Estimating and Enhancing Learners' Knowledge of Academic Vocabulary

Dr Ronald Boyle & Tim Kirk

United Arab Emirates University, UAE

Ronald.Boyle@uaeu.ac.ae

timkirk@uaeu.ac.ae

1.   Introduction

Students who are non-native speakers of English and who plan to study for a bachelor's degree in the medium of English need to be aware of the vocabulary requirements for such study, particularly if they are speakers of non-cognate languages such as Turkish or Arabic. This is because the connection between success in academic reading and vocabulary knowledge is well established (Corson 1997; Nation 2001:144-149). Research into the vocabulary of second language learners has focused on how many of the high-frequency words of the language students are likely to encounter in their reading and how many they need to know (Read 1988). Laufer (1997) and Nation (2001) both argue that non-native speakers in English-medium universities need at least 3,000 word families in order to understand 95% of the running words in a text. The research reported here was a response to the textual difficulty of a required reading course in 'great ideas' (e.g. Lao-Tzu, Plato and Machiavelli) which was introduced by the College of Humanities and Social Science (CHSS) at the UAE University. This innovation raised questions about the size of the students' vocabulary and revealed that little was known about this subject. The purpose of the research, therefore, was to determine what the vocabulary knowledge of a sample of female undergraduate students enrolled in the 'great ideas' course was and to attempt to correlate the students' vocabulary knowledge with their scores on the IELTS Reading test. A further objective was to provide students with information about their vocabulary at different stages of their education, to show them how fast their vocabularies grow, and to make them aware of what they need to learn.

2.   Knowing a word

2.1   Receptive and productive knowledge

Nation (2001:23-30) and others distinguish between a person's receptive and productive knowledge of a word. These in turn are sub-divided into the knowledge of the form, meaning and use of a word

(See Appendix 1). In this research, as a result of constraints affecting testing, we have adopted a minimum interpretation of 'knowing a word,' and that is merely the receptive knowledge of a word's meaning.

2.2   Defining a word

By a 'word,' most researchers mean a word family, which is defined as the stem, or lemma, of a word plus inflectional suffixes (e.g. view, views, viewed and viewing) and derivational affixes (e.g. viewer, review, preview and interview). Words are considered to belong to the same family if the meaning of a derived form can be inferred from the base word with minimal effort. A few researchers, however, treat the inflected and derived forms as separate words, and this explains the inflated estimates of vocabulary size which are reported in the literature (Goulden, Nation and Read 1990: 343-344; Read 2000:84).

3.   Testing breadth of knowledge

Breadth of vocabulary knowledge refers to the total number of words families that a person knows. In this study, breadth is measured by means of Nation's (2001) Vocabulary Levels Test (VLT) because it is 'the nearest thing we have to a standard test in vocabulary' (Meara 1996:38). The VLT has five levels: 2,000-word level, 3,000-word level, 5,000-word level, University level (See Coxhead 2000) and 10,000-word level. It is available at www.lextutor.ca/tests/levels/recognition/2-10k/.

The 2,000- and 3,000-word levels test students' knowledge of high frequency words, while the 5,000-word level is on the boundary between high-frequency and low frequency words (Nation 1990:261). In this study, we dispensed with the 10,000-word level after finding that it revealed little about the students whom we tested.

The VLT employs a word-definition matching format, as shown below.

1. apply	
2. elect	
3. jump	_____ choose by voting
4. manufacture	_____ become like water
5. melt	_____ make
6. threaten	

The VLT tests only content words and not function words. There are 36 words at each level and 18 definitions. The definitions are written using words from the level below, and the words in each group belong to the same part of speech. The purpose of this is to stop students using suffixes as clues as to which word they should select.

4.   The sample

The sample consists of 461 female students enrolled on HSR 250, Reading and Comprehension, the required 'great ideas' course for all students entering the College of Humanities and Social Sciences, UAE University. New entrants were tested at the start of three semesters, in September 2005, January 2006 and September 2006.

