

# The Vocabulary Knowledge of University Students in Saudi Arabia



Ahmed Al-Masrai  
Swansea University, UK



James Milton  
Swansea University, UK

This paper presents an empirical study that investigates the vocabulary knowledge of 92 Saudi university learners of English as a Foreign Language (EFL) near the start and near the end of their university studies. Two tests were used to measure the participants' vocabulary size: the well-established Eurocentres Vocabulary Size Test (EVST) (Meara & Jones, 1990) test and a newly created XK\_Lex (Al-Masrai, 2009), designed to make a more accurate measurement of vocabulary size. Results suggest that Saudi university students' vocabulary size is about 2000 to 3000 words on entry to university and around 5000 words nearer graduation. These figures thus emphasise that Saudi university learners' level is, on average, some way short of the kind of level associated with complete fluency in EFL. Thus, the study suggests further English language support for the graduated students.

## Introduction

Researchers, teachers, students and materials writers can all agree that mastering the vocabulary of a foreign language is essential if the second language as a whole is to be mastered. However, as Schmitt (2008) points out, the ideal approach for learning vocabulary is yet to be found. He attributes this to several factors, but particularly to the lack of clear

descriptions and guidelines in teaching materials. It is unclear exactly how much vocabulary should be learned, and which vocabulary should be learned at any stage of the process of learning. To discover this kind of information, it is important to have good tests of vocabulary knowledge to establish the kind of norms of progress that might inform teachers and learners alike. There is a substantial body of research in vocabulary testing, yet few studies (e.g., Milton, 2006a, 2006b) have dealt with assessing and monitoring the progress of learners' vocabulary knowledge. However, measuring learner's vocabulary knowledge and showing their progress should prove very useful. Therefore, this paper seeks to establish the norms of vocabulary knowledge and progress among EFL learners at a Saudi Arabian university, and to compare these figures with learners elsewhere. This information should better inform the process of learning and teaching EFL in Saudi Arabia.

## Background

### The importance of vocabulary knowledge

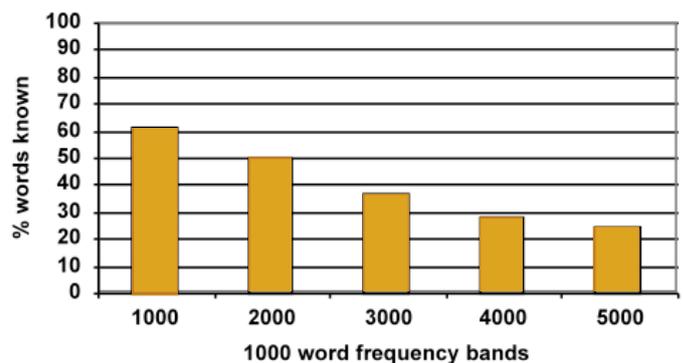
Over the past two decades or so, research in vocabulary testing, teaching and learning has increased in volume. This research shows that vocabulary is considered an essential part almost in every aspect of language knowledge (Daller,

Milton & Treffers-Daller, 2007). The literature reveals a strong relationship between EFL learners' vocabulary knowledge and language skills in general and reading comprehension in particular. EFL/ESL language abilities are heavily dependent on learners' vocabulary knowledge (Shen, 2008). Nation (1990) notes in particular the importance of vocabulary size as an overall predictor of reading performance. Since vocabulary knowledge is at the heart of foreign language ability and development, then assessing its level and progress is believed to be very useful. Once a teacher knows the size and type of vocabulary learners know, s/he can try to implement certain methods to improve their vocabulary level. If learners' vocabulary knowledge appears low then steps can be taken to increase it, which should have a beneficial effect on overall language performance. Or, if a teacher knows after testing her/his learners' vocabulary knowledge that they have problems with knowing most of the high frequency words, then s/he can give priority to teaching these words.

