Textbook Input as Agent of Low Vocabulary Uptake by EFL Learners

Thamer Alharthi and Ahmad Alfotais

Department of European Languages and Literature, King Abdulaziz University, Jeddah, Saudi Arabia
Department of Language and Linguistics, University of Essex, Colchester, United Kingdom
talharthi@kau.edu.sa

Abstract : For most people 'learning, working and communicating in a global context' begins in a FL classroom and one of the major tasks of the FL student is the learning and retaining of vocabulary. There are growing concerns about the limited vocabulary knowledge in Saudi EFL contexts (AlSaif, 2011). Given that vocabulary is probably the most crucial aspect of optimized FL learning, it has a prominent place in language teaching textbooks and materials. However, as O'Loughlin (2012, p. 256) suggests, "There is little published information investigating the vocabulary presented in course-books". The present study seeks to shed light on the role of vocabulary input in textbooks and asks whether it is responsible for the low vocabulary uptake by Saudi EFL learners at tertiary level. The study examines the impact of some potential factors including frequency, recycling and word class on the students' achievements in vocabulary knowledge. A total of 40 third-year university English majors' lexical knowledge was assessed, using achievement vocabulary tests to obtain an overall picture of the size of their vocabulary. The results show that the learners' vocabulary comprises less than the 2,000 most frequently used words in English, with adjectives learned better than other parts of speech. This confirms findings reported in the literature (Milton, 2007) that there is a linear relationship between the frequency of words and students' knowledge, such that they retain high frequency words better than low frequency words. A number of ELT course books were then analyzed using the VocabProfile software program (Cobb, 2009). The vocabulary content of textbooks provides minimal chances for learners to develop their vocabulary, as they recycled a small number of vocabulary items and did little to expand their vocabulary beyond these.

Introduction

No one denies the essentiality of vocabulary knowledge as a key yet challenging component, across all stages of second language acquisition. It is without doubt, however, that learners will often encounter unknown words through a variety of authentic sources, among which are textbook and teaching material input, to help them tackle such lexical learning challenge. Empirical studies have shown that Saudi students graduate from high school with a relatively small number of words that they were supposed to acquire (Al-Akloby, 2001; Al-Hazemi, 1993; AlSaif, 2011). The aim of the current study is to discover how vocabulary input in textbooks relates to the Saudi learners' low level of vocabulary size. To the best of our knowledge, this is a pioneering study in the target context as it investigates the role of a yet unexamined potential factor, textbooks' vocabulary input, in the vocabulary learning of Saudi EFL learners at university level. The study also aims at measuring the vocabulary size of the participants in relation to the textbooks as well as examining the effect of other learnability factors such as word frequency, part of speech, and recycling.

Al-Hazemi (1993) investigated the vocabulary knowledge of students who had just finished their high school education. His work revealed that Saudi high school leavers and military cadets acquire about 1,000 words out of the 2,000 most frequently used words in English. A few years later, Al-Akloby (2001), relying on conclusions obtained by Al-Hazemi (1993), conducted an exploratory study to examine the factors behind the Saudi learners' failure to learn English vocabulary in Saudi public schools without formally establishing the actual size of their vocabulary knowledge, i.e. without employing vocabulary measurement tools. He reported a number of factors contributing to Saudi EFL public schools' vocabulary learning failure such as: inadequate use of vocabulary learning strategies, insufficient presentation of vocabulary, limiting the lexical presentation in textbooks to pronunciation and meaning, ineffective use of vocabulary recycling and testing, and some socio-psychological factors (e.g. attitude, motivation, classroom anxiety, and parental encouragement). Recently, AlSaif (2011) further confirmed previous findings where Saudi EFL learners in both public schools and university showed poor vocabulary knowledge. He concluded that the insufficient presentation of vocabulary in Saudi schools' textbooks might be responsible for learners not achieving the

vocabulary needed for adequate comprehension of written or spoken text.

