The Comparative Development of Vocabulary Breadth and Depth, and Academic Vocabulary for ESP/EAP Learners

Is'haaq Akbarian[*]
Assistant Professor, University of Qom, Qom, Iran

Abstract
This small-scale study investigates (a) whether academic vocabulary compares in development with vocabulary breadth and depth for Iranian ESP/EAP learners and, if the answer is positive, (b) whether this trend of development happens across proficiency levels. Fifty-seven graduate students served as the subjects who were also divided into high and low groups based on whether they had acquired the most frequent 2,000 words. Multiple regression analysis results show much shared variance between breadth and depth tests, and academic vocabulary test for the participants as a whole group. Therefore, as learners' vocabulary breadth and depth increase, so does their academic vocabulary. A similar finding is also observed for the high group and low group. However, the finding for the low group is contrary to our expectation. The results suggest *more* systematic vocabulary development for the high group, *less* for the participants as a whole group, and *least* in the low group. This investigation has some implications for language, and more particularly vocabulary, instruction for ESP/EAP purposes in Iran.

Keywords: Academic Vocabulary; Breadth and Depth of Vocabulary Knowledge; Language Proficiency; ESP/EAP Learners

Introduction

Different university disciplines require a high level of reading proficiency in English for the students to further succeed in their academic studies. Consequently, in contexts of learning English as a foreign language (EFL), e.g. Iran, university entrance examinations and English programs of different university disciplines are mainly concerned with reading proficiency. This naturally culminates in an emphasis on building a large vocabulary repertoire to cope with this requirement. Of course, researchers (Bernhardt & Kamil, 1995; Nation, 2001, 2006; Qian, 1999,

[*] *E-mail address*: ishaaq.akbarian@gmail.com, i-akbarian@qom.ac.ir
Correspondence address: Department of English Language and Literature, University of Qom, Iran; PO Box No. 37185-396, Qom, Iran

2002) have empirically shown that vocabulary is a good, if not the best, predictor of reading. To further show the significance of this issue, below we only refer to one very recent study that clearly confirms the strong prediction of vocabulary in reading comprehension. It suggests that the emphasis on vocabulary at these universities points to a promising direction in learning English for specific or academic (ESP/EAP) purposes.

In their study, Laufer and Ravenhorst-Kalovski (2010) explored the relationship between second language (L2) learners' vocabulary size, lexical text coverage that their vocabulary provides, and their reading comprehension. The researchers measured vocabulary size by Vocabulary Levels Test (henceforth VLT), lexical coverage by the newest version of Vocabulary Profile, and reading comprehension by a standardized national test. Based on their data, they suggested two thresholds: a minimal one, 4,000–5,000 word families, that results in 95 per cent coverage (including proper nouns) and an optimal one, knowledge of 8,000 word families, that provides a lexical text coverage of 98 per cent (including proper nouns). They also found that 64 per cent of the variance in the score on reading comprehension was accounted for by vocabulary. This finding "corroborates once more the earlier claims that vocabulary may be the major factor in reading comprehension." (p. 26). This study, which investigates the role of vocabulary in reading comprehension from two prevailing perspectives, i.e. the relationship between L2 vocabulary knowledge and success in reading comprehension, and vocabulary knowledge and lexical coverage (vocabulary threshold), offers a clear and tenable case for the significance of vocabulary in language education in general and in different university disciplines in particular.

Review of Literature

Research in the mere area of vocabulary knowledge is approached from different perspectives. One prolific line of research, among others, is concerned with assessing the two dimensions of vocabulary knowledge, i.e., breadth and depth of vocabulary knowledge. Meanwhile, the area of vocabulary in academic texts, or academic vocabulary, has drawn the attention of researchers, providing fresh insights and understanding of the nature of vocabulary knowledge (Xue & Nation, 1984; Coxhead, 2000; Nation & Coxhead, 2001). Breadth of vocabulary knowledge is concerned with the quantity of word knowledge. In other words, how many words does a language learner know? Depth of vocabulary knowledge, however, relates to the quality of word knowledge, or how well a learner knows a

