The Role of Vocabulary Size in Assessing Second Language Vocabulary

Kevin John Zimmerman
Brigham Young University - Provo

Follow this and additional works at: https://scholarsarchive.byu.edu/etd
Part of the Linguistics Commons

BYU ScholarsArchive Citation
https://scholarsarchive.byu.edu/etd/578

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen amatangelo@byu.edu.
THE ROLE OF VOCABULARY SIZE
IN ASSESSING SECOND LANGUAGE PROFICIENCY

by

Kevin J. Zimmerman

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Arts

Department of Linguistics
Brigham Young University
April 2004
of a thesis submitted by

Kevin J. Zimmerman

This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

_________________________ __________________________________________
Date     C. Ray Graham, Chair

_________________________ __________________________________________
Date     Dee Gardner

_________________________ __________________________________________
Date     Diane Strong-Krause
As the chair of the candidate’s graduate committee, I have read the dissertation of Kevin J. Zimmerman in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date C. Ray Graham
Chair, Graduate Committee

Accepted for the Department

Alan Manning
Graduate Coordinator

Accepted for the College

Van Gessel
Dean, College of Humanities
This study investigated the role of vocabulary size in assessing learners from various L1 backgrounds and for institutional placement. Participants included native speakers of Spanish, Korean, Japanese, and Mandarin, who took a test that estimated the size of their productive English vocabulary. First, the vocabulary size of students from each language group was compared and the relationship between vocabulary size and institutional placement level was examined. Then, scores were analyzed to determine how cognates and loan words influenced the participants’ performance on the vocabulary test. Further, students’ vocabulary size scores were correlated with their placement scores that evaluated reading, writing, speaking, listening, and grammar. Finally, the vocabulary size of new students was compared to returning students within the same institutional placement level.

Within the same institutional placement levels, speakers of Spanish and Korean had larger English vocabularies than speakers of Japanese and Chinese. It was also found...
that there is a clear distinction between students at varying institutional placement levels in terms of their vocabulary size. The cognate analysis revealed that students learn cognates faster than non-cognates. The correlation analysis revealed that vocabulary size correlated most with speaking, and listening, followed by grammar, then reading and writing. Finally, the cross-sectional analysis indicated that the vocabulary size of newly placed students was generally larger than that of continuing students.
ACKNOWLEDGEMENTS

I would like to thank my Father in Heaven, of whose guiding influence I was well aware in the process of writing this thesis. Thanks to my beloved wife Roseli, whose unwavering belief in my abilities and love for me were a tremendous source of motivation and support. Thanks to Dr. Graham, who gave me the idea for this project and to Dr. Strong-Krause, who aided me in carrying out the study. Thanks to Dr. Gardner, who encouraged me to change the organization and focus of my thesis. His love for words and their meanings has been infectious. Thanks to Farrah Keeler for administering the pilot study, and to the ELC writing teachers for administering the test. Thanks to Dr. Eggett, who helped me interpret the results of the statistical analyses. Finally, thanks to Randy Robertson, who provided valuable suggestions for revising earlier drafts of this thesis.
TABLE OF CONTENTS

Table of contents .................................................................................................................. viii
List of tables and figures..................................................................................................... xi

Chapter 1: Introduction ......................................................................................................... 1

Chapter 2: Review of Literature .......................................................................................... 4

  Testing Breadth vs. Depth ................................................................................................. 5
  Testing Receptive vs. Productive Knowledge ................................................................. 6
  The Role of L1 on Vocabulary Test Performance ......................................................... 7
  The Role of Vocabulary Size in Assessing for Institutional Placement ....................... 9

Correlation Studies ............................................................................................................. 10

  Reading comprehension ................................................................................................. 11
  Writing ability .................................................................................................................. 11
  Listening comprehension ................................................................................................. 13
  Summary of correlation studies ...................................................................................... 13

Newly Placed vs. Continuing Students ............................................................................. 14

Vocabulary Size Tests ....................................................................................................... 15

  Eurocentres Vocabulary Size Test 10KA ................................................................ 16
  Vocabulary Levels Test ................................................................................................. 17
  Productive Vocabulary Levels Test .............................................................................. 19

Research Questions ........................................................................................................... 22
Chapter 3: Methodology ................................................................. 24

Instruments .................................................................................. 24

The Productive Vocabulary Levels Test ....................................... 24

Cognate questionnaire ................................................................. 26

Procedures .................................................................................. 27

Pilot test ..................................................................................... 27

Primary test ................................................................................ 27

Participants ................................................................................ 28

Statistical Analyses ..................................................................... 29

Chapter 4: Findings and Discussion ............................................. 31

Hypothesis 1 ................................................................................ 32

Hypothesis 2 ................................................................................ 36

Hypothesis 3 ................................................................................ 40

Hypothesis 4 ................................................................................ 41

Chapter 5: Conclusion ............................................................... 44

Implications ............................................................................... 49

Limitations ................................................................................ 50

Directions for future research .................................................. 51

References ............................................................................... 53
Appendix A: Integrated levels version of the
Productive Vocabulary Levels Test, Part A ........................................59
Appendix B: Integrated levels version of the
Productive Vocabulary Levels Test, Part B .................................61
Appendix C: Student questionnaire and instructions (Spanish) .........................64
Appendix D: Student questionnaire and instructions (Japanese) .........................66
Appendix E: Student questionnaire and instructions (Korean) .............................68
Appendix F: Student questionnaire and instructions (Mandarin) .........................70
Appendix G: Cognate questionnaire .........................................................72
Appendix H: Cognates and loanwords on the
Productive Vocabulary Levels Test (Asian languages) .........................77
Appendix I: Cognates and loanwords on the
Productive Vocabulary Levels Test (Spanish) ......................................78
Appendix J: Student questionnaire ............................................................80
Appendix K: Instructions to the teacher .....................................................81
Appendix L: Instructions for section 2 .......................................................82
LIST OF TABLES AND FIGURES

Table 1: Mean Scores on the Productive Vocabulary Levels Test by L1 and Level ......32
Table 2: Analysis of Variance for L1 and Level ..................................................34
Table 3: Institutional Placement Levels Differentiated by Vocabulary Size Scores ......35
Table 4: Estimated Productive Vocabulary Size at Varying
Word Frequencies by Institutional Placement Level
as Measured by the Productive Vocabulary Levels Test .........................36
Table 5: Number and Percentage of Items on the Productive Vocabulary
Levels Test Eliciting Cognates at Each Word Frequency Level ..................37
Table 6: Percentage Correct of Questions
Eliciting Cognates and Nonognates for All Levels .................................37
Table 7: Mean Percentage Scores on Cognates and Noncognates by Level ..........38
Table 8: Correlations of Vocabulary Size with Language Skills .........................41
Table 9: Mean Productive Vocabulary Size Scores of
Newly Placed and Continuing Students by Level ......................................42
Table 10: Analysis of Variance for Months Enrolled .......................................43
Table 11: Test of Differences Between Newly Placed and Continuing Students .........43
Figure 1: Mean Scores on the Productive Vocabulary Levels Test by L1 and Level .....33
Figure 2: Cross-sectional Growth of Cognates and Noncognates by Level ............39
Interest in vocabulary testing has been mounting over the past decade or so, and recent studies of vocabulary size have heightened the importance of vocabulary assessment in language testing (Laufer, 1992, 1997, 1998; Laufer and Nation, 1995, 1999; Meara, 1992, 1996, 2000; Nation 1983, 1990, 1993; Read, 1988, 1998, 2000; and Schmitt 1994, 1997, 1999). Several studies have explored the role of students’ L1 in their performance on vocabulary tests, including the role of cognates (García, 1988, as cited in Nagy, 1993; García, 1991; Hancin-Bhatt, 1994, Meara, 1992; Read, 1998; Tadamitsu, 2001). There has also been some investigation into the role of vocabulary size in assessing for institutional placement (Laufer, 1998; Laufer and Nation, 1999).

A study of word frequency counts reveals that knowledge of the 2,000 most frequent word families constitutes a threshold of the words required for basic oral communication (Schonell, Meddleton, Shaw, Routh, Popham, Gill, Mackrell, and Stephens, 1956). These 2,000 word families make up about 87% of written texts and about 80% of typical academic texts (Nation, 1990). The studies looking at the relationship between vocabulary size and reading comprehension support the threshold hypothesis, which states that there is a threshold of vocabulary necessary for comprehending written texts (Alderson, 1984). In general, no fewer than 1 in 20 words should be unknown to the reader in order to comprehend a given text. In terms of vocabulary size, this means that a student ought to know about 3,000 word families to comprehend most texts (Laufer, 1996, 1997, as cited in Qian, 1999). About 4,000-5,000 word families are necessary for comprehending more academic texts, such as an economics textbook (Sutarsyah, Nation, and Kennedy, 1994). In addition, Hazenberg
and Hulstijn (1996, as cited in Beglar, 2000) estimate that a base of at least 10,000 word families, half of what an average adult native speaker knows, is needed to study at the university level.

A principle underlying this study is that vocabulary provides the “enabling knowledge” required to be successful in other areas of language proficiency (Laufer and Nation, 1999). Indeed, words are the primary carriers of meaning (Vermeer, 2001), and there is growing evidence that the more extensive one’s vocabulary, the higher their language proficiency will be. The size of a student’s vocabulary has been found to correlate closely with reading comprehension (Beglar, 1999; Laufer, 1992; Qian, 1999) as well as with writing ability (Astica, 1993; Beglar, 1999; Laufer, 1998; Laufer and Nation, 1995; Linnarud, 1986).

Vocabulary testing has also been found to be a useful tool in diagnostic or placement exams. Tests of vocabulary size can discriminate between groups of learners (Meara, 1992) and aid in admissions (Laufer, 2002), as well as help in placing students into appropriate institutional placement levels within a program (Laufer and Nation, 1999; Schmitt, 1994). If used for diagnostic purposes, vocabulary size tests can allow teachers to identify and remedy deficiencies in their students’ vocabularies (Schmitt, 1994).

The current study attempts to confirm some of the findings of previous studies regarding the role of vocabulary size in assessing learners from various L1 backgrounds and for institutional placement. In addition, this study applies vocabulary testing in a new way by measuring a perceived English proficiency gap between newly placed and continuing students, an issue explored by Brown (1980). Participants include students
whose L1s are Spanish, Korean, Japanese, and Mandarin studying at an intensive English language program in Utah.

The word “cognate,” as used in this study, encompasses two concepts. First, it is used in the traditional sense, which is any two words that are aurally or orthographically similar and are derived from the same ancestor language (e.g., similar words between English and Spanish). The word “cognate” is also used in this study to encompass the usual meaning of “loanword,” which is any two words that are aurally or orthographically similar because of borrowing (e.g. English words borrowed by the three Asian languages in this study: Korean, Japanese, and Mandarin).

Two main questions will guide the present research. First, what is the role of vocabulary size in assessing learners from various L1 backgrounds? Second, what is the role of vocabulary size in assessing for institutional placement?
CHAPTER 2: REVIEW OF LITERATURE

There were two main purposes of the present study. The first was to investigate the role of vocabulary size in assessing learners from various L1 backgrounds. The second was to investigate the role of vocabulary size in assessing for institutional placement. To establish the research questions of this study, the following literature review will examine issues involved in vocabulary testing in general, in assessing learners from various L1 backgrounds, and in assessing for institutional placement. Finally, three widely used vocabulary size tests will be considered.

*Testing Breadth vs. Depth*

One of the general issues that has emerged regarding vocabulary testing is the distinction between vocabulary *breadth*, or size, and vocabulary *depth*. Vocabulary breadth tests measure how many word families a person knows, and usually only the primary meaning of the words. Three widely used breadth tests are the Eurocentres Vocabulary Size Test 10KA (EVST, Meara and Jones, 1990), the Vocabulary Levels Test (VLT; Nation, 1983, as cited in Laufer and Nation, 1991), and the Productive Vocabulary Levels Test (PVLT; Laufer and Nation, 1995). Vocabulary depth tests, on the other hand, measure how well a person knows the words, including the synonymy, polysemy, and collocations of the words. One widely used depth test is Paribakht and Wesche’s (1993) Vocabulary Knowledge Scale (VKS, as cited in Waring, 2002).

The division between vocabulary breadth and depth, however, is not as distinct as it may seem on the surface. Qian (1999) concluded from his vocabulary study of 77 Korean and Chinese speakers that a knowledge of primary word meaning can often include a knowledge of synonymy, polysemy, and collocation, especially if the words
share these elements in the students’ L1. Thus, the constructs of depth and breath of vocabulary partially overlap (Qian, 1999).

It is possible that students may seem to know a word presented on a vocabulary breadth test yet be unable to understand or use its various inflections and derivations (Beglar, 1999; Mezynsky, 1983), or to use the word in particular contexts (Mezynsky, 1983). Thus, a student who knows the word *obey*, may be able to produce *obeys*, *obeying*, and *obeyed*, which follow regular inflectional patterns, but may not be able to produce the words *obedience*, *obedient*, *obediently*, and *disobey*. Indeed, there is little empirical evidence that if students know the base form of a word that they will also know inflected or derived forms of that word (Beglar and Hunt, 1999). There is also little evidence that a student who understands the primary meaning of a word will understand the meaning of that word in a fixed, multiword expression.

