The Reading Matrix
Vol. 7, No. 3, December 2007

VOCABULARY, PROFICIENCY AND READING COMPREHENSION

Maryam Golkar and Mortaza Yamini
mgolkar@um.ac.ir
myamini@rose.shirazu.ac.ir

Abstract

This study set out to empirically determine the reliability and validity of the Vocabulary Levels Tests, both the passive and productive versions. Furthermore, attempt was made to investigate the nature of the students' vocabulary knowledge with regard to their passive and active knowledge of the L2 words as a whole and at different word frequency levels. Moreover, the relationships between these two types of vocabulary knowledge and the learners' proficiency level and reading comprehension ability were studied. And finally, it was scrutinized if there were any significant differences between the High and Low proficient learners and also English majors and non-majors' passive and active vocabularies. Three tests, the Vocabulary Levels Test, the Productive Version of the Vocabulary Levels Test, and a TOEFL test, were administered to a group of 76 Iranian undergraduate students majoring in engineering and English Language and Literature. The results proved the Vocabulary Levels Tests to be reliable and valid tests of vocabulary size. The learners' passive and active vocabularies were also found to be highly correlated as a whole and at each separate word-frequency level. Passive vocabulary was always larger than active vocabulary at all levels; however, the gap between the two increased at lower word-frequency levels. In addition, there was a high correlation between the learners' vocabulary knowledge on the one hand and proficiency and reading comprehension ability on the other hand. It was also found that there was a statistically significant difference between the vocabulary knowledge of High proficient and Low proficient groups and also between the English majors and non-majors. The High proficient group and the English majors had greater passive and active vocabulary knowledge than their corresponding Low proficient group and the non-majors.

Introduction

Vocabulary as "the building block of language" (Schmitt, Schmitt, & Clapham, 2001, p. 53), is considered by some to be "the single most important aspect of foreign language learning" (Knight, 1994). Learners also regard learning vocabulary as one of the most important and at the same time difficult aspects of learning a language (Laufer, 1986). Yet, for a long time, this aspect of language research was largely neglected (Harlech-Jones, 1983; Laufer, 1986; Read, 1988). According to Laufer (1986), the majority of researchers studied grammar and phonology as these were more amenable to making generalizations in contrast

with vocabulary, which does not lend itself so easily to making abstractions and generalizations. However, there has recently been a renewed emphasis on the importance of lexical knowledge and vocabulary acquisition research (Herman, 2003; Jones, 1995; Laufer, 1986; Read, 1988; Zareva, 2005). Bachman (2000) speaks of the revival of vocabulary acquisition research and is hopeful that this will result in "new insights into the nature of vocabulary and a broadened view of its role in language use" (p. 9). In fact, there have been so many researchers turning their attention to vocabulary that according to Zahar, Cobb and Spada (2001), this line of research is no longer the "neglected area" that Meara (1980) spoke of.

Research on vocabulary acquisition involves having an understanding of what 'knowing a word' means, and then based on the definition of the concept of word, one can use appropriate tools and procedures to measure vocabulary knowledge (Bogaard, 2000). Vocabulary knowledge is not "an all-or-nothing phenomenon" (Laufer, 1998; Laufer & Paribakht, 1998), but involves degrees of knowledge (Meara, 1990). Lexical knowledge is "a continuum consisting of several levels and dimensions of knowledge," from just a familiarity with the word to the ability to use it correctly in free production (Laufer & Paribakht, 1998, p. 367)

Cronbach (1942, as cited in Bogaard, 2001, p. 491), refers to five aspects of vocabulary knowledge: "generalization (knowing the definition), application (knowledge about use), breadth of meaning (knowing different senses of a word), precision of meaning (knowing how to use the word in different situations), and availability (knowing how to use the word productively)." Nation (1990) proposed four aspects of vocabulary knowledge: form (spoken, written), position (grammatical, collocations), function (frequency, appropriateness), and meaning (conceptual, associative) (cited in Laufer & Paribakht, 1998). Laufer's proposal for the different dimensions of lexical knowledge includes "form (Phonological, graphic, morphological), syntactic behavior, meaning (referential, associative, pragmatic) and relations with other words (paradigmatic and syntagmatic)" (Laufer & Paribakht, 1998, p. 368).

Henriksen (1999) classifies lexical knowledge into three categories: partial vs. precise, shallow vs. deep and receptive vs. productive, with the use of word being an overriding control process. Zareva, Schwanenflugel and Nikolov (2005) investigated this three-dimensional model to see which dimension(s) was (were) more revealing of the overall state of learner's vocabulary knowledge at different proficiency levels. They found vocabulary size, word frequency effects, number of associations, and within-group consistency of participants' associative domain to be better predictors of language proficiency than native-like commonality of associations. In another study, Zareva (2005) tried to identify the smallest set of variables that could most efficiently predict lexical knowledge and determine the extent to which this three-dimensional model could account for the variance in word knowledge. She found that verifiable self-report could singly account for much of the variance in vocabulary knowledge and could be taken as its single best predictor. Moreover, the results showed that the model as a whole could virtually explain all the variance in the vocabulary knowledge of learners at different levels of language proficiency.

As for the assessment of vocabulary knowledge, different vocabulary tests have been devised, each tapping a specific aspect of lexical knowledge (Meara & Buxton, 1987; Schmitt, 1998). Some tests are more concerned with measuring the learners' breadth of knowledge, that is, the size of their vocabulary, while other tests try to address the learners'

depth of knowledge, which is the quality of their vocabulary knowledge or how well they know particular words (Greidanus & Nienhuis, 2001; Hunt, 1998; Laufer & Nation, 1995; Read, 1988; Read, 1993; Wesch & Paribakht, 1996; Vermeer, 2001; Wolter, 2001). The two types of word knowledge, however, have been found to be highly correlated (Schmitt & Meara, 1997).

Read (1988) introduced the checklist method of measuring vocabulary knowledge (also called the yes/no method), which simply presents the learners with a list of words and asks them to check the words they know. He claims that, although the test is not useful for assessment purposes, it can be used for research as it is an "economical way of surveying knowledge of a large number of words" (p. 21). Read (1993) proposed a new test format, the word associates format, which he believes assesses the depth of the learners' vocabulary knowledge while at the same time covering a large number of words in a simple straightforward manner. Laufer and Nation (1995) presented another new test, the Lexical Frequency Profile, which measures "the proportion of high frequency general service and academic words in learners' writing … seen as being a measure of how vocabulary size is reflected in use" (p. 305). Wesche and Paribakht (1996) devised the Vocabulary Knowledge Scale, which they claim assesses the learners' depth of vocabulary knowledge. Both the word associates and the Vocabulary Knowledge Scale were subsequently used in many research projects to determine the participants' knowledge of words (e.g., Nassaji, 2004; Zareva, 2005).