word (Milton, 2009). As Nassaji (2004, p. 112) states, researchers have indicated "the complexity and multi-dimensionality of word knowledge and have suggested that knowing a word well should mean more than knowing its individual meanings in particular contexts." A word includes many aspects to be gradually learned by language learners, such as its pronunciation, spelling, register, stylistic and morphological features (Haastrup & Henriksen, 2000; Nation, 1990; Richards, 1976), and knowledge of the word's syntactic and semantic relationships with other words in the language, including collocational meanings and knowledge of antonymy, synonymy, and hyponymy (Chapelle, 1994; Henriksen, 1999; Read, 2000). Academic vocabulary, however, is a sub-area of breadth and, to some extent, depth of vocabulary knowledge. According to Coxhead (2007), academic vocabulary accounts for around 10 per cent of the words in an academic text such as a textbook. These words occur across a wide range of academic subject areas, such as Biology, History, Marketing, and Commercial Law. However, they do not occur in the first 2,000 words of English that occur most frequently. The first most frequent 2,000 words account for around 75 per cent to 80 per cent of a textbook, depending on the level of technical vocabulary in the book (Coxhead, 2007), so it appears that the first most frequent 2,000 words plus academic words give good return for learning.

The list of academic vocabulary, based on Coxhead (2000), includes 570 word families. Because 570 word families in Academic Word List (AWL) are a lot to learn at once, the list is divided into ten sub-lists that are arranged by frequency. Sub-list One contains the first most frequent 60 words in AWL, Sub-list Two contains the next most frequent 60 words, and so on. However, Sub-list Ten has only 30 words. The words in the lists were selected if they "occurred at least 100 times in an academic corpus of 3.5 million words of varied genres and in at least 15 of the 28 disciplines within the four broad subject groupings of the corpus: arts, commerce, law, and science" (Coxhead, 2000, p. 221). The list of academic words helps to distinguish English for academic purposes (EAP) from general English and "sets an agenda for focused language learning" (Hyland & Tse, 2007, p. 238). It is worth emphasizing here that the Academic Vocabulary Test (henceforth AVT) was originally part of VLT, but was taken out for the purpose of this study and used as a separate test.

Since the focus of the paper is on academic vocabulary and vocabulary breadth and depth, we will now survey studies that focus on these variables and the

interaction between them and, inevitably, a mere mention of some studies that involve reading comprehension and vocabulary. Therefore, our emphasis, here, is on vocabulary, and *not* on reading.

Research on non-native speakers' vocabulary size mainly focuses on what minimum number of words international students need to know for the demands of their studies. Sutarsyah, Nation, and Kennedy (1994) estimate that knowledge of 4,000–5,000 English words will be a prerequisite for understanding an undergraduate economics textbook. Hazenberg and Hulstijn (1996) argue that a non-native speaker of Dutch in the first year at a university in the Netherlands needs a vocabulary of 10,000 Dutch words to be able to deal with reading materials. Read (2000, p. 83) argues that non-native speakers need to recognize at least 95 per cent of the words in a text for efficient reading. Nation (1990) and Laufer (1992, 1997) argue that achieving at least the 3,000-word level is necessary to meet this target. Read (2000) states that the two studies he conducted in Indonesia show that first year learners fall short of this target. The aforementioned study by Laufer and Ravenhorst-Kalovski (2010) also corroborates the earlier findings, suggesting two thresholds of lexical coverage, i.e. minimal (95 per cent) and optimal (98 per cent) in "adequate" reading comprehension.

According to Boyle and Kirk (2006, p. 14), however, research on vocabulary size has almost all been undertaken "among multilingual groups on university campuses in English-speaking countries". One exception is the research into the vocabulary knowledge of a homogeneous national or linguistic group conducted by Nurweni and Read (1999) in Indonesia. They found that university students in an English-medium program knew 1,226 word families on average. Aside from this study, "no more recent research appears to exist" (Boyle & Kirk, 2006, p. 14).

Other studies in the area of size and depth look at vocabulary from different perspectives. For instance, with Nation's Range program, also called VocabProfile, Laufer and Nation (1995) tried to produce a Lexical Frequency Profile (LFP) of student compositions. They examined two sets of compositions, written by 65 students, merely based on the number of words that the students had used from the first most frequent 1,000 words, the second most frequent 1,000 words, words from University Word List (UWL, that is now called academic vocabulary), and a last group considered low frequency words or unknown words. The researchers intended to see to what extent such a profile would correlate with the students'

scores on Productive Vocabulary Levels Test (PVLT) with the reliability of 0.86, using KR-21. The results showed that the students with larger vocabulary used fewer high frequency words and more low frequency words in comparison with the students who used smaller vocabulary. There was a correlation of 0.70 and 0.60 between vocabulary scores from UWL section of PVLT and the number of academic words the students used in the two sets of compositions, thus accounting for 49 per cent and 36 per cent of the shared variance between UWL section of PVLT and the academic words used in the compositions.