It is likely that many language teachers have observed that a large L2 vocabulary does not necessarily imply depth of knowledge about those words. Read (2000) contends, however, that while size tests “may seem superficial, they can give a more representative picture of the overall state of the learner’s vocabulary than an in-depth probe of a limited number of words.” Qian (1999) came to a similar conclusion after he administered the Vocabulary Levels Test (VLT) and the Word Associates Format to 44 Korean speakers and 33 Chinese speakers. He found that the scores of the two tests were closely and significantly correlated at .78 for the Korean speakers and .82 for the Chinese speakers.

Vermeer (2001) agrees that there is no good reason to suppose that vocabulary size tests are any less valuable than vocabulary depth tests, since growth in vocabulary breadth and depth are dependent on the same thing – frequency of input. She
administered a breadth test and a depth test to 1,600 Dutch monolingual and bilingual 4- and 7-year olds. She found that the scores had significant ($p < .01$) correlations of .73, .83, and .93 on three different measures, and concluded that there is no real distinction between vocabulary breadth and depth. She explained that “a deeper knowledge of words is the consequence of knowing more words, or that, conversely, the more words someone knows, the finer the networks and the deeper the word knowledge.” It is clear, therefore, that the literature still tends to distinguish between vocabulary breadth and depth, although recent studies have begun to blur the distinction.

An advantage to using a breadth test is that it can assess a wide range of words in a relatively short amount of time, making it especially useful for assessing students with widely varying proficiencies.

*Testing Receptive vs. Productive Knowledge*

A second issue that has emerged regarding vocabulary testing is the distinction between *receptive* vocabulary and *productive* vocabulary. Students have receptive knowledge of a word if they are able to retrieve the appropriate meaning when reading or hearing the word (both receptive language skills). The EVST, VLT, and VKS are receptive vocabulary tests. Student’s have a productive knowledge of the word if they are able to produce the appropriate word when speaking or writing (productive skills). The PVLT is a productive vocabulary test.

Laufer (1998) divides productive vocabulary into *controlled* production and *free* production. This distinction differentiates students’ ability to produce a word when required to do so by a teacher or test (controlled) and their actual production when writing or speaking on their own (free; Laufer and Nation, 1999). Laufer (1998) found
that students may not feel confident enough to use the word when writing or speaking, but instead opt for another word that is perhaps more frequent, or to circumlocute to communicate the same meaning.

One’s receptive vocabulary is usually larger than one’s productive vocabulary. For instance, a learner may recognize the word *impede* and even be able to provide a synonym when encountering it within a context, but be unable to recall the word while writing or speaking (Laufer, 2002). Laufer (1998) used the PVLT to assess the productive vocabulary of twenty-six 16-year olds and twenty-two 17-year olds, then compared the scores to their scores on the original VLT, which assesses receptive vocabulary knowledge. She found that receptive vocabulary size exceeded productive vocabulary size by 200 word families, or 10 percent, for the 16-year olds and by 950 word families, or 32 percent, for the 17-year olds. Similarly, Schmitt (1997) found that scores on the receptive VLT were 19 to 25 percent higher than on an equivalent productive vocabulary test.

*The Role of L1 on Vocabulary Test Performance*

One factor that may contribute to varying rates of English vocabulary acquisition is the students’ L1. It is not atypical in many programs to see, for example, speakers of Spanish in the same classroom as speakers of Mandarin. Many Indo-European languages, such as French, German, Greek, Italian, Portuguese, and Spanish, share many cognates with English, and these cognates are usually academic and low-frequency words such as *inherent, homogeneous, impede, and obscure* (Meara, 1992). This means that speakers of these languages may be able to guess the meaning of these low-frequency words and appear to have a sizeable vocabulary in English while they lack a knowledge of basic
vocabulary (Meara, 1992; Read, 1998). On the other hand, students whose mother
tongue has little lexical overlap with English, such as Korean or Mandarin, may require a
great deal more time to acquire the words that some of their classmates seem to learn
almost effortlessly. After a semester or two, despite high motivation to learn and diligent
studying, these students who were placed into the same institutional placement level at
the beginning of their course are still studying in the same classroom together, but their
true language abilities vary widely. Understandably, administrators are often reluctant to
make slower learners repeat a level once they have already covered the material,
especially since they pay the same tuition as their classmates who have made greater
gains in the language.

To what extent cognates aid students in guessing is unclear, however. García
(1988, as cited in Nagy, 1993; García, 1991) found that some native Spanish speakers
learning English failed to guess the meaning of cognates that were identical semantically
and orthographically to common Spanish words. Hancin-Bhatt (1994) tested 4th year
and 8th year Spanish-speaking students learning English and found that the ability to
recognize cognate relationships and to use them in determining meaning increased with
age and proficiency faster than noncognates. In a similar study, Nagy (1993), who tested
the reading comprehension of 74 Spanish/English bilingual students in Grades four, five,
and six, also found that the students with larger English vocabularies were more likely to
recognize cognates.

Since García’s 1998 study, Tadamitsu (2001) has examined the role of cognates
in performance on the VLT. Tadamitsu administered versions A and B of the 1-2,000,
2,000-3,000, and 3,000-5,000 word levels to 196 Japanese EFL students and found that
over 40% of his participants scored higher on the 2,000-3,000 word level on version A and about 28% of his participants scored higher at the 1-2,000 word level on version B. Tadamitsu argues that a major question yet to be answered is whether the number of cognates included in the test is representative of the total number of cognates at each level. He suggests that the percentage of cognates in learners’ L1 be considered when revising the VLT for a specific group of learners. Instead of revising the test, however, it would be possible to simply interpret the scores of examinees from varying L1s.

The Role of Vocabulary Size in Assessing for Institutional Placement

Two recent studies have examined the role of vocabulary size in assessing for institutional placement. First, Laufer (1998) used scores from the PVLT to compare the productive vocabulary size of twenty-six 16-year olds to that of twenty-two 17-year olds. She found that the 16-year olds knew approximately 1,700 words, and the 17-year olds knew about 2,550 words, a difference of about 850 words. Warring (1999), however, voiced the concern that nothing is known about these two groups other than that they differed by a single school year. No information was given regarding their general academic level, motivation, attitude toward English, and so on.

The following year, Laufer and Nation (1999) conducted a study on the validity and reliability of the PVLT scores. Their aim was to see whether groups at varying English proficiencies would produce significantly different scores at each word frequency level, and whether the test could distinguish between the different institutional placement levels. They tested twenty-four 10th graders, twenty-three 11th graders, eighteen 12th graders, and fourteen 1st year university students studying in the English department. The scores consistently increased with each higher grade, indicating that the test is a valid
measure of vocabulary growth. They also found significant correlations between four parallel versions of the test at the $p = .0001$ level. The reliability of the PVLT scores had a KR21 score of .86.

A major oversight in reporting their data, however, is that they never mention where the study took place or what the native languages of the students were. The reader is left to presume that the study took place in Haifa, Israel, where Laufer lives, and that the participants were native speakers of Hebrew, but this is only an assumption. They also claim that the class grade was a good indication of the participants’ English proficiency, yet they provide no other measures of the participants’ proficiency, such as grades or scores on other standardized English tests.

**Correlation Studies**

Vocabulary size has been correlated with general intelligence (Hutson, 1983) and with language proficiency in general (Hever, 1995). If the correlation between language skills and vocabulary size is high, a vocabulary size test may be a good placement tool (Laufer, 2002; Meara, 1992; Schmitt, 1994). No empirical studies were found correlating vocabulary size with speaking, which is surprising given the primacy of speech. There were also no studies found correlating vocabulary size with grammar scores, although grammar has been found to correlate poorly with other language skills. Meara (1992) does mention that scores on the EVST correlate poorly with speaking and well with grammar skills, but does not give the actual correlation scores. The following are summaries of studies that have correlated vocabulary size with reading, writing, and listening.
Reading comprehension.

Laufer (1992) administered the VLT and the Eurocentres Vocabulary Test to 92 university freshmen whose native language was either Arabic or Hebrew. She found that the scores on these tests correlated with reading comprehension scores at .50 and .75 respectively. Qian (1999) found more robust correlations in his study of 44 Korean speakers and 33 Chinese speakers using version A of the VLT and the Reading Comprehension section from the TOEFL. The correlation between the two tests was .78.

In the same year, Beglar (1999) conducted an almost identical study with 496 Japanese high school students using four versions of the 2,000 word level from the VLT and 464 students with four versions of the UWL from the same test. He found that scores on the Reading Comprehension subsection of the TOEFL test correlated with scores on versions A and B of the 2,000 word level at .66 and .62, respectively, and with scores on versions A and B of the UWL section at .67 and .71, respectively.

Qian’s findings are 11 and 7 percentage points lower than Beglar’s, respectively. Since these studies were both correlational, this difference between the findings of the two studies cannot be explained by the native language of the participants or their proficiency. Regardless of the difference in the correlations, they both point to a statistically significant and positive correlation between vocabulary size and reading comprehension.

Writing ability.

The following five studies applied varying methods, but found substantial correlations between vocabulary size and writing ability.

In the study mentioned above, Beglar (1999) also compared the TOEFL Structure
and Written Expression subsection with two of the four versions of the VLT, versions A and B, and found that the correlations were .61 and .65 respectively, slightly lower than the correlations found with reading comprehension.

In a rather different experiment from Beglar’s, Astica (1993) used a scoring technique similar to the ESL Composition Profile, which analyzes content, organization, vocabulary, language use, and mechanics, to score 210 writing samples. A multiple regression analysis revealed that 84% of the variance could be accounted for by vocabulary. Similarly, Linnarud (1986) analyzed compositions written in Swedish by native and non-native speakers of Swedish. She correlated the compositions in terms of total number of words for each composition, number of words per sentence, lexical individuality, and lexical sophistication. Like Astica, she concluded that vocabulary size was the single largest factor in writing quality, but did not provide actual correlation numbers.

Laufer and Nation (1995) used Nation’s Range program, also known as VocabProfile, to produce a Lexical Frequency Profile (LFP) of student compositions. The LFP examined two sets of compositions written by 65 students solely on the basis of the number of words that the students used from the first 1,000 most frequent words, second 1,000 most frequent words, words from the UWL, and a last group considered low frequency or unknown words. The aim of the study was to determine to what extent such a profile would correlate with the students’ scores on the PVLT. The reliability of the PVLT scores in the study was .86, using KR-21. They found that the students who had larger vocabularies used fewer high frequency words and more low frequency words than the students with smaller vocabularies. Vocabulary scores from the UWL section of
the PVLT correlated with the number of academic words students used in the two sets of compositions at .70 and .60.

In a later study, Laufer (1998) conducted a similar study in which she correlated participants’ scores on the PVLT with the type of vocabulary they used in writing a composition, as analyzed by the LFP. Twenty 16-year olds who had studied English for 6 years and twenty-two 17-year olds who had studied English for 7 years took part in the study. She found that the 17-year olds had larger vocabularies than the 16-year olds. Unlike her former study, she found that the 17-year olds, who had the larger vocabularies, did not use a greater breadth of words while writing their compositions. She attributes this to the students’ lack of incentive to use less frequent vocabulary, opting instead to use safer and more familiar high-frequency words in their writing.

*Listening comprehension.*

Very few studies were found that correlate vocabulary size and listening comprehension. Beglar and Hunt (1999) compared listening comprehension with scores on four versions of the 2,000 word level of the VLT with 496 Japanese high school students and on four versions of the UWL level with 464 students. They used the Listening Comprehension subsection of the TOEFL as their instrument. They found correlations of .44 and .45, respectively, which is lower than the correlation scores that they had found with reading and writing in the same study.

*Summary of correlation studies.*

The research reviewed here indicates that vocabulary size, as measured by the tests used in these studies, is more highly correlated with reading comprehension and writing ability, and less so with listening comprehension. This pattern may in part be
explained by the fact that the VLT was developed using words derived from literacy-based frequency lists. In addition, vocabulary size tests are nearly always administered in written form, demanding literacy skills more than oral skills.

Although the correlation studies discussed above suggest that vocabulary size is a substantial contributor to proficiency in each of the language skills, Schmitt (1999) argues that using vocabulary scores to predict future language learning achievement is limited, since other factors such as motivation, aptitude, and learning opportunities contribute to a student’s ability to succeed in learning a second language. This appears to be true, but these other factors will not be discussed in depth in the present study.

**Newly Placed vs. Continuing Students**

Students who come from abroad to enroll in intensive language programs are usually required to take a placement exam to determine at which institutional placement level they will begin their ESL studies. These exams usually assess reading, writing, speaking, listening, and grammar proficiency. As the semesters go by, the students often graduate to the next levels, often without having to pass another placement test. Even though all the students are exposed to the same grammar lessons and read the same books, the language skills of these students develop at markedly different rates. New students enter the program and are placed at the appropriate level, but their language skills often seem superior to their classmates who have continued from the previous semester. Why is there this perceived English proficiency gap between newly placed and continuing students?

One possible factor contributing to this disparity is the tendency for returning students to be “socially promoted.” That is, they have completed the assignments and
have had good attendance, so the natural thing is to move them to the next level so they
can learn something new. However, learning a language takes time, and often more
information is covered within a semester than students are able to fully acquire.

The perceived English proficiency gap between newly placed and continuing
students appears to be widespread among ESL and EFL programs that use a placement
test to place new students into the appropriate institutional placement levels. Brown
(1980) has stated that many students who continue studying for more than one semester
would likely have to repeat their classes if they had to take the placement tests again.
Brown evaluated the test scores of 319 ESL learners at UCLA and found that the placed
students scored 6.71 (out of 50) points higher on a cloze test and 9.82 (out of 100) points
higher on the final examination than continuing students. This difference between the
newly placed student and the continuing student may mean that programs are expecting
students to make gains faster than they are capable in the time available. One possible
way to measure the perceived English proficiency gap between newly placed and
continuing students is to assess their vocabulary size.