5.   Previous research

Research on vocabulary size is mostly done among multilingual groups on university campuses in English-speaking countries, and there are only a few reports of work undertaken elsewhere. Cobb (1999), for example, estimates that students entering Sultan Qaboos University in Oman in the mid-1990s knew between 500 and 1,000 word families. Arnaud and Sauvignon (1997) report that Finnish secondary school leavers (speakers of a non-Indo-European language) acquire around 1,500 word families over a seven-year period of study, and that European language learners in secondary school only move beyond the 2,000-word level upon entering university. The only detailed piece of research which is readily available was undertaken by Nurweni and Read (1999) in Indonesia. They found that the 314 university students whom they tested in an English-medium programme knew, on average, 1,226 word families after five years of secondary school language instruction.

6.   Results and discussion

6.1   The UAE University sample

The calculation of the number of word families, following Read (1993), involves adding the mean scores of the 2,000-, 3,000- and 5,000-word levels, and finding what percentage of the total

possible score the mean represents. The average of the whole sample was 1,919 word families, and the averages for individual semesters were:

Semester	Sep 2005	Jan 2006	Sep 2006	Whole sample
Word families	1,888	1,881	1,957	1,919
N	81 students	156 students	224 students	461 students

These averages are significantly higher than those reported from Oman (Cobb 1999) or Indonesia (Nurweni and Read 1999), and they compare favourably with the standards of European secondary school leavers. Nevertheless, they are still well below the 3,000-word threshold level required for university study (Laufer 1997).

Country	Oman (Cobb)	Indonesia (Nurweni & Read)	Finland (Arnaud & Sauvignon)	UAE (Boyle & Kirk)
Word families	500-1,000 (estimate)	1,226	± 1,500	1,919

6.2   The University level scores

The expected declining pattern of scores from the 2,000-word level through the 3,000-word level and the 5,000-word level is interrupted by the results of the University level scores. The University level consists of Latinate words based on Coxhead's (2000) Academic Word List (AWL). These slightly higher scores may reflect the attention paid to the AWL in the UAE University's pre-university programme, but Read (2000:121), who found similar results in a similar situation, is wary of drawing this conclusion.

Word level	2,000	3,000	5,000	University
Mean score (out of 18)	9.74	6.95	4.04	4.42

6.3   Correlation of Vocabulary Knowledge and IELTS Reading Scores

Qian (2002) used the VLT and found that the learners' scores were good predictors of success on the reading section of TOEFL. This project hoped to find a similar correlation between the VLT results and scores on the reading section of IELTS. However, there was no measurable correlation. UAEU students take IELTS at the end of their English preparatory programme, with a target of IELTS band 4.5 required to enter the faculties. Thus the vast majority of our sample, coming from the same English language training course, reported previous IELTS reading scores within a narrow range of bands 4-5. This one-band variance could not be meaningfully correlated with the 18-item scores of the Vocabulary Level tests. However, inside each IELTS reading half-band there might be a range of 3-4 raw points, meaning that someone at the top end of IELTS reading band 5 might have up to 12 raw points more than someone at the bottom end of IELTS reading band 4. In future, if permission can be obtained to access raw IELTS reading scores rather than the reported bands, a correlation with the results of the Vocabulary Levels test might appear.

7.   Pedagogical issues

After an earlier presentation, members of the audience asked us to address a number of issues that are relevant to vocabulary learning. Three of these issues are summarized below.

7.1   Repetition

Nation's (1982) survey of research suggested that as many as sixteen repetitions may be required before the form and basic meaning of a word can be retrieved easily from the mental lexicon. School textbooks are generally not designed to allow systematic repetition to this extent.