Qian (1999) showed that the scores on VLT and Word Associates Test (henceforth WAT) were closely and significantly correlated at 0.78 for the 44 Korean-speaking participants and 0.82 for the Chinese-speaking ones. He concludes that breadth is as valuable as depth to vocabulary knowledge.

The studies surveyed above indicate that the researchers have merely attempted an investigation into the number of words known by EFL learners or how much variance vocabulary knowledge provides, and also the correlation between the breadth and depth of vocabulary knowledge or academic vocabulary. Though Hazenberg and Hulstijn's (1996) study is relevant to our research, it might not be compared, as expected, with other studies since, as Laufer and Ravenhorst-Kalovski (2010) argue, this research study was done "with the Dutch, not English vocabulary. Besides, the list of the threshold vocabulary was created on the basis of a dictionary. Usually dictionaries list lexical items, not word families. The [word] families figure would be considerably lower" (p. 19). The surveyed studies have not specifically addressed the issue of the development of these aspects comparatively. No particular study has been conducted to shed light on whether Iranian university students majoring in different disciplines first develop the most frequent 2,000 words prior to academic words, or the reverse, or whether these academic words compare with their depth of vocabulary knowledge. As a result, a study is warranted to tackle this question with ESP/EAP learners in an EFL situation, like Iran which is the context of our study.

Statement of the Problem

Coxhead (2007) and Schmitt, Schmitt, and Clapham (2001) observe that academic vocabulary do not occur in the first 2,000 words of English that occur most frequently. In other words, this list is deliberately designed not to include the first 2,000 words of English. The words in the list occur across the levels after the

second most frequent 1,000 words. Logically then, an ESP/EAP learner, following a systematic and sound English program, should first learn the first most frequent 2,000 words of English and then academic vocabulary, and not the reverse in the sense that the learner learns a lot of academic vocabulary before passing this threshold. A reverse pattern of development (that is, learning many academic vocabularies before passing the first most frequent 2,000 words of English) might reveal an indication of poor proficiency and, thus, a concern for the stakeholders. The motivation for this study, therefore, lies in the attempt to evidence this trend of development in vocabulary acquisition for Iranian ESP/EAP learners. More specifically, the following questions guide the current research:

1. Does academic vocabulary compare in development with vocabulary breadth and depth for Iranian ESP/EAP learners?
2. If the answer is positive, does such a development occur across proficiency levels?

Method

Participants

Fifty-seven Iranian graduate ESP/EAP learners, studying at the University of Qom, took part in the study (Table 1). These available participants were in the first year of graduate level. For the purpose of the study, the participants were once considered as a whole group and then divided into high and low proficient groups based on whether they had acquired the most frequent 2,000 words in VLT. Those acquiring the level were included in the high proficient group and those not acquiring the level were included in the low proficient one. The reason for adopting this strategy was the fact that the academic words appear after the most frequent 2,000 words in VLT, as discussed below.

Table 1
Profile of the Participants

Field of Study (Major)	Gender		Number of Students
	Male	Female	
Physics	---	7	7
Mathematics	14	14	28
Electronic commerce	11	11	22
Total	25	32	57

Instruments

To gather data, the following three tests were used: Vocabulary Levels Test, Academic Vocabulary Test, and Word Associates Test.

Vocabulary Levels Test (VLT): Version 2 of VLT, revised and validated by Schmitt et al. (2001), entailing a section on academic vocabulary, was used in this study. Each level of VLT contains 30 items. There are three definitions on the right and six words on the left. Participants must choose the right word that goes with each meaning, writing the number of that word next to its meaning. The maximum possible score is 90, with one point for each item at the three levels.

VLT a) includes levels of frequency and, as a result, difficulty, b) is a statistically reliable vocabulary measure (Read, 2000), c) is related to success in reading, writing, and general language proficiency as well as to academic achievement (Laufer, 1997), d) can provide efficient placement and admission in language teaching programs, e) easy to administer, and f) can be completed in a short time. For these reasons, VLT was considered as the basis for dividing the participants into two low and high groups in this study to compare the results across proficiency levels.

Akbarian (2008) performed Cronbach's Alpha on VLT for the four 2,000-, 3,000-, 5,000-, and 10,000-word frequency levels at 0.963, which is quite high. Based on an earlier study and observation, the 10,000-word frequency level was not included in this study for being beyond the language proficiency of the Iranian ESP/EAP learners. The following is an example of VLT.

Figure 1

A Sample of VLT

Participants must choose the right word that goes with each meaning. They must write the number of that word next to its meaning.

