Modern English Education, Vol. 18, No. 2, Summer 2017
http://dx.doi.org/10.18095/meeso.2017.18.2.02

The Effect of Korean University Students' Lexical Size and Coverage in Reading Comprehension in English

Heejeong Ji*
Pukyoung National University
Yousun Shin
Kyungsung University

Ji, Heejeong & Shin, Yousun. (2017). The effect of Korean university students' lexical size and coverage in reading comprehension in English. *Modern English Education, 18*(2), 25-43.

This study investigated whether Korean university students' vocabulary size and their proficiency levels affect their reading comprehension in an EFL environment. Additionally, the study attempted at exploring to what extent vocabulary knowledge determines 'adequate' reading when lexical coverage are considered. For this purpose, vocabulary size was measured by the revised version of Vocabulary Levels Test (Schmitt, Schmitt, & Clapham, 2001). Reading comprehension items were adopted from commercially published TOEIC practice test items. Lexical coverage was calculated by using Vocabulary Profile. A total of 50 students who have existing TOEIC scores took part in the study, and none of them majored in English. The combination of regression, correlation coefficient, and ANOVA analysis were carried out to analyze the collected data. The results revealed that vocabulary size and reading comprehension scores were highly correlated. Evidently, Korean university students need to reach approximately 5,000 word families in order to be an 'adequate' reader in an academic setting, which resulted in the word coverage of around 97% including proper nouns. Pedagogical implication and further research directions are also suggested.

**[lexical size/lexical coverage/L2 reading comprehension/
어휘크기/어휘점유율/제2언어읽기이해력]**

* First author: Heejeong Ji, Corresponding author: Yousun Shin

I. INTRODUCTION

Reading is widely recognized as one of the most important skills for academic success, both in first language (L1) and L2 environments (Johns, 1981; Rosenfeld, Leung, & Oltman, 2001; Sherwood, 1977; Snow, Burns, & Griffin, 1998). Among many factors, vocabulary knowledge is a good predictor of reading, if not the best (Bernhardt & Kamil, 1995; Laufer, 1992; Nation 2001, 2006; Qian, 1999, 2002; Ulijn & Strother, 1990). So far, it has been widely recognized that the acquisition of vocabulary is essential for successful second language use and plays an important role in the formation of complete spoken and written texts (Laufer & Nation, 1999; Read, 2000). Besides, Schmitt (2000) emphasizes that "lexical knowledge is central to communicative competence and to the acquisition of a second language" (p. 55). Likewise, a lot of studies have shown that L2 readers rely heavily on vocabulary knowledge and the lack of that knowledge is the main and the largest obstacle for L2 readers to overcome (Huckin & Bloch, 1993).

Therefore, it is necessary to ensure that L2 learners have sufficient vocabulary to read well (Grabe, 2009; Hudson, 2007; Koda, 2005). Many educators and researchers reached the consensus that L2 learners need sufficient vocabulary to be an adequate reader. If so, the question lies in how much vocabulary needs to read fluently enough to well understand and perform academic tasks in L2. To find out the exact answer mentioned above, we must understand lexical threshold, the minimal vocabulary which is necessary to be an adequate reader in L2. Much research has already focused on this issue encompassing 'lexical threshold' (Hsueh-Chao & Nation, 2000; Laufer, 1989, 1992; Laufer & Ravenhorst-Kalovski, 2010; Nation, 2006). However, the results of the studies were different depending on the target. Also, there are only a few studies on this topic for Korean university students.

Therefore, this study aimed at exploring the main issues about how vocabulary knowledge affects Korean EFL university students' reading comprehension in English when L2 learners' general proficiency levels represented by TOEIC scores are considered. As a result, the study is expected to enable Korean EFL researchers and educators to recognize the interplay of vocabulary knowledge and L2 proficiency in a higher educational context and to draw some empirical implications from the current study. It is expected that Korean EFL language educators will be able to take advantage of the informative findings to provide better teaching methodology in reading to L2 learners.

II. LITERATURE REVIEW

1. Vocabulary and Reading Comprehension

Reading comprehension is interrelated to many components, including those involving language proficiency (e.g., grammatical knowledge and awareness of discourse structure),

the text itself, (e.g., text length, text difficulty, and topic), and those concerning the reader (interest in topic, motivation, amount of exposure to print, purposes for reading, L1 reading ability, and inferencing ability) (Droop & Verhoeven, 2003; Grabe, 2009; E. Nam, 2012; van Gelderen et al., 2004; van Gelderen, Schoonen, Stoel, de Glopper, & Hulstijn 2007). Among them, however, vocabulary knowledge has been considered vital in reading. Previous research has confirmed that there exists a strong relationship between vocabulary knowledge and reading comprehension.