Literature Review

Vocabulary knowledge has been perceived by educators and researchers as the most essential building block in authentic communication (Milton, 2009). It is the first step in learning any language as can be seen from Wilkins's (1972, p. 111) strong and famous claim that, "while without grammar very little can be conveyed, without vocabulary nothing can be conveyed". Nation (2001) has identified three dimensions with nine multiple components which must be known about vocabulary, including written and spoken form, meaning and association links, and grammatical characteristics, collocations, and contextual constraints on use like register and frequency. For each component, Nation categorized lexical knowledge into receptive and productive vocabulary knowledge. It is of great importance to establish a good understanding of the amount of vocabulary a foreign language learner needs in order to comprehend written or spoken discourse, that is, text more than one sentence long. For university EFL learners, knowing the most frequently used 2,000 words will give them the competence to communicate effectively in speaking and writing modes (Schmitt, 2010). Therefore, at the tertiary level, the task of learning a large vocabulary is quite different for EFL learners and 3,000 words is crucial to handle authentic texts (Nation, 2001; Schmitt, 2000). It is not realistic or even possible to include all vocabulary in the instructional materials at once. However, it is worth considering the above pickup rate in the first two years of studying the target language.

Fellow English instructors at university level share with us their worries that an overwhelming majority of Saudi EFL learners have a language problem due mainly to their vocabulary which is prone to attrition at the end of a two-year university language course (Alharthi, 2014). When prompted for the possible reasons for their deficit in vocabulary size and learning, textbook materials are given as one likely major source of difficulty. In many EFL environments, the textbook materials are the only writing of English learners have available to them, and hence their main source of vocabulary learning. This understanding has been empirically supported by Al-Akloby's (2001) study. He highlighted a number of sources of vocabulary learning breakdown in the Saudi English language curricula and textbooks including insufficient presentation of vocabulary items, which was also found to be confined to pronunciation and meaning only. Moreover, it was found that unproductive use of vocabulary recycling and testing was another source of vocabulary learning failure. Similarly, Al-Hazemi (1993) and AlSaif (2011) concluded that Saudi EFL

school leavers' vocabulary size was estimated at around or below 1,000 words, a finding far from the size identified by previous research and the Saudi Ministry of Education's target as an acceptable vocabulary level for reading authentic texts, i.e. 3,000 words. These serious issues motivated us to address the following questions:

- What is the vocabulary size of EFL university level students?
- Do EFL university level students attain an adequate number of high-frequency words enabling them to cope with the reading materials they are presented with?

As Schmitt (2014, p. 942) succinctly notes, "The message about the need for a large vocabulary size to be able to function well in English seems to be taking hold". If it is the case that the study participants' vocabulary knowledge is below the 3,000 word level based on our local vocabulary test, the next step will involve examining the vocabulary textbooks introduced to our participants at King Abdulaziz University (KAU). The present study investigates the size of vocabulary knowledge that university students gain from vocabulary textbooks and how the selection and presentation of vocabulary in these textbooks may influence their learning. That is, the study examines the impact of some potential factors including frequency, recycling and word class on the students' achievements in vocabulary knowledge. This would provide further information about the learnability of English vocabulary.

The Study

Research objectives and questions

Our study is predominantly quantitative with a view to measuring the vocabulary size of Saudi EFL learners at the tertiary level. In accordance with the previous literature, we expect that the participants in this study would score low in the vocabulary test. The current study also intended to examine in depth textbook input as a possible factor that might influence the participants' poor vocabulary knowledge. The main purpose of the present study was to provide answers to the following questions:

- How much vocabulary out of the most frequently used 3,000 English words do Saudi EFL students know at KAU?
- What is the effect of word class on vocabulary learning?
- How much vocabulary do the EFL textbooks under investigation present?
- What is the frequency and distribution of these words?
- How much is vocabulary recycled in the EFL textbooks?

- What is the effect of recycling on vocabulary learning?
- Is there a significant relationship between frequency, recycling and word class and the participants' achievement scores in a test of word knowledge?

Sampling and Method

The participants in this study are 40 male third-year university students studying in the Department of European Languages and Literature. The students are all Saudi and Arabic is their mother tongue. Recruiting participants from third year was necessary since our test includes items taught to the participants in their first two years. Also, the test was administered only a few weeks before the end of their formal instruction.