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

Academic Vocabulary Test (AVT): AVT is identical to VLT in format. Therefore, it is included in VLT. As stated earlier, AVT was originally part of VLT but was taken out for the purpose of this study and used as a separate test. The maximum possible score on AVT was 30.

Word Associates Test (WAT): WAT (see Read, 1993) measures synonymy, polysemy, and collocation. It entails 40 items, with each item containing one stimulus adjective, and two boxes, each containing four words. Among the four words in the left box, one to three words can be synonymous to one aspect of, or the whole meaning of, the stimulus word. Also, there can be one to three words that collocate with the stimulus word among the four words in the right box. There are always four correct answers in each item, thus effectively reducing the chances of guessing. In scoring, each word correctly chosen was awarded one point. The maximum possible score, therefore, was 160 for the 40 items. The following is an example:

acute

hidden	often	rich	sharp	angle	hearing	illness	stones

The reliability of the test (KR-20), as reported by Read (1993), was 0.92. The split-half reliability, as reported by Qian (2002), was 0.89. The scores obtained from this measure were treated as the variable or measure of depth of vocabulary knowledge while those obtained from VLT were treated as the measure of breadth of vocabulary knowledge, and those obtained form AVT were considered as the measure of academic vocabulary in the data analyses.

Procedures

The participants were notified of the research purpose of the study. All the participants willingly and voluntarily answered the paper-based tests in their class time. They did not sit in one simultaneous session. Each group took the tests at a different time. The researcher clearly described the way to respond to the tests and reviewed the examples provided in the instruction sheets. The time needed to complete VLT, including AVT, was 30 minutes. The time allocated to WAT was 30 minutes, too.

Research Design and Data Analysis

To answer the research questions, a series of stepwise multiple regression analyses were applied to the data with a view of finding the more related variable from between the two variables of breadth and depth of vocabulary knowledge to the variable of academic vocabulary. It is expected that as the score on the variable of depth increases, the score on academic vocabulary will increase as well. This trend is expected between breadth and academic vocabulary. However, this positive correlation is not expected between breadth and academic vocabulary for the participants in the low group since they have not acquired the most frequent 2,000 words after which the academic vocabulary occur. Therefore, the correlation is not expected to resemble that of the participants as a whole or in the high group. The analysis was also performed with regard to the proficiency level of the participants. Using SPSS version 16.0, the overall alpha significance level was preset at $p < .05$ for all the statistical analyses conducted.

Results

In answering the first research question, the performance of the participating students on the three levels of VLT was taken into account whereas in answering the second research question, the participants were divided into high and low proficient groups based on whether they passed the cut-off score for the most frequent 2,000 words in VLT. N. Schmitt (personal communication, May 9, 2008) suggests that the cutting point for acquiring any of the levels on VLT is 24. It means that if the participants answer 24 (80 per cent) items on any of the levels on VLT correctly, they acquire the level. Given this suggestion, the participants scoring below 24 on the 2000-word level of most frequent words on VLT fell in the low group (N = 43) whereas those scoring at 24 or above were included in the high group (N = 14).

Results for the Participants as One Group

The descriptive statistics in Table 2 provide a general profile of the collected data for the participants as one group. The table shows that the participants answered more than half of the items in AVT, more than one third of the items in VLT, and nearly one third of the items in WAT. Maximum possible score equals the number of the respective items in each test:

Table 2
Descriptive Statistics for the Participants as One Group on the Three Tests

	MPS	Mean	Std. Deviation	SM	SSD	N
AVT	30	16.2281	8.13506	54.0936	27.11687	57
VLT	90	33.9825	18.29568	37.7583	20.32853	57
WAT	160	51.8246	25.12620	32.2904	15.70387	57

MPS = Maximum possible score, SM = Standardized mean, SSD = Standardized standard deviation

The lowest observed score in VLT was 3 whereas the highest one was 86. In fact, only three participants obtained the cutting point (24) for the 3,000-word level, one level above the 2,000-word level that is the basis for dividing the participants into high and low groups. It is no good news for graduate level education. The mean for the three levels of VLT is almost 34 while the median is 33, indicating a normally distributed set of scores, given the kurtosis (0.217) and skewness (0.623). Which test, i.e. vocabulary breadth or vocabulary depth, is more related to academic vocabulary? To that end, a stepwise multiple regression analysis was conducted on the resultant data. The purpose in applying the multiple regression analysis is to determine the more related variable to the variable of academic vocabulary. In other words, this analysis is intended to determine which of the two tests, VLT or WAT, is the more related variable, or test, to AVT. Before determining the more related variable to academic vocabulary, the Pearson correlation analyses of the variables under study are presented. Table 3 shows the correlations and the resultant effect sizes between VLT, WAT, and AVT.