Nation proposed the Vocabulary Levels Test (VLT) as a measure of learners' breadth of lexical knowledge in 1983 and its revised version later in 1990. According to Schmitt et al. (2001, p. 58), the VLT "provides a profile of a learner's vocabulary, rather than a single figure estimate of overall vocabulary size." This test measures passive vocabulary knowledge at 5 word-frequency levels (2000, 3000, 5000, the university word list, and 10,000 words). Each level has 6 clusters including 6 words and 3 definitions. The testees are required to match the words and the definitions. The words in each cluster are from the same part of speech and are de-contextualized so that no clues to the meaning are provided by form. Also, the words in each cluster are semantically distinct to make the test a sensitive test of vocabulary knowledge; that is, the format is sensitive to partial word knowledge (Beglar, 2000; Waring, 1996). In other words, learners with just a superficial knowledge of the basic meanings of the words can select the right word. To answer the test, learners do not need to have knowledge about other aspects of the word such as its grammatical form, collocation, function, and so on. They need not "differentiate between semantically related words or show awareness of shades of meaning" (Read, 1988, p. 18). The ratio of the nouns, verbs and adjectives in each level is 3: 2: 1. The VLT enjoys high reliability and is highly scalable; that is, if a learner reaches the criterion at a lower level, for example, the 5000-word level, (s)he has most probably mastered the 2000 and 3000 levels as well (Read, 1988). Read and Chapelle (2001) consider this test as a discrete, selective, context-independent test which is "designed to measure learners' vocabulary size as a trait without any reference to any particular context of use" (p. 8).

The productive version of the Vocabulary Levels Test was proposed by Laufer and Nation (1995). It has the same word levels and number of items. Each item in this productive version consists of a sentence with a missing word whose initial letters are provided. The letters are given to prevent the learners from producing an alternative form which might fit the context and to restrict them to producing the desired item (Waring, 1996). This productive

version has also been reported to be a valid measure of vocabulary size and can discriminate among learners of different proficiency levels (Laufer, 1998; Laufer & Nation, 1995). The Vocabulary Levels Tests, both passive and active, have been used in many studies to specify participants' breadth of vocabulary knowledge or to group them from this point of view (e.g., Cobb, 1999; Qian, 1999; Webb, 2005; Zahar, Cobb, & Spada, 2001).

The vocabulary size tests, however, have been criticized by Laufer and Goldstein (2004) as being unidimensional and assessing only passive recognition. They tried a computer adaptive version of the test, which could measure active and passive recall and recognition of vocabulary items. They found the four strengths of knowledge to be implicationally scaled, with the active recall being the most difficult and the passive recognition the easiest. They also noticed that the passive recall test was the best predictor of academic achievement, accounting for 42.6% of the total variance in the students' grades. The test, they believe, could tap the incremental nature of vocabulary knowledge and could be more efficiently used for placement and research purposes than the depth tests.

Regarding the relationship between vocabulary size, proficiency level and reading comprehension, word knowledge has been known as an important factor in language proficiency (Grabe, 1991; Hermann, 2003; Zareva et al., 2005), reading comprehension (August, Carlo, Dresler & Snow, 2005; Qian, 1999, 2002; Read, 1988) and subsequent vocabulary acquisition (Pulido, 2003; Verspoor & Lowie, 2003). Laufer (1998) asserts that "vocabulary correlates with holistic assessments of writing and general proficiency, and is the best single predictor of reading comprehension" (p. 1). Similarly, a number of other sstudies have shown that tests of breadth of vocabulary knowledge can very well predict success in reading, writing, general proficiency and academic achievement (Laufer & Goldstein, 2004; Saville-Troike, 1984; Nation & Meara, 2002). According to Nassaji (2004), depth of vocabulary knowledge facilitates inferencing, thus enhancing vocabulary acquisition through reading. Qian (1999) also argues for the superiority of depth over breadth of L2 vocabulary knowledge in predicting reading comprehension ability. Zareva et al. (2005), however, contend that the superiority of quality over size occurs only after the 5,000 to 6,000 threshold has been passed. Some other studies also support the existence of a threshold below which reading comprehension is seriously hampered, the threshold varying between the 2,000 word families (Hirsh & Nation, 1992), 3,000 word families or 5,000 individual word forms (Laufer, 1998) and 4,000-5,000 word families (Sutarsyah, Nation, & Kennedy, 1994).

Laufer (1998) argues that guessing the meaning of words is not possible unless one knows at least about 95% of the neighboring words. In her opinion, knowledge of 3,000 word families (5,000 lexical items) is the minimum threshold to make one able to read about 95% of the reading passage and achieve about 60% comprehension. She maintains that learners with fewer lexical items are poor readers and holds that knowing more than 5000 word families (8000 lexical items) correlates with 70% reading comprehension (Hunt, 1998). According to Henriksen (2000), some people consider 2000 words enough to read a text with 80% comprehension, but argues that the minimum requirement for satisfactory reading is 95%. In his opinion, a vocabulary of 10,000 words is necessary for academic studies, which is the estimate also obtained by Hazenberg and Hulstijn (1996) in a separate study on Dutch students. Hsueh-chao and Nation (2000) consider the 98% coverage of vocabulary as the minimum requirement to attain unassisted comprehension of the text. Nation and Waring (1997) propose a vocabulary of 15,000 to 20,000 as a prerequisite to native-like reading comprehension. And finally, Schmitt et al. (2001) hold that learning the most frequent 2000

words is necessary for oral communication, the 3000 level is needed for beginning to read authentic passages, and the 5000 word level makes reading authentic texts possible, allowing learners to guess the meaning of the unknown words from context. They believe that knowledge of the most frequent 10,000 words is needed for handling university study requirements.

Milton and Meara's (1995) research results indicate that the rate of vocabulary acquisition is much higher (2500) when learners are in the context where the language is being spoken than when they are learning words in the formal context of classroom. According to Hever and Cederberg (1997), students in the ordinary school system in Sweden have a vocabulary size of about 3000 words when they start their upper secondary studies and 5000-7000 words at the pre-university level. Cobb (1999) estimates the receptive vocabulary of commerce students at Sultan Qaboos University in Oman at 1,000 words. Laufer (1998) expects a vocabulary size of 3500 to 4000 for the graduates of the Israli high schools and compares this to the 18,000-20,000 word families known by English native speakers at the end of high school. There are, however, different estimates of the adult native speakers' vocabulary knowledge, ranging from 14,000 (Zareva et al., 2005), or 17,000 (Goulden, Nation & Read, 1990) to 20,000 active words (Henriksen, 2000). Children, as reported by August et al. (2005), are said to learn approximately 800 to 900 roots a year up to 12 years old, and have a 5000 to 7000 vocabulary repertoire in their L1 before they enter school. By the fifth grade, they are estimated to have an L1 vocabulary size of 8,400 words.

In view of the above-mentioned discussions, the present study attempted to examine the Vocabulary Levels Tests, both the passive and productive versions, to see if they were reliable and valid tests of vocabulary knowledge and at the same time to investigate the basic passive/receptive vocabulary and controlled active/productive use of vocabulary and the relationship between these two forms of word knowledge. Moreover, the relationship between the vocabulary level, proficiency and reading comprehension ability was explored. The study of the differences between the vocabularies of the High and Low proficient students and also English majors and non-majors was another concern of this research.

Research Questions

The specific questions to be answered in this study were as follows:

1. Are the Vocabulary Levels tests, passive and productive versions, reliable and valid tests of vocabulary knowledge?

2. What are the relationships between learners' passive and active vocabularies and the active-to-passive ratios? What are these relationships at each separate level of lexical knowledge?

3. Are there any significant differences between learners' passive and active vocabularies and the active-to-passive ratios in the tests as a whole and at each separate vocabulary level ?