*Vocabulary Size Tests*

There are currently three widely used vocabulary size tests available. The three
tests were investigated when selecting the instrument for this study. Before considering
the three tests, however, it is appropriate to discuss how they are constructed.

Current vocabulary size tests have followed certain guidelines pertaining to their
architecture. Vocabulary size tests that are used for proficiency or placement purposes
should include the broadest possible range of word families. The word families used in
vocabulary breadth tests are usually compiled by a computer that scans corpora and
orders the word families by frequency, from the most common to the least common word families. To ensure that test items are testing a single psychometric unit (vocabulary size), the other skills of reading, writing, speaking, and listening should be taxed as little as possible (Schmitt, 1994). Thus, words used in the definitions or surrounding text should always be more common than the word being tested (Schmitt, 1994).

An estimation of total vocabulary size can be attained in two ways. The first is based on sampling from a dictionary, and the second is based on corpus-derived lists of word families grouped by frequency (Nation, 1990). The dictionary sampling method involves selecting a dictionary that contains the number of word families that learners are expected to know, then testing a selection of those words. The word families tested may be the first word of every tenth page, for example. The problem with this method is that higher frequency words tend to have longer entries, and are thus more likely to end up on the test, which may skew the results (Nation, 1990).

The second method to estimate vocabulary size is to select word families according to their frequency in a corpus. Usually, these word families are grouped together into the first 1,000 most frequent words, the second 1,000 most frequent words, and so on. This kind of test has generally been used only with people with low English vocabularies, namely non-native speakers of English (Nation, 1990).

Eurocentres Vocabulary Size Test 10KA.

Meara and Jones’ (1990) Eurocentres Vocabulary Size Test 10KA (EVST) is a computerized vocabulary test. It was commissioned by the Eurocentres chain of European language schools, which needed a placement test that would minimize demand on the faculty’s time (Read, 2000). In the test, about one third of the words are nonsense
words, and the students’ responses to the nonsense words are used to adjust the final score. The real words are sampled from varying frequency groups from within the 10,000 most frequent words of English. Following is an example list of words from the EVST (Meara, 1992a, as cited in Read, 2000):

1 dring  2 take in  3 majority
4 modest  5 vowel   6 easy
7 within   8 obsolation  9 foundation

The test-taker sees the word on the screen, with the question “Do you know the meaning of this word?” The test-taker indicates ‘Yes’ with one key and ‘No’ with another. The test is efficient, usually taking less than ten minutes to sit, and the scores are easy to interpret. A popular feature of the test is that students can see their scores immediately (Read, 2000).

Many teachers and researchers, however, have their reservations about using the EVST. The most significant concern is that scores tend to be unreliable with low proficiency learners, since they lack the ability to distinguish between words and non-words, and tend to respond in unpredictable ways (Read, 2000). Second, the test does not require students to demonstrate knowledge of the meaning of the words. Finally, because the test is computerized, it cannot be administered easily to a large sample at one time.

_Vocabulary Levels Test._

The Vocabulary Levels Test (VLT; Nation, 1983, as cited in Laufer and Nation, 1991) is widely used to assess the breadth of vocabulary knowledge of ESL and EFL students. Meara (1996) has stated that the VLT was the “nearest thing we have to a
standard test in vocabulary,” at least up until he made the statement. The VLT uses the following format:

1 business
2 clock _______ part of a house
3 horse _______ animal with four legs
4 pencil _______ something for writing
5 shoe
6 wall

Eighteen word families are tested at each level. However, Nation argues that 36 word families are tested because the test-taker must check every word against the definitions in order to select the correct answer (Read, 2002).

Word families are tested from each of five different frequency groups: 1-2,000, 2,000-3,000, 3,000-5,000, words from the University Word List (UWL), and 5,000-10,000. Word families from the first 2,000-word level generally have a frequency of more than 332 occurrences in every 5 million words (Schmitt, 2001). Word families from the 2,000-10,000 word levels were taken from an even distribution across the Thorndike and Lorge list (1944), the General Service List (West, 1956), and data from Kucera and Francis (1967; Schmitt, 2001). The UWL (Xue and Nation, 1984, as cited in Beglar and Hunt, 1999), rather than being a list of 1,000 word families, contains 808 general academic words that account for roughly 8% of words not found in the first 5,000 most frequent word families (Beglar and Hunt, 1999), and for this reason is somewhat different from the other levels (Read, 1988). In terms of difficulty, words from the UWL fall between the 5,000 and 10,000 range (Laufer, 2002). (See Beglar, 1999, for more on the development of the VLT.)

The validity and reliability of several versions of the VLT have been verified (Beglar, 1999; Read, 1998; Schmitt, 2001). Beglar (1999) tested the reliability of the
VLT. In his study, 496 Japanese students ages 15 to 23 took four versions of the 2,000-word level test, and 464 participants took four versions of the UWL level test. He performed a Rasch analysis on these items and found that scores from the 2,000-word and UWL levels had very high IRT reliability at .98 and .95, respectively (a figure of over .70 is generally considered reliable). There were no misfitting items on the UWL test, and only one item (treasure) misfit, and two items (victory and ancient) overfit the model from the 2,000-word level. Beglar found no clear explanation for these problem items. The number of participants in Beglar’s study was large enough to conclude that the reliability of these findings is high.

Read (1988) reports that he administered the VLT at the beginning of the English Language Institute Proficiency Course to assist teachers in placement and in planning the course. At the end, the authors administered the test again and found that the reliability coefficients were very high for both pretest and posttest, at .94 and .91 respectively. Schmidt (2001) found similar results in his study with 801 participants, in which he compared scores from two versions of the test. Reliability coefficients ranged from .92 to .96 for different sections of the test.

In addition to being valid and reliable, the VLT is easily obtainable, as Nation has published it as an appendix several times (1990, 1993, 2001; also Schmitt, 2001).

*Productive Vocabulary Levels Test.*

The third vocabulary size test, a modified version of the VLT, is called the Productive Vocabulary Levels Test (PVLT; Laufer and Nation, 1995). The format is similar to a C-test, which provides half the word. The PVLT, however, provides only enough letters to “disambiguate the cue” (Laufer and Nation, 1999). The authors piloted
a version of the test to seven native speakers, and letters were added to each word until each item could be answered correctly by six or more of the seven native speakers (Laufer and Nation, 1999). The underlined spaces following each item are the same length, regardless of the number of letters needed to complete the word.

Laufer and Nation (1999) concluded their study by saying that “the Productive Vocabulary Levels Test is a very practical instrument.” Certainly, the test is easy to find, as they have published it as an appendix to two of their articles (1995; 1999), and it can also be found on the internet (http://www.er.uqam.ca/nobel/r21270/levels/). In addition, results are fairly easy to score, as answers are counted as right or wrong. Besides being easy to score, the test does not take long to complete. Beglar (1999), using a version of the VLT that integrated four versions of the 2,000 and university word levels, found that most students finished the 72 items in less than 30 minutes. The PVLT may take longer to administer since there are 90 items instead of 72, and because examinees are required to write the missing part of the words rather than circle the correct answer.

A few concerns about the PVLT have been raised. Meara (2000) argues that a sampling of only 18 word families from each level may be insufficient. He points out that at the 5,000-10,000 word level, only 1 in every 278 word families are being sampled, which may not give a very precise indication of a student’s vocabulary knowledge at that level. However, Laufer and Nation (1999) found that four parallel versions of the PVLT produced similar results, indicating that testing a sample of 18 word families should produce similar results to testing 54 words.

Another concern is that each item tests not only the word being elicited, but all the other words as well (Meara, 1992). This should not be a major concern because the
words in the surrounding text are always more frequent than the target word, and higher frequency words are generally acquired first (Anderson and Freebody, 1981; Laufer, 2001; Nation, 1990).

Yet another concern is that the UWL, which is used in the PVLT, has been upstaged by the newer Academic Word List (AWL; Coxhead, 2000), which lists fewer word families (570 instead of 808), yet provides better coverage of words found in academic texts (Schmitt, 2001).

In addition, the word families tested in the VLT and PVLT ostensibly represent English vocabulary across all frequency levels, but no consideration was given while designing the test to the percentage of cognates represented in the test in relation to other languages. For example, if it is found that 10% of the elicited words on the test are cognates with Russian, is 10% the true percentage of lexical overlap between English and Russian? As Read (1998) points out, “further research is needed to establish the reliability and validity of the [VLT] as a measure of vocabulary knowledge and to account for the variation in results according to the learner’s L1.”

In summary, the three most widely-used vocabulary breadth tests, the EVST, the VLT, and the PVLT, were considered when selecting the instrument for the current study. Of the three, only the PVLT measures productive vocabulary size. The PVLT also stands out in terms of its format because it requires the student to produce the word as they would if they were speaking or writing, and thus seems to more realistically mirror natural language use, whereas the EVST contains nonsense words and the VLT is multiple-choice. In addition, although the validity and reliability of the VLT has been studied more than the other two tests, the findings may be extended to the PVLT since
the same words are tested. Finally, the PVLT has been made available on the internet, and a pen and paper version of the test can easily be administered in class, which is not true with the computer-based EVST. Therefore, the PVLT was determined to be the most suitable instrument for the present study.

Like most vocabulary tests, the PVLT is administered in written form, but differs from most other vocabulary tests in that it requires students to produce the words instead of simply to recognize them. For this reason, one may find that scores on the PVLT will correlate more highly with writing scores than with reading comprehension. This remains to be seen, however, since, as Laufer’s (1988) study revealed, a student may be able to produce a word on the test but may avoid using the word while writing. In addition, writing raters usually consider not only the vocabulary, but also the organization, ideas, length, voice, grammar, and punctuation to score a student’s writing.

Research Questions

Based on the review of literature, the following questions and hypotheses have been established. Questions 1 and 2 (in part A) deal with assessing learners from various L1 backgrounds; Question 1 investigates the role of L1 and question 2 investigates the role of cognates. Question 3 (in part B) relates to the role of vocabulary size in assessing for institutional placement. Finally, question 4 (part C) inquires whether vocabulary size scores can differentiate newly placed and continuing students.

A. The role of vocabulary size in assessing learners from various L1 backgrounds

1. Is there any difference in vocabulary size scores, as measured by the PVLT, by native language and by institutional placement level?

2. Is there any difference in vocabulary size scores between cognate and noncognate items at each institutional placement level, as measured by the PVLT?
B. The role of vocabulary size in assessing for institutional placement

3. Is there any relationship between vocabulary size scores, as measured by the PVLT, and scores on the institutional placement exams in reading, writing, speaking, listening, and grammar?

C. Case study

4. Is there any difference between the English vocabulary size scores of newly placed students and continuing students within the same institutional placement level, as measured by the PVLT?
CHAPTER 3: METHODOLOGY

In the current study, the role of vocabulary size in assessing learners from various L1 backgrounds was investigated by measuring the vocabulary size of non-native English speakers at varying institutional placement levels. The role of cognates in the students’ performance on the vocabulary test was also explored. Vocabulary size scores were then correlated with other language skills. Finally, vocabulary assessment was applied in a novel way to measure the perceived English proficiency gap between newly placed and continuing students.

**Instruments**

*Productive Vocabulary Levels Test.*

The primary instrument used in this study was the Productive Vocabulary Levels Test (PVLT). The PVLT is a diagnostic test developed by Laufer and Nation (1995; see Appendixes A and B). The PVLT was chosen because it is a breadth test that is easy to administer and score. In addition, it requires students to produce the words rather than to simply recognize them. A correct answer thus shows a stronger command of the word.

The PVLT uses the following format:

I’m glad we had this opp_________ to talk.

Each of the five frequency levels of the PVLT is represented by 18 items on the test, making 90 questions total. Within each level, the items are presented in order of higher to lower frequency. Thus, the questions tend to increase in difficulty since lower frequency words tend to be more difficult. Because the words are taken from a leveled sampling, scores on the test provide a rough estimate of the students’ vocabulary size. For instance, if a student testing at the 1-2,000-word level gets 9 out of the 18 items
correct, it can be assumed that he or she knows roughly 500 out of the 1,000 word families from that level. Furthermore, since higher frequency words are generally acquired first, the rest of the words in each sentence are always more frequent than the word being tested.

Items are considered correct if students write the correct word and part of speech, even if there are mistakes in spelling or grammar. For example, in the item, “In order to be accepted into the university, he had to impr_______ his grades,” the ideal answer is improve. The words improves and improved, even with spelling mistakes, would be considered correct, since vocabulary is what is being tested. The words improvement, improvise, etc. would be marked as incorrect, since they belong to different parts of speech or a separate headword. Whether a student has satisfactorily mastered a level or not is determined by the administrator of the test, but a score of 85% to 90% at the 2,000 word level would indicate that the student can use the most frequent words of English.

Three changes were made to the format of the PVLT for this study. First, the spelling of a word in question 20 (see Appendix A, Part B) was changed from the British favourite to the American favorite. There were no other non-American spellings in the original PVLT.

The second change made to the PVLT was to include an example after the instructions (see Appendix A, part B and Appendix B). The example is an item taken from the 1-2,000 word section of a parallel version of the PVLT.