Laufer (1992) explored the connection between vocabulary size and reading with university students in Israel. The results showed that highly significant correlations were found between reading and vocabulary scores, which was 0.50 between reading and the Vocabulary Levels Test and 0.75 between reading and the Eurocentres Vocabulary test. Qian (1999) investigated the roles of depth and breadth of vocabulary in reading comprehension with 74 ESL learners who reached to a minimal vocabulary size of 3000 word families. Participants consisted of 41 Korean speakers and 33 Chinese speakers staying in Canada. The results showed that the correlation between VS (vocabulary size test) and RC (reading comprehension) was 0.78, and the one between DVK (depth of vocabulary knowledge) and RC was 0.82. Laufer and Ravenhorst-Kalovski (2010) studied the relationship between vocabulary size and reading comprehension again. 745 college students in Israel participated in this research. It revealed that the correlation was 0.8.

Within educational context in Korea, M. O. Min and S. W. Lee (2016) investigated the effect of size and depth of vocabulary knowledge of 40 Korean EFL university students in their reading comprehension. All of the participants were from English education department. The participants' approximate vocabulary size was 7000, more than 5000 word families. Also, it showed that the size and the depth of vocabulary knowledge were correlated with reading comprehension, except for academic vocabulary. The correlation coefficient of the 2000, 3000, 5000, and 10000 words with RC was 0.401, 0.505, 0.455, and 0.482, respectively. The coefficient correlation of the depth and RC was 0.486. All of vocabulary sizes and depth had a significant difference with RC.

2. Lexical Threshold and Coverage for Adequate Reading Comprehension

There are two well-known hypotheses on the relationship between L1 and L2 reading. One of them is the linguistic interdependence hypothesis and the other is the linguistic threshold hypothesis. The linguistic interdependence hypothesis insists that L1 reading ability transfer to L2 reading. On the other hand, the linguistic threshold hypothesis suggests that a certain level of L2 proficiency is required before the skills which were learned from L1 literacy transfers to L2 reading. Therefore, many L2 researchers has explored lexical threshold to determine to what extent 'adequate' reading comprehension takes place.

The definition of ‘adequate’ reading comprehension is not clear since it could be different, depending on the purpose or contexts. According to Laufer (1989), she set ‘adequate’ reading comprehension at a score of 55%, which was a passing score of the English for Academic Purposes course in which the participants registered. The results indicated that at 95% coverage, there were significantly more participants with a score of 55% and above than with a score below 55%. Laufer (1992) also found that a vocabulary level of 3000 word families needed for a score of 55%. That is, 3000 word families are set to be the minimal vocabulary threshold which is necessary for ‘adequate’ reading comprehension to take place.

In a more recent study, Laufer and Ravenhorst-Kalovski (2010) investigated lexical threshold again by estimating lexical coverage, vocabulary size, and reading comprehension. The results of their study suggested two lexical thresholds: one optimal and one minimal. The optimal threshold is that the knowledge of 8000 word families and the coverage of 98% including proper nouns lead to independent reading. On the other hand, the minimal one is the knowledge of 4000 to 5000 word families and about 95 % coverage (including proper nouns).

Hsueh-Chao and Nation (2000) defined ‘adequate’ reading comprehension as the score of 71%. It was calculated by averaging the scores that the learners in the 100% coverage group received on the two tests; one was a multiple choice reading comprehension test while the other was a written recall test. It was also shown that around 98% coverage of vocabulary is required for learners to gain ‘adequate’ unassisted comprehension of the text. Although a few learners at 95% and 90 % coverage seemed to overcome some of obstacles and gain close to adequate comprehension, they were a few. Strictly speaking, Nation (2006) advised that 8000 to 9000 word families are necessary for ‘adequate’ reading comprehension, if 98% coverage is needed for unassisted and independent reading comprehension.

Schmitt, Jiang, and Grabe (2011) examined earlier studies on lexical threshold for adequate reading comprehension, such as Laufer’s (1992) and Nation’s (2006). Their results advocated that 98% coverage is more reasonable for L2 readers in an academic setting. Furthermore, the result supported Nation’s (2006) study, indicating that L2 learners need 8000 to 9000 word families to read independently in the given context.

This study explored how much vocabulary and what percentage of lexical coverage in a text needs for Korean university students to reach ‘adequate’ reading comprehension and tried to confirm whether the result from the present study is in line with existing threshold studies. Specifically, the current study attempted at finding out pedagogical implications on lexical threshold in a Korean EFL situation.

This study addressed the following research questions:

1) To what extent do scores on Korean university students' vocabulary size and TOEIC contribute to predicting the performance on reading comprehension?
2) To what extent do Korean university students' vocabulary size, proficiency level, and reading comprehension interrelate with one another?
3) How much vocabulary knowledge is necessary for 'adequate' reading comprehension in a Korean EFL environment?
4) What percent of lexical coverage is required in the text to reach 'adequate' reading comprehension?

III. METHOD

1. Participants

A total of 50 students took part in this study. All of them were from a local university located near Busan Metropolitan City and they were required to take a TOEIC class as part of their degree program. Fourteen males and 36 female students participated in this study and the age of the participants ranged from 20 to 27 years old. Their existing TOEIC score varied from 320 to 780. They had studied English for more than ten years before they were admitted into the university. The participants' major varied from Engineering to Economics but none of them majored in English.