4. What are the relationships between learners' passive and active vocabularies and their proficiency level?

5. What are the relationships between learners' passive and active vocabularies and their reading comprehension ability?

6. Are there any significant differences between the passive and active vocabularies and the active-to-passive ratios of High proficient and Low proficient students?

7. Are there any significant differences between the passive and active vocabularies and the active-to-passive ratios of English majors and non-majors?

Methods

Participants

The participants in this study were 76 male and female undergraduate students with an age range of 20 to 30, studying at Shiraz University. They were the students of three intact classes. Forty-six of them were engineering students studying English for Specific Purposes as required course. The others (N = 30) were senior English language and literature majors. Fourteen of the engineering students and eight of the English majors had either not taken the tests seriously or were not present in the first session of the test administration and had not taken the vocabulary tests. They were thus excluded from the study reducing the number of participants to 32 engineering and 22 English students. Moreover, only 27 engineering and 13 English majors were present in the second session of the test administration and took the TOEFL test.

Instruments

Three tests were used in this study including the Vocabulary Levels Test (Version 2), the Productive Version of the Vocabulary Levels Test, and the TOEFL test.

The Vocabulary Levels Test (Version 2)

The Vocabulary Levels Test used in this study was the second version of the levels test revised and validated by Schmitt et al. (2001). It was originally produced by Nation in 1983 and was later revised by him in 1990. It provides an estimate of vocabulary size at 2000, 3000, 5000, and 10000 frequency levels and also provides an estimate of the size of the examinee's academic vocabulary. In this version, there are 10 clusters at each level and each cluster has six words and three definitions. So the test has 150 items. There are 15 nouns, 9 verbs, and 3 adjectives in each word level with a ratio of 3: 2: 1, so there are 75 nouns, 50 verbs, and 25 adjectives assessed in the whole test. The testees are supposed to match the definitions on the right in each cluster with the corresponding words on the left. The items are not contextualized so that no clues to the meaning are provided. The format of the test used in this research was modified so that an answer sheet could be provided; that is, the definitions were numbered from 1 to 150, and the words in the clusters were designated with small letters from "a" to "f". An example item is provided below:

a. business

b. clock	1. _____ part of a house
c. horse	2. _____ animal with four legs
d. pencil	3. _____ something used for writing
e. shoe	
f. wall	

Schmitt et al. (2001) conducted a study to establish item discrimination, item facility, reliability and validity of the test. The reliability of the different levels of version 2 ranged from .92 to .96. The mean facility index for the tests ranged from .78 to .29 decreasing as the target words move from the most frequent to less frequent items. The mean discrimination indices varied from .51 to .67 with no items having discrimination values below .30, which is an acceptable discrimination index. The different levels of the test were found to be 'highly scalable.' The results of factor analysis also showed that "the levels test is unidimensional, with the key measured trait being vocabulary knowledge" (Schmitt, et al., 2001).

The Productive Version of the Vocabulary Levels Test

The Productive Version of the Vocabulary Levels Test for Controlled Active Vocabulary Size (Laufer & Nation, 1995) also measures vocabulary knowledge at the same five levels as the passive version does. The test has 18 items in each level and a total of 90 items. Here, the testees read a sentence from which a word has been omitted. The beginning letters of the words have been provided to prevent the use of non-target words. The students are supposed to provide the missing letters in each word in the space provided on the answer sheet. This test had 13 common words with the passive version, 6 in the 2000 level, 3 in the 3000 level and 4 in the academic level. A sample item from the test looks like the following:

He was riding a bi_____.

Laufer and Nation (1995) stated that Vocabulary Levels Tests, both passive and productive versions, are established measures of lexical knowledge and they used the productive version of the test as the criterion against which the Lexical Frequency Profile, another vocabulary test, was validated.

TOEFL Test

TOEFL (Test of English as a Foreign Language) is a standardized test used to evaluate the English proficiency of those whose native language is not English. The TOEFL test utilized in this research was the 1980 version copyrighted by the Educational Testing Service. It has three sections. Section I--Listening Comprehension--includes 50 items; section II--Structure and Written Expression--includes 40 items; and Section III--Reading Comprehension and Vocabulary--includes 60 items. Due to limitations of time and for practical purposes, just the second and third sections of the test were administered to the participants.

Procedure

The tests were administered in two sessions. First the two vocabulary tests were administered in one session during the students' class hours. The productive version of the test was administered first to make sure that the students did not take benefit from the experience and knowledge gained from the passive version. As there were 13 words in common between the two tests, the students' exposure to the definition and meaning of the words in the passive version might have affected their performance on the productive version had they been given the passive version first.

When the passive version was given to the participants, they were asked to answer an item even if they were not definitely sure that their response was correct. However, they were prohibited from blind guessing, although they were informed that incorrect responses would get no negative points.

Two weeks later, the TOEFL test was administered under the standard procedures. The participants were given 25 minutes to answer Section II and 45 minutes to do Section III.

Results and Discussion

Estimating Passive and Active Vocabulary Levels

To estimate the testees' passive and active vocabulary levels, the procedure followed by Laufer (1998) was adopted:

1. Passive Vocabulary Level

((2000 passive score * 2) + 3000 passive score + Academic vocabulary score + 5000 passive score + ((3000 passive score + 5000 score) / 2) + ((5000 passive score + 10000 passive score) / 2 * 4) + 10000 passive score) / 330 * 10000

2. Active Vocabulary Level

((2000 active score * 2) + 3000 active score + 5000 active score + University Word List score + ((3000 active score + 5000 active score) / 2) + ((5000 active score + 10000 active score) / 2 * 4) + 10000 active score) / 198 * 10000

The descriptive statistics for the participants' passive and active vocabulary levels separately and their vocabulary level as a whole are displayed in Table 1.

Table 1
Participants' Passive and Active Vocabulary Levels and their Total Vocabulary Level

Tests	Minimum	Maximum	Mean	S.D.	variance
Passive Level	348.48	9196.97	4347.64	1930.12	3725370
Active Level	378.79	6893.94	2514.03	1605.60	2577968
Vocabulary Level	363.64	8032.83	3430.84	1715.59	2943238

It is evident that the learners enjoyed a higher passive (4347.64) than active (2514.03) vocabulary level. As mentioned by Laufer and Goldstein (2004), it shows that active knowledge

of the words is a more advanced form than passive knowledge. Figure 1 graphically displays the difference between the participants' passive and active vocabulary levels in the test as a whole.

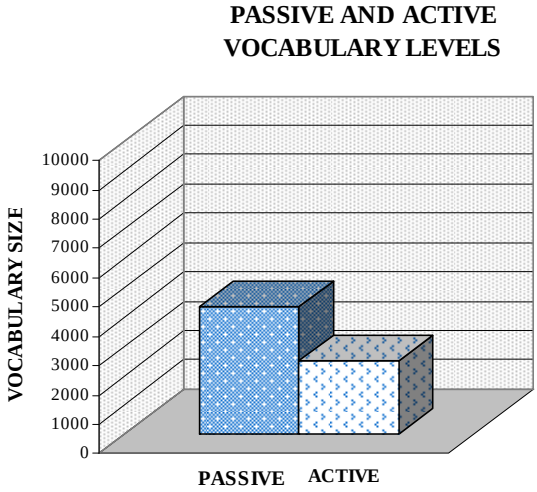

Figure 1. Passive and Active Vocabulary Levels

Research Question 1

Are the Vocabulary Levels tests, both the passive and productive versions, reliable and valid tests of vocabulary knowledge?