The third change was to the format of the PVLT. Questions from all five levels (1-2,000, 2,000-3,000, 3,000-5,000, UWL, and 5,000-10,000) were integrated in the following manner: question 1 of the revised test came from the 1-2,000 word level,
question 2 from the 2,000-3,000 word level, question 3 from the 3,000-5,000 word level, question 4 from the UWL, question 5 from the 5,000-10,000 word level, and then repeating in this manner with question 6 from the 1-2,000 word level, and so on. This was done to prevent more advanced students from feeling unchallenged in their answers in the high frequency sections of the test, and so that lower-proficiency students would not find the low frequency sections so difficult that they would give up.

Finally, the questionnaire and instructions for each part of the test were translated into Japanese (Appendix C), Korean (Appendix D), Mandarin (Appendix E), and Spanish (Appendix F) for the native speakers of these languages at levels one and two of the program. Students in levels 3-5 received the instructions in English.

*Cognate questionnaire.*

Another instrument used in this study was a cognate questionnaire (see Appendix G). The word “cognate,” as used in this study, refers to cognates as well as loanwords. Native speakers were asked to identify the underlined words of the PVLT that were orthographically or aurally similar to any word in their native language. All of the native speakers were teaching assistants or professors of their native language at Brigham Young University. Since there may be lexical differences depending on geographic location, the number of native speakers who were surveyed varied. Varying dialects of Japanese and Korean were not a concern, therefore three native speakers were deemed sufficient. Four Mandarin speakers were surveyed, two from Taiwan and two from mainland China, and nine Spanish speakers were surveyed, four from Argentina, two from Spain, and the other three from Columbia, Peru, and Mexico.
An example of identifying cognates was provided on the cognate questionnaire, and the researcher sat with the native speakers to answer questions if necessary. When there was not complete agreement between the native speakers of a given language regarding which words were cognates, an additional native speaker was consulted in order to make a final decision. Results of the cognate questionnaire are found in Appendixes H and I.

Procedures

Pilot test.

A pilot study was conducted at an intensive language program at Utah Valley State College in Orem, Utah. The test was administered to a class of 15 ESL students of varying L1 backgrounds. Everyone completed each of the two sittings of the test in approximately 20 minutes. No changes were made to the test itself as a result of the pilot study. On the questionnaire, however, question two was added to obtain the student’s gender, and for question five, the boxes with the months and years were added, replacing a simple line where students were supposed to write the date that they began studying at the program (see Appendix J).

Primary test.

Before administering the vocabulary test, the researcher met with the writing teachers to distribute copies of the test and to explain the purpose of the study. Each teacher also received instructions for administering the test (Appendix K). The vocabulary tests were administered to all institutional placement levels of the program at the beginning of the winter semester, January 2003, during the students’ writing classes. Since the test is rather long, containing 90 fill-in-the-blank type questions, the test was
divided into two parts (see Appendix A for Section 1 and Appendix B for Section 2) and administered over two days to minimize fatigue. Students received instructions before taking Section 2 of the test on the second day (see Appendix L).

Items on the PVLT were scored correct if the student wrote the correct word in the correct part of speech, even if there were spelling errors, and incorrect if the student wrote nothing, wrote the correct word but in the wrong part of speech, or wrote any other word. A second rater scored ten of the tests, resulting in an interrater reliability of .99.

Participants

The study was conducted at Brigham Young University’s English Language Center (ELC) in Provo, Utah. The ELC draws international students from all over the world, but mostly from South America, Korea, Japan, and China. Incoming students at the ELC take tests that were developed by the ELC and which assess their reading, writing, speaking, listening, and grammar skills. Based on their scores, they are placed into one of five levels. Level-one students are at a beginning novice level, and level-five students are at a high intermediate level of English proficiency.

Every student enrolled at the ELC took the PVLT. However, only the scores of Spanish, Korean, Japanese, and Mandarin speakers who completed sections A and B of the test were considered. The number of participants, therefore, was 173, but fewer students were considered for some analyses.

Statistical Analyses

Following are the four research questions, restated as hypotheses, followed by a description of the data analysis performed for each hypothesis. Each analysis was
performed at the $p = .05$ level.

A. The role of vocabulary size in assessing learners from various L1 backgrounds

1. There is no difference in vocabulary size scores between native speakers of Spanish, Korean, Japanese, and Mandarin, nor by institutional placement level, as measured by the PVLT.

A 2-way ANOVA was performed to find the interaction between each L1 and institutional placement level.

2. There is no difference in vocabulary size scores between cognate and noncognate items, as measured by the PVLT.

The 2-way ANOVA was performed, with the results being analyzed in 2 groups: cognate and noncognate items on the test. The analyses were performed separately for each language except Mandarin, which had only one cognate on the test. A Tukey-adjusted pairwise $t$ test was also performed to compare the students’ performance on items eliciting cognates with those eliciting noncognates.

B. The role of vocabulary size in assessing for institutional placement

3. There is no relationship between vocabulary size scores, as measured by the PVLT, and scores on institutional placement exams in reading, writing, speaking, listening, and grammar.

A Pearson correlation was computed to measure the correlation between the vocabulary score and the scores for reading, writing, speaking, listening, and grammar. Only scores of new ELC students, i.e. those who had enrolled in January, 2003, were considered for this analysis, since it was expected that returning students would have increased somewhat in proficiency since the time of their placement exams. The correlation significance was set at .05.
C. Case study

4. There is no difference between the English vocabulary size scores of newly placed students and continuing students within the same institutional placement level, as measured by the PVLT.

    A one-way ANOVA was performed for every ELC level. The predictor was the number of months studied at the ELC and the response was the raw scores of the vocabulary test. A Tukey-adjusted pairwise $t$ test was also performed to compare the vocabulary size of the newly placed and the continuing students.
CHAPTER 4: FINDINGS AND DISCUSSION

The current study investigated the English vocabulary size of 173 native-speakers of Spanish, Korean, Japanese, and Mandarin studying at an intensive English language program. The role of vocabulary size in assessing learners from various L1 backgrounds was investigated by looking at the role of L1, institutional placement level, and cognates in students’ performance on the test. In addition, the role of vocabulary size in assessing for institutional placement was explored by measuring the relationship between vocabulary size scores and other language skills. Finally, vocabulary size scores were applied in a novel way to measure a perceived difference between newly placed and continuing students.

As somewhat of a side note, one of the items on the PVLT was found to have two equally plausible answers (see Appendix B, item 9). The item reads:

There are several misprints on each page of this te__________.

The correct answer, according to the key, is text, although the word test works just as well. It was decided to follow the key and count only the answer text correct. However, it is recommended that the context for this item be disambiguated or that the item be replaced altogether to increase the reliability of the PVLT.

It was first necessary to obtain reliability scores for the instrument used in this study in order to know that the data obtained for the research questions were credible. Cronbach’s alpha was used to measure the internal reliability of the scores obtained on the PVLT. Cronbach’s alpha reports a coefficient that is a function of test items and the average intercorrelation among the items. The reliability of the PVLT, containing 90
questions, was .94. This indicates that the scores obtained from the PVLT were highly reliable, justifying investigation into the research hypotheses.

**Hypothesis 1**

Hypothesis 1 stated that there is no difference in vocabulary size scores by native language and by institutional placement level. Only the 1-2,000, 2,000-3,000, 3,000-5,000, and 5,000-10,000 sections of the PVLT were used in this analysis. The UWL section was not used because the frequency of some of the words in the UWL section overlap with the frequency of the words in the 5,000-10,000 section, even though none of the words tested on the PVLT are the same. Table 1 shows the mean scores on the vocabulary test for the four languages at each ELC placement level. Spanish and Korean speakers have nearly the same vocabulary size scores overall.

### Table 1

**Mean Scores* on the Productive Vocabulary Levels Test by L1 and Level.**

<table>
<thead>
<tr>
<th>Level</th>
<th>Spanish</th>
<th>Korean</th>
<th>Japanese</th>
<th>Mandarin</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.7</td>
<td>4.7</td>
<td>3.0</td>
<td>1.0</td>
</tr>
<tr>
<td>(10)</td>
<td>(3)</td>
<td>(4)</td>
<td>(2)</td>
<td>(1)</td>
</tr>
<tr>
<td>2</td>
<td>11.8</td>
<td>11.0</td>
<td>8.5</td>
<td>1.0</td>
</tr>
<tr>
<td>(10)</td>
<td>(5)</td>
<td>(2)</td>
<td>(2)</td>
<td>(1)</td>
</tr>
<tr>
<td>3</td>
<td>18.4</td>
<td>17.0</td>
<td>13</td>
<td>10.2</td>
</tr>
<tr>
<td>(53)</td>
<td>(19)</td>
<td>(22)</td>
<td>(7)</td>
<td>(5)</td>
</tr>
<tr>
<td>4</td>
<td>30.1</td>
<td>28.6</td>
<td>22.4</td>
<td>11.0</td>
</tr>
<tr>
<td>(53)</td>
<td>(28)</td>
<td>(14)</td>
<td>(8)</td>
<td>(3)</td>
</tr>
<tr>
<td>5</td>
<td>39.9</td>
<td>38.2</td>
<td>27.7</td>
<td>31.1</td>
</tr>
<tr>
<td>(47)</td>
<td>(20)</td>
<td>(13)</td>
<td>(7)</td>
<td>(7)</td>
</tr>
<tr>
<td>Overall</td>
<td>20.97</td>
<td>19.89</td>
<td>14.92</td>
<td>10.87</td>
</tr>
<tr>
<td>(173)</td>
<td>(75)</td>
<td>(55)</td>
<td>(26)</td>
<td>(17)</td>
</tr>
</tbody>
</table>

*Note. n indicated in parentheses.

*Out of 90 items.
Figure 1 shows a line graph of the numbers from Table 1. The mean scores of Spanish and Korean speakers are consistently close across institutional placement levels.

Figure 1: Mean Scores on the Productive Vocabulary Levels Test by L1 and Level.

A two-way ANOVA was performed to find the interaction between each L1 and institutional placement level. Table 2 reveals that the vocabulary test clearly discriminated between the students’ L1s and institutional placement levels. The total scores were significantly different from each other at the $p = .0001$ level, indicating that the scores derived from the PVLT are valid, since they made a clear distinction between placement levels. The hypothesis for question 1 must therefore be rejected. An alternative hypothesis is accepted that L1 and institutional placement level have significant but independent effects on vocabulary size, but that the language/level interaction is not significant.
Table 2

*Analysis of Variance for L1 and Level.*

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>F*</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>3</td>
<td>8.10</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Level</td>
<td>4</td>
<td>46.39</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>L1 x level</td>
<td>12</td>
<td>0.46</td>
<td>0.9344</td>
</tr>
</tbody>
</table>

*Adjusted F test.*  
\[ p = .05. \]

A post-hoc Tukey-adjusted t test was performed on the overall scores to find which languages were significantly different from the others at the \( p = .05 \) level. There was a significant difference between each language, except for Spanish and Korean, which had a \( p \) value of 0.99.

The closeness in mean vocabulary scores between speakers of Spanish and Korean is surprising given that Spanish shares many more cognates with English than does Korean. Therefore, lexical overlap cannot fully account for test performance, and other reasons must be explored. A post-hoc analysis revealed that, prior to the vocabulary test, the Koreans had been studying English for an average of 5 years and 4 months, while the Spanish speakers had been studying English for an average of 3 years and 4 months, two years less than the Koreans.

In addition to finding out which languages were significantly different from the others, another post-hoc Tukey was performed to find where the differences were between institutional placement levels (Table 3). The Tukey revealed that there is an increase in statistical significance between institutional placement levels as level increases, as shown in the following \( p \) values: between levels 1 and 2 (\( p = .08 \)), between levels 2 and 3 (\( p = .03 \)), between levels 3 and 4 (\( p = .0007 \)), and between levels 4 and 5 (\( p = .0001 \)). This increase in statistical significance is probably due in part to an increase in the number of students as level increases.
Table 3

*Institutional Placement Levels Differentiated by Vocabulary Size Scores.*

<table>
<thead>
<tr>
<th>ELC Level</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (10)</td>
<td>—</td>
<td>.0818</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>2 (10)</td>
<td>—</td>
<td>.0269</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>3 (53)</td>
<td>—</td>
<td>.0007</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 (53)</td>
<td>—</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 (47)</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. n indicated in parentheses.*

*p = .05*

The average vocabulary size at each level is shown in Table 4 in terms of actual word families. These numbers were derived by converting the raw scores into an estimated number of word families by multiplying the raw score by the appropriate number to reach the total number of word families in that level. For example, if a student knows 9 out of the 18 words tested at the 1-2,000 word level, his approximate productive vocabulary size would be 1,000 at that level. Thus, the average score at the 1-2,000 word level for level 1 was 2.2, which approximates 244 word families. As the totals column shows, mean productive vocabulary size increases steadily from 394 at level 1 to 3,589 at level 5, which is an average increase of 799 word families per level.
### Table 4

*Estimated Productive Vocabulary Size at Varying Word Frequencies by Institutional Placement Level as Measured by the Productive Vocabulary Levels Test.*