### Testing vocabulary knowledge

The importance of testing vocabulary knowledge goes in line with the importance of vocabulary knowledge itself. According to Pearson, Hiebert and Kamil, (2007, p. 282), "in order to teach vocabulary more effectively and better understand its relation to comprehension, we need first to address how vocabulary knowledge and growth are assessed." The significance of tests of this kind is that, provided they are well-constructed and give reliable and, most important of all, trustworthy estimates of EFL learners' vocabulary knowledge, it becomes possible to monitor the vocabulary knowledge and progress of learners (Milton & Hopkins, 2006). Without this information neither teachers nor learners can be aware of whether they have appropriate vocabulary, and sufficient volumes of vocabulary, for the tasks or the exams being undertaken. There are a few tests that, it can be argued, are reliable and valid in the area of vocabulary and these are based on the most frequent words in English. For example, the X\_Lex test can provide what are argued to be good estimates of knowledge of the 5000 most frequent words and can give accurate indications of comparative levels of knowledge (Milton, 2009). Results from these tests compare well with estimates of the vocabulary knowledge needed for comprehension derived from coverage estimates (e.g., Nation 2006). These methods recognise that

learners tend to learn vocabulary fairly predictably and that vocabulary uptake is likely to be strongly influenced by word frequency. Figure 1 demonstrates clearly how the X\_Lex test functions to give normative figures for test takers and shows the scale of knowledge in each of the first five 1000 word frequency bands in English.



*Figure 1:* Frequency profile for Greek learners of EFL (cited from Milton, 2006a, p. 32)

It can be seen from Figure 1 that learners tend to know more vocabulary in the first 1000 words level, less in the second thousand, and so on. Statistical analysis in Milton's study confirms the relationship between the frequency bands and the vocabulary size. This evidence supports the claim that the higher the frequency of a word the better the chance of its learnability, as a rule of thumb. The implication of this is that a test based on the most frequent words of English is likely to give a good overall estimate of a learner's vocabulary size, since knowledge outside these most frequent bands appears to be small.

### Normative figures for vocabulary knowledge

Research has shown how much vocabulary knowledge is needed to use a foreign language efficiently. In this section, we will shed some light on some figures reported in the literature as a threshold to perform well in a foreign language. However, most of the research concerning the relationship between vocabulary size and language proficiency has been carried out within the realm of reading (Staehr, 2008). A study by Laufer (1989) concluded that knowledge of at least 95% of the vocabulary of a text is required for comprehension. Further, Hu and Nation (2000) and Nation (2006) suggest that coverage of 98-99% of a text is necessary for reading

for pleasure. On the basis of coverage figures of this kind the literature suggests that learners might need to know around 8000-9000 word families in order to read authentic texts. Technical and academic texts might require a higher vocabulary and in a recent study, Nation and Webb (2011) claim that a learner would need a vocabulary size of around 20,000 words in order to read an academic text comfortably. The literature which indicates the volume of vocabulary needed for mastery in the other skill areas is less extensive but Nation (2006) suggests that a vocabulary size of 6000-7000 word families is needed to comprehend spoken discourse. This finding probably fits well with Milton (2009), which states that EFL learners need vocabulary knowledge of around 3000 words or more for basic communication, out of the most frequent 5000 words in English, besides the most predictable and formulaic exchanges in speech and listening. Regardless of the skill, however, it is generally assumed that the more words learners know, the better the chance they have of understanding while reading or listening in the foreign language.

## Normative figures for vocabulary knowledge: Saudi students

Since it was introduced in 1927 (Assah, 1969, cited in Al-Hazemi, 1993), EFL has become an established part of the Saudi curriculum. In the

Ahmed Al-Masrai is a lecturer in English language at King Abdulaziz Military Academy in Riyadh and is currently enrolled in a full-time PhD program at Swansea University, UK, researching the effects of L1 Arabic lexical organization on L2 English lexical learning.

