

The Relationship between Iranian EFL Learners' Reading Comprehension, Vocabulary Size and Lexical Coverage of the Text

Nader Assadi Aidinlou

Young Researchers and Elite club, Ahar Branch, Islamic Azad University, Ahar, Iran

Mitra Taghinezhad Vaskehmahalleh *

Young Researchers and Elite club, Ahar Branch, Islamic Azad University, Ahar, Iran

Abstract

This study was an attempt to investigate the relationship between EFL learners' vocabulary size, lexical text coverage and their reading comprehension. To fulfill the purpose of this study, 120 male and female students studying at Talesh Azad University were selected based on their performance on Nelson Proficiency Test. Then a Nelson reading proficiency test was administered after the Nelson proficiency test in order to homogenize learners in their reading proficiency. After that, learners' vocabulary size was measured by the Levels Test. Their lexical coverage of text also was measured by the newest version of Vocabulary Profile and reading comprehension by a teacher made test. Results of the study showed that there was strong relationship between reading comprehension, vocabulary size and lexical coverage of the text.

Keywords: reading comprehension, vocabulary size, lexical coverage

INTRODUCTION

Shaywitz (2003), points out reading process is the royal road to knowledge, it is essential to the success in all academic subjects. Furthermore, he states that reading comprehension is an important life skill and it is one of the most important domains in education, because it is the best predictor of success in higher education and job performance. Reading can be challenging, particularly when the material is unfamiliar, technical, or more complex. It is the common experience of EFL teachers that, most students fail to learn to read efficiently and adequately in the target language. Many students fail to conceptualize reading as a search for meaning, so they have a lot of problems during the study. They can read variety of texts with accuracy and fluency but cannot demonstrate an understanding of what they have read. This is often indicated by an inability to recall key information from the text, to retell the story or answer questions related to it. Students with reading problems tend to be less aware of text

structure and have poorer recall of textual ideas than good readers (Fitzgerald, 2003). It is very important that the reader is aware that the purpose of reading is to make meaning of the text being read, not just be able to decode it (Hedin & Conderman, 2010). Furthermore, Stanley (1984) claimed that, one of the L2 readers' problems is that they may view texts as samples of language rather than information and he claimed that L2 readers studied the text less efficiently than L1 readers. Even when students read a text at a basic level, the level of comprehension is typically shallow and lacks the necessary depth for adequate understanding (Best et al., 2005).

The prevalence of reading skill among other skills in academic context and the way EFL/ESL learners with different abilities focusing on their vocabulary knowledge react to the academic reading comprehension texts can be a source of variation in this regard (Nation 2011 & Laufer et.al, 2010). Concerning this issue, in recent decades studies on the relationship between vocabulary knowledge, the size or the number of the words a person knows and reading comprehension have drawn attention of the scholars in language learning and teaching research such as Coxhead, 2000; Hu and Nation, 2000; Kurnia, 2003; Laufer and Ravenhorst-Kalovski, 2010; Nation, 2006, 2013 and 2015; Schmitt et. al., 2011 in native and second language, and Amirian et. al., 2015; Farvardin, 2013; Mehrpour, Razmjoo, and Kian, 2011; and Rouhi, 2013 in EFL context.

Many of these studies have been conducted on the relationship between vocabulary knowledge (depth and breadth) and the learners' reading comprehension in the specific texts such as fictions, short stories, articles in newspapers, and the ESP texts in first language (L1) and second language (L2), and only on the ESP texts in foreign language context (FL). As in academic contexts English language is as a means for getting information from textbooks and online articles and papers, ESL/EFL learners need high proficiency in reading skill. However, EFL intermediate students even at universities have difficulties in reading and understanding English texts adequately. Regarding this problem, in a foreign context such as Iran where EFL learners at intermediate level particularly students with majors other than English language have problems in reading comprehension and complain about the unfamiliar vocabularies that impede their comprehension, the researcher of the present study attempts to measure the learners' actual vocabulary knowledge concerning the lexical size (breadth) of the learners at intermediate level. In addition, with respect to the significant role of learners' vocabulary size in reading comprehension, another variable, text coverage that is the percentage of words that a reader understands in a text, should be measured while focusing on the effect of vocabulary knowledge on learners' reading comprehension performances (Laufer et.al, 2010).

