Journal of Language and Translation
Volume 4, Number 1(7), (pp.41-54), 2014

Comparative Study of the Academic Vocabulary Content of Electronic Engineering Corpora, GE Materials and M.S. Entrance Examinations

Parviz Maftoon[1]; **Aris Hakhverdian**[2*]
[1] Department of Foreign languages, Islamic Azad University, Science and Research Branch, Tehran, Iran
[2] Department of Foreign languages, Islamic Azad University, Science and Research Branch, Tehran, Iran

Received: 10 February , 2013 Accepted: 15 May, 2014

Abstract
The importance of vocabulary learning has been underlined in the field of English for Academic Purposes (EAP) because non-English majors who require reading English texts in their fields of study have to expand their English vocabulary knowledge much more efficiently than ordinary ESL/EFL learners. Since academic vocabulary instruction in Iranian universities is realized through the use of General English (GE) textbooks, the present study pursued a dual purpose of evaluating Iranian engineering undergraduates' knowledge of academic vocabulary, as well as conducting a comparative corpus analysis of the academic vocabulary content of GE textbooks, electrical/electronic engineering Master's Entrance Examinations, and authentic electrical and electronic engineering texts. The participants were 520 engineering students from ten universities. The corpora under study comprised a total of 1,180,000 running words. The results revealed an inadequacy of academic vocabulary knowledge among Iranian engineering undergraduates, as well as the fact that GE textbooks for engineering students do not have sufficient coverage of academic vocabulary as compared to M.S. examinations and authentic electronic engineering corpora. There was also insufficient correspondence between academic words in M.A. entrance exams and the authentic texts.

Keywords: corpus analysis, academic vocabulary, EAP textbooks, M.S. entrance Examinations, electrical engineering corpus

INTRODUCTION

Several researchers have underlined the significance of vocabulary knowledge in the field of English for academic purposes (EAP) (Cobb, 1999; Hyland, 2006; Read, 2000; Sutrasyah, Nation, & Kennedy, 1994). The main reason for the emphasis is the crucial role vocabulary knowledge plays in reading academic texts and producing academic writing (Cobb, 2007). The need for more efficient vocabulary teaching in EAP, therefore, makes it necessary to take great care in the choice of vocabulary items to be included in an EAP syllabus.

Vocabulary knowledge is also known to have a direct relationship with reading ability. Anderson and Freebody (1981, cited in Nagy & Scott, 2006) mention the Instrumentalist Hypothesis, postulating a direct link between vocabulary knowledge and reading comprehension. Nation and Coady (1988) also summarize the connection between vocabulary knowledge and reading by saying that, generally, research indicates that not only is vocabulary knowledge important for reading, but also reading helps the acquisition of vocabulary knowledge. Reviewing research on the association between vocabulary and reading, they also point out that the number of difficult words has always been shown to be the most significant predictor of readability in general.

The importance of vocabulary knowledge is also acknowledged in the field of English for specific purposes (ESP). Qian (1999, 2000) has

*Corresponding Author's Email: arihak@yahoo.com

found that university students need to know a core vocabulary large enough to enable them to read and comprehend English texts in their corresponding fields of study. There has also been an extensive study on the need for EAP courses, which makes reference to the inadequacy of vocabulary knowledge among undergraduate students in Hong Kong's largest English-medium university (Evans & Green, 2007). Cohen et al. (1988, cited in Coxhead, 2000) state that such problems arise because most students are generally less familiar with such words than with technical vocabulary in their own fields of study. The reason is that academic words occur with lower frequency than general-service words do (Worthington & Nation, 1996; Xue& Nation, 1984). It can, therefore, be argued that vocabulary items which occur frequently in academic texts, but are less frequent than the 2000 most common words, can be categorized as academic vocabulary.

The importance of vocabulary knowledge and its role in facilitating reading comprehension, in general, and reading EAP/ESP texts, in particular, has prompted a number of researchers to compile vocabulary lists for pedagogic purposes. The earliest example is the General Service List of English Words (GSL) (West, 1953), a set of 2,000 headwords, each representing a word family. Xue and Nation (1984) compiled the University Word List (UWL), comprising 836 words, not included in the 2,000 words of the GSL, but common in academic texts. According to Nation (1990), the words on the UWL account for 8% of the vocabulary in a typical academic text.

However, the most recent attempt to compile a list of words with mostly academic use is The Academic Word List developed by Coxhead (2000). The list contains 570 word families not included in the most frequent 2,000 words of English. The basic idea behind the compilation of the list, according to Coxhead, is that it could serve as a basis for making sound decisions concerning the selection of vocabulary items worth focusing on during class time, as well as independent study activities.

Considering the use of word lists as a basis for vocabulary teaching, Honeyfield (1977) acknowledges the importance of frequency as a criterion for vocabulary selection, stating that if there is a relatively small number of words, known to be highly useful, it would be realistic to propose systematic selection and teaching of such vocabulary items. Much earlier, Richards (1974) had also referred to West's (1953) compilation of the GSL as an attempt to provide an objective basis for vocabulary selection.

Decisions concerning what words to teach can be said to have greater significance in EAP and ESP courses than in general English programs. One reason could be that students of disciplines other than English make an instrumental use of English, with less time to spend on the improvement of their knowledge of the language. Coxhead (2000) regards making principled decisions about choice of words to be taught as one of the most challenging aspects of vocabulary learning and teaching in EAP.

The present study, therefore, evaluates the academic vocabulary knowledge of Iranian engineering students, and compares the academic vocabulary coverage of authentic texts in electrical engineering, GE textbooks and the English section of electrical and electronic engineering M.S. entrance examinations.