The descriptive statistics and the reliability coefficients of the two vocabulary tests as a whole and at each separate word level and also the TOEFL test and its reading comprehension section are presented in Table 2. To estimate the reliability of the tests, the Kudar-Richardson Formula 21 (KR-21) was utilized. To facilitate the comparison between tests of different lengths, their reliabilities were adjusted for 100 items and are reported together with the observed reliability. Likewise, the means are reported in percentages to make comparison possible. As is observed in the table, the tests enjoyed high reliabilities. The adjusted reliability of the passive version, the productive version and the TOEFL test were .94, .95, and .93, respectively. The reliabilities of the scores obtained from the subtests were also high and ranged from .82 to .91.

The correlation coefficients between the Vocabulary Levels Tests (passive and productive) and the vocabulary part of the TOEFL test turned out to be .88 and .81, respectively. They were significant at p=.000. It shows that these two tests are valid tests of vocabulary knowledge. Furthermore, as Schmitt et al. (2001) report, "learners acquire more frequently-used words before they acquire less frequently-used ones" (p. 67). In other words, they found that the four frequency sections of the Vocabulary Levels Test (the passive version) are quite scalable; that is, if a learner has mastered one level, it can be assumed that (s)he has reached the criterion mastery at the higher frequency levels. A look at Table 2 reveals that the results obtained on this test administration are quite scaleable, too. The means of the different passive and active levels progressively decrease as one goes down to lower word frequency levels. The means for the passive version at the 2000, 3000, academic, 5000 and the 10000 word levels were 25.30, 19.57,

17.80, 10.33 and 3.19, and those for the productive version were 12.98, 7.08, 3.11, 2.24 and .89, respectively. The means of the academic sections neatly fit between the 3000 and 5000 word levels in both passive and productive tests in this research. However, Schmitt et al. (2001) claim that the words in "the Academic section are not primarily frequency driven" (p. 68) and that they can be placed anywhere between the 2000 level and the 10,000 level.

Table 2
Descriptive Statistics and Reliability Coefficients

Test	No items	Range	Min.	Max	Mean		S.D.	Variance	Reliability	
					Raw	%			Obs.[a]	100 It
Passive Test	**150**	**134**	**6**	**140**	**76.24**	**50.83**	**29.15**	**850**	**.96**	**.94**
2000 level	30	25	5	30	25.30	84.33	5.01	25.12	.89	.96
3000 level	30	29	1	30	19.57	65.23	7.88	62.02	.92	.97
Academic	30	28	0	28	17.80	59.33	7.06	49.79	.88	.96
5000 level	30	29	0	29	10.33	34.43	8.41	70.79	.94	.98
10,000 level	30	23	0	23	3.19	10.63	4.63	21.47	.90	.97
Active Test	**90**	**64**	**4**	**68**	**26.50**	**29.44**	**16.11**	**259.5**	**.94**	**.95**
2000 level	18	26	3	29	12.98	72.11	4.54	20.62	.86	.97
3000 level	18	29	0	29	7.08	39.33	5.81	33.71	.92	.98
5000 level	18	14	0	14	2.24	12.44	3.37	11.36	.88	.98
UWL[b]	18	14	0	14	3.11	17.28	3.97	15.76	.89	.98
10,000 level	18	6	0	6	.89	4.94	1.27	1.60	.81	.96
TOEFL Test	**100**	**72**	**16**	**88**	**47.73**	**47.73**	**17.6**	**310.3**	**.93**	**.93**
Section II	60	46	7	53	26.63	44.38	10.92	119.2	.89	.93

Note. a. Obs = observed.

b. UWL= The University Word List Level, or the Academic Vocabulary Level.

Moreover, Jafarpur (2001) uses decision consistency as another check for the validity and defines it as "the percent classification of subjects by the experimental test that corresponds to those by the criterion" (pp. 32-33). In this research, the top and bottom 27% of the students ranked according to their TOEFL scores were considered High and Low groups (N=11). The extent to which the two tests could correctly place the students in the High and Low proficiency groups is shown in Table 3.

Table 3
Decision Consistency

	High	Low
Passive	82%	91%
Active	64%	64%

On the whole, the passive vocabulary test could correctly determine the students' place in the High or Low proficiency groups in 86% of the cases and the controlled productive version

predicted proficiency level correctly with 64% accuracy. The passive version was superior to the productive version in this regard.

All in all, the results of the correlational analyses between the Vocabulary Levels tests and some previously established reliable and valid test such as the TOEFl, test and also the scalability and decision consistency of their results attest to the fact that both the passive and productive versions of the Vocabulary Levels Test are quite reliable and valid tests of vocabulary size.

Research Question 2

What are the relationships between learners' passive and active vocabularies and the active-to-passive ratios? What are these relationships at each separate level of lexical knowledge?

The mean of the passive vocabulary test in percentage was 50.83 and that of the active test 29.44. The active-to-passive ratio turned out to be 58 percent. Therefore, the learners' passive vocabulary was larger than their active vocabulary; in other words, the learners could not use all the words they knew productively. Only 58% of them were employed in production.

The Pearson product-moment correlation coefficients between the passive and active vocabulary scores and also between the passive scores and the active-to-passive ratios were estimated. The coefficients were .87 (p=.00, N= 54) and .30 (p=.03, N=54), respectively. There was a high correlation between the learners' passive and active vocabularies; in other words, as the learners' passive vocabulary increased, their active vocabulary increased as well. However, the low correlation between the passive vocabulary and the active-to-passive ratio, although statistically significant, indicates that students could not turn their passive vocabulary into active in the same proportion as their passive vocabulary grew. That is to say, their productive use of vocabulary items lagged behind their passive vocabulary knowledge.

The second part of this research question tries to clarify if the patterns observed for the relations in the whole test held at each separate vocabulary level as well. The correlations between the passive and active vocabularies at each separate word level were quite high and significant at the p=.000. They were .77 for the 2000 level, .80 for the 3000 level, .76 for the 5000 level, .77 for the 10,000 level and .52 for the academic level. This indicates that as the learners' passive vocabulary knowledge increased at each separate level, their active vocabulary increased as well. Table 4 displays the means of the passive and active tests at each separate vocabulary level and their corresponding active-to-passive ratios.

Table 4
Means and Ratios of the Passive and Active Vocabularies at Different Word Levels

Word Level	Passive (%)	Active (%)	Ratio (%)
2000	84.32	72.12	86
3000	65.25	43.31	66
Academic or UWL	59.32	17.28	29
5000	34.44	12.45	36
10,000	10.62	2.16	20

The Table shows that the amount of passive and active vocabularies decreases as one goes down the column to lower frequency word levels; however, the amount of active vocabulary decreases more abruptly than the amount of passive vocabulary. In other words, the gap between passive and active vocabularies increases at the lower frequency levels. Put another way, learners are less able to put their passive knowledge to use productively at lower frequency levels. A look at the ratio column also shows that the ratio decreases with a decrease in the word frequency level. The words at the 2000 word level are more frequent words and are more likely to be used by the learners in speech or writing than words at the 3000, 5000 or 10,000 levels. Similarly, words at the 3000 level have a higher possibility of being employed in production than those at the 5000 or 10,000 word levels. The only exception was the academic word level, which could not be used as well as the words at the 5000 word level. The university students participating in this study were to some extent familiar with the academic terminology; however, they could not use a large proportion of them in speech and writing. These words are usually formal words not likely to be used in everyday communication. The bar graph in Figure 2 depicts the decrease in the passive and active vocabulary levels as one goes from higher to lower word frequency levels and also the scalability of the different word levels.