2. Instruments

1) Vocabulary Levels Test

Participants' vocabulary size was measured by the revised version of Nation's Vocabulary Levels Test (Schmitt et al., 2001). The test contains 2000, 3000, 5000, 10000, and academic vocabulary. The reliability of each test is fairly high. The reliability of 2000, 3000, 5000, 10000, and academic vocabulary is .922, .927, .927, .924, and .960, respectively. Each vocabulary test per frequency contains 30 items. Each correct answer receives one point without penalty when choosing an incorrect answer or leaving one blank. The participants only took 2000, 3000, 5000, and academic vocabulary test, in which 10000 levels test were left out since it was considered too demanding for the participants to take. Therefore, a total score of the Vocabulary Levels Test is 120. The test is designed to match the three definitions with three of the six words on the left by writing the number of the word on the blank, next to its definition, as shown in the example:

1	ancient	
2	curious	not easy
3	difficult	very old
4	entire	related to God
5	holy	
6	social	

2) Reading Comprehension Test

Reading comprehension test consisted of four passages from TOEIC (The Test of English for International Communication) Practice Test [Educational Testing Service (ETS), 2016]. All of the participants were taking TOEIC class at the time of data collection so that they were familiar with TOEIC test formats. The passages selected included a form of a letter or an e-mail correspondence which contain a wide variety of common vocabulary used in business. The passages were ones without tables or graphics in order to minimize text features related to reading comprehension, such as a picture, diagram or graph.

By using Fletch Kincaid's grade level formula, the passages were checked with readability in tandem with text length. The Fletch Kincaid grade level formula shows a school grade level of the text in the USA by calculating the number of words per sentence, syllables per word. That is, this formula estimates the text difficulty, using the words used in the text. It provides additional information about which grade level of students can read a text of interest. It turned out that the readability of the passages was 5.3, 6.8, 8.6, and 11.1, respectively. Grade level 5.3 indicates that students in 5th grade in the USA are able to read the first text. Grade 6.8 means US 7th graders are capable of reading the second passage. Similarly, Grade 8.6 for 9th graders, and grade 11.1 for 11th graders. The text length ranged from 202 to 247 words in the passage. The four passages were not significantly different in terms of text length with each other.

Each passage includes five multiple-choice items, and a total score of the reading comprehension test was 20. Multiple-choice item formats were selected to evaluate the participants' reading comprehension objectively. Like the Vocabulary Levels Test, each correct answer receives one point without any penalty when choosing an incorrect answer or leaving one blank.

3) Lexical Coverage

The lexical coverage of four passages in reading comprehension test was calculated by using lexical profile available at Tom Cobb's site (http://lextutor.ca). Vocabulary Profilers break texts down by word frequencies based on BNC (British National Corpus). They

divide the words of texts into the BNC based 20 frequency levels plus off-list (mostly including proper nouns).

3. Data Collection Procedure and Data Analysis

Participants first filled in the questionnaire about background information, including age, educational level, gender, discipline, length of time of learning English and their existing TOEIC score. TOEIC scores were used to identify all of the participants into three proficiency groups: novice, intermediate, and advanced group. Then, they took the Vocabulary Levels Test (2000, 3000, 5000, and Academic Vocabulary) for 40 minutes. Next, they read four reading passages and answered the corresponding reading comprehension items for 25 minutes. The participants were allowed to return to the previously answered items when necessary. The data were collected during the fall semester, 2016.

The data were mainly analyzed by using SPSS (ver. 18.0). The combination of regression analysis using backward elimination, Pearson product-moment correlation coefficient, and a series of one-way ANOVA were conducted in order to answer the research questions.

IV. RESULTS

1. The Effects of Lexical Size and TOEIC Scores on Reading Comprehension

The first step of the study was to find out to what extent scores on L2 learners' vocabulary size and their proficiency determined by previous TOEIC scores contribute to predicting reading comprehension. All of the participants were divided into three proficiency levels to examine the effects of proficiency levels: novice, intermediate, advanced, which was identified with their previous TOEIC scores. As shown in Table 1, the participants whose TOEIC score ranged from 655 to 780 were identified as 'advanced' level (n = 13), while 'intermediate' level (n = 23) scored from 455 to 645 and 'novice' level (n = 14) scored from 320 to 450 as 'novice' level (n = 14).

The descriptive analysis of the TOEIC scores are presented in Table 1 and all the test scores consistently exhibited mean differences among L2 learners by proficiency levels. A one-way ANOVA was conducted to compare participants' performance on the three tests. Results indicated that significant mean differences were found in TOEIC scores, VLT total, and reading comprehension tests with $F(2,47) = 126.657$, $p < 0.001$,

$F(2,47) = 23.222$, $p < 0.001$, $F(2,47) = 16.460$, $p < 0.001$, respectively. Post-hoc tests using the Scheffe exhibited that all the mean differences were found between advanced and novice levels but there were no significant mean differences between advanced and intermediate levels except TOEIC scores.