matter of vocabulary teaching, the Saudi Ministry of Education (MOE) documented that students are expected to leave high school with a vocabulary size of around 3000 words. However, research has questioned whether the majority of the students reach this vocabulary size. Two particular studies have investigated Saudi learners' vocabulary knowledge: Al-Hazemi (1993) and Al-Bogami (1995). These two studies suggested that Saudi students leave high school with disappointingly low vocabulary knowledge. According to Al-Hazemi (1993), learners in his study scored between 800 and 2000 words with a mean score of around 1000 words. Al-Bogami's (1995) findings conform to Al-Hazemi's and suggest that Saudi students score very poorly in vocabulary size tests. It appears most learners leave

high school with vocabulary knowledge below the threshold for anything but the most limited communication (Nation, 2001). If learning in schools is poor, then this should have consequences for the study of English at a university. It should not be surprising to find low vocabulary size among learners at a university level, although it might be expected that university students would be a sub-group of the most able school learners. Limited knowledge would inevitably hinder learners' progress in their higher educational study and place a high learning burden on the students to gain the levels of knowledge expected of university graduates. We have some evidence of this since it appears that lexical deficiency hinders the reading and comprehending of the texts used at university (e.g., Al-Homoud & Schmitt, 2009). However, we lack direct evidence of the vocabulary knowledge among university learners and the scale of this deficiency if, indeed, it does exist.

## Aims and objectives

Based on the normative figures of vocabulary knowledge reported in the literature review section, this study will investigate the EFL vocabulary size of Saudi students majoring in English language at the university level. The main research questions addressed in this study are: (a) what is Saudi university students' EFL vocabulary size and (b) what implications does this have?

## Methodology

### Participants

This study involved 92 male Saudi learners of English from two different university levels. The first group comprised 55 participants from level 2 who had received 935 hours of English language teaching. The other group comprised 37 subjects from the final year (level 10) who had received around 2177 classroom hours of instruction. The participants selected for this study were majoring in English language and will become English language teachers after they have graduated. Level 2 participants were at the end of the second semester of their first year at the university when tested. The participants in the other group (Level 10) were also at the end of the second semester of their final year at the university, year five. The two measurements used in this study, which will be described in the following section, are expected to provide some indications of learners' vocabulary knowledge near their time of entry to the university and near their graduation.

## Instruments

Two types of vocabulary size tests were used to collect data for this study:

- ◆ XK\_Lex (Al-Masrai, 2009): a test designed for this study and intended to make more accurate estimates of learners' vocabulary
- ◆ The computerised EVST test (Meara & Jones, 1990)

Both measures are Yes/No tests and constructed to measure learners' breadth of vocabulary knowledge of the most 10,000 frequent words of English.

XK\_Lex differs from the EVST in the respect that the XK\_Lex is designed to monitor learners' vocabulary knowledge profile in the first ten 1000 word frequency bands and makes an overall estimate from sub-estimates at each of the ten bands. EVST is well documented in the literature but, crucially from the point of view of making accurate overall calculations, discounts from its estimate knowledge of vocabulary in the less frequent bands. If a learner scores low in any of the first frequent levels the test, systematically, stops testing the next vocabulary levels and assumes no knowledge in these infrequent bands. For this particular shortcoming the XK\_Lex is designed to fill this gap and to test whether there are significant volumes of vocabulary omitted from the EVST estimate.

XK\_Lex is constructed to test knowledge of the most frequently occurring 10,000 words in English and presents an estimate of the overall breadth knowledge of this vocabulary. The lexical items on which the test is based are drawn from Nation (1984) and Kilgarriff (2006) and are lemmatised. It is a paper-and-pencil Yes/No test that presents learners with 100 words listed in ten columns, ten words in each. Learners should choose each word they know. There are also 20 "pseudo" words that are designed to look and sound like real English words. Two of these 20 words are set in each of the ten frequency bands. As suggested by Anderson and Freebody (1983), pseudo words can reduce the chance that learners will overestimate their vocabulary knowledge. Claiming knowledge of non-words can lead to adjusting learners' final scores after deducting the marks of the non-words from the real English words to give a reasonable estimate of learners' vocabulary knowledge. The time needed to run one form of the test generally takes five to ten

minutes. In order to calculate learners' vocabulary knowledge to get a raw score of 10,000, the number of Yes responses to real words is added together and multiplied by 100. The number of Yes responses to pseudo words is then calculated and multiplied by 500. This number is deducted from the raw score to provide an adjusted final score. The participants were given two paper and pencil forms (A and B) of the XK\_Lex, using different words and pseudo words, and also sat the computerized EVST test after they had completed the XK\_Lex test.