Data Collection Instruments

Local vocabulary test

The current study measured the students' receptive vocabulary knowledge by employing a vocabulary test constructed by the researchers. This test was in a pencil and paper form, modeled on the Vocabulary Levels Test (VLT) by Nation (2001). While the VLT tests the learners' vocabulary knowledge of the most frequent five levels in English, the Receptive Vocabulary Test RVT was designed to measure how many vocabulary items of the most frequent 3,000 words the participants know from their textbooks. The target words were based on the contents of the participants' instructional materials. The examinees were presented with 60 multiple choice items where learners were asked to choose the correct definitions of the items from another corresponding 240 definitions, half of which were distractors. The first two word levels had seven clusters while the last one had six. Therefore, the first two 1,000 word frequency levels tested the knowledge of 21 items; however, the third 1,000 word level had 18 items due to there being fewer words from this level in the textbooks. Both corresponding definitions and distractors were taken from the textbooks. However, definitions consisted of words from a higher frequency band to ensure that the learners were tested for the target items. The test minimized guessing by arranging the items alphabetically, based on word length, while keeping all the words in each cluster from the same class removed any grammatical clues. Moreover, the test can be administered in a short period of time because the definitions are in short forms. Participants were also told that the results of the vocabulary test would not affect their current studies in any way.

The vocabulary test was scored as follows: a correct answer was given one mark and a wrong one zero.

Textbook materials

The students in the English department at KAU are presented with a number of modules during their four years of study to get a BA in English Language. In the first two years, they attend two intensive vocabulary modules, *Vocabulary 1* and *Vocabulary 2.* Vocabulary is also introduced in three reading modules during the first three semesters. The textbooks under investigation are Nelson's (1986) *Intermediate Vocabulary* and Nelson's (1986) *Elementary Vocabulary*; these are introduced during the first and second semester in the English Department at KAU. We will pay special attention to the presentation and recycling of target words in these materials.

Both of the textbooks were scanned and digitized. The researchers used the PDF-XChange Viewer to convert the scanned images into editable text. The converted text was later examined and corrected to avoid any misspelling and/or changes (e.g. some letters were converted into numbers such as *w0rd* or *scho1l*). We then extracted content words presented in exercises and wordlists in each classroom lesson. These words were put into two different text files for each textbook. After a final thorough revision, all the items in both text files were merged into a third file.

In three separate steps, the researchers pasted each text file into the Vocabprofiler British National Corpus BNC-20 programme available on Cobb's (2009) *Compleat Lexical Tutor* website. The VocabProfiler BNC-20 analyzes the text and divides the words into 20,000 word frequency bands. The output of the programme contains the number of word families, types, token, type coverage, and cumulative coverage. In order to check how a word is repeated in these textbooks, we used another programme from Cobb's (2009) website, namely, Text-Lex Compare. The researchers pasted the wordlist gathered form the first textbook into the first window then pasted the wordlist from the second textbook into the other window. The Text-Lex Compare analyses two or more texts and gives the degree of repetition of words in these texts. The output of this programme includes the number of tokens and word families recycled in reference to another text, as well as to the amount of not shared or new words in each text.

Results and Discussion

1. How much vocabulary out of the most frequent 3,000 English words do Saudi EFL students know at KAU?

The results revealed that the estimated vocabulary size of Saudi EFL students at KAU, a few weeks before the end of their second year of studies, was around 1,447 words out of the 3,000 most frequent words in English. Table 1 shows the number of known words for each of the first three BNC 1,000 word levels and for the total number of words. The standard deviations and the minimum and maximum vocabulary scores indicate high variation in the students'

vocabulary scores for the words at each frequency

level, as well as for all word levels together.

Table 1. Number of words known at each word level and in total

BNC word frequency level	Mean	Std. Deviation	Minimum	Maximum
1000 word level	636.25	257.36	190.48	1000.00
2000 word level	481.49	295.04	0.00	952.38
3000 word level	330.25	276.97	0.00	833.33
Total	1447.98	776.71	285.71	2738.10

The participants' vocabulary size is therefore in harmony with the previous findings obtained by Al-Hazemi's (1993) study of high school leavers but smaller than in AlSaif's (2011) study of university students. However, comparing the results of this study with Al-Hazemi's might not be appropriate. The subjects in Al-Hazemi's study were high school leavers and military cadets who had a chance to study English for a period of two hours a week while the participants in our study were third year university students who studied English intensively for four semesters. In contrast, AlSaif (2011) concluded that Saudi EFL university students know an average of 2,628.57 words by the end of their second year. The higher estimate of the vocabulary size of Saudi EFL university students in AlSaif (2011) might be interpreted as a result of the vocabulary size test he employed. The current study measures the learners' vocabulary knowledge of the most frequently used 3,000 words presented to them in their textbooks. On the other hand, AlSaif (2011) uses a yes/no test designed to measure students' vocabulary size out of the 5,000 most frequently used words in English. Such a test format might have overestimated the vocabulary knowledge of AlSaif's (2011) participants.