TABLE 1
Statistics of TOEIC, VLT, and Reading Comprehension Test

	Advanced (n = 13) *Mean (SD)*	Intermediate (n = 23) *Mean (SD)*	Novice (n = 14) *Mean (SD)*	*F*	*p*
TOEIC	690.38 (38.86)	547.39 (59.87)	386.07 (38.59)	126.657	.000**
Vocabulary 2000	26.08 (2.46)	23.78 (3.19)	17.71 (5.68)	17.189	.000**
Vocabulary 3000	21.62 (5.25)	19.30 (4.52)	10.36 (3.65)	24.873	.000**
Vocabulary 5000	18.54 (5.25)	14.83 (4.66)	10.36 (4.33)	10.159	.000**
Academic List	23.69 (5.41)	20.83 (4.43)	15.43 (5.93)	9.201	.000**
Vocabulary Total	89.92 (14.73)	78.74 (13.16)	53.86 (15.60)	23.222	.000**
Reading Score	15.69 (2.14)	14.17 (1.88)	11.21 (2.36)	16.460	.000**

**$p < .001$

Next, regression analysis using backward elimination was carried out to determine the predictive powers of L2 learners' vocabulary size and TOEIC on reading comprehension. This technique is considered either to identity significant predictors or to delete the predictors which are not useful for predicting the variables. The results revealed that Model 4 is statistically significant, $F(2, 49) = 21.235$, $p < 0.001$. TOEIC and Academic list showed a significant beta weight ($\beta = .418$, $p < 0.01$) and ($\beta = .357$, $p < 0.01$), respectively (β weights indicates the strength of the relationship between a predictor and the criterion variable).

As shown in Table 2, the adjusted R^2 value of the fitted model explains that there was an overlap between the combined variables and reading comprehension scores; the variables has about 45.2% explained variance in reading comprehension. The beta (β) weights indicates that reading comprehension scores are positively related to TOEIC and academic vocabulary, increasing by .418 for every extra point in TOEIC score while increasing .357 for every extra point in the academic vocabulary score. TOEIC is a slightly stronger predictor than academic vocabulary in predicting reading comprehension. Thus, the regression analysis concluded that the fitted model (Model 4) yields a significant effect for the combined contribution of TOEIC and Academic vocabulary with leaving all other variables eliminated.

TABLE 2
Model Summary for TOEIC and Academic Vocabulary Predicting Reading Comprehension (n = 50)

Model					Unstandardized Coefficients		Standardized Coefficients	t	p
		R	R^2	AdjR^2	B	Std. Error	Beta		
4	(Constant)				5.652	1.298		4.356	.000**
	TOEIC	.689[d]	.475	.452	.009	.003	.418	3.221	.002**
	Academic Vocabulary				.160	.058	.357	2.754	.008**

Note 1. Predictors: (Constant), TOEIC, Academic Vocabulary
Note 2. Dependent Variable: Reading Comprehension
*p < .05, **p < .001

2. The Relationships Among Learners' Lexical Size, Proficiency, and Reading Comprehension

The Pearson product-moment correlation coefficient was conducted on the data to compare the correlations among the variables. Table 3 revealed that the relationship between L2 learners' vocabulary size and their reading comprehension was highly correlated. The correlation coefficient of vocabulary 2000 and reading comprehension was .439 and found to be significant at p < .001. The correlation coefficient of vocabulary 3000, 5000, and Academic list was .528, .464, and .599, respectively. Among all the variables, academic vocabulary and reading comprehension had the highest correlation. Since TOEIC test focuses on business English, it is highly likely to contain less frequent academic vocabulary than other reading passages found in a daily life context.

TABLE 3
Correlations Among VLT, TOEIC, and Reading Comprehension (n = 50)

	2	3	4	5	6	7
1. TOEIC Score	.656**	.673**	.611**	.579**	.729**	.624**
2. 2,000 Level	1	.755**	.519**	.707**	.856**	.439**
3. 3,000 Level		1	.578**	.737**	.898**	.528**
4. 5,000 Level			1	.645**	.793**	.464**
5. Academic Level				1	.899**	.599**
6. Voca. Total					1	.592**
7. Reading Comp						1

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

3. The Lexical Size for 'Adequate' Reading Comprehension

In this study, 'adequate' reading comprehension was set as the score of 15 out of 20 (75%). The score was calculated by averaging the scores that the participants ranging from 550 to 785 points in TOEIC received in the reading comprehension test. The decision was based on the two factors that (1) all of the participants took TOEIC class at the university at the time of data collection, and (2) TOEIC scores from 550 to 785 points are classified as independent users (Threshold and Vantage) on the Common European Framework Reference (CEFR), which is used worldwide to describe L2 learners' language proficiency. That is, the participants within the score ranges seem to go beyond the threshold where adequate comprehension is likely to take place. For data analysis in this section, all of the participants were divided into two groups: one group with 'adequate' reading comprehension which indicated 15 points out of 20 or above in the test, and the other group with below 15. A series of one-way ANOVA were carried out to compare the two groups between adequate and below adequate levels.

