **forward** (adv, 562)
 267. **FREE** (adj, 368)
 268. **FRIEND** (n, 273)
 269. **front** (n, 802)
 270. **FULL** (adj, 269)
 271. **fully** (adv, 835)
 272. **function** (n, 771)

273.	FURTHER (adj, 380)	321.	husband (n, 842)
274.	further (adv, 808)	322.	IDEA (n, 233)
275.	future (n, 617)	323.	identify (v, 681)
276.	gain (v, 985)	324.	image (n, 779)
277.	game (n, 587)	325.	imagine (v, 923)
278.	garden (n, 817)	326.	immediately (adv, 826)
279.	GENERAL (adj, 407)	327.	impact (n, 688)
280.	generally (adv, 758)	328.	IMPORTANT (adj, 217)
281.	GIRL (n, 384)	329.	improve (v, 603)
282.	GIVE (v, 74)	330.	INCLUDE (v, 176)
283.	glass (n, 890)	331.	increase (n, 845)
284.	GO (v, 47)	332.	INCREASE (v, 305)
285.	God (n, 622)	333.	INDEED (adv, 496)
286.	GOOD (adj, 73)	334.	indicate (v, 738)
287.	GOVERNMENT (n, 235)	335.	individual (adj, 670)
288.	GREAT (adj, 118)	336.	individual (n, 761)
289.	ground (n, 513)	337.	industry (n, 553)
290.	GROUP (n, 191)	338.	INFORMATION (n, 260)
291.	GROW (v, 310)	339.	instead (adv, 538)
292.	growth (n, 932)	340.	intend (v, 901)
293.	guy (n, 696)	341.	INTEREST (n, 286)
294.	hair (n, 866)	342.	interested (adj, 825)
295.	half (n, 919)	343.	interesting (adj, 791)
296.	HAND (n, 155)	344.	international (adj, 893)
297.	HAPPEN (v, 261)	345.	Internet (n, 701)
298.	happy (adj, 695)	346.	introduce (v, 633)
299.	HARD (adj, 440)	347.	INVOLVE (v, 335)
300.	HEAD (n, 204)	348.	ISSUE (n, 341)
301.	HEALTH (n, 457)	349.	JOB (n, 307)
302.	HEAR (v, 245)	350.	join (v, 515)
303.	HEART (n, 488)	351.	JUST (adv, 75)
304.	heavy (adj, 937)	352.	KEEP (v, 157)
305.	help (n, 756)	353.	key (adj, 660)
306.	HELP (v, 185)	354.	kid (n, 736)
307.	HIGH (adj, 145)	355.	kill (v, 792)
308.	highly (adv, 892)	356.	KIND (OF) (n, 292)
309.	HISTORY (n, 495)	357.	KNOW (v, 61)
310.	hit (v, 885)	358.	knowledge (n, 550)
311.	HOLD (v, 188)	359.	lack (n, 979)
312.	HOME (n, 158)	360.	land (n, 735)
313.	hope (n, 920)	361.	language (n, 654)
314.	HOPE (v, 376)	362.	LARGE (adj, 178)
315.	hospital (n, 820)	363.	LATE (adj, 333)
316.	hot (adj, 872)	364.	LATER (adv, 251)
317.	HOUR (n, 284)	365.	LAW (n, 461)
318.	HOUSE (n, 192)	366.	lay (v, 760)
319.	huge (adj, 834)	367.	LEAD (v, 214)
320.	human (adj, 504)	368.	leader (n, 722)

