

Applying the Lexical Coverage Hypothesis to Establish the Suitability of EFL Reading Materials: A Case Study of the TOEFL (ITP)

John B. Collins¹

Abstract

This paper employed the lexical coverage construct, defined as the percentage of words in a text that a reader understands (Laufer & Ravenhorst-Kalovski, 2010, p. 16), to establish a) the vocabulary size needed to achieve adequate comprehension of the TOEFL (ITP) reading section and reading passages taken from two EFL textbooks, and b) whether the EFL textbooks provide opportunities for vocabulary instruction sufficient to achieve specific reading comprehension targets; in this case, a reading comprehension level of approximately 50%-55% (at the intermediate level) and 65% (at the upper-intermediate level) on the TOEFL (ITP). Using the RANGE software package and employing 95% and 98% lexical coverages, the results pointed to a gap of at least 2,500-4,000 word families between TOEFL reading texts and the two EFL textbooks. At 98%, the gap was approximately 6,000 word families. It was concluded that the two textbooks do not provide sufficient vocabulary instruction for students to meet the lexical demands of the TOEFL reading comprehension targets. The implications of this for textbook selection and vocabulary instruction are discussed.

Key terms: Lexical coverage, vocabulary size, reading comprehension, TOEFL (ITP), EFL reading textbooks

1. Introduction

This paper centers on the relationship between vocabulary knowledge and reading comprehension, and was carried out within the framework of the lexical coverage construct. The research described employs the lexical coverage construct to inform textbook and other teaching material selection decisions in terms of whether they provide sufficient vocabulary instruction for students to reach an adequate level of L2 reading comprehension. In this study, “adequate” is defined as the achievement of target TOEFL scores required for successful course completion at the institution where this research took place (Ritsumeikan Asia Pacific University, APU). At the intermediate level students are required to obtain a TOEFL (ITP) score of between 461 and 480, while at the upper-intermediate level they are expected to obtain a score of between 481 and 500 (Intermediate English B Syllabus, 2017; Upper Intermediate English B Syllabus, 2017). Using a TOEFL score calculation formula (Philips, 2003, p. 550), it is possible to convert these scores into approximate percentage scores, assuming a constant level of achievement across all three sections of the test. A score of 461 would require a comprehension score of between 50% and 55% and a score of 500

¹Lecturer, Ritsumeikan Asia Pacific University (APU), Beppu City, Oita, Japan.
Email: crm11427@apu.ac.jp

would require 65% comprehension. By comparing these figures to previous lexical coverage and vocabulary size studies (described below), it is possible to establish the approximate vocabulary size needed to reach the students' target TOEFL scores. Furthermore, by establishing the vocabulary size presented in the EFL textbooks, it is possible to determine the adequacy of these materials in terms of providing sufficient vocabulary instruction to reach these TOEFL targets.

1.1 The TOEFL and its use at APU

Since its development in the 1960's, the TOEFL has become one of the most widely recognized assessments of English proficiency, with over 4,500 testing sites in over 165 countries around the world (ETS, 2010). Despite being designed to meet the needs of North American colleges and universities (Spolsky, 1990, p. 99), the TOEFL has since become a benchmark English proficiency test recognized by companies and prospective employers around the world. At the institution where this study took place, the TOEFL has become a major component of the assessment framework and students are required to achieve a score of 500+ on the TOEFL (ITP). Not only are students required to sit the TOEFL during their time at APU, but their score also constitutes 25% of their final grade in the compulsory courses in the Standard Track¹. Preparing students for the TOEFL therefore constitutes a major component in the English program curricula and a central focus of classroom instruction.