Background

A number of researchers have investigated the coverage of theword families of Academic Word List (AWL)Coxhead (2000) in different corpora. Dutton (2006) used the AWL to measure the academic words coverage of the sample essays in the academic version of the IELTS and found that 7.17 % of the words in his corpus of 24,819 tokens came from the AWL, much lower than the 10% figure reported for Coxhead's (2000) academic corpus. Dutton (2006) also reports that his corpus contains only 345 out of the 570 AWL word families, what he attributes to the smaller size of the corpus comprising samples of IELTS Writing Task 2 and the range of topics they cover.

Chen and Ge (2007) conducted a study on the frequency and coverage of the AWL word families in a corpus of 50 medical research articles comprising 190425 running words (tokens). They found that the AWL words accounted for 10.70% of the words in their corpus and that there were only slight differences in the distribution of academic vocabulary in the different sections of the articles (Abstract, Introduction, Materials and Methods, Results, and Discussion). The authors concluded that academic vocabulary, with a high text coverage and even dispersion throughout medical research articles, is an important set of vocabulary items for students of medicine.

In a more recent study, Martinez, Beck, and

Panza (2009) measured the coverage of the AWL in agriculture research articles and found that only 92 word families appeared in their corpus of 826,401 words. Moreover, a qualitative study showed that certain words had genre/field-specific meanings and that many words of general use had a different meaning in academic texts in agriculture. The authors conclude that there is a need to produce field-specific academic word lists, which, in their opinion, should include all frequent academic lexical items necessary for the expression of the rhetoric of the specific research area.

Hsu (2009) used the AWL to compare the lexical content of a corpus of 20 (GE) textbooks (355,958 tokens) against a collection of 9 business-core textbooks (1,087,723 tokens). The results of her study showed that 49 to 415 words from the AWL have been used in the GE reading textbooks as opposed to 421 to 537 in the business textbooks. The business textbooks, on the other hand, use a smaller range of vocabulary items (4,000-5,000), compared to the GE textbooks (3,000-13,000). Hsu expresses hope that the indices in her study would help teachers and instructional materials writers to pay attention to continuity of vocabulary presentation in curriculum design while deciding on GE and business English materials.

As for the pedagogical use of the AWL, Coxhead and Nation (2001) suggest that the list may be used to teach vocabulary, following the acquisition of the 2,000 words of general usefulness and preceding the instruction of field-specific words. Providing further evidence for the need to teach academic vocabulary, Li Siu-leung and Pemberton (1994) state that since the 80s, in addition to the idea of reading as a "psycholinguistic guessing game," there has been an interactive view maintaining the importance of bottom up processes, such as focusing on vocabulary in reading comprehension. They provide empirical support for the interactive approach, stating that students mention vocabulary deficiency as one of their major problems in text comprehension. Li Siu-leung and Pemberton refer to Nation (1990), where he claims that there is a lexical threshold of 3,000 word families necessary for comprehension to take place. Otherwise, according to Nation, too much brain capacity will be taken up, and consequently, top down processes will not have a chance to take place.

Trimble (1985, cited in Li Siu-leung& Pemberton, 1994) also claims that neither the general vocabulary, nor technical words cause many problems; but rather it is the middle-frequency context-independent words that are problematic in reading comprehension. Coady (1993), on the other hand, maintains that extensive exposure alone is not enough; rather, there should be a systematic instruction of these middle-frequency academic vocabulary items. He concludes that one of the objectives of EAP/ESP courses should be helping students acquire academic words and bringing the level of their lexical repertoire up to the level necessary for good comprehension.

In line with the rationale mentioned above for academic vocabulary instruction, and following the preparation of academic vocabulary lists, such as the AWL, there have been various attempts to put the AWL to practical use. Some of these initiatives are in the form of software and Internet sites for learners and teachers. One such program, which can be bought on the Internet, is Garry's Vocabulary Teacher, available at http://www.cict.co.uk/software/gvd. The software uses various vocabulary activities andthe AWL as a basis for concordancing. Another example is a site by the very name of Academic Vocabulary which can be accessed at www.nottingham. ac.uk /~alzsh3/acvocab/. The site allows users to locate academic vocabulary in texts of their choice or make gaps in texts to prepare vocabulary exercises. The site also makes it possible to access concordancing tools to see academic vocabulary used in authentic contexts. The AWL has also been used by Cobb (2003) in providing vocabulary lists for his site,The Compleat Lexical Tutor (http://www.lextutor.ca/ vp/eng/), where learners can use the AWL, along with lists of most frequent 1,000 and 2,000 words in English, to assess the vocabulary levels of their own writings.

In addition to the above tools available on the Net, books have also been published which provide various types of vocabulary and reading activities centered around the AWL. Schmitt and Schmitt (2005, 2011) use the list as a starting point and through various exercises provide, at least, four encounters with each of the 570 headwords of the AWL. In another book, McCarthy and O'Dell (2008) include different sections that introduce key verbs and quantifying expressions and show how words typically combine with each other in academic English. The book also provides examples of using academic vocabulary in various academic situations to express notions and perform functions.

Null Hypotheses

Considering what was said above about establishing the need to teach academic vocabulary and the correspondence between academic word coverage of GE materials, M.S. entrance English examinations and subject related electrical engineering corpora, the following null hypotheses were formulated:

1-Iranian engineering undergraduates do not possess an adequate knowledge of
academic vocabulary.

2- There is no statistically significant difference between the relative percentage of

AWL words in Iranian GE textbooks and in academic texts.

3- There is no statistically significant difference between the relative percentage of

AWL words in Iranian electronics engineering M.S. entrance examinations and
in Iranian GE textbooks.

4- There is no statistically significant difference between the relative percentage of

AWL words in Iranian electronics engineering M.S. entrance examinations
and in electrical engineering corpora.

5- There is no statistically significant difference between the relative percentage of

AWL words in Iranian GE textbooks and in electrical engineering corpora.