The results show that there is a very strong positive correlation coefficient between VLT and AVT (r = 0.873, p = 0.000) and between WAT and AVT (r = 0.868, p = 0.000). Though *not* a particular focus of this study, a further analysis showed that the correlation between VLT and AVT with respect to sex is a bit higher for males (r = 0.939, p = 0.000) and not noticeably different for females (r = 0.828, p = 0.000). Thus, the more words Iranian ESP/EAP learners know, the more academic vocabulary they are cognizant of. This point seems to be even higher for males in comparison with females. Also, the correlation coefficient between WAT and AVT with respect to sex does not differ noticeably for males (r = 0.838, p = 0.000) and for females (r = 0.894, p = 0.000). The above correlations and the resultant effect sizes explain for about 75 per cent of the common variance between the tests involved.

Table 3
Correlations for the Participants as One Group

	VLT	WAT
AVT	.873	.868
Effect Size (R^2)	.762	.753
Sig. (1-tailed)	.000	.000
N	57	57

Following the correlation analyses, a stepwise multiple regression analysis was applied to the data to determine the amount of prediction afforded by the two tests in performing on AVT. Multiple regression analysis will produce the same amount of shared variance irrespective of which variable is entered as the independent or the dependent. The analyses were first focused on the examination of the magnitude of R^2 changes.

Table 4
Model Summary for the Participants as One Group

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.873[a]	.762	.758	4.00193	.762	176.404	1	55	.00
2	.906[b]	.821	.814	3.50931	.058	17.525	1	54	.00

a. Predictors: (Constant), VLT
b. Predictors: (Constant), VLT, WAT

The variable of VLT was entered in the model as the first variable for which the R^2 (adjusted) was reported at 0.758 ($p < .001$), suggesting that VLT provided a significant portion of prediction on AVT. As the second row of the table shows, the variables of VLT and WAT were entered into the model and the R^2 (adjusted) was increased to 0.814 ($p < 0.001$), indicating that WAT could afford an additional and significant 5.8 per cent of the criterion variance over and above the variable of VLT.

In order to find the more related variable to the variable of AVT, the resultant coefficients were further inspected. As Table 5 shows, the coefficients imply that VLT accounts for more changes in AVT. In other words, for one standard deviation of change in VLT, there will be 0.873 of a standard deviation change in

AVT or vice versa. However, the second model shows that the inclusion of the participants' performance on VLT and WAT in one row points to the more contribution of VLT in comparison to the contribution of WAT. In other words, for one standard deviation of change in VLT, there will be 0.487 of a standard deviation change in AVT or vice versa whereas for one standard deviation of change in WAT, there will be 0.455 of a standard deviation change in the test on AVT or vice versa.

Table 5
Coefficients for the Participants as One Group

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.035	1.126		2.696	.009
	VLT	.388	.029	.873	13.282	.000
2	(Constant)	1.229	1.077		1.141	.259
	VLT	.217	.048	.487	4.481	.000
	WAT	.147	.035	.455	4.186	.000

a. Dependent Variable: AVT

Results for High Proficient Group

The descriptive statistics in Table 6 give a general profile of the data collected for high proficient group. Also, the respective correlation analyses and the resultant effect sizes of the variables under study are reported in Table 7. The table shows that there is a high positive correlation between AVT and VLT ($r = 0.744$, $p = 0.001$) and between WAT and AVT ($r = 0.718$, $p = 0.002$). The correlations account for a shared variance of 0.553 and 0.516 in AVT, respectively.

Table 6
Descriptive Statistics for High Proficient Group

	MPS	Mean	Std. Deviation	SM	SSD	N
AVT	30	26.0714	3.56186	86.9048	11.87285	14
VLT	90	57.4286	13.31693	63.8095	14.79659	14
WAT	160	83.4286	14.37259	52.1429	8.98287	14

MPS = Maximum possible score, SM = Standardized mean, SSD = Standardized standard deviation

Table 7
Correlations for High Proficient Group

	VLT	WAT
AVT	.744	.718
Effect Size (R^2)	.553	.516
Sig. (1-tailed)	.001	.002
N	14	14

Subsequent to performing the correlation analyses and strength of association, a stepwise multiple regression analysis was applied to the data with a view of finding the more related variable to AVT for high proficient language learners. As Table 8 shows, the analyses were first focused on the examination of the magnitude of R^2 changes.