369. **LEARN** (v, 331)
 370. **LEAVE** (v, 123)
 371. **leg** (n, 973)
 372. **legal** (adj, 929)
 373. **LET** (v, 275)
 374. **letter** (n, 546)
 375. **LEVEL** (n, 240)
 376. **LIE** (v, 474)
 377. **LIFE** (n, 114)
 378. **light** (n, 503)
 379. **LIKE** (v, 242)
 380. **LIKELY** (adj, 470)
 381. **LINE** (n, 267)
 382. **list** (n, 719)
 383. **listen** (v, 836)
 384. **LITTLE** (adj, 174)
 385. **little** (adv, 552)
 386. **LIVE** (v, 211)
 387. **LOCAL** (adj, 249)
 388. **location** (n, 862)
 389. **LONG** (adj, 195)
 390. **LONG** (adv, 396)
 391. **look** (n, 645)
 392. **LOOK** (v, 92)
 393. **LOSE** (v, 302)
 394. **loss** (n, 729)
 395. **love** (n, 621)
 396. **love** (v, 574)
 397. **LOW** (adj, 301)
 398. **MAIN** (adj, 370)
 399. **maintain** (v, 699)
 400. **MAJOR** (adj, 447)
 401. **majority** (n, 999)
 402. **MAKE** (v, 35)
 403. **MAN** (n, 105)
 404. **manage** (v, 545)
 405. **management** (n, 876)
 406. **manager** (n, 921)
 407. **MARKET** (n, 471)
 408. **MATERIAL** (n, 490)
 409. **MATTER** (n, 332)
 410. **maybe** (adv, 822)
 411. **MEAN** (v, 165)
 412. **measure** (n, 878)
 413. **medium [often PL]** (n, 609)
 414. **medical** (adj, 949)
 415. **MEET** (v, 225)
 416. **meeting** (n, 618)
 417. **MEMBER** (n, 236)
 418. **mention** (v, 662)
 419. **method** (n, 506)
 420. **mile** (n, 882)
 421. **MIND** (n, 344)
 422. **MINUTE** (n, 408)
 423. **miss** (v, 734)
 424. **model** (n, 616)
 425. **modern** (adj, 704)
 426. **MOMENT** (n, 317)
 427. **MONEY** (n, 264)
 428. **MONTH** (n, 196)
 429. **MORNING** (n, 465)
 430. **MOTHER** (n, 374)
 431. **MOVE** (v, 220)
 432. **movement** (n, 583)
 433. **music** (n, 568)
 434. **NAME** (n, 259)
 435. **NATIONAL** (adj, 486)
 436. **natural** (adj, 598)
 437. **nature** (n, 530)
 438. **necessary** (adj, 505)
 439. **NEED** (n, 320)
 440. **NEED** (v, 117)
 441. **NEW** (adj, 87)
 442. **news** (n, 648)
 443. **nice** (adj, 991)
 444. **NIGHT** (n, 257)
 445. **normal** (adj, 776)
 446. **North** (n, 974)
 447. **note** (n, 731)
 448. **note** (v, 730)
 449. **notice** (v, 918)
 450. **NUMBER** (n, 148)
 451. **obtain** (v, 683)
 452. **obviously** (adv, 943)
 453. **occur** (v, 596)
 454. **OFFER** (v, 285)
 455. **OFFICE** (n, 373)
 456. **officer** (n, 677)
 457. **OLD** (adj, 160)
 458. **online** (adj, 987)
 459. **OPEN** (adj, 484)
 460. **OPEN** (v, 351)
 461. **operate** (v, 959)
 462. **operation** (n, 783)
 463. **opinion** (n, 957)
 464. **opportunity** (n, 548)