METHOD

Participants

The participants in the present study were engineering undergraduates, who took the Vocabulary Levels Test (VLT) (Schmitt, Schmitt, &Clapham, 2001) for the evaluation of their academic vocabulary knowledge, and consisted of a total of 520 (315 male, 205 female) students of computer (82), electronics (81), agriculture (27), software and process (55), hardware (52), mechanics(67),textile (105) andchemistry (51)from 10 different (3 state and 7 Azad) universities.

Instruments

The Vocabulary Levels Test *(VLT).*In order to evaluate the academic vocabulary knowledge of engineering undergraduates, the researchers used the Vocabulary Levels Test (VLT), which is the second version of the test revised and validated by Schmitt, Schmitt and Clapham (2001). Originally developed by Nation (1983) and later revised by him in 1990, the test provides an estimate of vocabulary size at 2000, 3000, 5000, and 10000 frequency levels, meanwhile giving an estimate of the test-takers' vocabulary size. The test consists of 10 clusters at each level with six words and three definitions to be matched. In this way, test-takers need to read the six words, as well as the three definitions in each cluster and then match the three definitions with the three correct words. In the same way, test-takers will answer the other clusters.When one level is answered, they can proceed to the next level. On the whole, there are 75 nouns, 50 verbs, and 25 adjectives assessed in the test. The items are decontextualized so that no clues are provided to their meanings. Schmitt et al. (2001) established the item discrimination, item facility, reliability, and validity of the test.

*The softwareRange32***.**In order to determine the proportion of academic and non-academic words in the three corpora of GE textbooks, M.S. entrance examinations and electrical engineering corpora, the researchers made use of the software Range32 (Heatley, Nation, &Coxhead, 2002), developed at Victoria University of Wellington. The software allows up to 32 text files to be compared for the frequency and range of occurrence of the words on the most common 2,000 and 3,000 words, as well as the words in the AWL.

Corpora

The researchers built three corpora for the purpose of the study. These were constructed using textbooks and examinations available in the market, as well as eBooks and web pages.

*The General English textbooks corpus.*In order to create the corpus of GEtextbooks, the researchers first selected the books which had the largest number of reprints and complemented the list with the textbooks used in the universities (See Appendix A fora list of the books used to build the corpus).

Examinations Corpus. The following M.S. entrance examinations were obtained for the construction of the engineering M.S. entrance examinations corpus: (1) Islamic Azad University - Electronics Engineering (2004, 2007, 2008, 2009) (2) state universities - Electrical Engineering (2003, 2004, 2005, 2006, 2007, 2008).

It is worth mentioning here that the Azad University entrance examinations and those of the state universities differed in the fact that the former incorporated much shorter English sections, consisting only of twenty-item multiple

choice tests of vocabulary. The state university entrance examinations, however, were made up of three sections of vocabulary, grammar, and reading comprehension.

In order to obtain more data for the entrance examinations corpus, the researchers attempted to locate the sources from which the reading comprehension texts in the state university examinations were extracted. The following are some of the sites and articles from which the passages were taken. It must be added that most of the passages were authentic and had been placed in the test booklets with little alteration (See Appendix B for the addresses of the sites).

Theelectrical engineering corpus. For the pupose of building the corpus, consultations were made with a number of university teachers in the field regarding the type of texts students of electronics needed to read as reference or complementary materials. On the basis of the teachers' advice, 11 Webpages and documents, nine articles, eight books, two theses, and one glossary were selected for building the corpus.

Procedure

Academic Vocabulary Knowledge Evaluation. To assess the academic vocabulary knowledge of Iranian engineering undergraduates, the researchers used the Vocabulary Levels Test (Schmitt et al., 2001). The students were 520 in number, and from 10 universities.

The researchers administered the tests toward the end of the EAP courses offered in the fall semester of 2010, during regular ninety-minute class time. The tests were given in the presence of the EAP teachers, who explained to their students that the result of the test was irrelevant to the final grades of the course, but it would help the teacher program for better vocabulary instruction during the rest of the semester.

Corpus Analysis. There were three corpora used in the present research. The researchers built these corpuses using the Web, as well as GE tests and textbooks available in the market.

The general English textbooks corpus. In order to create the GE corpus, the researchers first selected the books which had the largest number of reprints and complemented the list with the textbooks used in the universities where VLT was administered.

The books were scanned and saved in plain text format. The next step was cleaning the texts, which involved removing numbers, abbreviations, and proper nouns from the texts, as well as correcting misspellings originally in the books or resulting from the process of scanning. In order to do this, the program frequency33 (Heatley et al., 2002) was used to provide a frequency list of the words used in the books. To simplify the process, the researchers pooled the text files of all the books into one large file, and the frequency list was obtained.

The reason for constructing the frequency list was to help the researchers clean the corpus because most of the names, proper nouns, and misspellings were found among the low frequency items on the list. Table 1 shows part of the frequency list.

As can be seen in Table 1, items 1, 2, 12 and 13 seem to be made up of different words stuck together due to the process of scanning and file conversion. In such cases, the items had to be located in the text and rectified by adding spaces between the separate vocabulary items.

Item 5 appears to be the phonetic representation of the word European; included in the glossaries preceding reading passages in some of the textbooks. Such items were also located and deleted from the corpus. Item 7 seems to be an abbreviation and had to be deleted after being verified as such. Finally, items 11 and 16 are examples ofproper nouns which were deleted from the corpus. The corpus amounted to 252,974 running words after cleaning.