<table>
<thead>
<tr>
<th>Placement level</th>
<th>n</th>
<th>1-2,000</th>
<th>2,000-3,000</th>
<th>3,000-5,000</th>
<th>5,000-10,000</th>
<th>Totals* out of 10,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>244</td>
<td>6</td>
<td>33</td>
<td>111</td>
<td>394</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>491</td>
<td>42</td>
<td>92</td>
<td>90</td>
<td>715</td>
</tr>
<tr>
<td>3</td>
<td>53</td>
<td>812</td>
<td>146</td>
<td>246</td>
<td>417</td>
<td>1,621</td>
</tr>
<tr>
<td>4</td>
<td>53</td>
<td>1,103</td>
<td>290</td>
<td>412</td>
<td>263</td>
<td>2,068</td>
</tr>
<tr>
<td>5</td>
<td>47</td>
<td>1,356</td>
<td>425</td>
<td>580</td>
<td>1,228</td>
<td>3,589</td>
</tr>
</tbody>
</table>

*Mean scores are shown in Table 1. N = 173.

\[p = <.05.\]

**Hypothesis 2**

Hypothesis 2 states that there is no difference in vocabulary size scores between cognate and noncognate items at each institutional placement level, as measured by the PVLT. Table 4 shows the number and percentage of PVLT items eliciting cognates. Appendixes D and E show specifically which words on the test are cognates. Out of the 90 words elicited on the PVLT, 48 (or 53%) are cognate with words in Spanish, 9 (or 10%) with Korean, and 12 (or 13%) with Japanese. Words from the UWL, many of which have Latin origins, contain the highest percentage of cognates across languages.
Table 5

*Number and Percentage of Items on the Productive Vocabulary Levels Test Eliciting Cognates at Each Word Frequency Level.*

<table>
<thead>
<tr>
<th>Level</th>
<th>Spanish</th>
<th></th>
<th>Korean</th>
<th></th>
<th>Japanese</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>2,000</td>
<td>7</td>
<td>39</td>
<td>2</td>
<td>11</td>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td>3,000</td>
<td>9</td>
<td>50</td>
<td>3</td>
<td>17</td>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td>5,000</td>
<td>7</td>
<td>39</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>UWL</td>
<td>15</td>
<td>83</td>
<td>3</td>
<td>17</td>
<td>5</td>
<td>28</td>
</tr>
<tr>
<td>10,000</td>
<td>10</td>
<td>56</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Total</td>
<td>48</td>
<td>53</td>
<td>9</td>
<td>10</td>
<td>12</td>
<td>13</td>
</tr>
</tbody>
</table>

*Note.* Each level of the PVLT contains 18 questions, and the entire test contains 90 questions.

Table 6 reveals that students scored 103% better, or twice as high, on questions eliciting cognates as they did on noncognate items of the PVLT. The difference between cognates and noncognates was fairly consistent across languages.

Table 6

*Percentage Correct of Questions Eliciting Cognates and Noncognates for All Levels.*

<table>
<thead>
<tr>
<th>L1</th>
<th>n</th>
<th>% Cognates</th>
<th>% Noncognate</th>
<th>% Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish</td>
<td>75</td>
<td>41</td>
<td>19</td>
<td>108</td>
</tr>
<tr>
<td>Korean</td>
<td>55</td>
<td>44</td>
<td>24</td>
<td>92</td>
</tr>
<tr>
<td>Japanese</td>
<td>26</td>
<td>40</td>
<td>18</td>
<td>111</td>
</tr>
<tr>
<td>Overall</td>
<td>156</td>
<td>43</td>
<td>21</td>
<td>103</td>
</tr>
</tbody>
</table>

*Note.* Appendixes H and I show the English words elicited on the Productive Vocabulary Levels Test that have cognates in the given languages.
For this analysis, the UWL section of the PVLT was included. The 2-way ANOVA reveals, as shown in Table 5, that the number of cognates acquired from one semester to the next exceeds the number of noncognates, and that the difference continues to grow as proficiency increases. When the numbers in the differences column are calculated as numbers of words and compared between levels, the cognates exceed noncognates by about 301 word families between levels 1 and 2, about 134 between levels 2 and 3, about 200 between levels 3 and 4, and about 180 between levels 4 and 5. A Tukey-adjusted pairwise *t* test was performed to compare the students’ performance on items eliciting cognates with those eliciting noncognates.

Table 7

*Mean Percentage Scores on Cognates and Noncognates by Level.*

<table>
<thead>
<tr>
<th>Level</th>
<th>N</th>
<th>Cognate</th>
<th>Noncogate</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>.09</td>
<td>.02</td>
<td>.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td>(.6287)</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>.22</td>
<td>.08</td>
<td>.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>**</td>
<td></td>
<td>(.0630)</td>
</tr>
<tr>
<td>3</td>
<td>53</td>
<td>.32</td>
<td>.13</td>
<td>.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>**</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>53</td>
<td>.47</td>
<td>.23</td>
<td>.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>**</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>47</td>
<td>.59</td>
<td>.32</td>
<td>.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>**</td>
<td>**</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Only speakers of Spanish (*n* = 75), Korean (*n* = 55), and Japanese (*n* = 26) were considered for this analysis. *N* = 156.

*p* = <.05. **p** = <.0001.

Caution must be used when interpreting these data, however. These numbers pertain to speakers from all three language backgrounds, Spanish, Korean, and Japanese,
As a group, with Spanish speakers accounting for nearly half the group. Therefore, it may be true that a Spanish speaker’s English vocabulary at level 1 consists mostly of words that are cognate with Spanish, but it is unlikely that more than three-fourths of a Japanese speaker’s English vocabulary at level 1 are cognates with Japanese. The analysis was not performed looking at each language group individually because the sample sizes were too small.

Figure 2 is a line graph representing the numbers in Table 7. The graph provides a visual representation that as proficiency increases, the acquisition of cognates continues to outpace noncognates.

*Figure 2. Cross-sectional growth of cognates and noncognates by level.*

As a group, students correctly answered a higher percentage of items with cognates as proficiency increases. A possible explanation for these results is that the higher proficiency students answered more questions eliciting cognates simply because
they understood the other words in the sentence.

It is interesting to note that even at level 5, students correctly answered only 59% of the questions eliciting cognates. Why would the students not guess at the word when the first few letters indicate that it could be a cognate? The answer may lie in the type of incorrect answers that were given. The incorrect answers indicated that many of the students may not have been reading the sentences, but instead were simply looking at the letters immediately before the blank and writing in the first word that came to mind. For example, in the following sentence,

The insect causes damage to plants by its toxic sec____________.

many of the Spanish speakers completed the word to spell secret instead of secretions, even though the word secretions is secreciones in Spanish. Unfortunately, such an approach to taking the test could undermine the validity of the results, since students may indeed know the correct word, or be able to guess it from the context, if they were to read the entire sentence.

Another possible explanation why many students completed the blank with the incorrect word, even if the correct answer was a cognate, was that they did not understand the context and simply guessed as best they could. Thus, the fact that words in the surrounding text are always more frequent than the target word may be of less help than one would hope, especially for low proficiency students.

Hypothesis 3

Hypothesis 3 stated that there is no relationship between vocabulary size scores, as measured by the PVLT, and scores on the institutional placement exams in reading, writing, speaking, listening, and grammar. The UWL section of the PVLT was not
included in this analysis. Table 8 shows the Pearson correlation scores between vocabulary size and reading, writing, speaking, listening, and grammar scores. Vocabulary size correlated most highly with speaking and listening (.66), followed by grammar (.62) and finally reading and writing (.60). The variance overlap ($r^2$), which is obtained by squaring the correlation number, indicates the strength of the correlation, or the degree to which vocabulary size scores and language proficiency scores overlap.

Table 8

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Speaking</th>
<th>Listening</th>
<th>Grammar</th>
<th>Reading</th>
<th>Writing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary</td>
<td>.66</td>
<td>.66</td>
<td>.62</td>
<td>.60</td>
<td>.60</td>
</tr>
<tr>
<td>$r^2$</td>
<td>.44</td>
<td>.44</td>
<td>.38</td>
<td>.36</td>
<td>.36</td>
</tr>
</tbody>
</table>

*Note. The top number is the Pearson $r$, and the bottom is the variance overlap ($r^2$).*

All $p$ values were significant at the <.0001 level. $N = 173$.

Results of the correlation analysis between vocabulary size and reading, writing, speaking, listening, and grammar were a bit different than expected. The literature indicated that scores on the vocabulary size test would correlate most highly with reading, followed by writing and speaking scores. This sounds reasonable since the PVLT is a written test, requiring the students to read, but also has blanks, requiring students to write. Contrary to this logic, vocabulary size scores actually correlated highest with speaking and listening. The differences between previously found correlations and those found in this study are probably best explained by the instruments used.

**Hypothesis 4**

Hypothesis 4 stated that there is no difference between the English vocabulary size scores of newly placed and continuing students within the same institutional
placement level, as measured by the PVLT. Only levels 3-5 were considered for this analysis, since sample sizes were small in levels 1 and 2. In addition, the UWL section was not used in this analysis. As shown in Table 9, there is a substantial difference in productive vocabulary size between newly placed students and continuing students. In this cross-sectional analysis, newly placed students generally have larger vocabularies than continuing students. In terms of actual numbers of words, the average newly placed student knows approximately 377 word families more than the student placed 4 months earlier, and 950 word families more than the student placed 8 months earlier.

Table 9

<table>
<thead>
<tr>
<th>Source</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELC mo.</td>
<td>0 4 8</td>
<td>0 4 8</td>
<td>0 4 8</td>
</tr>
<tr>
<td>Means</td>
<td>15.26 11.75 11.00</td>
<td>23.65 22.29 12.50</td>
<td>36.80 30.35 23.71</td>
</tr>
<tr>
<td>n</td>
<td>31 20 7</td>
<td>26 24 8</td>
<td>10 26 7</td>
</tr>
</tbody>
</table>

_Note._ ELC mo. indicates how many months participants had been studying at the ELC at the time of the vocabulary tests. Scores may be interpreted as numbers of words by moving the decimal two places to the right.

_N = 159._

One may use these numbers to estimate the average vocabulary growth per semester. Although this is a cross-sectional rather than a longitudinal analysis, one may subtract the vocabulary size from, for example, the level 4 students at 0 months (the newly placed students), from the level 5 students at 4 months (students returning after one semester), and so on across the table, and find that the students are learning an average of approximately 398 word families per semester.
A one-way ANOVA was performed to see if there is a statistical difference between the vocabulary size of 0-month, 4-month, and 8-month students. The results in Table 10 show that the three groups of students are statistically different from each other at the .05 level. The null-hypothesis must therefore be rejected, and an alternate hypothesis is accepted that there is a statistically significant difference between the vocabulary size scores of newly placed and continuing students within the same institutional placement level, as measured by the PVLT.

Table 10

*Analysis of Variance for Months Enrolled.*

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>$f^*$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Months at ELC</td>
<td>2</td>
<td>3.65</td>
<td>0.0281</td>
</tr>
</tbody>
</table>

*Adjusted F test.

$p = <.05$.

A post-hoc Tukey-adjusted $t$ test was performed on the overall scores to find which groups of students were significantly different from the others at the $p = .05$ level. Results showed that the vocabulary size of 0- and 4-month students is significantly higher than the vocabulary size of the 8-month students at the $p = .05$ level. However, there was no statistical difference between the 0- and 4-month students.

Table 11

*Test of Differences Between Newly Placed and Continuing Students.*

<table>
<thead>
<tr>
<th>Months at ELC</th>
<th>0</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>—</td>
<td>0.7616</td>
<td>0.0174*</td>
</tr>
<tr>
<td>4</td>
<td>—</td>
<td></td>
<td>0.0094*</td>
</tr>
<tr>
<td>8</td>
<td>—</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant at the $p = .05$ level.
CHAPTER 5: CONCLUSION

In the current study, the role of vocabulary size in assessing second language proficiency and placement has been explored. Participants included 173 native speakers of Spanish, Korean, Japanese, and Mandarin, who took a test that estimated the size of their vocabulary in English. Their vocabulary size scores were analyzed by measuring the students’ vocabulary size at varying institutional placement levels. The role of cognates in the students’ performance on the vocabulary test was also explored. Further, students’ vocabulary size scores were correlated with their placement scores that evaluated reading, writing, speaking, listening, and grammar. Finally, vocabulary assessment was applied in a novel way to measure the perceived English proficiency gap between newly placed and continuing students. Findings of this research, as well as how they agree with those of previous research, will be discussed below. This is followed by the implications of the findings. Limitations of the study will be considered, and finally, suggestions for future research will be proposed.

The reliability of the PVLT was .94. This is quite a bit higher than Laufer and Nation’s (1999) reliability figure of .86. There are two reasons for this difference. The main reason the reliability score in this study may be higher is because the sample size was larger, at 173 as opposed to 80. Second, the current study used Crombach’s alpha, whereas Laufer and Nation used KR21. Both are measures of internal consistency, but this may also be a cause for the difference. Nevertheless, scores from both the true-false test and the PVLT were found to be reliable, justifying further analysis of the data derived from them.

For hypothesis 1, it was found that, out of 90 items, speakers of Spanish had the
highest mean vocabulary score (20.97), followed closely by Korean (19.89), then Japanese (14.92), and Mandarin (10.87). There was no statistically significant language/level interaction, nor was there any statistical difference, between the vocabulary size of the Spanish and Korean speakers. The closeness in mean vocabulary scores between speakers of Spanish and Korean is surprising given that Spanish shares many more cognates with English than does Korean. As was shown in Question 2, out the 90 words elicited on the PVLT, 48, or 53%, were cognate with Spanish, whereas only 9 (or 10%) were cognate with Korean. Therefore, The equivalence in scores is not accompanied by equivalent lexical overlap, and other reasons must be explored.