In this regard, the present study not only seems to be in line with all the previous studies carried out in L1, L2, and EFL contexts that showed the relationship between lexical size and the learners' reading comprehension performances, it is also based on the assumption that the main purpose of emphasizing on reading skill in Iranian universities are training and promoting technical reading comprehension skills in English and instructing students how to deal with technical texts relevant to their

majors. One of the fundamental prerequisite in this regard would be the deployment of appropriate procedures to evaluate learners' the actual vocabulary knowledge concerning their lexical size, its relationship with the percentage of the text coverage, and reading comprehension ability.

RESEARCH QUESTIONS

Based on the theoretical frameworks of the study, the present study tried to address and find answers to the following questions.

- RQ1: Is there any relationship between EFL learners' reading comprehension, vocabulary size, and lexical coverage of the text?

METHOD

Design of the Study

In any research field, the approach adopted by the researcher to answer a research question or to test a hypothesis should be appropriate enough to yield the most plausible and reliable evidence. After surveying the area of language learning and teaching research, more specifically, language classroom research, and taking into consideration the aim of this research which is designed to assess the relationship between the learners' lexical knowledge and the other crucial factor such as text coverage, lexical size, and their reading comprehension scores. All in all, as it is clear, the design of this study is a descriptive one. Descriptive studies are conducted to demonstrate relationships between things in the world around you and also referred to as "correlational" or "observational" studies (Bickman & Rog, 1998).

Participants

A total of 180 male and female pre-intermediate level students took part in this study, all of them were students of university in Talesh-Iran were studying in different departments, majors other than English language, and were doing a general reading course for academic purpose, a course that aims at improving the learners' reading comprehension ability in English that makes them ready for taking the following academic ESP courses. Prior to the university, they studied English for seven years at high school. Many of them did not continue their study immediately after school. All of them were Iranian foreign language learners. There were no age and sex limitations in the study.

Sampling Procedure

As far as this study was concerned, 180 participants were randomly selected among 250 EFL learners of university. Then the researcher determined their homogeneity through Nelson proficiency test and 120 participants whose scores fall one SD above and below the mean were selected to take part in this study. This test was first piloted with 30 students to check its reliability and then the test was implemented for the purpose of homogenizing the sample of the study and to make sure that the study enjoys

homogeneous and identical participants with respect to the participants' English language proficiency.

Instruments

The Nelson proficiency test, new version of vocabulary Levels Test (Nation 2006; revised by Beglar, 2009), a Nelson reading comprehension test, a vocabulary profiler (IBM program) to measure text coverage, and two reading comprehension tests of different text types were used in this study.

Nelson Proficiency Test

Nelson test was administered to the experimental and control groups to ensure their homogeneity regarding their proficiency level. The validity and reliability of the Nelson test have been estimated several times before by other researchers and it is considered as highly valid test of English proficiency (Shahivand & Pazhakh, 2012). The test was implemented for the purpose of homogenizing the sample of the study and making sure that the study enjoyed homogeneous and identical participants with respect to their English language proficiency. The reliability of the test then was calculated as 0.87 based on Cronbach's alpha coefficient method which is an acceptable reliability.

Nelson Reading proficiency test

The first instrument was a Nelson reading proficiency test (Brown, Fishco, & Hanna 1993). This test consisted of 30 multiple choice items in which three passages were used. Learners were asked to answer the questions in 45 minutes. This test was first piloted with 30 students with similar characteristics to that of the main participants of the study to check its reliability and then the test was implemented for the purpose of homogenizing the sample of the study and to make sure that the study enjoys homogeneous and identical participants with respect to the participants' English language proficiency. The reliability of the test then was calculated as 0.79 based on KR-21 method which is an acceptable reliability.

Vocabulary Levels Test

Learners' vocabulary size was measured by the new version of Vocabulary Levels Tests, 20 vocabulary frequency lists including 20,000 most frequent words (Nation 2006, Nation & Beglar, 2007) which was validated by Beglar (2009). According to the level of the participants, intermediate, the researcher chooses items from the 2000, 3000, and 5000 most frequent words. Each word in Vocabulary Level Test represents a word family, (the word, its inflections and derivations). Each frequency level test includes 10 items and each item represents knowledge of 100 words. Every correct answer receives one point, an incorrect answer or no answer receives 0 points. Due to the participants' level, intermediate, they were not given the 10,000 level as it is too difficult for them. Test is not a precise measure of vocabulary size, it can be considered as a tool to measure the learners' knowledge of items from particular levels.