Table 1
GE Books Word Frequency Sample

WORD	FREQUENCY
1- YEARIN	2
2-YEARONWITHDAYS	2
3-YELLING	2
4-YELLOWISH-RED	2
5-YER-O-PI-AN	2
6-YIELDS	2
7-YKVS	2
8-YOGA	2
9-YOGHURT	2
10-YOGURT	2
11-YORKSHIRE	2
12-YOUDID	2
13-YOUMAY	2
14-YOUORDERED	2
15-YOURSELVES	2
16-YUAN	2

The M.S. Examinations Corpus.Similar to the procedure followed for the GE textbooks, the English sections of the test booklets were scanned and saved in plain text format. The materials from the internet sites and articles, however, were relatively harder to save in text format because many were pdf articles requiring passwords, which had to be converted to word or text formats. The main problem was with the pdf files that contained pictures and graphs. The researchers obtained the software ABBY FineReader 10 to convert the pdf articles into word and saved them in text format. The size of the corpus comprising M.S. examinations and their complete sources thus reached 77,053 running words after cleaning.

The electrical engineering corpora.In order to construct the corpus,the researchers consulted with university teachers in the field concerning which texts students of electronics needed to read as complementary materials.

The electrical engineering corpora was also cleaned and subjected to word frequency analysis. The main difference of this corpus with the previous corpora was that in this case there were a larger number of words, abbreviations, and symbols that had to be dealt with. The final count of the tokens in the cleaned corpus reached 851,063.

Analysis of Corpora for *AWL Coverage***.** In order to determine the proportion of academic and non-academic words in the three corpora of GE textbooks, M.S. entrance examinations and electrical engineering corpus, the researchers, as mentioned before, made use of the software Range32.Table 2 is an example of the output obtained from the Program, presenting a vocabulary profile of the GE textbooks corpus.

The word lists 1, 2, 3, in the first column of Table 2 correspond to the 2,000, 3,000, and academic vocabulary lists, respectively. The fourth box in the first column (not on the list) refers to words which are not on any of the above lists, these may be less frequent words, proper names, or misspellings. The purpose for cleaning the texts, in fact, was to make sure that the number of the words in this group would not falsely grow, and thus provide a distorted proportion of academic vocabulary to other words in the corpus. The third column represents the percentage of types, where all repeated words are considered as an instance of one type. The fourth column shows the percentage of word families, which contain a base word, its inflected forms and transparent derivations.

Table 2
Vocabulary Profiling of EAP Textbooks Corpus

WORD LIST	TOKENS/%	TYPES/%	FAMILIES
1	198654/78.53	3038/27.11	970
2	20463/ 8.09	2049/18.29	879
3	11959/ 4.73	1301/11.61	487
not on the lists	21898/ 8.66	4817/42.99	?????

RESULTS

Testing Hypothesis I

Table 3 shows the descriptive statistics of the academic section of the test administered to 520 engineering students.The data in the table indicates that the mean score obtained by the participants is far below 24, the cutoff score for the level. However, to be able to generalize the result to the population of Iranian engineering students, the researchers had to first examine the normality of the distribution of the scores and then determine if the results can be generalized.

Even a visual inspection of the histogram in Figure 1 clearly shows that the distribution of the scores is positively skewed and departs from normal.Therefore, it was necessary to use a non-parametric test to show that the difference between the mean of the population and the cutoff point of 24 is significant.

Table 3
Descriptive Statistics for Academic Vocabulary

	N	Minimum	Maximum	Mean		Std. Deviation	Variance
	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic
Engineering VLT Academic Level	520	.00	30.00	8.5577	.25588	5.83496	34.047
valid N (listwise)	520						

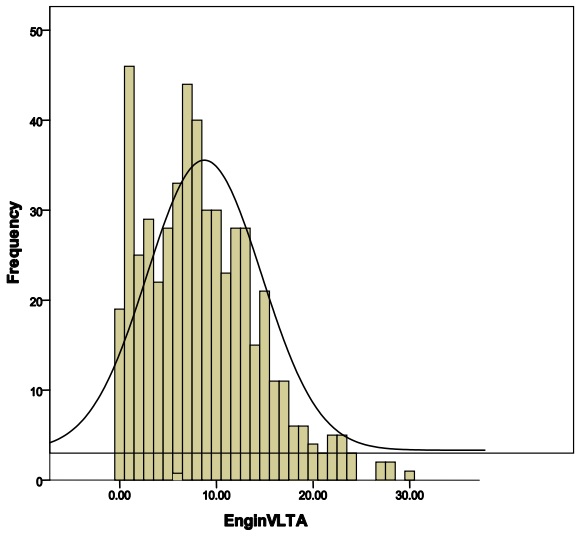


Figure 1VLT Results for Engineering Students

For this purpose, the following set of hypotheses was formed:

H0 : $\mu = 24$

H1 : $\mu \neq 24$

In order to test the above hypotheses, a new distribution of (X-24) was formed and subjected to a runs test to test the following set of hypotheses:

H0 : $\mu = 0$

H1 : $\mu \neq 0$

The results of the runs test, in Table 4, show that the mean of the population of (X-24) is significantly different from zero ($p < 0.05$). The negative value of Z further shows that the mean in the population is lower than zero. Therefore, the first hypothesis was maintained with a high degree of confidence. It could be concluded that Iranian engineering students do lack an adequate knowledge of academic vocabulary.

It is also revealing to know if there is any correlation between the students' level of academic vocabulary knowledge and their scores on the 2,000- word level of the (VLT). The strength or weakness of such a correlation would lend credit to, or refute, claims concerning the necessity of instruction in academic vocabulary regardless of the student's knowledge of the 2,000 most common words in English.

For this purpose, the researchers conducted the Spearman's Rho correlation test on the results of the 2,000-word and academic words sections of the VLT test. Table 5presents the result of the test and shows a moderate correlation between the participants' knowledge of themost common 2,000 words in English and the participants' academic vocabulary knowledge, $r(520) = .516$, $p < .01$. The results could be considered as an indication of the probability that even if the participants had an adequate knowledge of the first 2,000 words in English, theydo still need instruction in academic vocabulary.