According to Table 8, only the variable of VLT was entered in the model as the only variable for which the R^2 (adjusted) was reported at 0.516 ($p < .002$), suggesting that VLT provided a significant portion of prediction on AVT for the participants in high proficient group. WAT is not shown to be the more related variable to AVT at this level.

Table 8
Model Summary for High Proficient Group

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.744[a]	.553	.516	2.47848	.553	14.849	1	12	.002

a. Predictors: (Constant), VLT

Table 9
Coefficients for High Proficient Group

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.
		B	Std. Error	Beta		
1	(Constant)	14.648	3.038		4.822	.000
	VLT	.199	.052	.744	3.853	.002

a. Dependent Variable: AVT

Results for Low Proficient Group

The descriptive statistics, reported in Table 10, depict a general profile of the data for low proficient group. Again, their respective correlation analyses and the resultant effect sizes of the variables under study are reported in Table 11, showing that there is a strong positive correlation between VLT and AVT (r = 0.769, p = 0.000) and between WAT and AVT (r = 0.743, p = 0.000). The correlations for the tests account for a shared variance of 0.591 and 0.552 in AVT, respectively. The result between VLT and AVT is contrary to our expectation since the correlation should be low, not high, for the low group as stated in method section.

Table 10
Descriptive Statistics for Low Proficient Group

	MPS	Mean	Std. Deviation	SM	SSD	N
AVT	30	13.0233	6.44201	43.4109	21.47335	43
VLT	90	26.3488	12.18681	29.2765	13.54090	43
WAT	160	41.5349	18.34374	25.9593	11.46484	43

MPS = Maximum possible score, SM = Standardized mean, SSD = Standardized standard deviation

Table 11
Correlations for Low Proficient Group

	VLT	WAT
AVT	.769	.743
Effect Size (R^2)	.591	.552
Sig. (1-tailed)	.000	.000
N	43	43

After computing the correlations and determining the resultant effect sizes for low proficient group, a series of stepwise multiple regression analyses were performed on the results of this group. The variable of VLT was entered in the model as the first variable for which the R^2 (adjusted) was reported at 0.581 ($p <$ 0.000), suggesting that VLT provided a noticeable amount of prediction on AVT for low proficient group. Next, the variables of VLT and WAT were entered into the model and the R^2 (adjusted) was increased to 0.665 ($p < 0.002$), indicating that WAT could afford an additional and significant 9 per cent of the criterion variance over and above the variable of VLT (Table 12).

Table 12
Model Summary for Low Proficient Group

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	Df2	Sig. F Change
1	.769[a]	.591	.581	4.16792	.591	59.335	1	41	.000
2	.825[b]	.681	.665	3.72885	.090	11.224	1	40	.002

a. Predictors: (Constant), VLT
b. Predictors: (Constant), VLT, WAT

Table 13
Coefficients for Low Proficient Group

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.312	1.529		1.513	.138
	VLT	.406	.053	.769	7.703	.000
2	(Constant)	.229	1.503		.152	.880
	VLT	.259	.064	.490	4.020	.000
	WAT	.144	.043	.409	3.350	.002

a. Dependent Variable: AVT

The first model shows that the variable of VLT accounts for more changes of the test on AVT. In other words, for one standard deviation of change in VLT, there will be 0.769 of a standard deviation change in AVT or vice versa. However, the second model shows that the two variables of VLT and WAT were included as the more related variables, but VLT accounts for more change. In other words, for one standard deviation of change in VLT, there will be 0.490 of a standard deviation change in the test on AVT or vice versa whereas for one standard deviation of change in the variable of WAT, there will be 0.409 of a standard deviation change in the test on AVT or vice versa.

Discussion

This study has shown that almost 95 per cent of the members of the sample of Iranian graduate students responding to our tests have not reached a threshold of most frequent 3,000 words which, according to Nation (2006), provide at least 89