2. What is the effect of word class on vocabulary learning?

The mean scores for each part of speech (i.e. nouns, verbs, and adjectives) show that adjectives are the most commonly known vocabulary items, followed by nouns and then verbs. This seems to contradict

current views and also the findings of Horst & Meara (1999) and Alharthi (2014). A possible reason for this finding is that the distribution of correct answers in one or more word classes is not normal and, therefore, the mean correct scores are not the right indicators of central tendency for the scores of the words from one or more word class. The Kolmogorov-Smirnov and Shapiro-Wilk normality tests as well as histograms and Q-Q plots revealed that scores for verbs and adjectives in the vocabulary test were not normally distributed. Table 2 displays the median scores for each part of speech.

Table 2: Mean and median scores for each part of speech

Part of speech	Mean	Median	Std. Deviation
Nouns	46.47	44.44	18.96
Verbs	43.52	36.11	25.39
Adjectives	52.77	47.22	23.99

As shown in Table 2, the median scores still suggest that adjectives are the most commonly known vocabulary items by the participants followed by nouns and then verbs. In addition, the unequal distribution of parts of speech in the RVT might also explain the order of the most commonly known parts of speech obtained in the current study. Table 3 summarizes the results of the RVT mean scores for each part of speech in each frequency band as well as the number of words for each part of speech in each frequency band and in total.

Table 3: Mean scores (%) for each part of speech and word frequency

Part of speech	Number of items		Frequency band	Mean	Std. Deviation
	All levels	Each level			
Nouns	33	12	1000	62.04	28.77
		12	2000	47.22	29.21
		9	3000	34.26	32.47
Verbs	9	3	1000	62.97	32.64
		3	2000	43.52	32.68
		3	3000	22.22	28.73
Adjectives	16	4	1000	80.56	25.44
		6	2000	52.32	35.67
		6	3000	35.65	32.41

The lack of equal presence of different parts of speech in the vocabulary test might provide more chances for error in nouns than in the remaining parts of speech. For example, the test has 12 nouns and 3 adjectives from the first 1,000 word frequency band. If participants respond correctly to only one adjective, they score 33.33% in adjectives from the first 1,000 word frequency level. On the other hand, participants have to know four nouns to achieve the same score. However, for this explanation to make sense, verbs should be the most commonly known part of speech, not adjectives, since only nine verbs were tested in the RVT. It could be argued that nine items would be insufficient for testing word knowledge of a certain part of speech, and vocabulary size tests should include a bigger number to test the vocabulary knowledge of that part of speech. In addition, the researchers thought of the possibility that the textbooks under investigation might have offered more adjectives than members of the other word classes. In order to find out the number of each part of speech in these textbooks, the researchers tried tagging the word lists from the textbooks with CLAWS4. The analysis showed that the wordlists generated from the textbooks total about 53.84% nouns, 33.45% verbs, and 12.73% adjectives. This finding does not explain the vocabulary test results. Based on the percentage of each part of speech represented in the textbooks, the order of the most commonly known words by part of speech should be in the order nouns, verbs, and adjectives. The present findings are in line with Laufer's (1997) alternative argument that part of speech has no clear effect on learnability.

3. How much vocabulary do the EFL textbooks under investigation contain? What is the frequency and coverage distribution of these words?

The analysis of each textbook shows up the lack of sufficient presentation of the 2,000 most frequent words in English. The cross-check of the first textbook, Nelson's (1986) *Elementary Vocabulary*, with Vocabprofile BNC-20 shows that there are 1,116 tokens or 760 word families from the BNC. Only 544 tokens or 304 word families are from the first 1,000 frequency band while 280 tokens or 215 word families are from the second 1,000 frequency band. The distribution of words in the remaining frequency bands decreases with successive frequency bands. For example, there are 141 tokens or 115 word families from the third 1,000 frequency band while there are only 38 tokens or 35 word families from the fifth frequency band. In sum, the number of words presented in the first textbook from the 2,000 most frequently used words is 824 tokens or 519 word families which together comprise 73.84% out of the 20 most frequent 1,000 BNC bands.