As shown in Table 4, significant mean differences were found in the variables of vocabulary 5,000, Academic List, and Vocabulary total scores, with $F(1, 48) = 7.514$, $F(1, 48) = 8.549$, $F(1, 48) = 7.112$, respectively. The results implicated that L2 learners with adequate reading demonstrated distinctive patterns different from ones who have not reached at an adequate reading level.

TABLE 4
Descriptive Analysis and ANOVA Results for Vocabulary Size Necessary for Adequate Reading (n = 50)

Variables	Adequate Reading Group (n = 22) Mean (SD)	Non-adequate Reading Group (n = 28) Mean (SD)	F	p
2,000 Level	24.14 (3.944)	21.54 (5.541)	3.461	.069
3,000 Level	19.00 (5.354)	16.14 (6.830)	2.593	.114
5,000 Level	16.82 (5.654)	12.75 (4.835)	7.514*	.009*
Academic List	22.64 (3.947)	18.04 (6.489)	8.549*	.005*
Vocabulary Total	82.59 (15.346)	68.46 (20.771)	7.112*	.010*

Note. Adequate Reading = at or over 15, non-adequate reading = below 15
*p < .05

The Figure 1 below visually illustrated the score differences between the two groups. It is noticeable to observe consistent score patterns between the groups across all the vocabulary levels.

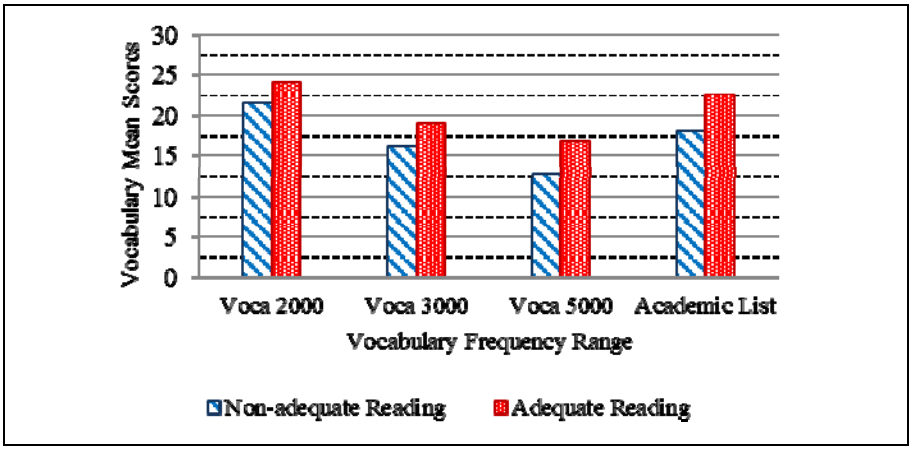

FIGURE 1 Score Patterns of VLT Between the Groups by Adequate Reading

4. Lexical Coverage and Size for 'Adequate' Reading

Finally, this section intended to examine how much vocabulary knowledge and what percentage of lexical coverage are required to be an 'adequate' reader. For that purpose, in Table 5, the lexical coverage was calculated by averaging lexical coverage from four passages used in the reading comprehension test.

TABLE 5
Lexical Coverage of Reading Comprehension Test by BNC Frequency Lists

Frequency Level	Coverage % Passage 1	Coverage % Passage 2	Coverage % Passage 3	Coverage % Passage 4	Average Cumulative Coverage %
K-1	71.9	75.39	74.04	67.82	72.29
K-2	7.62	9.95	11.06	13.86	82.91
K-3	2.86	4.71	2.98	6.44	87.16
K-4	2.86	2.09	2.98	3.96	90.13
K-5	--	--	1.7	1.49	90.93
K-7	--	--	0.43	--	91.04
K-8	0.95	--	0.43	--	91.38
K-11	0.48	--	--	--	91.50
OFF	13.33	7.85	6.38	6.44	100.00

Note. Proper nouns shown in the 'off list'

Since the contents of the passages were letters or e-mail correspondences between business

partners, many proper nouns were included in the 'off list.' The percentage of proper nouns within the 'off list' was recalculated and shown in Table 6. Then the average percentage of proper names was added into lexical coverage all along the second column in Table 7.