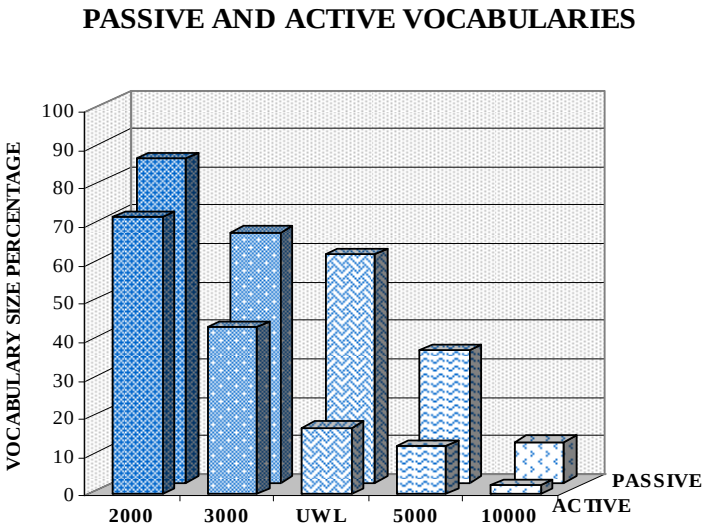

Figure 2. Passive and Active Vocabularies at Separate Word Levels

Research Question 3

Are there any significant differences between learners' passive and active vocabularies and the active-to-passive ratios in the tests as a whole and at each separate vocabulary level?

A number of paired samples t-tests were utilized to check the significance of the differences between the learners' passive and active vocabularies as a whole and at each separate frequency level. As the lengths of the tests and subtests were different, the means and standard deviations are expressed in percentage terms. Table 5 displays the results.

Table 5
Paired Samples T-Tests between Passive and Active Scores wholly and at Different Levels

Pairs	Mean	S.D.	t-value	d.f.	Sig. (2-tailed)
1. Passive Vocabulary Level Active Vocabulary Level	50.83 29.44	19.44 17.90	16.28	53	.000
2. Passive 2000 Active 2000	84.32 72.11	16.71 25.33	5.49	53	.000
3. Passive 3000 Active 3000	65.25 43.31	26.25 32.26	8.28	53	.000
4. Passive 5000 Active 5000	34.44 12.45	28.05 18.72	8.78	53	.000
5. Passive Academic Active UWL	59.32 17.28	23.52 22.06	13.78	53	.000
6. Passive 10,000 Active 10,000	10.62 2.16	15.45 7.03	5.67	53	.000

As is observed, there was a significant difference between the learners' passive and active vocabularies in the tests as a whole and at each separate word level (p = 000). The learners' passive vocabulary was always larger than their active vocabulary. In addition, the means of active-to-passive ratios at each two subsequent levels were compared. As the place of the Academic or UWL level is not certain, the means of this level were compared with the means of the 3000, 5000 and 10,000 word levels. The results are shown in Table 6.

Except for the t-value of the A/P 5000 and A/P UWL pair, the difference between the means in all the other pairs was significant at p=.000. This further confirms that at higher frequency levels, a higher proportion of the words are used productively. As one moves to lower frequency levels, this proportion decreases to a statistically significant degree. Although the participants' passive and active academic vocabulary size fell between the 3000 and 5000 word levels, the amount these words were put into practice by the learners was akin to the 5000 word level, indicating that a lower proportion of these words was used in communication.

Table 6
Paired Samples T-Tests Between Active-to-Passive Ratios at Each two Subsequent Levels

Ratio	Mean	S.D.	t-value	d.f.	Sig. (2-tailed)
1 A/P 2000 A/p 3000	.51 .37	.12 .21	6.43	53	.000
2 A/P 3000 A/P 5000	.37 .18	.21 .23	5.38	53	.000
3 A/P 5000 A/P UWL	.18 .16	.23 .25	.44	53	.668
4 A/P UWL A/P 10,000	.16 .05	.25 .13	4.30	543	.000

5 A/P 5000 **A/P 10,000**	.18 .5	.23 .13	3.82	53	.000
6 A/P 3000 **A/P UWL**	.37 .16	.21 .25	6.29	53	.000

Research Questions 4 & 5

What is the relationship between learners' passive and active vocabularies and their proficiency levels?

What is the relationship between learners' passive and active vocabularies and their reading comprehension ability?

To estimate the relationship between the learners' passive and active vocabularies on the one hand and their proficiency and reading comprehension ability on the other hand, the Pearson Product-Moment correlation was used. The scores of the students on the TOEFL test were taken as an indication of their proficiency level, and the scores on Section III of the TOEFL test, claimed to measure the "ability to understand non-technical matter" (Educational Testing Service, 1980, p. 5), were considered to be indicative of their reading comprehension ability. However, the first 30 items in this section are vocabulary items, which might thus inflate the correlations obtained. Therefore, a second correlation coefficient was estimated for the relationship between the passive and active vocabularies and the second 30 items of section III, i.e., the reading comprehension items. The descriptive statistics and the results of the correlations are shown in Tables 7 and 8.

In all cases, there were substantial correlations between each pair of variables significant at p=.000. Therefore, there was a close relationship between the learners' passive and active vocabulary knowledge and their proficiency level (r = .91 and r = .88, respectively) and also between their passive and active vocabularies and reading comprehension ability (r = .75 and r = .80). In other words, those with a larger vocabulary repertoire enjoyed higher proficiencies as well as being better able to read and comprehend texts. Moreover, active vocabulary turned out to be more highly correlated with reading comprehension. The reason might be that passive knowledge takes much practice and experience in language to turn into active, and so those with a higher active vocabulary have had a higher amount of practice in reading texts, too; hence their better reading comprehension ability. As expected, the correlations with Section III of the TOEFL test were higher when it included the vocabulary part, especially so for the passive test (r = .89 and r = .88, respectively) as the TOEFL vocabulary items are multiple-choice items checking the learners' passive recognition of the words.

Table 7
Descriptive Statistics of the Groups Taking the TOEFL Test

	No. of students	Range	Min.	Max	Mean		S.D.	Variance
					Raw	%		
Passive	40	134	6	140	73.5	49	30.94	957.08
Active	40	64	4	68	24.70	27.44	16.70	278.88
TOEFL	40	72	16	88	47.73	47.73	17.61	310.26
Section III	40	46	7	53	26.63	44.38	10.92	119.22

Reading Items	40	30	2	26	11.35	43.65	5.86	34.34

Table 8
Correlation Coefficients Between Passive and Active and Proficiency and Also Reading Ability

	Proficiency	**Section III**	**Section III Reading Comprehension Items**
Passive Vocabulary	.91	.89	.75
Active Vocabulary	.88	.88	.80

Research Question 6

Are there any significant differences between learners' passive and active vocabularies and the active-to-passive ratios of High proficient and Low proficient students?