## Results

### Saudi University Students' vocabulary size

Tables 1 and 2 summarize the vocabulary knowledge of two levels of university learners.

**Table 1**

*Summary of the high level learners' scores*

Test format	N	Minimum	Maximum	Mean	Std. Deviation
EVST	37	2400	7450	4198.65	1201.532
XK-Lex A	37	2400	7800	5262.16	1265.647
XK-Lex B	37	2400	7600	5186.49	1188.407

Table 1 summarizes the means, the minimum, the maximum scores, and the standard deviation of the vocabulary level of the learners in the final year at the English language department in KSU. The results highlighted in this table show that the mean score of vocabulary size of the Level 10 students, in both forms of the XK\_Lex, are almost the same. The estimate made by EVST is lower.

**Table 2**

*Summary of the low level learners' scores*

Test format	N	Minimum	Maximum	Mean	Std. Deviation
EVST	55	550	2450	1680.91	520.549
XK-Lex A	55	1200	5400	3025.45	895.714
XK-Lex B	55	1300	4300	2907.27	759.106

Figures in Table 2 summarize the level of the second year students in KSU. These scores are, as might be expected, lower than the Level 10 students' scores. The difference between the scores on each test is statistically significant: EVST  $t(90) = 13.765$  sig <

.001; XK\_LexA  $t(90) = 9.768$ , sig. .001; XK\_LexB  $t(90) = 11.233$ , sig. < .001. The scores produced by the two forms of the XK\_Lex tests are again very similar while the EVST test produces a noticeably lower estimate of vocabulary size.

Parallel forms and split half analysis, together with the frequency profile these results demonstrate, suggest that the tests are valid and are performing reliably (Al-Masrai, 2009).

## Discussion

We began this paper by pointing out that vocabulary is an essential area of knowledge for successful foreign language acquisition, and it is vocabulary size in particular that appears to impact a learner's language skills. Therefore, the major aim of this study was to investigate Saudi university students' vocabulary knowledge in order to establish their levels of vocabulary knowledge and to make judgements as to whether their knowledge is at a level suitable for university study through the medium of English. Two tests were used to test learners' vocabulary size-- the well-established EVST and the newly created XK\_Lex tests.

Results from this study reveal that Saudi university learners know, on average, between 1650 and 3000 words in English around the time of entry to university. By the time they graduate, their knowledge has increased and the results suggest they know, on average, between about 3000 and 5000 words in English. This range of scores is produced by differences between the EVST scores which are usually lower and the XK\_Lex scores, which usually provide higher estimates. It was mentioned earlier that EVST assumes vocabulary is learned in strict frequency order and therefore discounts knowledge of much infrequent vocabulary. It assumes no knowledge at all in these areas, once vocabulary scores in a frequency band drop below the highest of levels. However, it appears that learners know considerable infrequent vocabulary which EVST does not include in its estimate. It is assumed, therefore, that EVST will underestimate learners' vocabulary knowledge and the XK\_Lex will provide a better calculation of overall vocabulary size since it includes estimates in these levels. It can be assumed, therefore, that learners probably know approximately 2000 to 3000 of the most frequent 10,000 words of English when they enter university and something

over half of this range of words, about 5000 words, when they leave.