The post hoc investigation revealed that the Korean speakers had been studying an average of two years longer than the Spanish speakers. In addition, it could be that there is a greater focus on mastering vocabulary in English classes in Korea than in Spanish-speaking countries, and that the Korean students continue to apply word-learning strategies that they learned at school. Further, the educational culture in Korea may value memorization more than in Latin countries. There may also be a substantial difference in the amount of time that the Korean-speaking students spend studying English vocabulary compared to the Spanish-speaking students. Unfortunately, answers to these questions were not explored since these results were not foreseen.

It was also discovered that the average estimated productive vocabulary size was 394 word families at level 1, increasing to 3,589 word families at level 5, with an average difference of 799 word families per level. These numbers do not mean that the average student is learning 799 word families per semester. Based on the results from question 4, which compares the vocabulary size of newly placed and continuing students, it can be
estimated that the average ELC student learns roughly 398 word families per semester, which is about three and a half months long. This would translate to about 1,592 word families per year, which is double Laufer’s (1998) findings, in which there was a 850 word difference between 10th and 11th graders in Israel. However, given that ELC students are exposed to English outside of the classroom, it is expected that they would learn more words in a year than would students in an EFL context, as was the case in Laufer’s study.

It is also important to remember that students’ receptive vocabulary is most likely 19 to 25 percent higher than their productive vocabulary (Schmitt, 1997), which would bring the receptive vocabulary size of level 5 students to somewhere between 4,271 and 4,486. These figures fall within the range of the 4,000-5,000 word families that Sutarsyah, Nation, and Kennedy (1994) estimate is necessary for comprehending an academic text. Indeed, there are students from the ELC who pass the TOEFL each semester and go on to study at an American university. These students, however, have a major lexical handicap compared to their native speaker classmates, who know about 18,000-20,000, or nearly five times as many, word families (Nation, 1990). Certainly these students must compensate for their smaller vocabularies through an admirable work ethic and a high level of determination.

It was also found that students are not mastering the first 2,000 most frequent word families, even by level 5. At level 5, students knew of an average of only 1,356 out of the 2,000 most frequent word families. The remainder of their English vocabulary consisted of 2,233 word families. This means that more than half of the students’ productive vocabulary size consists of words less frequent than the first 2,000 word
families, which sheds some doubt on the claim that higher frequency words are generally acquired first (Anderson and Freebody, 1981; Laufer, 2001; Nation, 1990).

For hypothesis 2, it was found that Spanish, Korean, and Japanese students are learning cognates at approximately the same rate, as well as noncognates at approximately the same rate. This is a strong indication that non-native speakers of English are equally likely to learn cognates, regardless of their L1.

It was also found that students at higher institutional placement levels correctly answered a higher percentage of questions eliciting cognates than students at lower institutional placement levels. This supports similar findings by García (1988, 1991), Hancin-Bhatt (1994), and Nagy (1993). Admittedly, it is a leap in logic to assume that higher proficiency students correctly answered more cognate questions than lower proficiency students because of an increased ability to recognize cognates. The most likely reason the higher proficiency students correctly answered more items eliciting cognates is simply because they understood the rest of the sentence, whereas the lower proficiency students did not.

For hypothesis 3, it was revealed that the relationship between vocabulary size and listening, reading, and writing scores are different from correlations found in previous studies. The correlation between vocabulary size and listening comprehension in this study was .66, which is a much stronger correlation than that found by Beglar (1999) of .44 and .45. The correlation between vocabulary size and reading (.60) was a bit lower than other findings. Laufer (1992) had found correlations of .50 and .75, Qian (1999) of .78, and Beglar (1999) of .66, .62, .67, and .71. The writing correlation (.60) was also a bit low compared to previous findings. Beglar (1999) had found correlations
of .61 and .65, Astica (1993) of .84, Laufer and Nation (1995) of .60 and .65, Laufer (1998) of .62 and .78, and Linnarud (1986) who said that vocabulary size was the single largest factor in writing quality, but did not provide actual correlation numbers.

These higher correlations with speaking and listening might be explained by the primacy of these skills, compared to the cognitively more complex and learned tasks of reading, writing, and grammar. There are many skills involved in reading, writing, and grammar that the PVLT is not designed to evaluate. Reading requires not only recognition of written words, but also such skills as skimming, scanning, interpreting graphs and glosses, inferring, and a well developed connection between orthography and phonemics. Good writing requires skills such as prewriting, producing a clear thesis statement and topic sentences, using transitions, and control of punctuation. Finally, grammatical concepts are even more abstract than reading or writing. Thus, although these correlations were not expected, they are understandable.

Finally, in hypothesis 4, it was confirmed that newly placed students have larger vocabularies than continuing students. On average, newly placed students know about 370 word families more than the student placed 4 months ago, and 950 more word families than the student placed 8 months ago. These findings are similar to those of Brown (1980), who evaluated the test scores of 319 ESL learners at UCLA and found that the placed students scored 6.71 points higher (out of 50) on a cloze test and 9.82 (out of 100) points higher on the final examination than continuing students ($p < .01$). Given that vocabulary provides the enabling knowledge to perform other language skills, this finding may go far to explain the English proficiency gap between newly placed and continuing students that many teachers have intuitively observed.
Implications

Several implications for second language learners and teachers arise from the findings of this study. First, it has been suggested in the literature that vocabulary size tests may be used as part of a battery of placement tests (Laufer, 2002; Meara, 1992; Schmitt, 1994). Correlations obtained from this study warrant this. However, it is not recommended that a vocabulary test replace the other proficiency tests that are common at English language programs. The correlation scores in this and other studies mentioned in the review of literature are simply not high enough. Correlation scores would need to begin in the high .80s or low .90s for administrators to be sufficiently confident that a vocabulary test could properly place students.

Another implication is centered in the finding that there is an English proficiency gap between newly placed students and continuing students. Given that different students learn at different rates, a one-time placement exam upon entering a language program results in classes filled with students with a wide range of proficiencies. It is widely believed that it benefits the teacher and the students if the students within a class are at relatively equal language proficiencies. It would be appropriate, therefore, to require students to pass an end-of-semester exam to correctly place continuing students. The disadvantage of doing this is that students may find it unfair that they all pay the same tuition, yet some may advance in the program while others are held back. A possible solution to this dilemma is to divide the levels into subgroups, for example, 1a, 1b, 2a, 2b, and so forth, so that smaller levels of achievement can be recognized. Students whose end-of-semester exam scores do not merit a full level advancement can
feel like they are progressing (and in fact they are, although perhaps more slowly than others), while reviewing and mastering the same material with a different textbook.

Much research indicates that extensive reading is a primary source of acquiring new vocabulary (breadth) and deepening understanding of existing vocabulary (depth). This is the general opinion of administrators at the ELC as well. ELC students are required to read approximately twenty to thirty pages per day, and most of the reading is narrative in nature. Based on the findings here, continuing students are not acquiring new vocabulary through this extensive reading fast enough to match the vocabulary size of their newly placed classmates. It is recommended, therefore, that other methods for teaching vocabulary within the classroom be explored at the ELC. In addition, more expository texts should be required in the reading classes since there is a difference between expository and narrative texts with regard to types and use of vocabulary. Reading more expository texts would also seem more appropriate preparation for many of the students, whose primary objective is to pass the TOEFL and to enter an American university.

Limitations

Four limitations of the current study were identified. First, the correlations found between vocabulary size and the reading, writing, speaking, listening, and grammar scores from the ELC placement test cannot be duplicated outside of the ELC. This is because the ELC has its own placement tests that are not used elsewhere, as the TOEFL or other standardized tests are.

A second limitation is that the present study was cross-sectional rather than longitudinal. Parallel versions of the PVLT are available and could have been
administered at the end of the 2003 winter semester in order to measure vocabulary growth at the end of the semester. These parallel versions have been reported to be highly reliable (Laufer, 1999). However, ELC teachers were reluctant to administer the vocabulary tests at the end of the semester because they claimed it took too much class time. Indeed, students took longer to take the tests than the pilot test had indicated. The teachers reported that the two sittings took approximately thirty minutes each, or about half the class time on two consecutive days.

The third limitation to this study is that the findings cannot be generalized to students in EFL settings, where the classroom is their only exposure to English, since this study was carried out in the United States where students have exposure to English outside the classroom.

Finally, this study has measured students’ vocabulary size, and therefore no claims may be made regarding their depth of vocabulary knowledge. Indeed, it would be enlightening to assess the state of the students’ vocabulary using a battery of vocabulary tests that would measure depth in addition to breadth. Sitting such a battery of tests, however, is time-consuming and was not feasible in the present study.

Directions for future research

Four ideas for future research evolved out of this study. First, it should be determined whether the elicited words on the PVLT are representative with regard to the ratio of cognates to noncognates existing between English and other languages. In other words, this study found that 53% of the English words elicited on the PVLT are cognate with Spanish, but is there truly a 53% lexical overlap between Spanish and English? What about other languages? If the test does not accurately represent a proper ratio of
cognates and noncognates, this should be understood so that scores of the PVLT could be
adjusted to be more comparable across languages.

Second, it was mentioned earlier that reading comprehension, and not the ability
to recognize cognates, may explain why higher proficiency students scored higher on
questions eliciting cognates. Further research could seek to understand whether reading
comprehension or cognate recognition has a greater influence on students’ ability to
correctly answer questions eliciting cognates.

Third, there does not appear to be any available vocabulary depth test that tests
knowledge of multiword expressions such as phrasal verbs, idioms, compound words, or
lexical phrases. Such multiword expressions are common in English, and often carry
unique meanings that the individual words themselves do not. Such a test could be
constructed based on a frequency analysis of a corpus. The test would also need to be
piloted and evaluated for validity and reliability.

Finally, it would be useful for ESL and EFL students and their teachers to know
what vocabulary range is needed to pass the TOEFL. The minimum required TOEFL
scores vary depending on the university, but fall somewhere around a score of 500 for the
paper test, with subscores of at least 50 on each section, or 173 on the computer test, with
subscores of at least 17 on each section. Such information could greatly increase
students’ motivation to proactively increase their vocabulary size.
REFERENCES


Beglar, D. & Hunt, A. (1999). Revising and validating the 2,000 word level and university word level vocabulary tests. Language Testing, 16, 131-162.


Appendix A

Integrated levels version of the Productive Vocabulary Levels Test, Part A

**Instructions:** There are 30 questions. Fill in the blank to complete the word.
Correct spelling is *not* required. The example has been done for you.

**Example:** He was riding a bicycle.

1. I'm glad we had this opp__________ to talk.
2. He has a successful car__________ as a lawyer.
3. Soldiers usually swear an oa__________ of loyalty to their country.
4. There has been a recent tr__________ among prosperous families toward a smaller number of children.
5. The baby is wet. Her dia__________ needs changing.
6. There are a doz__________ eggs in the basket.
7. The thieves threw ac__________ in his face and made him blind.
8. The voter placed the ball__________ in the box.
9. The ar__________ of his office is 25 square meters.
10. The prisoner was released on par__________.
11. Every working person must pay income t__________.
12. To improve the country's economy, the government decided on economic ref__________.
13. They keep their valuables in a vau__________ at the bank.
14. Phil__________ examines the meaning of life.
15. Second year university students in the US are called soph__________.
16. The pirates buried the trea__________ on a desert island.
17. She wore a beautiful green go__________ to the ball.
18. A bird perched at the window led__________.
19. According to the communist doc__________, workers should rule the world.
20. Her favorite flowers were or__________.
21. Her beauty and ch__________ had a powerful effect on men.
22. The government tried to protect the country's industry by reducing the I
importance of cheap goods.
23. The kitten is playing with a ball of yarn.
24. Spending many years together deepened their intimacy.
25. The insect causes damage to plants by its toxic secretion.
26. Lack of rain led to a shortage of water in the city.
27. The children's games were amusing at first, but finally got on the parents' nerves.
28. The thieves have forced an entrance into the building.
29. He usually read the sports section of the newspaper first.
30. The evacuation of the building saved many lives.
Appendix B

Integrated levels version of the Productive Vocabulary Levels Test, Part B

Instructions: There are 60 questions. Fill in the blank to complete the word. Correct spelling is not required. The example has been done for you.

Example: He was riding a bicycle.