Table 4
VLT Runs Test

Test Values	.0000
Total Cases	520
Number of Runs	13
Z	-5.609
Asymp. Sig. (2-tailed)	.000

Table 5
Correlation Between 2000-word and Academic Level

Spearman's rho	Engineering VLT 2000	Correlation Coefficient	1.000	.516**
		Sig. (2-tailed)	.	.000
		N	520	520
	Engineering VLT academic	Correlation Coefficient	.516**	1.000
		Sig. (2-tailed)	.000	.
		N	520	520

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

Testing Hypothesis II

The second hypothesis in the present research concerned the correspondence between the percentage of academic words found in a corpus of GEcourse books in Iran, on the one hand, and the (AWL) coverage found in academic texts.

Firstly, the researchers used the program RANGE to study the frequency of the occur rence of academic words in the corpus of eight GE textbooks. As mentioned earlier, the program determines the frequency of the first 1,000 most common words in English, the second 1,000 most common words, and the words, from the AWL in a given text. The results obtained are presented in Table 6.

As Table 6shows, the relative percentage of academic words in the EAP corpus is 4.73,

which is much lower than the approximate coverage of 10 % reported for AWL in different studies (Chen &Ge, 2007; Coxhead, 2000; Coxhead& Nation, 2001). To verify the statistical significance of the observed difference of coverage, the researchers performed a binomial test of significance, setting the cut point at 10%.

Table 6
AWL Coverage of EAP Corpus

WORD LIST	TOKENS/%	TYPES/%	FAMILIES
First 1,000	198654/8.53	3038/27.11	970
Second 1,000	20463/8.09	2049/18.29	879
Academic	11959/4.73	1301/11.61	487

Table 7 clearly indicates the significance of the difference between the AWL percentages found in the GE books and the 10 % observed in various academic texts, $p < .05$. Hence, the second hypothesis is rejected.

Table 7
Significance of AWL Percentage Difference

		Category	N	Observed Prop.	Test Prop.	Asymp. Sig. (2-tailed)
percentage	Group 1	<= 10	11959	.05	.50	.000[a]
	Group 2	> 10	241015	.95		
	Total		252974	1.00		

a. Based on Z Approximation.

Testing Hypothesis III

The third hypothesisconcerned the correspondence between the AWL coverage of the English section of electrical engineering M.S. entrance examinations and that of the GE textbooks corpus.

In order to test the hypothesis, the frequency of academic words in the M.S. examinations corpus was established using the program RANGE; the results are shown in Table 8

Table 8
AWL Frequency in M.S. Entrance Examinations Corpus

WORD LIST	TOKENS/%	TYPES/%	FAMILIES
First 1000	52959/68.73	2172/30.12	864
Second 1000	5194/ 6.74	926/12.84	507
Academic	9280/12.04	1352/18.75	526
Not in the lists	9620/12.48/.	2760/38.28	?????

The third row of Table 8 shows the frequency of the academic vocabulary as being 9280, which is 12.4 % of the total number of the tokens in the corpus. In the next step, the researchers used the Chi Square test of independence to test the statistical significance of the difference between the proportions of academic words in the two corpora of GE textbooks and M.S. entrance examinations. Tables 10 and 11 below show the results.

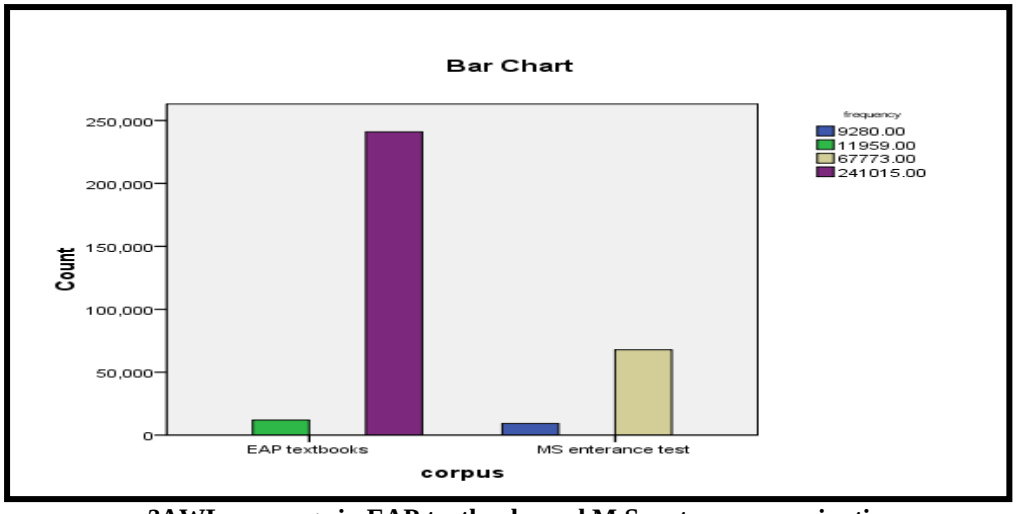
The columns under the frequency section of Table9 show, in a sequence, the frequency of AWLandnon-AWL items in the two corpora being investigated. Table 10, in addition, shows that the difference in the proportion of frequencies is significant$\chi^2(3, N = 330027) = 330027.000$, $p = .000$. It can also be observed in Figure 2 that despite its much larger size, the GE textbooks Corpus has only slightly more academic words.