These estimates probably fit well with estimates of the scale of vocabulary knowledge of learners in Saudi schools. It is suggested (e.g., Al-Hazemi, 1993) that Saudi learners probably finish school with, on average, about 1000 words in English. If it is assumed that only the most able of these learners progress to study English at university, then an estimate of about 2000 or 3000 words around the time of entry is believable. This figure suggests that university learners, while progressing through school, make the kind of progress that learners of EFL in other environments make, and acquire some 3 to 4 words per contact hour. Table 3 makes estimates of the approximate rate of lexical uptake per contact hour that learners in this study make before and during their university courses. Milton and Meara (1998) suggest that this rate of vocabulary uptake is typical of learners in Europe and the Far East. Laufer (2010) produces a slightly lower estimate of 2 to 3 words per contact hour, but worked with learners more predominantly in the Middle East and this may be because such learners have comparatively little exposure to the language outside the classroom. It appears that learners in this study acquire vocabulary at a slightly slower rate, about two and half words per contact hour, during their study at university and while this rate still falls within the norms described above, it is not immediately obvious why progress should be slow.

**Table 3**

*The mean scores, classroom hours and vocabulary uptake rate per contact hour*

Test	Level	N	Mean	Classroom Hours	Vocabulary Uptake Average per hour
XK_Lex	2	55	3025.45	935	3.24
	10	37	5262.16	2177	2.42

These figures also suggest that students of English at university are some ways away from the levels of knowledge which suggest they can study authentic texts independently, or have reached the levels of fluency associated with, say, CEFR levels C1 and C2 which might require scores approaching 9000. The average score of approximately 5000 words, obtained by learners in this study, probably indicate a CEFR level of about B2 (Milton & Alexiou, 2009). It will

be recalled that Laufer's (1989) estimate that about 5000 words provides 95% coverage of normal texts sufficient, if only just, for comprehension, was made with learners at this B2 level.

Participants in this study were selected from students at university level, majoring in English. Those learners are assumed to be qualified to teach English to learners in Saudi public schools at various levels.

James Milton is Professor of Applied Linguistics at Swansea University, UK. A long-term interest in measuring lexical breadth and establishing normative data for learning and progress has led to extensive publications including *Modelling and Assessing Vocabulary Knowledge* (CUP, 2007 with Michael Daller and Jeanine Treffers-Daller) and *Measuring Second Language Vocabulary Acquisition* (Multilingual Matters, 2009).

It may be the case that students with this level of English knowledge will become successful teachers of English particularly at the lower levels of the EFL curriculum. However, it must be borne in mind that such learners are not fluent and will require, if they are to become and to stay fully professional teachers in this area, considerable support and continued language enhancement. It suggests that continued and in-service training would be a real asset to such teachers and would help raise the levels of English in the school sector.

### Conclusion

The results of this study suggest that it is possible to make good estimates of the levels of vocabulary knowledge among learners in Saudi universities. The scores the learners obtain on these tests suggest that while English major students at Saudi universities make relatively good progress in acquiring English vocabulary at school and university, they nonetheless appear to fall short of the kind of standards of knowledge expected by the Saudi curriculum authorities. If these learners have smaller vocabularies in English than are expected, they will probably be less able overall in English than might be expected or desired. Knowledge of 3000 words or so on entry to university suggests that these learners will be far from fluent and will struggle to understand authentic texts without considerable support. The teachers of English in Saudi universities will need to reflect and respect this level of knowledge in their

teaching practices if the learners' time at university is to be optimised. A vocabulary size score of around 5000 words at about the time of graduation suggests learners will be competent rather than fluent users of English and if these learners are to become successful teachers of English as a foreign language, they will need language support and in-service training during their professional careers.