1. He takes cr_______ and sugar in his coffee.
2. The lawyer gave some wise coun_______ to his client.
3. The small hill was really a burial mou_______.
4. Because of the doctors' strike, the cli_______ is closed today.
5. For many people, wealth is a prospect of unimaginable felic_______.
6. The rich man died and left all his we_______ to his son.
7. Many people in England mow the la_______ of their houses on Sunday morning.
8. We decided to celebrate New Year's E_______ together.
9. There are several misprints on each page of this te_______.
10. She found herself in a pred_______ without any hope for a solution.
11. Pup_______ must hand in their papers by the end of the week.
12. The farmer sells the eggs that his he_______ lays.
13. The soldier was asked to choose between infantry and cav_______.
14. The suspect had both opportunity and mot_______ to commit the murder.
15. The deac_______ helped with the care of the poor of the parish.
16. This sweater is too tight. It needs to be stret_______.
17. Sudden noises at night sca_______ me a lot.
18. This is a complex problem that is difficult to comp_______.
19. They insp_______ all products before sending them out to stores.
20. The hurricane whi_______ along the coast.
21. Ann intro_______ her boyfriend to her mother.
22. France was proc_______ a republic in the 18th century.
23. The angry crowd sho_______ the prisoner as he was leaving the court.
24. A considerable amount of evidence was accum_______ during the investigation.
25. Some coal was still smol__________ among the ashes.
26. Teenagers often adm__________ and worship pop singers.
27. Many people are inj__________ in road accidents every year.
28. Don't pay attention to this rude remark. Just ig__________ it.
29. The victim's shirt was satu__________ with blood.
30. The dead bodies were mutil__________ beyond recognition.
31. If you blow up that balloon any more it will bu__________.
32. Suddenly he was thru__________ into the dark room.
33. The management held a secret meeting. The issues discussed were not
disc__________ to the workers.
34. He is irresponsible. You cannot re__________ on him for help.
35. She was sitting on a balcony and bas__________ in the sun.
36. In order to be accepted into the university, he had to impr__________ his grades.
37. He perc__________ a light at the end of the tunnel.
38. We could hear the sergeant bel__________ commands to the troops.
39. It's impossible to eva__________ these results without knowing about the research
   methods that were used.
40. For years waves of invaders pill__________ towns along the coast.
41. The telegram was deli__________ two hours after it had been sent.
42. Children are not independent. They are att__________ to their parents.
43. The boss got angry with the secretary and it took a lot of tact to soo__________ him.
44. He finally att__________ a position of power in the company.
45. The rescue attempt could not proceed quickly. It was imp__________ by bad
   weather.
46. The differences were so sl__________ that they went unnoticed.
47. She showed off her sle__________ figure in a long narrow dress.
48. We do not have adeq__________ information to make a decision.
49. The story tells about a crime and subs__________ punishment.
50. I wouldn't hire him. He is unmotivated and indo__________.
51. The dress you're wearing is lov__________.
52. She has been changing partners often because she cannot have a sta__________
    relationship with one person.
53. She is not a child, but a mat__________ woman. She can make her own decisions.
54. In a hom__________ class all students are of a similar proficiency.
55. Computers have made typewriters old-fashioned and obs__________.
56. He wasn't very popu__________ when he was a teenager, but he has many friends
    now.
57. You must wear a bathing suit on a public beach. You're not allowed to bath
    na__________.
58. The prisoner was put in soli__________ confinement.
59. The urge to survive is inh__________ in all creatures.
60. Watch out for his wil__________ tricks.
Appendix C

Student Questionnaire (Spanish)

Esta es una evaluación para ver qué tan grande es su vocabulario. Será usada en una investigación muy importante. No será calificado por esto, pero por favor conteste cada pregunta lo mejor que pueda. Usted tomará la Sección 1 el día de hoy, y la Sección 2 durante la siguiente clase. Su maestro le dará sus resultados. Su participación es voluntaria.

Por favor, conteste todas las siguientes preguntas antes de comenzar la evaluación de vocabulario.

1. Escriba su nombre (en letra de molde)_________________________.

2. ¿Cuál es su sexo?
   a. Masculino
   b. Femenino

3. ¿Cuál es su lengua nativa?_______________________

4. ¿Dónde se crió? Ciudad:_______________, País:_____________________

5. ¿Cuándo comenzó sus estudios en el English Language Center?
   ☐ Septiembre (Otoño)  ☐ 1999
   ☐ Enero (Invierno)    ☐ 2000
   ☐ Abril (Verano)      ☐ 2001
   ☐ 2002
   ☐ 2003

6. ¿En qué nivel fue colocado al comenzar en el English Language Center?
   a. Nivel 1
   b. Nivel 2
   c. Nivel 3
   d. Nivel 4
   e. Nivel 5

7. ¿En qué nivel está ahora?
   a. Nivel 1
   b. Nivel 2
   c. Nivel 3
   d. Nivel 4
   e. Nivel 5

8. ¿Cuál es su edad?___

9. ¿Por cuánto tiempo ha estudiado inglés? ______años, ______meses

10. ¿Cuánto tiempo ha vivido en los Estados Unidos? ______años, ______meses

Muchas gracias! Puede comenzar la evaluación de vocabulario.
Instructions for Section 1, Part B (Spanish)

Part A

Instrucciones: Hay 39 preguntas. Circule la “T” si la oración es verdadera, y la “F” si la oración es falsa. Circule el signo “?” si no entiende la oración.

Ejemplo: We can stop time. T F ?

Instructions for Section 1, Part B (Spanish)

Part B

Instrucciones: Hay 30 preguntas. Llene el espacio en blanco para completar la palabra.

No se requiere ortografía perfecta. El ejemplo ha sido hecho para usted.

Ejemplo: He was riding a bicycle.

Instructions for Section 2 (Spanish)

Escriba su nombre (en letra de molde, por favor) ____________________________

Apellido                        Nombre

Esta es la Sección 2 de la evaluación de vocabulario. Será usada para una importante investigación. No será calificado por esto, pero por favor conteste lo mejor que pueda. Su maestro le dará sus resultados. Su participación es voluntaria.

Puede comenzar ahora la evaluación.

Directions for Section 2 (Spanish)

Instrucciones: Hay 60 preguntas. Llene el espacio en blanco para completar la palabra.

No se requiere ortografía perfecta. El ejemplo ha sido hecho para usted.

Ejemplo: He was riding a bicycle.
Appendix D

Student Questionnaire (Japanese)

ここではあなたの英語の語彙力について調べます。ここでの結果はとても大切なリサーチのため
に使われます。あなたがこのセクションで行った結果に対する成績などはありませんが、どう
ぞ一つ一つの質問に真剣に答えて下さい。今日は始めにセクション1を行います。セクション2は
次のクラスで行います。それぞれのセクションの結果はあなたの先生を通してあなたに渡されます。
このテストへの参加は自由です。

このセクションを始める前に下記の質問すべてに答えて下さい。

1. 名前（ブロック体で書いて下さい）
   ……………………………………………………………………………………………………………………………………………………………………………………………
   姓  名

2. 性別
   a. 男
   b. 女

3. 母国語
   ……………………………………………………………………………………………………………………………………………………………………………………………

4. 生まれ育った場所
   町／市………………………………………国名………………………………………

5. いつこの ESL で勉強をし始めましたか？
   □ September (Fall) 9月（秋学期） 1999
   □ January (Winter) 1 月（冬学期） 2000
   □ April (Summer) 4 月（夏学期） 2001
   □ ……………………………………

6. この ELC で勉強を始めた時のレベルは何でしたか？
   a. レベル 1
   b. レベル 2
   c. レベル 3
   d. レベル 4
   e. レベル 5

7. 現在のあなたのレベルは何ですか？
   a. レベル 1
   b. レベル 2
   c. レベル 3
   d. レベル 4
   e. レベル 5

8. 年齢

9. どのくらい英語を勉強していますか？ 年…………勅…………

10. どのくらいアメリカに住んでいますか？ 年…………勅…………

ご協力ありがとうございました。それではこのセクションを始めて下さい。

Instructions for Section 1, Part A (Japanese)

Part A

解答の仕方：ここでは39問の質問に答えてもらいます。もしあなたの読んだ文が正しいと思う
ならば“T”に丸をつけて下さい。もしその文が間違っていると思うならば“F”に丸をつけて下
さい。もしその文が理解できなければ“_”に丸をつけて下さい。

例：We can stop time. T (F) ?
Instructions for Section 1, Part B (Japanese)

Part B
解答の仕方：ここでは 30 問の質問に答えてもらいます。言葉を完成するためにそれぞれの空欄を埋めて下さい。正しくスペルしてるかどうかは心配していただけないですか。例を参考にして下さい。

例： He was riding a bike ______.

Instructions for Section 2 (Japanese)

名前（ブロック体で書いて下さい）------------------------------------------姓 名

このセクション2では単語についての分析をします。ここでの結果はとても大切なるサーチのために使われます。あなたがこのセクションで行った結果に対しての成績などはありませんが、どうぞ一つ一つの質問に真剣に答えて下さい。このセクションの結果はあなたの先生を通してあなたに渡されます。このテストへの参加は自由です。

それではセクション2を始めて下さい。

Directions for Section 2 (Japanese)

解答の仕方：ここでは 60 問の質問に答えてもらいます。言葉を完成するためにそれぞれの空欄を埋めて下さい。正しくスペルしてるかどうかは心配していただけないですか。例を参考にして下さい。

例： He was riding a bike ______.
Appendix E

Student Questionnaire (Korean)

영어단어를 얼마나 아는지 알아보기 위한 평가입니다. 이 평가의 결과는 중요한 연구의 필요합니다. 이 평가에 참여해서 점수를 받지는 않지만 최선을 다해서 평가에 참여해주십시오. 오늘 1부 평가를 하고 다음 수업시간에서 2부 평가를 보겠습니다. 여러분의 선생님을 통해서 평가의 결과를 알려드리겠습니다. 여러분의 참여는 자율적 선택입니다.

단어 평가를 시작하기 전에 밑의 질문을 모두 대답해주십시오.

1. 당신의 이름은 무엇입니까?(영어로 쓰십시오) 성씨: _________________ 성씨 이름
2. 성이 무엇입니까?
   c. 남
   d. 녀
3. 당신의 본언어는 무엇입니까?(영어로 쓰십시오) _______________________
4. 당신이 태어나 성장한 곳은 어디입니까?(영어로 쓰십시오) 도시________________ 나라____________
5. 처음 ELC에서 영어 공부 시작한것은 언제 부터입니까?
   □ 가을학기 □ 겨울학기 □ 여름학기 □ 봄학기
6. 처음 ELC에서 영어를 배우기 시작한 등급은 무엇입니까?
   a. Level 1
   b. Level 2
   c. Level 3
   d. Level 4
   e. Level 5
7. 지금의 당신의 등급은 무엇입니까?
   a. Level 1
   b. Level 2
   c. Level 3
   d. Level 4
   e. Level 5
8. 몇살입니까? _____
9. 태어나자 뒤로에서 몇 년 몇 개월 동안 영어 공부를 하셨습니까? _____ 년, _____ 개월
10. 미국에서 얼마나 동안 살아습니까? _____ 년, _____ 개월

감사합니다. 지금부터 평가를 시작해주십시오.
Instructions for Section 1, Part A (Korean)

Part A


견본: We can stop time. T F ?

Instructions for Section 1, Part B (Korean)

Part B

지시: 여기에 30개의 문제가 있습니다. 견본대로 올바른 단어로 끝까지도록 빈칸에 채워주십시오.

견본: He was riding a bicycle______

Instructions for Section 2 (Korean)

당신의 이름은 무엇입니까?(영어로 써주십시오) _____________________________

영어단어 2 부 평가 입니다. 이평가의 결과는 중요한 연구의 필요합니다. 이평가에 참여해서 점수를 받지는 않지만 최선을 다해서 평가에 참여해주시기 바랍니다. 여러분의 선생님을 통해서 평가의 결과를 알려드리겠습니다. 여러분의 참여는 자율적 선택입니다.

지금부터 영어단어 2부 평가를 시작해주십시오.

Directions for Section 2 (Korean)

지시: 여기에 60개의 문제가 있습니다. 견본대로 올바른 단어로 끝까지도록 빈칸에 채워주십시오.

견본: He was riding a bicycle______
Appendix F

Student Questionnaire (Mandarin)

這是你字彙能力多寡的評估。這是為了很重要的研究，你不會收到這個成績，但是請你盡最大能力來回答每個問題。今天你會完成第一部分，下一次你會完成第二部分。你的老師會給你結果。您的參與是自願的。在你做字彙評估之前，請你回答以下所有的問題。

1. 你的姓名？(請拼出名字) 姓名
2. 你的性別？
   a. 男
   b. 女
3. 你的母語是什麼語言？
4. 你在哪裡長大？ 城市名： 国家：
5. 你什麼時候開始在英語中心學習？
   □ 九月(秋季) 1999
   □ 月(冬季) 2000
   □ 四月(夏季) 2001
   □ 2002
   □ 2003
6. 你剛開始在英語中心的階層是哪一個？
   a. (第一級)
   b. (第二級)
   c. (第三級)
   d. (第四級)
   e. (第五級)
7. 你現在在哪一個階層？
   a. (第一級)
   b. (第二級)
   c. (第三級)
   d. (第四級)
   e. (第五級)
8. 你今年幾歲？
9. 你學習英文多久了？ 年 月
10. 你在美國住多久了？ 年 月

謝謝你！你可以開始進行字彙評估。

Instructions for Section 1, Part B (Mandarin)

Part A

說明：共有 39 個題目，如果句子是正確的，請註明“T”。如果句子是錯誤的，請圈選“F”。如果您不明白題目，請圈選“?”。

範例： We can stop time. T F ?
Instructions for Section 1, Part B (Mandarin)

Part B

說明：此部分共有 30 個題目。請將空格填滿完成整個單字。不需要正確的拼字。完成的範例已提供。

範例：He was riding a bicycle______.

Instructions for Section 2 (Mandarin)

你的姓名？(請拼出名字) __________________________________________.