TABLE 6
Lexical Coverage by Proper Nouns

	Coverage % Passage 1	Coverage % Passage 2	Coverage % Passage 3	Coverage % Passage 4	Coverage % Average
proper name	11.43	5.23	5.53	6.44	7.16

Table 7 presents learners' vocabulary size, lexical coverage including proper nouns, and reading comprehension scores. Laufer and Ravenhorst-Kalovski's (2010) way of estimating learners' approximate vocabulary size was employed in this section. It was calculated on the basis of the 2000, 3000, and 5000 level of the test that the participants took. If a learner received 28 on the 2000 level, 22 on the 3000 level, and 8 on the 5000 level, his score would be 28 (1,000 level)+28 (2,000 level)+22 (3,000 level)+15 (4,000 level)+8 (5,000 level)=101 out of 150 (The learner may have at least 28 on the 1000 level if he has 28 on the 2000 level. Next, 15 on the 4,000 level comes from the average of the score of 22 on the 3000 level and 8 on the 5,000 level.). The score of 101 represents 3,366 word families (=101 X 5,000 / 150). Also, According to Laufer and Ravenhorst-Kalovski (2010), learners are divided into vocabulary levels as follows: anyone who received a score representing between 500 and 1,500 words was placed at the K1 level, those with a score representing 1,500–2,500 words were placed at the K2 level, those with 2,500–3,500 words were at the K3 level, those with 3,500–4,500 words were at the K4 level, and learners with vocabulary above 4,500 words were placed at the K5 level. The results were shown in Table 7.

TABLE 7
Lexical Size, Lexical Coverage, and Reading Comprehension

Approximate Vocabulary Size	Lexical Coverage %	Reading Score: Mean (*SD*)	Reading Percentage	Number of Students
1,000	79.45	10 (1.4)	50	2
2,000	90.07	11.64 (2.5)	58.2	11
3,000	94.32	13.81 (2.1)	69.05	21
4,000	97.29	16.53 (2.2)	82.65	15
5,000	98.09	17(0.0)	85	1

The results indicate that the difference between 1,000 and 2,000 words resulted in the highest in the coverage (10.62%), but it did not lead to the highest in reading

comprehension percentage (8.2%). On the contrary, it was the lowest in reading comprehension percentage, except for 5,000 words. It can be partially explained by the fact that the majority of the 1,000 frequent words are function words, not content words. Likewise, words at 1,000 and 2,000 levels are much more frequently used in the text, but the difference is not significant in comprehending the text. The difference between 2,000 and 3,000 words led to 4.25% increase in the coverage, and it resulted in 10.85% increase in reading comprehension. In other words, additional 1,000 words increased 2.97% in the coverage and 13.6% in reading comprehension test. It means that a small increase in coverage 2.97% from 3,000 to 4,000 words was linked to the best improvement in reading comprehension score (13.6%).

The lexical coverage from 2,000 to 4,000 was getting smaller, but the reading percentage was getting higher. It implicates that less frequent words can play an essential role in comprehending the text. The results also suggest that the more vocabulary learners obtain, the better chance of understanding the text they have in an EFL environment. There was 0.8% increase in coverage, and 2.35% increase in reading comprehension test between 4,000 and 5,000 words. This result might be caused by a lack of overall number of participants who participated in the study. There is only one participant belonging to 5,000 levels.

Combined to data from Table 7, it was also attempted to examine lexical threshold for 'adequate' reading. The two conditions to meet the criteria for reading an "adequate" reading score, which is determined to obtain 15 or more, are as following: (1) acquiring at least 4,000 word families or more words, and (2) reaching at around 97% coverage in the text. This result is similar to Laufer and Ravenhorst-Kalovski's (2010) study, suggesting that the minimal threshold for independent reading is the lexical knowledge of 4,000 to 5,000 word families and about 95 % coverage including proper nouns.

In Figure 2, the relationships among vocabulary size, lexical coverage, and reading score are visually represented. The two lines representing lexical coverage and reading score rapidly converge at 4,000 level and the two lines are getting more close at 5,000 level. Thus, it can be concluded that the minimum threshold for 'adequate' reading in our given context is more than 4,000 vocabulary level and around 97% coverage in the text. The interpretation of 97% coverage, however, needs to be cautious because the number seemed to be inflated due to the nature of the texts, which included a large proportion of proper nouns such as company names, person's names, or different types of goods.