### References

- Al-Bogami, A. (1995). *Teaching English vocabulary to EFL male students at intermediate and secondary public schools in Riyadh*. (Unpublished master's thesis). King Saud University, Riyadh, Saudi Arabia.
- Al-Hazemi, H. (1993). *Low level EFL vocabulary tests for Arabic speakers*. (Unpublished doctoral dissertation.) University of Wales, Swansea, UK.
- Al-Homoud, F. & Schmitt, N. (2009). Extensive reading in a challenging environment: A comparison of extensive and intensive reading approaches in Saudi Arabia. *Language Teaching Research*, 13, 383-401.
- Al-Masrai, A. (2009). *Measuring the English vocabulary size of Saudi university students: Validating a new 10,000 word vocabulary size test*. (Unpublished master's thesis). Swansea University, Swansea, UK.
- Anderson, A. & Freebody, P. (1983). Reading comprehension and the assessment and acquisition of word knowledge. *Advances in Reading/Language Research*, 2, 231-256.
- Assah, A. (1969). *Miracle of the desert kingdom*, Dublin: Cahill & Co., Ltd.
- Daller, H., Milton, J. & Treffers-Daller, J. (Eds.). (2007). *Modelling and assessing vocabulary knowledge*, Cambridge, UK: Cambridge University Press.
- Hu, M. & Nation, P. (2000). Vocabulary density and reading comprehension. *Reading in a Foreign Language*, 23, 403-430.
- Kilgarriff, A. (2006). BNC database and word frequency lists. Retrieved from <http://www.kilgarriff.co.uk/bnc-readme.html#lemmatised>
- Laufer, B. (1989). What percentage of text is essential for comprehension? In C. Lauren & M. Nordman (Eds.) *Special language: From humans thinking to thinking machines* (pp. 316-323). Clevedon, UK: Multilingual Matters.
- Laufer, B. (2010). Form-focussed instruction in

- second language vocabulary learning. In R. Chacón-Beltrán, C. Abello-Contesse, M. Torreblanca-López & M. López-Jiménez (Eds.) *Further insights into non-native vocabulary teaching and learning* (pp.15-27). Clevedon, UK: Multilingual Matters.
- Meara, P. & Jones, G. (1990). *Eurocentre's Vocabulary Size Test. User's Guide*. Zurich, Eurocentres.
- Milton, J. (2006a). Language lite? Learning French vocabulary in school. *Journal of French Language Studies*, 16, 187-205.
- Milton, J. (2006b). X-Lex: The Swansea vocabulary levels test. In C. Coombe, P. Davidson, P. & D. Lloyd, (Eds.), *Proceedings of the 7th and 8th current trends in English language testing (CTELT) conference*. Dubai, UAE: TESOL Arabia.
- Milton, J. (2009). *Measuring second language vocabulary acquisition*. Bristol, UK: Multilingual Matters.
- Milton, J., & Alexiou, T. (2009). Vocabulary size and the Common European Framework of Reference for languages. In B. Richards, M. Daller, D. Malvern, P. Meara, J. Milton, & J. Treffers-Daller (Eds.), *Vocabulary studies in first and second language acquisition* (pp. 194-221). Basingstoke, UK: Palgrave.
- Milton, J. & Hopkins, N. (2006) Comparing phonological and orthographic vocabulary size: Do vocabulary tests underestimate the knowledge of some learners. *The Canadian Modern Language Review*, 63, 127-147.
- Milton, J., & Meara, P. (1998). Are the British really bad at learning foreign languages? *Language Learning Journal*, 18, 68-76.
- Nation, P. (1984). *Vocabulary lists: Words, affixes and stems*. English University of Wellington, New Zealand: English Language Institute.
- Nation, P. (1990). *Teaching and learning vocabulary*. New York: Heinle & Heinle.
- Nation, I. S. P., (2001). *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press.
- Nation, P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, 63, 59-82.
- Nation, I. S. P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Boston, MA: Heinle.
- Pearson, P. D., Hiebert, E. H., & Kamil, M. L. (2007). Vocabulary assessment: What we know and what we need to learn. *Reading Research Quarterly*, 42, 282-296.
- Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, 12, 329-363.
- Shen, Z. (2008). The roles of depth and breadth of vocabulary knowledge in EFL reading performance. *Asian Social Science*, 4, 135-137.
- Staehr, L. (2008) Vocabulary size and the skills of listening, reading and writing. *Language Learning*, 36, 139-152.

\* \* \* \* \*

## TESOL Arabia offers

**International Travel Grants,  
Dr Lisa Barlow Memorial  
Travel-In Grants,  
Professional Course  
Development Grants,  
and Research Grants as a  
benefit to all members.**

Apply online at  
[www.tesolarabia.org](http://www.tesolarabia.org)