Milton (2009, p. 207) suggests that presenting a large amount of vocabulary in textbooks would enable learners to learn more vocabulary autonomously because it appears that the "more vocabulary that is presented in course books, the more vocabulary learners seem to acquire. Learners do not appear to get overloaded in this area of acquisition". Clearly, both the amount and how commonly the words are used in English are not adequate and fall far short of the suggested figure by Nation (2001). The first textbook is designed as an elementary proficiency textbook; however, this does not justify the presentation of only about 41.2% of the 2,000 most frequently used words which might not help learners achieve adequate comprehension.

The examination of the second textbook, Nelson's (1986) *Intermediate Vocabulary*, shows similar results. The total number of words in the second textbook is 1,263 tokens or 1,006 word families. Strikingly, only 367 tokens or 251 word families are from the first 1,000 frequency band while only 273 tokens or 204 word families are from the second 1,000 frequency band. In total, a dismaying figure of 640 tokens or 455 word families from the 2,000 most frequently words of English are presented in the first textbook. The representation of the 2,000 most frequently used words of English in the second textbook, which claims to be targeting intermediate proficiency learners, is also extremely insufficient. Only 32% of the 2,000 most frequently used words are represented, a percentage which is even lower than the one in the first textbook. Similar to the first textbook, the distribution of the words in the remaining frequency bands decreases for each band of less frequent words.

Merging the lexical input of both textbooks into one file and cross-checking it with the Vocabprofiler BNC-20 gives a complete picture of the amount of vocabulary these textbooks present. The examination revealed that these textbooks offer 2,379 tokens or 1,493 word families. More specifically, taken together these textbooks present 1,464 tokens or 782 word families from the 2,000 most frequently used words. Moreover, the textbooks include only 344 tokens or 257 word families from the third 1,000 frequency band. To draw a conclusion based on the test scores, there seems to be a direct impact of the presentation of new words and their selection based on frequency bands. In relation to the most frequent 3,000 words, the percentage of vocabulary presented in these textbooks is 38.29% from the first 1,000 frequency level while 23.25% and 14.46% words are presented from the second and third 1,000 frequency bands, respectively. The frequency distribution in these textbooks might be responsible for the pattern we arrived at above, where most known words come from

the first 1,000 frequency band. The examination of the textbooks shows that the Saudi EFL students at KAU might not have enough vocabulary presented to them. It is possible to conclude that these textbooks are not appropriate for this level of English students and would not enable them to have a sufficient comprehension of written text. Moreover, the two textbooks are used in 75 classroom teaching hours during the first year which means students need to learn 31.72 words per hour in order to learn all the words in the textbooks. This might well be another possible cause of the apparently low vocabulary knowledge of our participants. There should be more time allotted for vocabulary learning thus providing more chances of learning.

4. How much vocabulary recycling is in the EFL textbooks? What is the effect of recycling on vocabulary learning?

In order to investigate the recycling of the textbooks under investigation, the researchers copied the vocabulary of each textbook and inserted it into the programme Text-Lex Compare, available on the *Compleat Lexical Tutor* website (Cobb, 2009). Text-Lex Compare calculates recycling by counting the number of word repetitions in one text file as well as processing two text files at the same time. The mean percentage of the repeated or recycled words shared in the two textbooks is 22.09%, with 279 tokens having been repeated more than once. In total, the number of recycled words which occur in both textbooks is 469 out of 2379 tokens, none of which is from our test items. Therefore, we cannot run a statistical analysis to measure the correlation between recycling and correct test answers. However, in a recent study, AlSaif (2011) found a positive correlation (0.590) between recycling and learnability, which supports the hypothesis that the more often a word is encountered in a textbook, the more likely it is to be learned. Therefore, the lack of recycling in the textbooks might explain our participants' poor achievement scores in the vocabulary test.

5. Is there a significant relationship between frequency, recycling and word class and the participants' achievement scores in word knowledge?