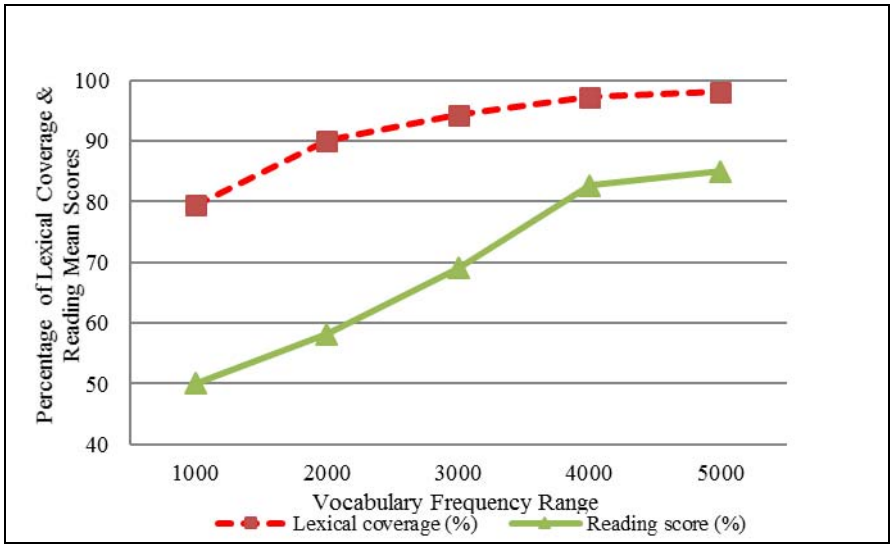

FIGURE 2 Relationships Among Lexical Size, Lexical Coverage, and Reading Score

V. CONCLUSION AND IMPLICATIONS

In this article, the relationships among vocabulary size, their proficiency, and reading comprehension were examined for 50 Korean EFL university learners as a whole group, and also divided into advanced, intermediate, and novice groups of learners with respect to their proficiency level measured by a general language proficiency test. To address the first research question, the present study explored the combined contributions of TOEIC and Academic vocabulary level knowledge to reading comprehension. The R^2 value in the regression analysis suggested that 45.2% of the variance in the reading score is accounted for by the combined factors of general language proficiency represented by TOEIC scores and vocabulary knowledge, specifically academic vocabulary level, indicating the integrated effects of general reading skills and academic vocabulary on the reading comprehension.

This result is incongruent with the earlier claims that vocabulary may be the most essential factor in reading comprehension (Anderson & Freebody, 1981; Laufer & Nation, 1999; Nassaji, 2003; Stahl, 2003). Rather, the findings of the study parallel with those who reported that vocabulary is not the only factor affecting comprehension. A large number of variables have an effect in reading, including those involving language proficiency (e.g., grammatical knowledge and awareness of discourse structure), the text itself, (e.g., text length, text difficulty, and topic), and those concerning the reader (interest in topic,

motivation, amount of exposure to print, purposes for reading, L1 reading ability, and inferencing ability) (Droop & Verhoeven, 2003; Grabe, 2009; E. Nam, 2012; van Gelderen et al., 2004; van Gelderen et al., 2007). Therefore, L2 educators and professionals need to have interest in other factors, such as grammatical knowledge, motivation, or background knowledge as well as vocabulary knowledge in a Korean EFL environment.

For the second research question, this study attempted to investigate the relationship between Korean university students' vocabulary size and their reading comprehension. The correlation analysis showed that Korean university students' vocabulary size and reading comprehension were highly correlated, like many other studies by Laufer (1992), Laufer and Ravenhorst-Kalovski (2010), Qian (1999), and M. O. Min and S. W. Lee (2016). Among various vocabulary levels, the correlation coefficient of the academic vocabulary was the highest. It means that university students need to improve academic vocabulary for their L2 reading in an academic setting. Also, it was found that L2 learners' proficiency level positively affected their reading comprehension. Based on the findings from the 1st and 2nd research questions, it could be argued that the difference in their reading comprehension scores can be substantially related to the difference in the size of vocabulary knowledge, in particular, academic vocabulary level.

Regarding the third research question, the current study attempted to investigate how the reading comprehension test scores were linked to lexical coverage and to their vocabulary knowledge. The comparison analysis suggested that there were significant mean differences at 5,000 level, academic vocabulary, and vocabulary total scores. It means that Korean university students should possess at least 5,000 word families or more academic vocabulary to be an 'adequate' reader in an educational context.

To find out the answers for the last research questions, it is examined how two factors of lexical threshold – lexical coverage and vocabulary knowledge – were related to reading comprehension. In other words, the study explored how much vocabulary knowledge and what percentage of lexical coverage is required for 'adequate' reading. The result advocated that more than 4,000 word families and around 97% coverage are necessary for 'adequate' reading. The finding was quite similar to Laufer and Ravenhorst-Kalovski's (2010), whose minimal threshold for independent reading is the knowledge of 4,000 to 5,000 word families and about 95 % coverage including proper nouns. On the contrary, this study seemed to yield less strict results in comparison to Nation's (2006) study in which a threshold of 8,000-9,000 word levels is combined with 98% of lexical coverage of the texts.

As Laufer and Ravenhorst-Kalovski (2010) mentioned, it is not legitimate to claim that adequate reading comprehension is not able to take place unless L2 learners reached the lexical threshold, or the threshold will lead to automatic increment in reading comprehension. However, without doubt, the findings of the current study are in accordance with previous research (Huckin, 1995; Laufer & Nation, 1999; Nation, 2001;

Schmitt, 2000) that they highlighted vocabulary knowledge as one of the essential factors for reading achievement. In sum, language professionals and instructors need to raise a great concern about increasing L2 learners' vocabulary levels in order for them to meet a considerable challenge in their academic future.