465. **option** (n, 700)
 466. **ORDER** (n, 222)
 467. **original** (adj, 678)
 468. **outside** (con, 594)
 469. **OWN** (adj, 115)
 470. **page** (n, 763)
 471. **PAPER** (n, 469)
 472. **parent** (n, 580)
 473. **PART** (n, 124)
 474. **PARTICULAR** (adj, 353)
 475. **PARTICULARLY** (adv, 475)
 476. **partner** (n, 915)
 477. **PARTY** (n, 346)
 478. **PASS** (v, 390)
 479. **past** (adj, 630)
 480. **pattern** (n, 752)
 481. **PAY** (v, 228)
 482. **PEOPLE** (n, 79)
 483. **PER CENT** (n, 409)
 484. **perform** (v, 859)
 485. **performance** (n, 671)
 486. **PERHAPS** (adv, 291)
 487. **PERIOD** (n, 352)
 488. **PERSON** (n, 329)
 489. **personal** (adj, 517)
 490. **phone** (n, 612)
 491. **physical** (adj, 857)
 492. **pick** (v, 629)
 493. **picture** (n, 561)
 494. **piece** (n, 564)
 495. **PLACE** (n, 139)
 496. **place** (v, 526)
 497. **PLAN** (n, 433)
 498. **plan** (v, 669)
 499. **play** (n, 942)
 500. **PLAY** (v, 230)
 501. **POINT** (n, 171)
 502. **point** (v, 697)
 503. **police** (n, 637)
 504. **POLICY** (n, 472)
 505. **political** (adj, 533)
 506. **poor** (adj, 601)
 507. **popular** (adj, 954)
 508. **population** (n, 911)
 509. **POSITION** (n, 355)
 510. **possibility** (n, 852)
 511. **POSSIBLE** (adj, 221)
 512. **possibly** (adv, 983)
 513. **post** (n, 989)
 514. **potential** (adj, 904)
 515. **POWER** (n, 271)
 516. **practice** (n, 511)
 517. **prepare** (v, 623)
 518. **presence** (n, 976)
 519. **present** (adj, 560)
 520. **present** (v, 559)
 521. **pressure** (n, 718)
 522. **prevent** (v, 854)
 523. **previous** (adj, 635)
 524. **PRICE** (n, 489)
 525. **principle** (n, 883)
 526. **private** (adj, 536)
 527. **PROBABLY** (adv, 318)
 528. **PROBLEM** (n, 183)
 529. **PROCESS** (n, 339)
 530. **PRODUCE** (v, 336)
 531. **product** (n, 682)
 532. **production** (n, 856)
 533. **programme** (n, 639)
 534. **project** (n, 586)
 535. **property** (n, 684)
 536. **protect** (v, 858)
 537. **prove** (v, 571)
 538. **PROVIDE** (v, 177)
 539. **PUBLIC** (adj, 381)
 540. **public** (n, 962)
 541. **publish** (v, 784)
 542. **pull** (v, 770)
 543. **purpose** (n, 581)
 544. **push** (v, 853)
 545. **PUT** (v, 150)
 546. **QUALITY** (n, 478)
 547. **QUESTION** (n, 210)
 548. **quickly** (adv, 607)
 549. **RAISE** (v, 451)
 550. **RANGE** (n, 483)
 551. **RATE** (n, 393)
 552. **REACH** (v, 328)
 553. **READ** (v, 316)
 554. **ready** (adj, 706)
 555. **REAL** (adj, 357)
 556. **realise** (v, 843)
 557. **REALLY** (adv, 190)
 558. **REASON** (n, 246)
 559. **RECEIVE** (v, 334)
 560. **recent** (adj, 565)