7.2   Spaced learning

Bahrick *et.al.* (1993:321) argue that 'the powerful long-term effects of widely spaced retrieval constitute an important, but unexploited contribution of memory research to education.' While English language textbooks may incorporate some form of spaced learning for grammar, there is little evidence to suggest that they pay the same attention to vocabulary. Pimsleur's memory schedule (Nation 2001:78) is an exponential scale which shows us how, according to research, eleven repetitions might be spaced.

Pimsleur's memory schedule

1	2	3	4	5	6	7	8	9	10	11
5	25	2	10	1	5	1	5	25	4	2
secs		mins		hrs		days			mths	years

7.3 Intentional learning

L1 learners acquire most of their vocabulary incidentally through reading, but L2 learners do not have thousands of hours of exposure to a foreign language. According to Zahar, Cobb and Spada (2001), three periods of English a week would result in the incidental learning of about 70 new words a year. Therefore, acquisition through reading must be supplemented by direct instruction. The communicative approach to language teaching over the past thirty years, however, has not favoured direct instruction (Focus on FormS) of grammar or vocabulary (Laufer 2006).

8. Pedagogical applications

In addition to questions about how we teach vocabulary, there were concerns about what we teach, and Nation (2001) reports research which shows that two of the most effective means of improving academic vocabulary are through a study of Greek and Latin roots and affixes and through the use of mnemonic devices such as the 'keyword' method.

8.1 Graeco-Latin roots and affixes

The first 1,000 word families of English are predominantly words of Germanic origin. After that, words of Graeco-Latin origin constitute two-thirds of the lexicon (Stockwell and Minkova 2001), and these are the words that are required for academic study (Corson 1997; Coxhead 2000). As skill in the morphological analysis of roots and affixes promotes L1 vocabulary growth (White, Power and White 1989), and as morphologically related words may be linked in the mental lexicon and therefore more easily retrieved (Corson 1997), there is value in learning roots and affixes. While educated speakers of cognate European languages can recognize most Latinate words and many of their affixes, speakers of non-cognate languages such as Turkish and Arabic have no such advantage

and need instruction in Greek and Latin roots and affixes in order to improve their knowledge of academic vocabulary.

An interesting finding in our research was that the Arabic-speaking students in the UAE University sample knew proportionately more of the Germanic words than of the Graeco-Latin words. Part of the reason for this may be the widespread use in the UAE of spoken English and its Germanic vocabulary, and the students' vocabulary may therefore reflect a better knowledge of what Cummins (1981) calls Basic Interpersonal Communication Skills, the skills we need for spoken social interaction, rather than the Cognitive Academic Language Proficiency which a person needs for university study (www.iteachilearn.com/cummins). Hence, while our test results suggest that our students benefit from a study of Coxhead's (2000) Academic Word List, it would seem that still more time and effort needs to be devoted to Latinate words.

8.2   Mnemonics

Ellis and Beaton (1997), Hulstijn (1997) and Nation (2001) discuss the proven success of the little-used 'Keyword' mnemonic method of vocabulary learning and its psycholinguistic underpinnings. Hulstijn (1997:210) believes that the method has attracted little attention because "it can be successfully applied with only a minority of vocabulary items, i.e. with words referring to objects that can be perceived visually." This would tend to exclude the abstract Graeco-Latin words, but Pierson (1989) argues that etymology can reveal the concrete basis of many of these words. These concrete ideas can serve a mnemonic purpose. If you know, for example, that the Latin root *port* means 'carry,' it should be easier to understand, remember and retrieve related words such as 'import,' 'support,' 'portfolio' and even 'opportunity.'

9.   Conclusion

Undergraduates are generally aware that a good vocabulary is essential to successful reading, but if their L1 is Turkish or Arabic they will need help in identifying the words that are of greatest benefit to them. Their learning burden will be reduced if they can recognize the morphological connections between many academic words and if they devote time to the study of Graeco-Latin roots and affixes.

Appendix 1

What is involved in knowing a word?