姓名

這是字彙評估的第二部分。這是為了很重要的研究，你不會收到這個成績，但是請你盡最大能力來回答每個問題。你的老師會給你結果。您的參與是自願的。

你可以開始字彙評估。

Directions for Section 2 (Mandarin)

說明：此部分共有 60 個題目。請將空格填滿完成整個單字。不需要正確的拼字。完成的範例已提供。

範例：He was riding a bicycle______.
Appendix G

Cognate questionnaire

Name: _____________________________   E-mail: _____________________________

City/country of birth: ____________/____________;  Native language: ______________

Years in US: _____; Years studied at university in the US: _____ / outside the US: _____

Instructions: If the underlined word is at all similar to a word in your native language, write the word on the line to the right of the sentence. (Use the Roman alphabet, as in English). Circle Y if the underlined word can have the same meaning in your native language. Circle N if the underlined word has a different meaning in your language.

Example (from Portuguese):
There is a copy of the original report in the file.  ____cópia____ Y N

1. I'm glad we had this opportunity to talk.  ____________ Y N
2. He has a successful career as a lawyer.  ____________ Y N
3. Soldiers usually swear an oath of loyalty to their country.  ____________ Y N
4. There has been a recent trend among prosperous families toward a smaller number of children.  ____________ Y N
5. The baby is wet. Her diaper needs changing.  ____________ Y N
6. There are a dozen eggs in the basket.  ____________ Y N
7. The thieves threw acid in his face and made him blind.  ____________ Y N
8. The voter placed the ballot in the box.  ____________ Y N
9. The area of his office is 25 square meters.  ____________ Y N
10. The prisoner was released on parole.  ____________ Y N
11. Every working person must pay income tax.  ____________ Y N
12. To improve the country's economy, the government decided on economic reform.  ____________ Y N
13. They keep their valuables in a vault at the bank.  ____________ Y N
14. Philosophy examines the meaning of life.  ____________ Y N
15. Second year university students in the US are called sophomores.  
16. The pirates buried the treasure on a desert island.  
17. She wore a beautiful green gown to the ball.  
18. A bird perched at the window ledge.  
19. According to the communist doctrine, workers should rule the world.  
20. Her favorite flowers were orchids.  
21. Her beauty and charm had a powerful effect on men.  
22. The government tried to protect the country's industry by reducing the importation of cheap goods.  
23. The kitten is playing with a ball of yarn.  
24. Spending many years together deepened their intimacy.  
25. The insect causes damage to plants by its toxic secretions.  
26. Lack of rain led to a shortage of water in the city.  
27. The children's games were amusing at first, but finally got on the parents' nerves.  
28. The thieves have forced an entrance into the building.  
29. He usually read the sports section of the newspaper first.  
30. The evacuation of the building saved many lives.  
1. He takes cream and sugar in his coffee.  
2. The lawyer gave some wise council to his client.  
3. The small hill was really a burial mound.  
4. Because of the doctors' strike, the clinic is closed today.  
5. For many people, wealth is a prospect of unimaginable felicity.  
6. The rich man died and left all his wealth to his son.  
7. Many people in England mow the lawn of their houses on Sunday morning.  
8. We decided to celebrate New Year's Eve together.  
9. There are several misprints on each page of this text.
10. She found herself in a **predicament** without any hope for a solution. 

11. **Pupils** must hand in their papers by the end of the week. 

12. The farmer sells the eggs that his **hen** lays. 

13. The soldier was asked to choose between infantry and **cavalry**. 

14. The suspect had both opportunity and **motive** to commit the murder. 

15. The **deacon** helped with the care of the poor of the parish. 

16. This sweater is too tight. It needs to be **stretched**. 

17. Sudden noises at night **scare** me a lot. 

18. This is a complex problem that is difficult to **comprehend**. 

19. They **inspect** all products before sending them out to stores. 

20. The hurricane **whipped** along the coast. 

21. Ann **introduced** her boyfriend to her mother. 

22. France was **proclaimed** a republic in the 18th century. 

23. The angry crowd **shoved** the prisoner as he was leaving the court. 

24. A considerable amount of evidence was **accumulated** during the investigation. 

25. Some coal was still **smoldering** among the ashes. 

26. Teenagers often **admit** and worship pop singers. 

27. Many people are **injured** in road accidents every year. 

28. Don't pay attention to this rude remark. Just **ignore** it. 

29. The victim's shirt was **saturated** with blood. 

30. The dead bodies were **mutilated** beyond recognition. 

31. If you blow up that balloon any more it will **burst**. 

32. Suddenly he was **thrust** into the dark room. 

33. The management held a secret meeting. The issues discussed were not **disclosed** to the workers.
34. He is irresponsible. You cannot rely on him for help.  
35. She was sitting on a balcony and basking in the sun.  
36. In order to be accepted into the university, he had to improve his grades.  
37. He perceived a light at the end of the tunnel.  
38. We could hear the sergeant bellow commands to the troops.  
39. It's impossible to evaluate these results without knowing about the research methods that were used.  
40. For years waves of invaders pillaged towns along the coast.  
41. The telegram was delivered two hours after it had been sent.  
42. Children are not independent. They are attached to their parents.  
43. The boss got angry with the secretary and it took a lot of tact to soothe him.  
44. He finally attained a position of power in the company.  
45. The rescue attempt could not proceed quickly. It was impeded by bad weather.  
46. The differences were so slight that they went unnoticed.  
47. She showed off her slender figure in a long narrow dress.  
48. We do not have adequate information to make a decision.  
49. The story tells about a crime and subsequent punishment.  
50. I wouldn't hire him. He is unmotivated and indolent.  
51. The dress you're wearing is lovely.  
52. She has been changing partners often because she cannot have a stable relationship with one person.  
53. She is not a child, but a mature woman. She can make her own decisions.  
54. In a homogeneous class all students are of a similar proficiency.  
55. Computers have made typewriters old-fashioned and obsolete.
56. He wasn't very **popular** when he was a teenager, but he has many friends now.

57. You must wear a bathing suit on a public beach. You're not allowed to bath **naked**.

58. The prisoner was put in **solitary** confinement.

59. The urge to survive is **inherent** in all creatures.

60. Watch out for his **wily** tricks.
Appendix H

*Cognates and loanwords on the Productive Vocabulary Levels Test (Asian languages)*

<table>
<thead>
<tr>
<th>Level</th>
<th>Korean</th>
<th>Japanese</th>
<th>Mandarin</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,000</td>
<td>cream A-1</td>
<td>charm A-21</td>
<td>dozen B-6</td>
</tr>
<tr>
<td></td>
<td>stretched B-16</td>
<td>cream B-21</td>
<td></td>
</tr>
<tr>
<td>3,000</td>
<td>career A-2</td>
<td>career A-2</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>gown A-17</td>
<td>gown A-17</td>
<td></td>
</tr>
<tr>
<td></td>
<td>lawn B-7</td>
<td>reform A-12</td>
<td></td>
</tr>
<tr>
<td>5,000</td>
<td>eve B-8</td>
<td>eve B-8</td>
<td>---</td>
</tr>
<tr>
<td>UWL</td>
<td>clinic B-4</td>
<td>area A-9</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>motive B-14</td>
<td>clinic B-4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>trend A-4</td>
<td>secretion A-29</td>
<td></td>
</tr>
<tr>
<td>10,000</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>

Note. Cognates and loanwords are listed in alphabetical order. The letter following each word indicates that the item appears in part A (Appendix A) or part B (Appendix B), and the number indicates the item number.
Appendix I

*Cognates and loanwords on the Productive Vocabulary Levels Test (Spanish)*

<table>
<thead>
<tr>
<th>Level</th>
<th>Word</th>
<th>Part-item number</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,000</td>
<td>admire</td>
<td>B-26</td>
</tr>
<tr>
<td></td>
<td>cream</td>
<td>B-1</td>
</tr>
<tr>
<td></td>
<td>dozen</td>
<td>A-6</td>
</tr>
<tr>
<td></td>
<td>introduced*</td>
<td>B-21</td>
</tr>
<tr>
<td></td>
<td>opportunity</td>
<td>A-1</td>
</tr>
<tr>
<td></td>
<td>popular</td>
<td>B-56</td>
</tr>
<tr>
<td></td>
<td>pupils</td>
<td>B-11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3,000</td>
<td>acid</td>
<td>A-7</td>
</tr>
<tr>
<td></td>
<td>career</td>
<td>A-2</td>
</tr>
<tr>
<td></td>
<td>council</td>
<td>B-2</td>
</tr>
<tr>
<td></td>
<td>importation</td>
<td>A-22</td>
</tr>
<tr>
<td></td>
<td>nerves</td>
<td>A-27</td>
</tr>
<tr>
<td></td>
<td>perceived</td>
<td>B-37</td>
</tr>
<tr>
<td></td>
<td>proclaimed</td>
<td>B-22</td>
</tr>
<tr>
<td></td>
<td>reform</td>
<td>A-12</td>
</tr>
<tr>
<td></td>
<td>stable</td>
<td>B-52</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5,000</td>
<td>adequate</td>
<td>B-48</td>
</tr>
<tr>
<td></td>
<td>cavalry</td>
<td>B-13</td>
</tr>
<tr>
<td></td>
<td>comprehend</td>
<td>B-18</td>
</tr>
<tr>
<td></td>
<td>entrance</td>
<td>A-28</td>
</tr>
<tr>
<td></td>
<td>ignore</td>
<td>B-28</td>
</tr>
<tr>
<td></td>
<td>mature</td>
<td>B-53</td>
</tr>
<tr>
<td></td>
<td>solitary</td>
<td>B-58</td>
</tr>
<tr>
<td>UWL</td>
<td>accumulated</td>
<td>B-24</td>
</tr>
<tr>
<td></td>
<td>area</td>
<td>A-9</td>
</tr>
<tr>
<td></td>
<td>clinic</td>
<td>B-4</td>
</tr>
<tr>
<td></td>
<td>doctrine</td>
<td>A-19</td>
</tr>
<tr>
<td></td>
<td>evaluate</td>
<td>B-39</td>
</tr>
<tr>
<td></td>
<td>homogeneous</td>
<td>B-54</td>
</tr>
<tr>
<td></td>
<td>inherent</td>
<td>B-59</td>
</tr>
<tr>
<td></td>
<td>intimacy</td>
<td>A-24</td>
</tr>
<tr>
<td></td>
<td>inspected</td>
<td>B-19</td>
</tr>
<tr>
<td></td>
<td>motive</td>
<td>B-14</td>
</tr>
<tr>
<td></td>
<td>philosophy</td>
<td>A-14</td>
</tr>
<tr>
<td></td>
<td>saturated</td>
<td>B-29</td>
</tr>
<tr>
<td></td>
<td>section</td>
<td>A-29</td>
</tr>
<tr>
<td></td>
<td>subsequent</td>
<td>B-49</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>B-9</td>
</tr>
<tr>
<td>10,000</td>
<td>deacon</td>
<td>B-15</td>
</tr>
<tr>
<td></td>
<td>evacuation</td>
<td>A-50</td>
</tr>
<tr>
<td></td>
<td>felicity</td>
<td>B-5</td>
</tr>
<tr>
<td></td>
<td>impeded</td>
<td>B-45</td>
</tr>
<tr>
<td></td>
<td>indolent</td>
<td>B-50</td>
</tr>
<tr>
<td></td>
<td>mutilated</td>
<td>B-30</td>
</tr>
<tr>
<td></td>
<td>obsolete</td>
<td>B-55</td>
</tr>
<tr>
<td></td>
<td>orchid</td>
<td>A-20</td>
</tr>
<tr>
<td></td>
<td>predicament</td>
<td>B-10</td>
</tr>
<tr>
<td></td>
<td>secretions</td>
<td>A-25</td>
</tr>
</tbody>
</table>

Note. Cognates and loanwords are listed in alphabetical order. The letter following each word indicates that the item appears in part A (Appendix A) or part B (Appendix B), and the number indicates the item number.

*False cognate.
Appendix J

Student questionnaire

This is an evaluation to see how big your English vocabulary is. This is for very important research. You will not receive a grade for this, but please answer every question the best you can. You will take Section 1 today, and Section 2 during the next class. Your teacher will give you your results. Your participation is voluntary.

Please answer all of the following questions before you begin the vocabulary evaluation.

1. What is your name? (please print) ______________________________________
   Last   First

2. What is your gender?
   a. Male
   b. Female

3. What is your first language? _______________

4. Where did you grow up? City:_______________, Country:_______________

5. When did you begin studying at the English Language Center?
   - September (Fall)  1999
   - January (Winter)  2000
   - April (Summer)  2001
   - 2002
   - 2003

6. What level were you in when you began at the English Language Center?
   a. Level 1
   b. Level 2
   c. Level 3
   d. Level 4
   e. Level 5

7. What level are you in now?
   a. Level 1
   b. Level 2
   c. Level 3
   d. Level 4
   e. Level 5

8. How old are you? _____

9. How long have you studied English? _____ years, _____ months

10. How long have you lived in the United States? _____ years, _____ months

Thank you! You may begin the vocabulary evaluation.
Appendix K

Instructions to the teacher

This is a two-part vocabulary size test that is to be administered in each reading class. It is in two parts because it is long. Please administer Section 1 the first day, and Section 2 the second day. The students will need about 25 minutes to finish the test, but give them as long as they need. Please make sure that every student completely fills out the questionnaire before taking the test. While they are taking the test, DO NOT help them, but tell them to do their best.

Kevin Zimmerman will score the tests and you will be given the results of the test.

Thank you!
Appendix L

Instructions for Section 2

What is your name? (please print) _____________________________________ Last                                  First

This is Section 2 of the vocabulary evaluation. This is for very important research. You will not receive a grade for this, but please answer every question the best you can. Your teacher will give you your results. Your participation is voluntary.

You may begin the vocabulary evaluation.