Lexical Coverage of Texts

To measure the lexical coverage of the texts, the researcher used a new version of the vocabulary profiler which matches a text to 20 vocabulary frequency lists created on the basis of the British National Corpus (BNC). This lexical profiler (IBM program) is available at Paul Nation's website (<http://www.victoria.ac.nz/lals/staff/paulnation/nation.aspx>) and at Tom Cobb's site (<http://lextutor.ca>).

Two reading comprehension texts with different genres entered into the computer, the program produced a list for each reading text, and it showed what percentage of the text is covered by each word frequency list. Words that were not included in the 20,000 most frequent vocabulary appears in a list as "off list" words. Recently, a special function is added by Tom Cobb on <http://lextutor.ca> site which allows the user to analyze a text in a way that all proper nouns included in the texts which are personal and geographical names (e.g. Smith, Paris) are reclassified and appeared in the first thousand most frequent words. It is based on the assumption that these proper nouns do not belong to the lexicon of a particular language, and if the reader is not familiar with them, the comprehension problems cannot be considered as the lack of lexical knowledge or lexical unfamiliarity. However, proper nouns which are as regular words, (e.g. Eiffel Tower, Syntactic Argumentation) appear in their related frequency lists. Therefore, the researchers analyzed each text twice: once with the new function that most of the proper names appear in the first most frequent word list (K1); and once without it, in this case most proper names appear in the "off list" words. Then the researcher calculated the difference in the number of the tokens of proper names between the two "off list" lists and converts it into percentage out of the total number of tokens. In this way, the percentage of proper names in the text obtained.

Reading Comprehension Test

Reading comprehension test first being revised by the researcher and two qualified English professors, the test was first piloted among 30 students for the purpose of calculating the reliability of the test. The results represented that the mean was 29.66 and the SD was 4.97. The reliability of the test then was calculated as 0.81 based on KR-21 method which is an acceptable reliability.

Procedure

The entire study took two week (four sessions). The first session devoted to Nelson proficiency test. Then Vocabulary Level Test was administered to the selected participants based on Nelson proficiency test (pre-intermediate level learners) for session two. In the third session, a Nelson reading comprehension test to observe the students' reading comprehension ability.

As the last measurement (session four) to observe the students' abilities in reading comprehension of different text types, each reading comprehension text type (narrative, informative) which contains reading passages followed reading comprehension questions, some measures the learner's comprehension of general English knowledge,

some questions focus on the understanding of words, and some on the understanding of global textual information(explicit and implicit information) were given to the participants, the obtained scores on reading comprehension tests with narrative genre and reading comprehension test with argumentative genre the were compared by the vocabulary profiler compared with the same data received to show when we used different text types at the same level concerning the learners’ vocabulary knowledge by focusing on the most frequency general words, whether the same results were obtained, that is; the scores of the participants in reading comprehension test were higher, lower, or equal.

Analysis: Analyzing Reading Test Based on Vocabulary Size and Lexical Coverage

The English reading texts with different genres were analyzed in terms of the percentage of coverage of each BNC (British National Corpus) frequency list. Table 1 shows the coverage of 10 lists. The proper nouns have not been categorized by the special function to be included in the K1 list, but are distributed among all lists. Thus the proper nouns which are personal and geographical names are included in the “off list” words.

Table 1. Coverage of the Reading Comprehension Test by BNC frequency lists

Frequency level	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10-20	List
Coverage 0% (test 1)	80.15	9.30	2.49	2.49	0.74	0.80	0.39	0.48	0.17	1.12	2.19
Coverage 0% (test 2)	79.55	.92	3.11	2.58	1.09	1.13	0.69	1.11	0.21	0.98	2.63
Average Cumulative Coverage	78.58	86.66	90.56	92.81	94	94.7	96.1	96.47	96.59	95.63	~100
										OFF	

Table 1 presents the percentage of proper names calculated following their categorization as the first 1,000 words. If it is assume that the proper names are familiar to the learner, then the 95% coverage can be achieved with knowledge of 4,000 words, which cover almost 93% and the proper nouns which cover an additional 2.1%. 98% coverage can be reached by knowledge of 7,000–8,000 and the proper nouns .