Form

Spoken	R	What does the word sound like?
	P	How is the word pronounced?
Written	R	What does the word look like?
	P	How is the word written and spelled?
Word parts	R	What parts are recognizable in this word?
	P	What parts are needed to express the meaning?

Meaning

Form & meaning	R	What meaning does this word form signal?
	P	What word form can be used to express this meaning?
Concept & referents	R	What is included in the concept?
	P	What items can the concept refer to?
Associations	R	What other words does this make us think of?
	P	What other words could we use instead of this one?

Use

Grammatical functions	R	In what patterns does the word occur?
	P	In what patterns must we use this word?
Collocations	R	What words or types of words occur with this one?
	P	What word or types of words must we use with this one?
Constraints on use (register, frequency ..)	R	Where, when, and how often would we expect to meet this word?
	P	When, where, and how often can we use this word?

Note: In column 3, R = receptive knowledge, P = productive knowledge

Nation (2001:27)

References

Arnaud, P.J.L. and Sauvignon, S.J. 1997. Rare words, complex lexical units and the advanced learner. In J. Coady and T. Huckin, (Eds.). *Second Language Vocabulary Acquisition.* Cambridge: Cambridge University Press. pp. 157-173.

Bahrick, H.P., Bahrick, L.E., Bahrick, A.S, and Bahrick, P.E. 1993. Maintenance of foreign language vocabulary and the spacing effect. *Psychological Science*, 4 (5), 316-321.

Cobb, T.M. 1999. Applying constructivism: A test for the learner-as-scientist. *Educational Technology Research and Development,* 47(3): 15-31.

Corson, D. 1997. The learning and use of academic English words. *Language Learning*, 47 (4), 671-718.

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

Cummins, J. 1981. Age on arrival and immigrant second language learning in Canada: a reassessment. *Applied Linguistics*, 2 (2): 132-149.

Ellis, N. C. and Beaton, A. 1993. Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning, 43* (4), 559-617.

Goulden, R., Nation, P. and Read, J. 1990. How large can a receptive vocabulary be? *Applied Linguistics, 11* (4), 341-363.

Hulstijn, J.H. 1997. Mnemonic methods in foreign language vocabulary learning. In J. Coady and T. Huckin (Eds.). *Second Language Vocabulary Acquisition.* Cambridge: Cambridge University Press. pp. 203-224.

Laufer, B. 1997. The lexical plight in second language reading: Does it change over time? In J. Coady and T. Huckin (Eds.). *Second Language Vocabulary Acquisition.* Cambridge: Cambridge University Press. pp. 20-34.

Laufer, B. 2006. Comparing focus on form and focus on forms in second-language vocabulary learning. *The Canadian Modern Language Review, 63* (1), 149-166.

Meara, P. 1996. The dimensions of lexical competence. In G. Brown, K. Malmkjaer and J. Williams (Eds.). *Performance and Competence in Second Language Acquisition.* Cambridge: Cambridge University Press.

Nation, I.S.P. 1982. Beginning to learn foreign vocabulary: A review of the research. *RELC Journal, 13,* 14-36.

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

Nurweni, A. and Read, J. 1999. The English vocabulary of Indonesian university students. *English for Specific Purposes 18*, 161-175.

Pierson, H. 1989. Using etymology in the classroom. *ELT Journal*, 41 (3), 57-63.

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

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

Read, J. 2000. *Assessing Vocabulary.* Cambridge: Cambridge University Press.

Stockwell, R. and Minkova, D. 2001. *English Words: History and Structure.* Cambridge: Cambridge University Press.

White, T.G., Power, M. A., and White, S. 1989. Morphological analysis: Implications for teaching and understanding vocabulary growth. *Reading Research Quarterly,* 24, 283-304.

Zahar, R., Cobb, T. and Spada, N. 2001. Acquiring vocabulary through reading: Effects of frequency and contextual richness. *The Canadian Modern Language Review, 57* (4), 541-572.