561. **recently** (adv, 632)
 562. **recognise** (v, 948)
 563. **record** (n, 575)
 564. **red** (adj, 712)
 565. **reduce** (v, 528)
 566. **refer** (v, 726)
 567. **reference** (n, 995)
 568. **reflect** (v, 848)
 569. **refuse** (v, 909)
 570. **regard** (v, 619)
 571. **region** (n, 863)
 572. **relate** (v, 737)
 573. **relation** (n, 926)
 574. **relationship** (n, 579)
 575. **REMAIN** (v, 263)
 576. **REMEMBER** (v, 364)
 577. **remove** (v, 694)
 578. **REPORT** (n, 423)
 579. **report** (v, 567)
 580. **represent** (v, 674)
 581. **RESEARCH** (n, 493)
 582. **response** (n, 716)
 583. **rest** (n, 636)
 584. **RESULT** (n, 231)
 585. **result** (v, 953)
 586. **return** (n, 788)
 587. **RETURN** (v, 395)
 588. **reveal** (v, 818)
 589. **RIGHT** (adj, 386)
 590. **RIGHT** (adv, 412)
 591. **RIGHT** (n, 306)
 592. **rise** (v, 549)
 593. **risk** (n, 720)
 594. **ROAD** (n, 482)
 595. **role** (n, 529)
 596. **ROOM** (n, 289)
 597. **rule** (n, 592)
 598. **RUN** (v, 200)
 599. **sale** (n, 827)
 600. **SAME** (adj, 120)
 601. **save** (v, 703)
 602. **SAY** (v, 32)
 603. **scale** (n, 877)
 604. **scene** (n, 965)
 605. **scheme** (n, 754)
 606. **SCHOOL** (n, 232)
 607. **sea** (n, 966)
 608. **seat** (n, 992)
 609. **section** (n, 557)
 610. **security** (n, 759)
 611. **SEE** (v, 57)
 612. **seek** (v, 522)
 613. **SEEM** (v, 130)
 614. **sell** (v, 576)
 615. **SEND** (v, 354)
 616. **SENSE** (n, 365)
 617. **series** (n, 643)
 618. **serious** (adj, 744)
 619. **serve** (v, 512)
 620. **SERVICE** (n, 209)
 621. **set** (n, 653)
 622. **SET** (v, 193)
 623. **share** (n, 875)
 624. **share** (v, 668)
 625. **shop** (n, 675)
 626. **SHORT** (adj, 394)
 627. **show** (n, 595)
 628. **SHOW** (v, 126)
 629. **SIDE** (n, 219)
 630. **sign** (n, 813)
 631. **significant** (adj, 711)
 632. **similar** (adj, 531)
 633. **simple** (adj, 527)
 634. **SIMPLY** (adv, 431)
 635. **SINGLE** (adj, 428)
 636. **SIT** (v, 322)
 637. **site** (n, 797)
 638. **SITUATION** (n, 450)
 639. **size** (n, 588)
 640. **skill** (n, 805)
 641. **slightly** (adv, 881)
 642. **SMALL** (adj, 166)
 643. **SOCIAL** (adj, 383)
 644. **society** (n, 543)
 645. **solution** (n, 879)
 646. **son** (n, 620)
 647. **sort (of)** (n, 510)
 648. **sound** (n, 870)
 649. **sound** (v, 988)
 650. **source** (n, 572)
 651. **South** (n, 833)
 652. **space** (n, 584)
 653. **SPEAK** (v, 330)
 654. **SPECIAL** (adj, 377)
 655. **specific** (adj, 847)
 656. **SPEND** (v, 347)

657.	staff (n, 600)	705.	tend (v, 867)
658.	stage (n, 582)	706.	TERM (n, 299)
659.	STAND (v, 247)	707.	test (n, 649)
660.	standard (n, 753)	708.	THING (n, 119)
661.	start (n, 864)	709.	THINK (v, 91)
662.	START (v, 186)	710.	thought (n, 570)
663.	STATE (n, 294)	711.	throw (v, 713)
664.	state (v, 960)	712.	TIME (n, 45)
665.	statement (n, 831)	713.	TOGETHER (adv, 238)
666.	station (n, 905)	714.	top (adj, 774)
667.	STAY (v, 442)	715.	top (n, 673)
668.	STEP (n, 485)	716.	total (adj, 789)
669.	stone (n, 984)	717.	town (n, 563)
670.	STOP (v, 362)	718.	trade (n, 910)
671.	STORY (n, 499)	719.	traditional (adj, 927)
672.	strategy (n, 757)	720.	training (n, 638)
673.	street (n, 733)	721.	travel (v, 930)
674.	strength (n, 990)	722.	treat (v, 790)
675.	STRONG (adj, 454)	723.	treatment (n, 723)
676.	structure (n, 786)	724.	tree (n, 764)
677.	STUDENT (n, 382)	725.	trouble (n, 994)
678.	STUDY (n, 420)	726.	TRUE (adj, 418)
679.	study (v, 924)	727.	truth (n, 907)
680.	style (n, 917)	728.	TRY (v, 154)
681.	SUBJECT (n, 441)	729.	TURN (v, 172)
682.	success (n, 610)	730.	TYPE (n, 366)
683.	successful (adj, 838)	731.	UNDERSTAND (v, 311)
684.	suddenly (adv, 889)	732.	unit (n, 812)
685.	suffer (v, 900)	733.	USE (n, 258)
686.	SUGGEST (v, 297)	734.	USE (v, 76)
687.	summer (n, 777)	735.	useful (adj, 951)
688.	SUPPORT (n, 419)	736.	user (n, 775)
689.	SUPPORT (v, 397)	737.	USUALLY (adv, 413)
690.	suppose (v, 749)	738.	VALUE (n, 379)
691.	SURE (adj, 459)	739.	variety (n, 982)
692.	surface (n, 928)	740.	various (adj, 502)
693.	SYSTEM (n, 206)	741.	version (n, 944)
694.	TABLE (n, 400)	742.	VIEW (n, 326)
695.	TAKE (v, 46)	743.	village (n, 969)
696.	talk (n, 968)	744.	visit (n, 950)
697.	TALK (v, 348)	745.	visit (v, 709)
698.	task (n, 816)	746.	voice (n, 520)
699.	teach (v, 977)	747.	WAIT (v, 399)
700.	teacher (n, 931)	748.	WALK (v, 487)
701.	TEAM (n, 491)	749.	wall (n, 611)
702.	technique (n, 946)	750.	WANT (v, 106)
703.	technology (n, 664)	751.	WAR (n, 438)
704.	TELL (v, 111)	752.	WATCH (v, 452)