The top and bottom 27% of the students according to their scores on the TOEFL test were considered as the High and Low proficient learners, respectively. Then, the two groups' passive and active vocabularies as a whole and at separate vocabulary levels as well as their active-to-passive ratios were compared using independent samples t-tests. The aim was to see if there were any significant differences between the vocabulary knowledge of the High proficient and Low proficient students and also the proportion to which each group was able to produce the words in speech and writing. Results are shown in Table 9.

Table 9
Independent Sample T-Tests Between High and Low Proficient Groups

Test	**Low Proficient**		**High Proficient**		**Difference**	
	Mean(%	S.D.(%)	Mean(%	S.D.(%)	T-value	Sig. (2-
Total Passive	**27.21**	**10.27**	**73.15**	**10.80**	**10.22**	**.000**
2000	63.94	21.18	95.76	10.76	4.74	.000
3000	28.79	11.76	89.09	10.76	12.55	.000
Academic	33.64	18.88	81.52	8.61	7.65	.000
5000	8.79	6.01	72.42	22.71	8.98	.000
10,000	.91	1.56	29.09	21.56	4.32	.001
Total Active	**10.61**	**4.11**	**49.60**	**16.32**	**7.68**	**.000**
2000	41.41	17.77	96.46	22.54	6.78	.000
3000	10.10	8.54	75.76	33.82	6.24	.000
UWL	.50	1.68	35.86	24.77	4.72	.001
5000	1.01	2.25	34.34	26.85	4.10	.002
10,000	.00	.00	6.06	10.95	1.83	.097
Total Ratio	**26.82**	**14.39**	**40.09**	**9**	**2.56**	**.019**

2000	40.45	10.49	60.64	12.90	4.02	.001
3000	26.55	26.99	48.09	19.23	2.16	.043
Academic or UWL	1	3	26.36	17.79	4.65	.001
5000	18.18	40.45	28	19.27	.727	.476
10,000	.00	.00	10.45	17.68	1.96	.078

There was a significant difference between the High and Low proficient learners' passive vocabulary knowledge in the test as a whole (p<.000) and at each separate word level (p<.001). Their active vocabularies were also significantly different in the test as a whole (p<.000) and at all but one frequency level (p<.002). So, the High proficient language learners knew a lot more vocabulary items than their Low proficient counterparts and were also more at ease using the words they knew in speech and writing. Only the difference between their active vocabularies at the 10,000 word level was not significant (p=.097). An investigation of the participants' performance on this part of the test revealed that the non-significant result was due to the "floor effect," that is to say, the active 10,000 word level was too difficult for the High and Low proficient learners alike, so neither group could perform well on this section, hence the non-significant difference observed. Figures 3 and 4 illustrate these differences graphically.

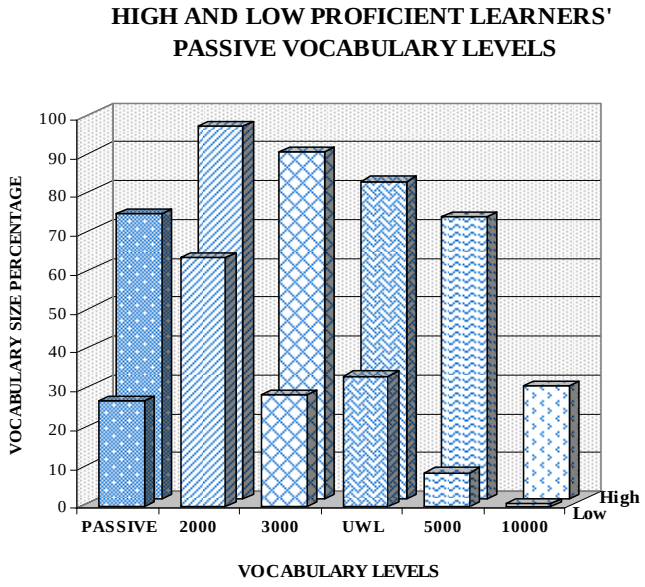

Figure 3. Differential Passive Vocabulary Levels of High and Low Proficient Learners

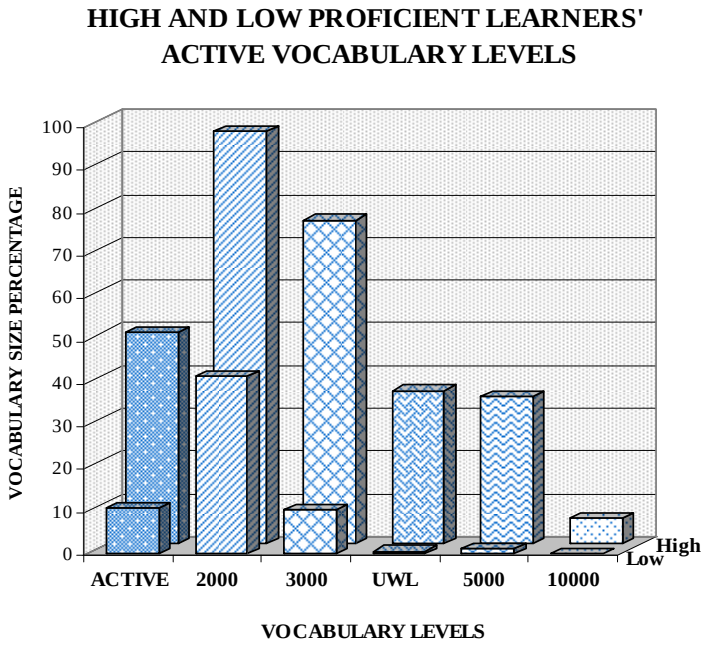

Figure 4. Differential Active Vocabulary Levels of High and Low Proficient Learners

Meanwhile, the comparison of active-to-passive ratios of the High and Low proficient groups shows that the students in the High proficient group were much more able to use their vocabulary knowledge productively. However, the difference was not significant at the 5000 and 10,000 word levels. The proportion of the words used in production at these two levels was low even for the High proficient group, therefore, this proportion did not show any difference from that of the Low group. In other words, at these two levels, the proportion to which the High proficient language learners put their word knowledge into use was not any different from that of the Low group, who did not have even much passive knowledge of the words to be able to use them productively.

On the whole, it can be concluded that the High proficient group enjoyed higher passive and active vocabulary knowledge and was better able to put this knowledge to use in production. That is to say, a higher proportion of their passive vocabulary knowledge in general and at different word frequency levels was utilized in production.

Research question 7

Are there any significant differences between learners' passive and active vocabularies and the active-to-passive ratios of English majors and non-majors?

Regarding the difference between the English majors and non-majors, a significant difference was observed between their means on the passive and active versions of the Vocabulary Levels Test and also between their active-to-passive ratios. The results are displayed in Table 10.

Table 10
Independent Samples T-Tests Between Majors and Non-majors

	Non-majors		Majors		t-Value	Sig. (2-tailed)
	Mean (%)	S.D. (%)	Mean (%)	S.D. (%)		
Passive	40.06	15.08	66.48	13.58	6.58	.000
Active	19.30	11.21	44.19	15.44	6.87	.000
Ratio	29	11	40	10	3.54	.001

As expected, the students whose major was English Literature had significantly higher passive and active vocabulary levels and were better able to use the words they knew in speech and writing. Figure 5 can very well demonstrate the differences between the two groups.