Since the aim of the paper was to find out the relationship between vocabulary size, coverage and reading scores, the researcher presented the combined data on the coverage data with the data on learners’ vocabulary size and the reading score. As mentioned earlier, in the section three on measuring vocabulary size, it divided the learners by intervals of 1,000 words. In Table 2 the BNC list is replaced with learners’ vocabulary size. If, for example, 5,000 words cover 94 % of a text, then learners with knowledge of 5,000 words can understand a similar percentage of this text. As mentioned in the section on measuring reading comprehension, the raw scores of reading are out of 20.

Table 2. Vocabulary size, lexical coverage and Reading comprehension Test

Approximate Vocabulary Size	1,000	2,000	3,000	4,000	5,000	6,000	7,000	8,000
Lexical Coverage	78.50	83.60	89.88	91.00	92	91	94.63	93
Narrative test Mean	13.23	12.68	12.83	14	15.90			14.97
(SD)	(3.65)	(.027)	(2.399)	(4.48)	(3.86)			(2.75)
No of students	33	23	18	28	10			8

Testing Research Hypotheses

RQ1: Is there any relationship between EFL learners' reading comprehension, vocabulary size, and lexical coverage of the text?

Table 3. Pearson Correlation; reading comprehension, vocabulary size, and lexical coverage of the text

		Reading Comprehension	
Vocabulary Size	Pearson Correlation	.537**	
	Sig.(2-tailed)	.000	
	N	120	
Lexical Coverage	Pearson Correlation	.493**	.473*
	Sig.(2-tailed)	.010	.015
	N	120	120

** . Correlation is significant at the 0.01 level (1-tailed).

As table 3, there was significant relationship between reading comprehension, vocabulary size, and lexical coverage of the texts. Thus the first null hypothesis as, "There is not any relationship between EFL learners' reading comprehension, vocabulary size and lexical coverage of the text" was rejected.

DISCUSSION AND CONCLUSION

The aim of the study was to explore the relationship between text coverage, vocabulary size of the learners, and reading comprehension. Considering the proposed questions of this study, the answers to each of them will be presented and discussed below.

RQ1: Is there any relationship between EFL learners' reading comprehension, vocabulary size, and lexical coverage of the text?

Pearson correlation was run to test the first null hypothesis; the reading tests' scores were correlated with learners' vocabulary size. As it is evident from the results, there was significant relationship in the scores of the participants in the reading comprehension test with their vocabulary size and lexical coverage of the text. More specifically the first null hypothesis, "There is not any relationship between EFL learners' reading comprehension, vocabulary size and lexical coverage of the text" was rejected. The results found in this study are in accordance with corpus-based studies (e.g. Nation, 2006), which show that the less frequent the vocabulary, the smaller the portion of text coverage.

However, as the result of the study indicated, the relationship between vocabulary size and reading implies that even a small increase in lexical coverage may be just as beneficial to reading as a larger increase in coverage. The results are in accordance with Laufer and Nation's (2001) study which, explored the relationship between vocabulary size and speed of decoding word meaning and found that speed on a particular word frequency level increased only when learners' vocabulary size progressed far beyond that level. This means that the participants with a large vocabulary read more fluently the frequent words in the text, which may have given them an overall advantage over the learners with a smaller vocabulary, who had not yet attained a similar level of fluency.

Hu and Nation (2000) also investigated the relationship between lexical coverage and reading comprehension. The conclusion of their study showed that 98% is the lexical coverage for adequate comprehension. The results are also in line with the results of the study conducted by Laufer (1992). His concludes have practical implications for syllabus designers to set vocabulary goals on the basis of the comprehension level expected of learners. Thus, here again, we can see how the notion of vocabulary threshold is contingent upon what is considered "reasonable" or "adequate" comprehension .

In the present study, the researchers combined data on the lexical coverage of two different kinds of texts with learners' vocabulary level. Since the texts that were analyzed for coverage were of similar nature and practically identical difficulty to the texts learners were examined on, it could find out how the reading scores on the tests were associated with coverage and with learners' vocabulary knowledge. Hence, the study contains elements from Laufer (1989, 1992), Hu and Nation (2000) and Nation (2006).

It cannot be claimed that reasonable reading comprehension cannot occur if learners have not reached the lexical coverage, or that the coverage will automatically yield good reading comprehension. In our data there were learners who did not fit the general pattern of "better vocabulary leading to better reading." The general reading skills of these students may have affected the reading score more than their vocabulary knowledge. As for the relationship between vocabulary size and coverage, there are texts, can be reached with a smaller vocabulary than suggested here. Conversely, in some texts with a large proportion of technical and jargon vocabulary, the above coverage may require the knowledge of more low frequency words than suggested in the paper. However, when people read in the area of their expertise, they are usually more familiar with the jargon than with general vocabulary (Cohen et al., 1979). Therefore, when researching reading for general and academic purposes, it is useful to look at academic argumentative prose of general nature.