per cent lexical coverage of written discourse. Graduate students in Iran are selected through a nationwide entrance examination. Based on the results of this examination, the candidates gain admission into different Iranian universities through a norm-referenced procedure. The university under consideration is similar to many, if not to all of the Iranian universities with respect to the educational level of the accepted graduate students. Given that, if the study is replicated and similar results are found with larger samples, then the findings will be indicative of the poor quality of the English language learning programs for graduate students in EAP/ESP contexts, including the years past. This means that these students will not be able to understand even an undergraduate economics textbook for which Sutarsyah et al. (1994) predict that an estimated knowledge of 4,000–5000 English words would be needed to comprehend. Neither will the students reach the minimal threshold of 4,000–5,000 word frequency resulting in the coverage of 95 per cent (including proper nouns), as suggested by Laufer and Ravenhorst-Kalovski (2010), for adequate reading comprehension. These participants have not reached the level of 3,000 words and some of them only know at least about 86 per cent of the running words in a text with a knowledge of about the most frequent 2,000 words (Nation, 2006) which is not adequate for correct guessing of the meaning of the unknown words in a text (Nation, 1990; Laufer, 1992; Read, 2000). Familiarity with 95 per cent of the running words is required for correct guessing (Read, 2000). Still, some scholars put the figure even higher (Hu & Nation, 2000; Laufer & Ravenhorst-Kalovski, 2010). The finding in this study, if confirmed by further studies, implies that graduate students will be handicapped in dealing with original texts in English.

The strong correlation between AVT, VLT, and WAT indicates that they overlap one another to a large extent. This is not something new. But probably on that basis, we can go on to suggest teaching these constructs in combination. Means and standard deviations have also been standardized to give a better comparison between the results in these three tests, of course showing a better gain on AVT. This might probably be due to thematic education in English classes that expose the learners to many academic words.

One corollary of VLT is that if a respondent answers about 15 out of 30 items in any level, then it can be implied that he or she has approximately mastered about 500 words (50 per cent) of the most frequent 1,000 words of that level (Laufer & Nation, 1999). Our sample has scored a mean of 16 in AVT, i.e. the words that

appear after the most frequent 2,000 words and run through the upper levels. In other words, the participants as a whole group have responded to 54 per cent of the words in AVT without even acquiring the cutting point for the most frequent 3,000 words in English. Probably, "when instruction is based on thematic areas, it is possible for learners' lexical profiles to look distorted" (Milton, 2007, as cited in David, 2008, p. 178). The standardized mean in VLT for the participants as a whole is even lower (Table 2), indicating a great challenge for the students in compensating for the words that they should have learned to correspond to this percentage in AVT. It is expected that the higher the scores in VLT increase, the higher the scores in academic words should appear. The proportion of the increase in these tests should not be the same. How they want to compensate for the gap is a big concern.

The study has also shown that there is a high correlation between VLT and AVT on the one hand, and VLT and WAT ($r = 0.675$, $p = 0.004$) for high proficient group on the other hand. It indicates that the participants in the sample of high proficient group in this study, scoring higher in academic vocabulary and above the most frequent 2,000 words after which academic vocabulary start to appear, have acquired vocabulary more systematically both from the quantitative perspective and the qualitative perspective. This is not surprising since the more words one knows, the finer the network knowledge among the words one can acquire as a result. In other words, the quality accompanies the quantity of word knowledge for the three tests involved.

Similarly, a high correlation is also observed between VLT and AVT on the one hand, and VLT and WAT ($r = 0681$, $p = 0.000$) for the participants in the low proficient group on the other hand. This is rather surprising for the reason that the participants in this group have not yet acquired the most frequent 2,000 words in VLT after which academic words appear. What is surprising for the low group is the mean (13.02) for AVT. It means that these participants approximately know about 43.4 per cent of the words in academic vocabulary list. As Coxhead (2007) and Schmitt et al. (2001) state, academic words do not occur in the first 2000 words of high frequency. As stated earlier, the list is deliberately designed not to include the academic words. This percentage of academic words known prior to acquiring the most frequent 2,000 words seems to be rather odd. It is an indication of the assumption that the participants in the low group have not followed a systematic and well-planned approach to learning vocabulary.

There might be other interpretations for the failure in systematic growth or irregularity in their vocabulary development. One interpretation for the irregular vocabulary growth profile of the sample in our study, especially the participants in low proficient group, is the difficulty in assessing beginners or lower intermediates' vocabulary with the use of word-frequency criteria. "Their vocabulary has not yet extended sufficiently beyond a basic, high-frequency core" (Richards, Malvern, & Graham, 2008, p. 201).