- 753. **WATER** (n, 277)
- 754. **WAY** (n, 82)
- 755. **wear** (v, 613)
- 756. **website** (n, 860)
- 757. **WEEK** (n, 184)
- 758. **weight** (n, 794)
- 759. **WELL** (adv, 81)
- 760. **West** (n, 887)
- 761. **WHITE** (adj, 463)
- 762. **WHOLE** (adj, 345)
- 763. **wide** (adj, 539)
- 764. **wife** (n, 602)
- 765. **WIN** (v, 480)
- 766. **window** (n, 615)
- 767. **wish** (v, 537)
- 768. **WOMAN** (n, 198)
- 769. **wonder** (v, 839)
- 770. **WORD** (n, 215)
- 771. **WORK** (n, 122)
- 772. **WORK** (v, 113)
- 773. **worker** (n, 725)
- 774. **WORLD** (n, 152)
- 775. **worth** (adj, 940)
- 776. **WRITE** (v, 213)
- 777. **wrong** (adj, 631)
- 778. **YEAR** (n, 54)
- 779. **YOUNG** (adj, 224)

Appendix G: Ethics Approval



Research Ethics

Western University Non-Medical Research Ethics Board NMREB Amendment Approval Notice

Principal Investigator: Dr. Farahnaz Fazl
Department & Institution: Education/Faculty of Education, Western University

NMREB File Number: 107207
Study Title: Measuring Productive Depth of Vocabulary Knowledge of the Most Frequent Words

NMREB Revision Approval Date: April 29, 2016
NMREB Expiry Date: October 31, 2016

Documents Approved and/or Received for Information:

Document Name	Comments	Version Date
Revised Western University Protocol	Received April 19, 2016	
Revised Letter of Information & Consent	Main Study	2016/04/04
Revised Letter of Information & Consent	Main Study EFL	2016/04/04
Other	Third Vocabulary Test, Persian Received April 7, 2016	

The Western University Non-Medical Science Research Ethics Board (NMREB) has reviewed and approved the amendment to the above named study, as of the NMREB Amendment Approval Date noted above.

NMREB approval for this study remains valid until the NMREB Expiry Date noted above, conditional to timely submission and acceptance of NMREB Continuing Ethics Review.

The Western University NMREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the Ontario Personal Health Information Protection Act (PHIPA, 2004), and the applicable laws and regulations of Ontario.

Members of the NMREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB.

The NMREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000941.