There are some limitations in this study. Above all, the study only employed 50 participants which made it impossible to generalize the results to larger population. Additionally, it is hard to compare this result to other studies since the current study only employed Korean EFL learners and their language proficiency was less narrowly ranged based on TOEIC scores (from 320 to 780). It seems that language proficiency provides some alternative interpretation in this case. Accordingly, there might be a possibility that the minimum criteria for 'adequate' reading can be changeable if a further study includes more participants with full range of language proficiency. Future research needs to be followed by instructional effort of vocabulary learning in order to yield more empirical results associated with the findings from this study. By employing a wide range of strategies and tasks to improve L2 learners' vocabulary size, further research will be able to better determine to what extent vocabulary knowledge is adequately required for L2 learners to be successful in academic setting, in particular. Despite the limitations, the results of the study offer valuable insights into the relationships among vocabulary size, coverage, and L2 reading ability which broaden our knowledge base on how combined knowledge of vocabulary size and coverage affects L2 learners' success in reading within an academic context.

REFERENCES

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In 1. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77-117). Newark, DE: International Reading Association.

Bernhardt, E. B., & Kamil, M. L. (1995). Interpreting relationships between L1 and L2 reading: Consolidating the linguistic threshold and the linguistic interdependence hypotheses. *Applied Linguistics*, *16*(1), 15-34.

Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second language learners. *Reading Research Quarterly*, *38*, 78-103.

Educational Testing Service. (2016). *ETS TOEIC practice test RC*. Seoul: YBM.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York: Cambridge University Press.

Hsueh-Chao, M. H., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language, 13*(1), 403-430.

Huckin, T. (1995). Second language reading and vocabulary learning. Norwood, New Jersey: Ablex Publishing Corperation.

Huckin, T., & Bloch, J. (1993). Strategies for inferring word-meanings in context: A cognitive model. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 153-178). Norwood, New Jersey: Ablex Publishing Corperation.

Hudson, T. (2007). *Teaching second language reading*. New York: Oxford University Press.

Johns, A. M. (1981). Necessary English: A faculty survey. *TESOL Quarterly, 15*(1), 51-57.

Koda, K. (2005). *Insights into second language reading*: *A cross-linguistic approach*. New York: Cambridge University Press.

Laufer, B. (1989). What percentage of text lexis 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. (1992). How much lexis is necessary for reading comprehension? In H. Bejoint & P. Arnaud (Eds.), *Vocabulary and applied linguistics* (pp. 126-132). Basingstoke & London: Macmillan.

Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, *16*(1), 33-51.

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language, 22*(1), 15-30.

Min, MiOk, & Lee, SeongWon. (2016). The size and depth of vocabulary knowledge of Korean EFL university students and its role in reading comprehension. *English Language Teaching, 28*(3), 91-106.

Nam, Eun-hee. (2012). A study of an effective English reading class at university through analysis of students' reading motivation and perceptions in reading activities. *Modern English Education, 13*(1), 281-301.

Nassaji, H. (2003). L2 vocabulary learning from context: Strategies, knowledge sources, and their relationship with success in L2 lexical inferencing. *TESOL Quarterly, 37*(4), 645-670.

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

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review, 63*(1), 59-82.

Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review*, *56*(2), 282-308.

Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*,

52(3), 513-536.

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.

Rosenfeld, M., Leung, S., & Oltman, P. K. (2001). *The reading, writing, speaking, and listening tasks important for academic success at the undergraduate and graduate levels*. Princeton, NJ: Educational Testing Service.

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.

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.

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the Vocabulary Levels Test. *Language Testing*, *18*(1), 55-88.

Sherwood, R. I. (1977). A survey of undergraduate reading and writing needs. *College Composition and Communication*, *28*(2), 145-149.

Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children committee on the prevention of reading difficulties in young children.* Washington, DC: National Research Council.

Stahl, S. A. (2003). Vocabulary and readability: How knowing word meanings affects comprehension. *Topics in Language Disorders*, *23*(3), 241-247.

Ulijn, J. M., & Strother, J. B. (1990). The effect of syntactic simplification on reading EST texts as L1 and L2. *Journal of Research in Reading*, *13*(1), 38-54.

van Gelderen, A., Schoonen, R., de Glopper, K., Hulstijn, J., Simis, A., Snellings, P., Stevenson, M. (2004). Linguistic knowledge, processing speed, and metacognitive knowledge in first- and second-language reading comprehension: A componential analysis. *Journal of Educational Psychology*, *96*(1), 19-30.

van Gelderen, A., Schoonen, R., Stoel, R., de Glopper, K., & Hulstijn, J. (2007). Development of adolescent reading comprehension in language 1 and language 2: A longitudinal analysis of constituent components. *Journal of Educational Psychology*, *99*(3), 477-491.

Examples in: English
Applicable Languages: English
Applicable Levels: Tertiary level

Heejeong Ji
Pukyoung National University
45 Yongso-ro, Nam-Gu. Busan, 48513, Korea
H.P.: 010-5508-5331
E-mail: 6825331@hanmail.net

Yousun shin
Kyungsung Univeristy
48434, 309, Suyeong-ro, Nam-gu, Busan, Korea
H.P.: 010-2510-9305
E-mail: jdream914@gmail.com

Received 15 March 2017
Reviewed 24 April 2017
Accepted 12 May 2017