Therefore, given that (a) the university undergraduate students in Iran compete in a nationwide entrance examination for graduate level that evaluates their knowledge of English mainly on the basis of content subjects in English and (b) the books offered to them in English instruction follows a theme-based or thematic approach, that are in fact selected from the prospective of content subject materials, these participants are as a consequence urged to learn academic words sooner without acquiring the frequency level of those words after which academic words start to appear. Our results, therefore, confirm that vocabulary breadth and depth of Iranian ESP/EAP learners do not compare in development with their academic vocabulary. This finding is also observed across proficiency levels much more clearly. This implies the lack of systematic and organized development or growth in the vocabulary component of the graduate learners' language proficiency. However, if this gap and discrepancy is bridged with systematic and principled approach to teaching vocabulary, then "the student's specific target corpus" will be a proper venue for gaining good results in learning academic vocabulary (Hyland & Tse, 2007, p. 251) and providing opportunities for the incidental and intentional learning of vocabulary in general. However, where exactly academic vocabulary should start in vocabulary instruction for ESP/EAP learners is a nice piece of investigation that awaits an in-depth evaluation.

Acknowledgements
I express my gratitude to my colleagues involved and the students participating in this investigation. I also acknowledge the financial support of the University of Qom (Registered No. 3245). An earlier version of this study was presented in International Language Assessment Conference in Armenia (ILACA), American University of Armenia, July 1–4, 2009.

Received 12 January 2010
Accepted 30 February 2010

References

Akbarian, I. (2008). *The role of vocabulary knowledge in predicting performance on reading comprehension item types*. Unpublished doctoral dissertation, University of Tehran, Tehran, Iran.

Bernhardt, E. B., & Kamil, M. L. (1995). Interpreting relationships between L1 and L2 reading: Consolidating the linguistic threshold and the linguistic interdependence hypotheses, *Applied Linguistics*, 16, 15–34.

Boyle, R., & Kirk, T. (2006). Investigating the relationship between students' vocabulary knowledge and their academic reading ability. *Proceedings of U.A.E. University Research Conference*, *7*, 12–16.
Retrieved from: *http://sra uaeu.ac.ae/Conference_7/Proceedings/pdf/UGRU/UGRU_12.pdf.*

Chapelle, C. (1994). Are C-tests valid measures for L2 vocabulary research?, *Second Language Research,* 10, 157–187.

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

Coxhead, A. (2007). *Five practical tips to improve writing skills using the Academic Word List. Oxford students' dictionary of English* (R10-R14). Oxford: Oxford University Press.

David, A. (2008). Vocabulary breadth in French L2 learners, *Language Learning Journal*, 36, 167–180.

Haastrup, K., & Henriksen, B. (2000). Vocabulary acquisition: Acquiring depth of knowledge through network building, *International Journal of Applied Linguistics,* 10, 221–240.

Hazenberg, S., & Hulstijn, J. (1996). Defining a minimal receptive second language vocabulary for non-native university students: An empirical investigation, *Applied Linguistics*, 17, 145–63.

Henriksen, B. (1999). Three dimensions of vocabulary development, *Studies in Second Language Acquisition,* 21, 303–317.

Hu, M., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension, *Reading in a Foreign Language*, 13, 403–430.

Hyland, K., & Tse, P. (2007). Is there an "Academic Vocabulary"?, *TESOL Quarterly*, 41, 235–253.

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. Arnaud & H. Benoit (Eds.), *Vocabulary and applied linguistics* (pp. 126–32). London: Macmillan.

Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 20–34). New York: Cambridge University Press.

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

Laufer, B., & Nation, I. S. P. (1999). A vocabulary size test of controlled productive ability, *Language Testing*, 16, 36–55.

Laufer, B., & Ravenhorst-Kalovski, G. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension, *Reading in a Foreign Language*, 22, 15–30.

Milton, J. (2009). *Measuring second language vocabulary acquisition*. Bristol: Multilingual Matters.

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, 107–134.

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening?, *The Canadian Modern Language*, 63, 59–82.

Nation, I. S. P., & Coxhead, A. (2001). The specialised vocabulary of English for academic purposes. In J. Flowerdew & M. Peacock (Eds.), *Research perspectives on English for academic purposes* (pp. 252–267). Cambridge: Cambridge University Press.

Nurweni, A., & Read, J. (1999). The English vocabulary knowledge of Indonesian university students, *English for Specific Purposes*, 18, 161–175.

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

Qian, D. D. (1999). Assessing the roles of depth and breath of vocabulary knowledge in reading comprehension, *The Canadian Modern Language Review,* 56, 283-307.

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

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.

Richards, B., Malvern, D., & Graham, S. (2008). Word frequency and trends in the development of French vocabulary in lower-intermediate students during Year 12 in English schools, *Language Learning Journal*, 36, 199–213.

Richards, J. C. (1976). The role of vocabulary teaching, *TESOL Quarterly,* 10, 77-89.

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

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

Xue, G., & Nation, I. S. P. (1984). A university word list, *Language Learning and Communication*, 3, 215–229.