As mentioned above, the examination of the two textbooks shows that no test item was recycled. Therefore, it is not possible to investigate the effect of recycling on vocabulary learnability. Consequently, we carried out analyses to examine the impact of frequency and part of speech on learnability. There was only a negative correlation between the participants' test scores and frequency *(r = -0.595, p < 0.01)*, which is expected since the less frequent a word is, the harder it is to learn it. We can conclude, based on these results, that frequency stands out as the only factor affecting the learnability of vocabulary. The remaining difficulty factors, such as part of speech and recycling, have no clear impact on the extent to which a word becomes easier or harder to learn.

Conclusion

Measuring the Saudi EFL learners' vocabulary size reveals similar findings compared to previous studies (Al-Hazemi, 1993; AlSaif, 2011). The vocabulary size of our participants is around 1,447 words out of the most frequent 3,000 words. Frequency alone shows a statistically significant effect on words learned, whereas part of speech and recycling do not have a significant impact. There is a significant correlation between frequency and test scores. This is predicted since frequently used words are easier to learn than infrequently used words. In general, frequency accounts for 35.4% of test scores variance.

The examination of the vocabulary textbooks reveals a lack of sufficient lexical input. The two textbooks offer 1,493 word families in total, with around half that number coming from the 2,000 most frequently used words, and an even smaller number of words in bands of less frequently used words. In addition, recycling is almost absent in these textbooks because only 19.71% of words have been recycled. In fact, not one word from the vocabulary test was among recycled words which prevented us from further investigation of the effect of recycling on learnability. To conclude, the findings of this study are in agreement with Milton (2009), where only frequency shows a clear effect on the difficulty of vocabulary learning.

References
1. Al-Akloby, S. (2001). *Teaching and learning English vocabulary in Saudi Arabian public schools. An exploratory study of some possible reasons behind students' failure to learn English vocabulary.* University of Essex, Unpublished PhD thesis.
2. Al-Hazemi, H. (1993). *Low-level EFL vocabulary tests for Arabic speakers.* University College of Swansea, Unpublished PhD Thesis.
3. Alharthi, T. (2014). The impact of attrition on vocabulary knowledge among Saudi teachers. *English Language Teaching,* 7(4), 111-118.
4. AlSaif, A. (2011). *Investigating Vocabulary Input and Explaining Vocabulary Uptake among EFL Learners in Saudi Arabia.* Swansea University, Unpublished PhD Thesis.
5. Cobb, T. (2009). *The Complete Lexical Tutor.* (Version 6.2) [Computer software]. Retrieved 14.7.2012, from http://www.lextutor.ca/.
6. Horst, M. & Meara, P. (1999). Test of a model for predicting second language lexical growth

through reading. *Canadian Modern Language Review/La Revue canadienne des langues vivantes, 56*(2), 308-328.

7. Laufer, B. (Ed.). (1997). *What's in a word that makes it hard or easy: Some intralexical factors that affect the learning of words*. Cambridge University Press: Cambridge.

8. O'Loughlin, R. (2012). Tuning into vocabulary frequency in coursebooks. *RELC,* 43, (2), 255-269.

9. Milton, J. (2007). *Lexical profiles, learning styles and the construct validity of lexical size tests.* In Modelling and Assessing Vocabulary Knowledge (Eds.) H. Daller, J. Milton and J. Treffers-Daller (pp. 47-58). Cambridge: Cambridge University Press.

10. Milton, J. (2009). *Measuring second language vocabulary acquisition.* Cambridge University Press: Cambridge.

11. Nation, P. (2001). *Vocabulary learning in another language*. Cambridge University Press: Cambridge.

12. Nelson, T. B. J. (1986a). *Elementary Vocabulary.* Thomas Nelson and Sons Ltd: Malaysia.

13. Nelson, T. B. J. (1986b). *Intermediate Vocabulary.* Thomas Nelson and Sons Ltd: Malaysia.

14. Schmitt, N. (2000). *Vocabulary in Language Teaching.* Cambridge University Press: Cambridge.

15. Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual.* Palgrave Macmillan: New York.

16. Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning,* 64(4), 913-951.

17. Wilkins, D. A. (1972). *Linguistics in language teaching* Arnold: London.

8/26/2015