ACKNOWLEDGEMENT

This work was supported by the Young Researchers and Elite Club, Ahar Branch, Islamic Azad University, Ahar, Iran.

REFERENCES

- Amirian, M. R., Salari, S., & Heshmatifar, Z. (2015). A Validation Study of the Newly-developed Version of Vocabulary Size Test for Persian Learners. *International Journal of Education and Research* (3), 8, 359.
- Beglar, H., & Pazhak, K. (2009). Choosing Words to Teach. In E. H. Hiebert & M. L. Kamil (Eds.), *Teaching and learning vocabulary: Bringing research into practice* (pp. 209–222). Mahwah, NJ: Lawrence Erlbaum Associates.
- Best, R. M., Rowe, M. P., Ozuru, Y., & McNamara, D. S. (2005). Deep-level comprehension of science texts: The role of the reader and the text. *Topics in Language Disorders*, 25, 65–83.
- Bickman, A., & Rog, I. (1998). Size and Sequence in Vocabulary Development. In E. H. Hiebert & M. L. Kamil (Eds.), *Teaching and Learning Vocabulary: Bringing research into practice* (pp. 223–242). Mahwah, NJ: Lawrence Erlbaum Associates.
- Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, 34, 213–238.
- Farvardin, M.T., & Koosha, M. (2013). The Role of Vocabulary Knowledge in Iranian EFL Students' Reading Comprehension Performance: Breadth or Depth? *Theory and Practice in Language Studies Journal*, 1, (11), 1575–1580.
- Fitzgerald, J. (2003). Multilingual reading theory. *Reading Research Quarterly*, 38, 118–122.
- Goulden, R., Nation, P., & Read, J. (1990). How Large Can a Receptive Vocabulary Be? *New York, Newbury: Funk and Wagnall*.
- Hedin, L. R., & Conderman, G. (2010). Teaching students to comprehend informational text through rereading. *The Reading Teacher*, 63(7), 556–565.
- Hu, M., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13, 403–430.
- Kurnia, N. (2003). *Retention of Multi-Word Strings and Meaning Derivation From L2 Reading*. Unpublished Doctoral Dissertation, Victoria University of Wellington, New Zealand.
- Laufer, B. & Ravenhorst-Kalvoski, G. (2010). Lexical Threshold Revisited: Lexical Text Coverage, Learners' Vocabulary Size and Reading Comprehension. *Reading in a Foreign Language Journal*, 22 (1), 15–30.
- Mehrpour, S., Razmjoo, S. A., & Kian, P. (2011). The Relationship between Depth and Breadth of Vocabulary Knowledge and Reading Comprehension among Iranian EFL Learners. *Journal of English language teaching and learning*, 2(222), 97–127.
- Nation, I. S. P. (2001). *Learning Vocabulary in another Language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2013). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, 63, 59–82.
- Nation, I. S. P. (2011). *Teaching and Learning Vocabulary*. Boston: HeinleCengage Learning.
- Nation, P., & Beglar, D. (2006). A vocabulary size test. *The Language Teacher*, 31(7), 9–13.

- Nation, P., Coxhead. A. , & Sim. D. (2015). *Measuring the Vocabulary Size of Native Speakers of English in New Zealand Secondary Schools*. Newsland Association for Research in Education.
- Rouhi, M., & Mousapour, G. (2013). EFL Learners' Vocabulary Knowledge and Its Role in Their Reading Comprehension Performance. *Second and Multiple Language Acquisition Journal*, 1 (2), 39-48.
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The Percentage of Words Known in a Text and Reading Comprehension. *The Modern Language Journal*, 95(1), 26-43.
- Shaywitz, S. (2003). *Overcoming dyslexia*. New York: Alfred A. Knopf.
- Spearritt. D.(1972). Identification of Sub-Skills of Reading Comprehension by Maximum Likelihood Factor Analysis. *Research Report Series*, 1.
- Stanley, Zh, (1984).The Roles of Depth and Breadth of Vocabulary Knowledge in EFL Reading Performance. *Asian Social Science Journal*, 4(12), 138-142.