Table 9
AWL Corpus/Frequency Cross tabulation

		Count				
		frequency				Total
		M.S. Exams AWL	M.S. Exams Non-AWL	Textbooks AWL	Textbooks Non-AWL	
corpus	EAP textbooks	0	0	11959	241015	252974
	M.S. entrance examinations	9280	67773	0	0	77053
Total		9280	67773	11959	241015	330027

Table 10
Chi-Square Test of Independence for Textbooks and Entrance Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	330027.000[a]	3	.000
Likelihood Ratio	358700.443	3	.000
Linear-by-Linear Association	241084.728	1	.000
N of Valid Cases	330027		



***2*AWL coverage in EAP textbooks and M.S. entrance examination**

Considering the above, it is possible to reject the null hypothesis of no significant difference between the AWL coverage of EAP textbooks and M.S. entrance examinations with a 95% degree of confidence.

Testing hypothesis IV

In order to test the hypothesis of no significant difference between the relative percentage of AWL words in Iranian electrical engineering M.S. entrance English proficiency examinations and in electrical engineering corpora,the relative frequency of academic words in electrical engineering corpora was determined and was compared to that of M.S. entrance examinations corpus using a chi square test of independence. Tables 11, 12, and 13 show the results of the above procedures, respectively.

Table 11
AWL Frequency in electrical engineering corpora

WORD LIST	TOKENS/%	TYPES/%	FAMILIES
First 1000	586305/68.89	2764/20.77	925
Second 1000	56391/ 6.63	1595/11.99	551
Academic	89071/10.47	1914/14.39	551
Not in the lists	119296/14.02	7032/52.85	?????

Table 12
Corpus/ Frequency Crosstabulation

		Count				
		frequency				Total
		M.S.Exa M.S. AWL	M.S.ExaM .S.Non-AWL	Engineering corpora AWL	Engineering Corpora AWL	
corpus	M.S. entrance exams	9280	67773	0	0	77053
	Engineering Corpora	0	0	89071	761992	851063

Table 13
Chi-Square Test of Independence for Entrance Examinations and Electrical Engineering Corpora

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	928116.000[a]	3	.000
Likelihood Ratio	531042.164	3	.000
Linear-by-Linear Association	406123.185	1	.000
N of Valid Cases	928116		

The results of the Chi Square test, shown in Table 14, also reveal a significant difference between the AWL coverage of M.S. entrance examinations and electrical engineering corpora, $\chi^2(3, N = 928116) = 928116.$, $p = .000$. It can be concluded, therefore, that the AWL coverage of M.S. entrance examinations and electronics and electrical engineering corpora are significantly different, and hence the fourth hypothesis is rejected with 95% level of confidence.

Testing Hypothesis V

The last hypothesis formed in the corpus analysis section of the present research concerns a comparison of the AWL coverage of GE textbooks and that of electrical engineering corpora.

Despite the obviously large difference between the AWL coverage of the two corpora (10.47% for the corpora and 4.73% for EAP textbooks corpus), a Chi Square test of independence was performed in order to demonstrate the statistical significance of the difference. Table 14 shows the frequencies of AWL and non-AWL words in the two corpora. Table 15demonstrates that the difference in the coverage is significant $\chi^2(3, N = 1104037) = 1104037$, $p = .000$. Therefore, it is possible to reject the null hypothesis of no significant difference between the AWL coverage of EAP materials and electrical engineering corpora $p < .05$.

Table 14
EAP Textbooks and Engineering Corpora/frequency Crosstabulation

		frequency				Total
		Text-books AWL	Textbooks Non-AWL	Engineering corpora AWL	Engineering corpora Non-AWL	
corpus	EAP textbooks	11959	0	241015	0	252974
	Engineering	0	89071	0	761992	851063
Total		11959	89071	241015	761992	1104037

The result of the Chi Square test demonstrates that AWL coverage of the three corpora are significantly different from each other, however, GE textbooks provide far less exposure to academic vocabulary than what would be needed to tackle M.S. entrance examinations or electrical engineering corpora.

.

Table 15
Chi-Square Test for EAP Textbooks and Electrical Engineering Corpora AWL Coverage

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1104037.000	3	.000
Likelihood Ratio	1188450.714	3	.000
Linear-by-Linear Association	585903.322	1	.000
N of Valid Cases	1104037		

DISCUSSION

The results of this study showed that Iranian engineering undergraduates scored an average of 8.55 on the VLT, which was significantly lower than the cutoff point of 24 set by the authors of the test. It must be mentioned that the total mean score was not the result of extremely low marks received by the students of only a number of the universities; the students of all the universities alike had scored below the cutoff point.

The findings of the present study are also in keeping with other investigations of academic vocabulary in a university context (Fan, 2000; Read, 2000; Santos 2004), where undergraduates are found to lag behind in their knowledge of academic vocabulary. In the absence of qualitative studies on the existence of any systematic decision making processes for vocabulary instruction in Iranian universities, it could be conjectured that the Iranian undergraduates' insufficient knowledge of academic vocabulary might have its causes in the lack of attention toward teaching such vocabulary items duringGEcourses or simply that academic vocabulary items are less frequent in non-academic texts. Academic vocabulary, therefore, have to receive special attention in EAP materials development.

Corpus analysis of EAP materials also revealed that the corpus of GE textbooks has a significantly lower percentage of AWL coverage (4.73%) than the required 10 %. It may, however, be argued that GE textbooks cannot, and should not, be expected to provide the same coverage of AWL as does an academic textmainly because GE materials address readers of a much lower English proficiency level. Despite the logic in the above statement, one should not forget that non-English undergraduates have little contact with academic English outside their studies and EAP courses are their sole opportunity to prepare for tackling non-technical vocabulary in ESP materials and authentic texts in English. Hence, providing exposure to academic vocabulary in form of tasks and activities seems to be necessary to ensure a minimum receptive knowledge of academic words.