Alireza Barouni Ebrahimi

RESEARCH	Applied Linguistics	
INTERESTS	Second Language Acquisition	
	Vocabulary and Language Assessment	
	English for Academic Purposes	
UNIVERSITY	PhD in Education	2017
EDUCATION	Curriculum & Applied Linguistics Studies	
	Western University, London, Ontario, Canada	
	MA in TEFL, Teaching English as a Foreign Language	2010
	Azad University, Esfahan, Iran	
	BA in TEFL, Teaching English as a Foreign Language	2006
	Azad University, Esfahan, Iran	
CERTIFICATION	Adult ESL Teacher Certificate of Accreditation (OCELT)	2016
	TESL Ontario 2016	
	CELPIP Instructor Training Certificate	2016
	Paragon Testing Enterprises	
	IELTS Examiner Certification	2016
	Cambridge English Language Assessment	
	IELTS Teacher Training Certificate	2016
	British Council, George Brown College, Toronto	
	Western Certificate in University Teaching and Learning	2015
	Western University, London, Canada	
	TESL Canada Federation Professional Certificate	2014
	Standard Level Three [Permanent]	

	Teaching Certificate , Teacher Training Course	2003
	Kian Farda Institute, Esfahan, Iran	
	Advanced English Graduation Certificate	2001
	Pooyesh Language Teaching Center, Esfahan, Iran	
AWARDS AND ACHIEVEMENTS	University of Western Ontario Doctoral Scholarship	2012 - 2016
	TESL Teacher Complimentary Registration Award	2014
	TESL Canada Conference 2014 at the University of Regina	
	Explore Program Bursary	2013
	French Language Immersion program	
	Université de Saint-Boniface, Winnipeg, Manitoba	
WORK EXPERIENCE	ESL Instructor	2017- current
	George Brown College, Toronto, Canada	
	ESL 9085: CELPIP General Test Preparation	
	ESL 9077: IELTS Exam Preparation	
	<ul style="list-style-type: none"> • Teach reading, listening, speaking, and writing skills to CELPIP and IELTS candidates from a variety of educational and vocational backgrounds • Provide students with clear explanations and instructions, opportunities for pair and group activities, effective feedback, and individual consultation • Meet with students during office hours to discuss their achievement, strengths, weaknesses, and concerns 	
	IELTS Examiner	2016 - current
	IELTS Centre CA278, Precise Solutions, Toronto, Canada	
	Teacher Assistant	2014-2016
	Faculty of Education, Western University, London, Canada	
	<ul style="list-style-type: none"> • Diverse Traditions: Approaches to Educational Research, 9678 Teaching Assistantship • Teaching English as a Second Language, 5413S Teaching Assistantship • Understanding Second Language Learning and Teaching, 9300A Teaching Assistantship • Facilitated the course coordinator's teaching in the classroom and was a guest speaker for some sessions 	

- Assessed students' improvement through grading in-class and at-home assignments
- Provided opportunities for individual consultation
- Taught outlining, paragraph and essay writing
- Provided constructive feedback on students' assignments and opportunities for individual consultation

ESL Writing Professor

2014

Fanshawe College, London, Ontario, Canada

WRIT 1034: Reason and Writing

- Developed international college students' writing skills, e.g. essay writing, prompt writing, proofreading, etc.
- Taught writing techniques to international first and second year college students with a variety of educational and vocational backgrounds
- Taught APA documentation styles for in-text citations, end-of-text citations, and essay formats
- Lectured on writing challenges: spelling, grammar, word choice, parallelism, run-on sentences, sentence fragments, punctuation, etc.
- Provided students with clear explanations and instructions, opportunities for pair/group activities, effective feedback, and individual consultation
- Met with students during office hours to discuss their achievement, strengths, weaknesses, and concerns

Research Assistant

2012- 2014

Faculty of Education, Western University, London, Canada

- Subject-specific Vocabulary and Expressions of Academic Subjects: A Corpus-based Study of Mathematics and Geography Research Assistantship
 - Annotated the 1, 2, 3 and 4 word bundles with annotation software and reported how the annotation went through the process and what problems existed.