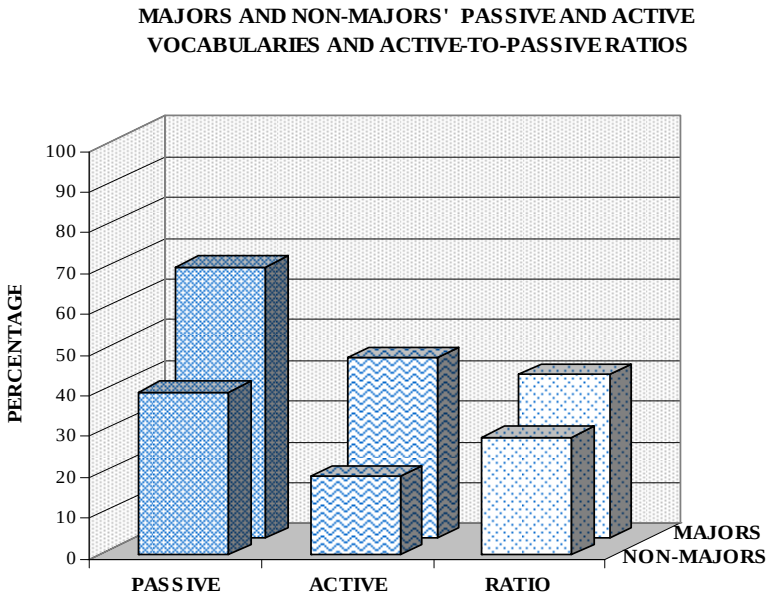

Figure 5. Differential Passive, Active and Active-to-Passive Ratios of Majors and Non-majors

Pedagogical Implications and Suggestions for Further Research

Research like this might pave the way toward a better understanding of the nature of the learners' vocabulary knowledge and might then prove to be helpful to pedagogy. The reliability and validity of the Vocabulary Levels Tests having been established empirically, language teachers as well as second language acquisition researchers can use the tests for both evaluation and research purposes. In addition, teachers can take benefit from the results as a better understanding of the relationship between active and passive vocabularies and the important role it has in language proficiency and reading comprehension ability is a prerequisite to devising

techniques to enable students enhance their passive knowledge of the words and learn how to make active use of that knowledge in the real act of communication.

Considering the gap between passive and active vocabularies and the fact that the gap widens at higher proficiency levels, teachers need to promote methods that are specifically useful in activating the passive knowledge of the students. It is a misconception to believe that the receptive vocabulary, which is partially developed through reading, will automatically develop into productive proficiency. As Nation and Meara (2002) suggest, the way to acquiring fluency in the use of words is either through repetition, what they name "the well-beaten path approach" (p. 43), or through making as many associations as possible between the new word and the words already known. Although vocabulary can be acquired incidentally through exposure to the words in the context of the materials the students read or listen to (Huckin & Coady, 1999; Elley, 1989; Joe, 1998; Hermann, 2003; Paribakht & Wesche, 1999, Wesche & Paribakht, 1999; Webb, 2005; Wode, 1999), it is a very slow process (Cobb, T., 1999; Hulstijn, 2002; Zahar, Cobb & Spada, 2001). To facilitate the process and speed it up, more systematic attention needs to be paid to the explicit teaching of the words in the language classroom (Horst, Cobb & Meara, 1998; Joe, 1998; Nation & Meara, 2002; Zahar, Cobb & Spada, 2001; Zimmerman, 1997). There is need for direct instruction to bring the differences between the way the words are used to the student's attention so that they will notice these differences when they later see the words in context and internalize the word usage. Material developers can also use the findings to devise better materials to help teachers in their efforts.

However, further research is needed to improve the characteristics of the tests as measuring and research tools and to determine the effect of different teaching and learning strategies on the development of the learners' passive and active vocabularies. As an example, the effect of guessing on the students' performance on the passive version of the Vocabulary Levels might be a good topic for scrutiny. It might also be investigated which techniques are more helpful in enhancing the students' recall and retention of the words they have learned.

Conclusion

One of the primary concerns of this research was the examination of the reliability and validity of the Vocabulary Levels tests, both the passive and productive versions. The study of the relationships between the students’ passive and active vocabularies was another focus of this research. Also, the relationships between the learners’ vocabulary knowledge and their proficiency level and reading comprehension ability were explored. Furthermore, attempt was made to determine the ratios of active to passive vocabularies of learners on the whole and at each separate level of vocabulary knowledge. And finally, the differences between the passive and active and also the ratios of active to passive vocabularies of High and Low proficient students and also majors and non-majors were investigated.

The results of the correlation between the Vocabulary Levels Tests and the vocabulary section of the TOEFL test, which is an already well-established reliable and valid test, proved that the Vocabulary Levels Tests enjoyed high degrees of reliability and were valid tests of vocabulary size. The tests produced highly consistent scores and measured a unitary trait, namely, vocabulary knowledge.

Furthermore, there was a positive relationship between the participants' passive and active vocabularies. In other words, as the students' passive knowledge of words increased, there was a related increase in their active vocabulary as well. However, the students' passive vocabularies

were almost always larger than their active vocabularies as a whole and at different frequency levels. That is to say, the gap between the learners' passive and active vocabularies increased at lower frequency levels, with a smaller number of low frequency words being used productively.

Vocabulary size also turned out to be highly correlated with both proficiency level and reading comprehension ability. Those who had a larger vocabulary repertoire were more proficient as well and were better able to read texts with comprehension. A comparison of the vocabulary size of High proficient and Low proficient students showed that in both groups, passive vocabulary was larger than active vocabulary. Moreover, the High proficient participants had a larger vocabulary repertoire and used a higher percentage of their passive vocabulary in production. The difference between the vocabulary size of the English majors and non-majors was also significant. The English majors' vocabulary size was larger than the non-majors', and they were better able to use the words communicatively.

References

August, D, Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research and Practice, 20* (1), 50-57. [online] <http://www.daneshyar.org/products.php?sh=&proid=12&serid=0&pid=&status=on&sort=db> [Sept. 2006]

Bachman, L. F. (2000). Modern language testing at the turn of the century: Assuring that what we count counts. *Language Testing, 17* (1), 1-42.

Beglar, D. (2000). Estimating vocabulary size. *JALT Testing & Evaluation SIG Newletter, 4(*1). 2-3.

Bogaards, P. (2000). Testing L2 vocabulary knowledge at a high level: The case of the Euralex French Tests. *Applied Linguistics, 21* (4), 490-516.

Cobb, T. (1999). Breadth and depth of lexical acquisition with hands-on concordancing. *Computer Assisted Language Learning, 12*(4), 345-360.

Educational Testing Service (1980). *TOEFL workbook.* Princeton, NJ: Author.

Elley, W. B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly, 24*, 174-187.

Goulden, R., Nation, P., & Read, J. (1990). *Applied Linguistics, 11*(4), 341-363.

Grabe, W. (1991). Current developments in second language reading research. *TESOL Quarterly, 25*, 375-406.

Greidanus, T., & Nienhuis, L. (2001). Testing the quality of word knowledge in a second language by means of word associations: Types of distractors and types of associations. *Modern Language Journal, 85*, 567-577.

Harlech Jones, B. (1983). ESL proficiency and a word frequency count. *ELT Journal, 37*(1). 62-70.