The results of the study also clearly indicate the large gap between the academic words coverage of GE textbooks (4.73%) and electrical engineering M.S. entrance examinations (12.04). The possible cause of the mismatch could be little concern by the test developers for the content validity of the examination, as far as vocabulary, in general, and academic vocabulary, in particular, are concerned. In other words, test developers must ensure that M.S. examinations do not require students to exhibit a greater knowledge of academic vocabulary than their undergraduate courses have provided them with. There might, of course, be the belief among the administrators of M.S. programs that candidates who have acquired adequate proficiency in English by their own efforts are more eligible to enter the program.

The researchers in this study have not been able to locate any previous research concerning the AWL coverage of any tests; most studies concentrate on the AWL coverage of different corpora of academic disciplines and have aimed to investigate the validity of the AWL as a useful list of academic words (Cobb & Horst, 2002; Chen &Ge, 2007; Sutarsyah et al., 1994; Vongpumivitch et al., 2009).

A comparison of the figures, in Tables 9 and 12 alsoindicates the large gap between the academic content of GE textbooks on the one hand, and that of the M.S. examinations and electrical engineering corpora, on the other. The reasons for the difference can be sought in the possible efforts of test makers to select

authentic passages and make the comprehension of these passages a partial criterion for eligibility to enter the graduate program. The question arising here is whether it would be fair to expect students to comprehend texts for which they have not been prepared, at least, as far as the academic vocabulary content is concerned. Consultations with the professors of electronics engineering have revealed to the researchers that, concerning the discipline-related section of the test, candidates are expected to demonstrate knowledge they have gained in their undergraduate courses. However, regarding the comprehension of English texts in their fields of study, the students are tested for knowledge, or application thereof, which they have hardly achieved in undergraduate programs. There have been no studies investigating the correspondence between the AWL content of university entrance examinations with that of authentic texts in a certain discipline, except for a paper by Aziez, (2011), in which he reports the AWL coverage of Indonesia's English National Examination texts to be 3.85%; and compares it with the much lower (2.47%) AWL percentage found in junior and senior high school English textbooks.

IMPLICATIONS

The results of the corpus analysis section of the present research revealed the fact that there is a considerable lack of correspondence between the academic words coverage of GE textbooks in Electronics engineering, on the one hand, and those of the M.S. entrance examinations and original texts, on the other. The researchers believe that the gap can be considered as an instance of unfairness to both candidates of M.S. entrance examinations, as well as undergraduates wishing to use original texts as sources of supplementary information. This, of course, is not to suggest that knowledge of academic vocabulary is the panacea for all reading problems of university students. However, regarding the prevalence of AWL words in texts from a variety of aca demic disciplines and the findings of the pre sent research, students would certainly benefit from mastery of academic vocabulary.

The AWL (Coxhead, 2000) can be used as a basis for selection of vocabulary items to be included in different tasks and activities giving information about academic words, including their meanings and the way they are used in context. More frequently used members of each AWL word family can also be introduced to help students understand the meanings of these derivations or inflected forms. It would also be possible to group AWL items and their family members on the basis of notions and functions they represent. Collocations of academic words with other verbs or nouns can be included in the course books. There is no doubt that university students would benefit doubly if academic vocabulary is presented within the context of their respective fields of study.

APPENDIX A

List of the General English books used to build the EAP materials corpus

1- Azabdaftari, B., Melkonian, Z., Salimi, E. J., Rahimpour, M., &Yaghmaie, J. (1989). English for students of agriculture. Tehran: SAMT.
2-Bidhari, P., Fallahi, M., Haghani, M., &Maftoon P. (2009) English for the students of engineering. (24th ed.). Tehran: SAMT
3- Birjandi, P. (2006). Basic English for university students. Tehran: SAMT.
4- Birjandi, P., &Mosallanejad, P. (2003). Auto run English: An intensive reading course for university students. Tehran: Mahdavi Publications.
5- Iranmehr, A., Erfani, S. M., &Davari, H. (2005). General English for university students. Tehran: Honar Publications.
6- Pourgiv, F., Sadighi, F., &Yamini, M. (1998). Reading for general English. Tehran: SAMT.
7- Sadeghi. A. (2003). English for computer users. Tehran: AndisheyeBartar Publications.
8- Shabani, M. B., &KarimiAlvar, N. (2005). General English plus. Tehran: Ayeezh Publications.

APPENDIX B

List of the sites reading passages for M.S. examination were obtained from

1. http://spectrum.ieee.org/computing/hardware/bursting-tech-bubbles-before-they-balloon/2.

Retrieved last May 13, 2010

2. http://elib.dlr.de/12232/1/is-multibody-simulation-software-suitable-kort%C2%81m-vaculin.pdf. Retrieved last May 13, 2010
3. http://www.physorg.com/news156440026.html. Retrieved last May 13, 2010
4. http://cordis.europa.eu/ictresults/index.cfm?section=news&tpl=article&BrowsingType=Features&ID=80430. Retrieved last May 13, 2010
5. http://web.princeton.edu/sites/ehs/osradtraining/biologicaleffects/page.htm.Retrieved last May 13, 2010

References

Aziez, F. (2011). Examining the vocabulary levels of Indonesia's English national examination texts. *Asian EFL Journal*. Professional Teaching Articles–CEBU. 5, 16–29.

Chen, Q., &Ge, G. (2007). A corpus–based lexical study on frequency and distribution ofCoxhead's AWL word families in medical research articles. *English for Specific Purposes, 26*(2), 502–514.

Coady, J. (1993). Research on ESL/EFL vocabulary acquisition: Putting it in context. In T.Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 3–23). Norwood, NJ: Ablex Publishing Corporation.

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

Cobb, T. (2007). Computing the vocabulary demand of L2 reading. *Language Learning and Technology. 11*(3), 38–64.