EAP Educator and EFL Teacher Trainer

2010-2012

Paya Language Institute, Esfahan, Iran

- IELTS and TOEFL instructor and undergraduate and graduate student ESL teacher
- Provided individual consultation in order to further assist new students in taking courses that addressed their specific needs or weaknesses
- Taught grammar, reading, listening, speaking, and writing to adult learners from a variety educational and vocational backgrounds
- Designed and developed a teacher training course for English as a Foreign Language (EFL) teachers who wished to have more innovative teaching ideas for the four language skills (Listening, Speaking, Reading and Writing)

EAP Educator and ESL Teacher

2009-2012

Sadr Institute of Higher Education, Esfahan, Iran

- IELTS/TOEFL/ESL instructor
- Taught grammar, reading, listening, speaking, and writing to adult learners from a variety of educational and vocational backgrounds
- Developed long range and daily lesson plans
- Created, adopted, and adapted student-centered communicative activities on a variety of topics
- Assessed students' improvement through developing, administering, and grading quizzes, tests, and in-class and at-home assignments
- Provided constructive feedback on students' assignments and opportunities for individual consultation
- Provided students with clear explanations and instructions, opportunities for group activities, effective feedback, and individual consultation

EAP Educator and ESL Teacher

2008

Pouya Nahadin Company, Esfahan, Iran

- Prepared students for overseas academic programs
 - Assisted students with what to study, how to apply, where to apply, and how to be prepared for IELTS or TOEFL examinations and equipped them with academic vocabulary and grammar
 - Provided individual consultation in order to further assist new students in taking courses that addressed their specific needs or weaknesses

ESL Teacher

2001-2008

Sadr Institute of Higher education, Esfahan, Iran

Kian Farda Institute, Esfahan, Iran

- Taught grammar, reading, listening, speaking, and writing to adult learners from a variety of educational and vocational backgrounds
- Developed long range and daily lesson plans
- Created, adopted, and adapted student-centered communicative activities on a variety of topics
- Assessed students' improvement through developing, administering, and grading quizzes, tests, in-class and at-home assignments
- Provided constructive feedback on students' assignments and opportunities for individual consultation
- Provided students with clear explanations and instructions, opportunities for group activities, effective feedback, and individual consultation

CONFERENCES

TESL Ontario Approved Professional Development Credit

TESL Ontario Conference 2016

The Sheraton Center, Toronto, Ontario, Canada, November 2016

Measuring Productive Depth of Vocabulary Knowledge of the Most Frequent Words

EUROSLA 26

University of Jyväskylä, Jyväskylä, Finland, August 24-26, 2016

Measuring Productive Depth of Vocabulary Knowledge

TESL Ontario Conference 2015

The Sheraton Center, Toronto, Ontario, Canada, November 12-13, 2015

National Interdisciplinary Conference on Current Issues of English Language Teaching and Learning

Ahwaz, Iran, March 4-6, 2015

Measuring Depth of Vocabulary Knowledge

TESL Canada Conference 2014

The University of Regina, Regina, Saskatchewan, Canada, May 8-10, 2014

TESL Ontario Approved Professional Development Credit

TESL London, Spring Conference

London, Ontario, Canada, May 3, 2014

Beliefs and their effects on performances

5th International Conference of Cognitive Science

Tehran, Iran, May 7-9, 2013

Depth of vocabulary knowledge and effects on reading comprehension

Innovative Approaches to Second Language Teaching (IASLT),

Western University, London, Ontario, Canada, April 19, 2013

Vocabulary Knowledge Performances

Western University Research Day

Western University, March, 2013

Synforms perception by Persian-speaking learners of English

2nd International Conference on Philology, Literature, and Linguistics

Athens Institute for Education and Research, Athens, Greece, July 13-16, 2009

Learning styles and modality of input presentation

EUROSLA 17, University of Newcastle, UK, September 13, 2007

PUBLICATIONS

Word knowledge: Aspects, viewpoints and performances

Asian EFL, 14(3) September 2012, 265-298

Aspects of word knowledge: Receptive versus productive performances

Iranian EFL Journal, 8(1), February 2012, 333-358

Perception and word recognition in listening and reading comprehension *Research Week Proceedings of Islamic Azad University, Iran, October, 2009*

The design of the basic beginner course

Language Forum, 32(1-2), Jan-Dec 2006