Hazenberg, S., & Hulstijn, J. H. (1996). Defining a minimal receptive second-language vocabulary for non-native university students: An empirical investigation. *Applied Linguistics, 17* (2), 145-163.

Henriksen, B (1999). Three dimensions of vocabulary development. *Studies in Second language acquisition, 21*, 303-317.

Henriksen, B. (2000). Theme: A word's a word. 12-13, [online] http://inet.dpb.dpu.dk/infodok/sprogforum/Espr3/henriksenb.html [SEPT. 2006]

Hermann, F. (2003). Differential effects of reading and memorization of paired associates on vocabulary acquisition in adult learners of English as a second language. *TESL-EJ: Teaching English as a second or foreign language, 7*(1), 1-16. [online] <http://www-writing.berkeley.edu/tesl-ej/ej25/a1.html> [Sept. 2006]

Hever, B., & Cederberg, S. (1997). The vocabulary tests are designed to estimate the approximate size of a person's English vocabulary. [online] <test bank – vocabulary.htm> [August 2002]

Hirsh, D. & Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language, 8*, 689-696.

Horst, Cobb & Meara (1998). Beyond a clockwork orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language, 11*, 207-223.

Hsueh-chao, M. H., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language, 13*(1), 403-430.<http://voc/ipsco/ac/dst/99/comp.html>. [September 2003]

Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language: A review. *Studies in Second Language Acquisition, 21*, 181-193.

Hulstijn, J. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research, 18*(3), 193-223.

Hunt, A. (1998). Laufer seminar. [online] **http://langue.hyper.chubu.ac.jp/jalt/pub/tlt/98/sep/laufer.html.** [August 2002]

Jafarpur, A. (2001). A comparative study of a C-Test and a cloze test. *Theoretische Grundlagen und praktische Anwendungen, 4, 21-41.*

Joe, A. (1998). What effect do text-based tasks promoting generation have on incidental vocabulary acquisition? *Applied Linguistics, 19*, 357-377.

Jones, F. R. (1995). Learning an alien lexicon: A teach yourself case study. *Second Language Research, 11*(2), 95-111.

Knight, S. (1994). Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. *Modern Language Journal*, 78(3), 285-299.

Laufer, B. (1986). Possible changes in attitude towards vocabulary acquisition research. *IRAL, XXIV*(1), 69-75.

Laufer, B. (1998). The development of passive and active vocabulary in a second language: same or different. *Applied Linguistics, 19*(2), 255-271.

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

Laufer, B. & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics, 16* (3), 307- 322.

Laufer, B., & Paribakht, T.S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning, 48*(3). 365-391.

Meara, P. (1980). Vocabulary acquisition: A neglected area of language learning. *Language Teaching and Linguistics: Abstracts, 13,* 221-246.

Meara, P. (1990). A note on passive vocabulary. *Second Language Research, 6*(2). 155-157.

Meara, P., & Buxton, P. (1987). An alternative to multiple-choice vocabulary tests. *Language testing, 4,* 142-151.

Milton, J., & Meara, P. (1995). How periods abroad affect vocabulary growth in a foreign language. *International Review of Applied Linguistics, 107,* 17-34.

Nassaji, H. (2004). The Relationship between depth of vocabulary knowledge and L2 learners' lexical inferencing strategy use and success. *The Canadian Modern Language Review, 61*(1), 107-134.

Nation, P., & Meara, P. (2002). Vocabulary. In N. Schmitt(Ed.), *An introduction to applied linguistics* (pp. 35-54). New York: Oxford University Press Inc.

Nation, P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt, & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagog*y (pp. 6-19). Cambridge: Cambridge University Press.

Paribakht, T. S. & Wesche, M. (1999). Reading and "incidental" L2 vocabulary acquisition: An introspective study of lexical inferencing. *Studies in Second Language Acquisition, 21,* 195-224.

Pulido, D. (2003). Modeling the role of second language proficiency and topic familiarity in second language incidental vocabulary acquisition through reading, *Language Learning, 53,* 233-284.

Qian, D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *The Canadian Modern Language Review, 56,* 282-308.

Qian, D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performace: An assessment perspective. *Language Learning, 52,* 513-536.

Read, J. (1988). Measuring the vocabulary knowledge of second language learners. *RELC Journal, 19*(2). 12-25.

Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing , 10* (3), 355-371.

Read, J., & Chapelle, C. A. (2001). A framework for second language vocabulary assessment. *Language Testing, 18* (1), 1-32.

Saville-Troike, M. (1984). What really matters in second language learning for academic achievement? TESOL Quarterly 18 (2), 199-219.

Schmitt, N. & Meara, P. (1997). Researching vocabulary through word knowledge framework: Word associations and verbal suffixes. *Studies in Second Language Acquisition, 19,* 17-36.

Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language learning, 48*, 281-317.

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the Vocabulary Levels Test. *Language Testing, 18* (1), 55-88.

Sutarsyah, C., Nation, P., & Kennedy, G (1994). How useful is EAP vocabulary for ESP? A corpus based case study. *RELC Journal, 25*(2), 34-50.

Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics, 22*, 217-254.

Verspoor, M. & Lowie, W. (2003). Making sense of polysemous words, *Language Learning*, 53, 547-586.

Waring, R. (1996). A comparison of the receptive and productive vocabulary sizes of some second language learners. [online] <***http://www1.harenet.ne.jp/~waring/papers/vocsize.html***> [August 2002]

Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition, 27,* 33-52.

Wesche, M., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: depth versus breadth. *Canadian Modern Language Review, 53*, 13-40.

Wesche, M., & Paribakht, T. S. (1999). Introduction: Incidental L2 vocabulary acquisition: Theory, current research and instructional implications. *Studies in Second Language Acquisition, 21*, 175-180.

Wode, H. (1999). Incidental vocabulary acquisition in the foreign language classroom. *Studies in Second Language Acquisition, 21*, 243-258.

Wolter, B. (2001) Comparing the L1 and L2 mental lexicon. *Studies in Second Language Acquisition, 23*, 41-69.

Zahar, R., Cobb, T. & Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *The Canadian Modern Language Review, 57*(4). [online]
<http://www.utpjournals.com/product/cmlr/574/574-Zahar.html> [Sept. 2006]

Zareva, A. (2005). Models of lexical knowledge assessment of second language learners of English at higher levels of language proficiency. *System, 33*, 547-562.

Zareva, A., Schwanenflugel, P.' & Nikolova, Y. (2005). Relationship between lexical competence and language proficiency: variable sensitivity. *Studies in Second Language Acquisition*, 27(4), 567-596.

Zimmerman, M. (1997). Do reading and interactive vocabulary instruction make a difference? An empirical study. *TESOL Quarterly, 31*(1), 121-140.

Mortaza Yamini, Ph.D., is an assistant professor at Shiraz University. He has taught at undergraduate and graduate levels and is now mostly offering graduate courses at the Department of Foreign Languages and Linguistics. His main interests are teaching and testing English as a foreign language, language learning strategies and reading comprehension. myamini@rose.ac.ir.

Maryam Golkar is a Ph.D. candidate at Shiraz University and a lecturer at Ferdowsi University in Mashad, Iran. Her main research interests are reading comprehension, vocabulary acquisition and language testing and teaching. Her address is:
Department of English Language and Literature, Univerisity of Mashad, Mashad, Iran
mgolkar@um.ac.ir