Cobb, T., & Horst, M. (2002). Is there room for an Academic World List in French?. Retrieved May13,2010,from http://www.lextutor.ca/cv/awl_F.htm

Cobb, T. (2003). The Web vocabulary profiler. Retrieved June 35, 2008 from http://www.lextutor.ca/vp/eng/

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

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

Dutton, M. (2006). Word lists tagging for teachers : Investigating the correlationof the general service list and academic word list to IELTS task 2 writing. *Proceedingsof the 19th Annual EA Education Conference*. Retrieved January 15, 2010 from http://www.englishaustralia.com.au/index.cg

Evans, S., & Green, C. (2007). Why EAP is necessary: A survey of Hong Kong tertiary students. *Journal of English for Academic Purposes, 6*(1), 3–17.

Fan, M. (2000). How big is the gap and how to narrow it? An investigation into the active and passive vocabulary knowledge of L2 learners. *RELC Journal, 31*(2), 105– 119.

Heatley, A., Nation, I. S. P.,&Coxhead, A. (2002). RANGE and FREQUENCY programs. Retrieved January 21, 2010 from http://www.vuw.ac.nz/lals/staff/Paul_Nation

Honyfield, J. (1977). Word frequency and the importance of context in vocabulary learning. *RELC Journal, 8*, 35–42.

Hsu, W. (2009). Measuring the vocabulary of college general English textbooks and English-medium textbooks of business core courses. *Electronic Journal of Foreign Language Teaching. 6*(2), 126–149. Retrieved March 22, 2010, from http://e-flt.nus. edu.sg/v6n22009/hsu.pdf

Hyland, K. (2006). *English for academic purposes: An advanced resource book.*London: Routledge.

Li Siu-leung, E., & Pemberton, R. (1994). An investigation of students' knowledge ofacademic and sub-technical vocabulary. In L. Flowerdew& A. K. K. Tong (Eds.),*Entering text* (pp. 183–196). Hong Kong: The Hong Kong University of Science & Technology.

Martinez, I. A., Beck, S. C., &Panza, C. B. (2009). Academic vocabulary in agriculture research articles: A corpus-based study. *English for Specific Purposes, 28*, 183–198.

McCarthy, M., & O'Dell, F. (2008). *Academic vocabulary in use*. Cambridge: CUP.Nagy, W., & Scott, J. (2006). The state of vocabulary research in the mid-1980's. In K. A. Dougherty Stahl and M. C. McKenna (Eds.).*Seeking understanding in how to teachreading: Selected works by Steven Stahl and a memorial Festschrift* (pp. 217–225).

New York: Guilford Publishing.Nation, I. S. P. (1983). Testing and teaching vocabulary. *Guidelines*, *5*(1), 12–25. Retrieved December 25, 2010, from http://www.victoria.ac.nz/lals/staff/Publications/paul-nation/1983-Testing-and teaching.pdf

Nation, I. S.P.*(*1990*).Teaching and learning vocabulary*. Boston: Heinle&Heinle.

Nation, I. S. P., &Coady, J. (1988). Vocabulary and Reading. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 97–110). London:

Longman.Qian, D. (1999, November). The value of vocabulary knowledge measures in

predicting ESL reading comprehension [Abstract]. Paper presented at AAAL 1999, Stamford , CT , USA.

Qian, D. (2000, March). The relationship between vocabulary knowledge and reading comprehension: A TOEFL 2000 study [Abstract]. Paper presented at the TESOL 2000 Convention, Vancouver, Canada.

Read, J. (2000). *Assessingvocabulary*. Cambridge: CUP.

Richards, J. C. (1974). Word lists: Problems and prospects. *RELC Journal*, *5*(2), 69–84.Santos, M.G. (2004). Some findings on the academic vocabulary skills of language-minority community college students. Focus on Basics, 6(D), 7–9. RetrievedJune16,2010,from www.ncsall.net/?id=175

Schmitt, D. & Schmitt, N. (2005). *Focus on vocabulary: Mastering the academic wordlist*. White Plains, NY: Longman Pearson Education.

Schmitt, D. and Schmitt, N. (2011). *Focus on vocabulary 2: Mastering the academic wordlist.* (2nd ed.)London: Longman.

Schmitt, N., Schmitt, D., &Clapham, C. (2001). Developing and exploring the behavior of a new version of the vocabulary levels test. *Language Testing,18*(1), 55–88.

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

Vongpumivitch, V., Huang, J., & Chang, Y. (2009). Frequency analysis of the words in the academic word list (AWL) and non-AWL content words in applied linguistics

research papers. *English for Specific Purposes, 28*(1), 33–41.

West, M. (1953). *A general service list of English words*. London: Longman.

Worthington, D., & Nation, I. S. P. (1996). Using texts to sequence the introductionof new vocabulary in an EAP course. *RELC Journal, 27*(2), 1–11.

Xue, G., & Nation, I. S. P. (1984). A university word list. *Language Learning and Communication 3*(2), 215–229

Parviz Maftoon is Associate Professor of teaching English at Azad University, Science and Research Campus, Tehran, Iran. He received his PhD degree from New York University in 1978 in TESOL. His primary research interests concern EFL writing, second language acquisition, SL/FL language teaching methodology, and language syllabus design. He has published locally and internationally and edited a number of research articles and books. He is currently on the editorial board of some language journals in Iran. Email: p_maftoon@srbiau.ac.ir

Aris Hakhverdian has a Ph.D. in TEFL from Islamic Azad University, Science and Research Branch in Tehran. He holds an M.A. in TEFL and a B.A. in English Translation, both from Islamic Azad University. He has been a faculty member of Garmsar Islamic Azad University since 1994. His main areas of interest are vocabulary learning; as well as simultaneous translation, which he has practiced since 2002.