1.2 Literature review

The following is a review of the literature surrounding the lexical coverage construct and studies which have employed this construct in order to establish the size of vocabulary required for the TOEFL reading comprehension section. Vocabulary knowledge is vital for reading to take place (Laufer, 1997, p. 20) and has been widely recognized as a strong predictor, if not the strongest predictor, of reading comprehension performance (Laufer & Ravenhorst-Kalovski, 2010, p. 16). This appraisal is supported by the considerable amount of research which centers on the construct of lexical coverage. Conversely, a high percentage of unknown words in a text has been identified as being markedly detrimental to reading comprehension (Hsueh-Chao & Nation, 2000, p. 422). Laufer (1989) concluded that a 95% coverage (one unknown word in twenty) is necessary to achieve a "reasonable" level of comprehension. The term "reasonable" here was defined as a reading comprehension score of 55%, the minimum passing grade where the study took place. Hsueh-Chao and Nation (2000) suggested that a lexical coverage of 98% (one unknown word in every fifty) is necessary for adequate reading comprehension. On further analysis to establish exactly what level of comprehension was considered adequate in the Hsueh-Chao and Nation study, Laufer and Ravenhorst-Kalovski (2010) calculated a figure of 71% reading comprehension (p. 18). A lexical coverage of 98% has also been supported by the findings of Schmitt, Jiang, and Grabe (2011). Laufer and Ravenhorst-Kalovski (2010) compared lexical coverage, vocabulary size and reading scores on the English Psychometric Tests, and concluded that an adequate level of reading comprehension is possible at an optimal lexical coverage of 98% (or 8,000 word

families) or alternatively at a minimal coverage of 95% (or 4,000-5,000 word families). This study also illustrated that even slight increases in lexical coverage can have a large impact on reading comprehension – indeed a slight gain in coverage of 0.8% (between the 5,000 and 6,000 levels) was associated with the largest gain in comprehension scores. The results of the Laufer and Ravenhorst-Kalovski study are shown in figure 1.

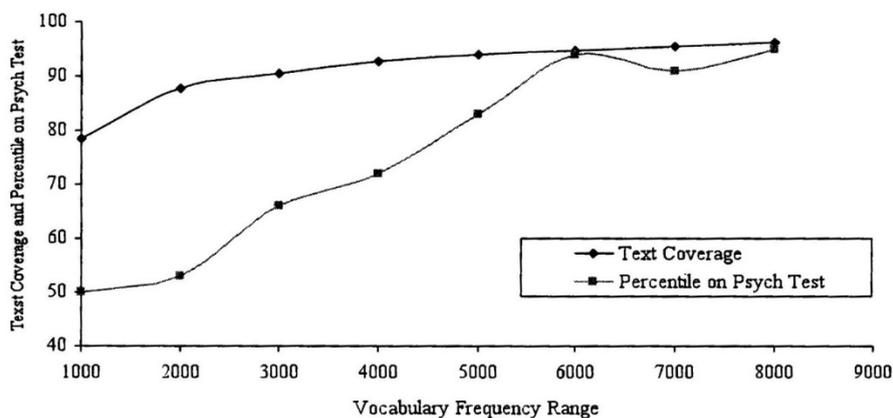


Figure 1: Text coverage and reading scores in relation to frequency range (Laufer & Ravenhorst-Kalovski, 2010, p. 24)

However, terms such as “adequate” and “reasonable” comprehension have no clear definition (Laufer & Ravenhorst-Kalovski, 2010, p. 16) since defining such terms is dependent on the context within which they are being applied. That is to say, adequate could mean 55% comprehension in one context but 85% in another. As described above, students at APU are required to reach a minimum reading comprehension score of between approximately 50% and 55% (at the intermediate level) and 65% (at the upper-intermediate level). In terms of the present study, therefore, the findings of most relevance are those of Laufer (1989) and Hsueh-Chao and Nation (2000). Laufer (1989) pointed to a required lexical coverage of 95% to achieve a 55% reading comprehension. This is essentially identical to the target TOEFL score (%) required of intermediate students. Likewise, the findings of Hsueh-Chao and Nation (2000) point to a required lexical coverage of 98% for a comprehension score of 71%. Given that upper-intermediate students require a TOEFL score of approximately 65%, it can be said that a 98% coverage is beyond the hypothetical threshold these students require. Based on these figures, upper-intermediate students would likely require a lexical coverage of between 96% and 97%.

1.2.1 The TOEFL reading section and vocabulary size

A number of studies have been carried out in order to establish the vocabulary size required to reach a 95% lexical coverage of TOEFL reading comprehension texts. Chujo and Nishigaki (2003)

determined that test-takers would require a minimum vocabulary size of 6,150 *lemma* [consisting of a base word and its inflected forms] (Read, 2000, p. 18) to reach the 95% threshold. Likewise, the findings of Chujo (2004) indicated a minimum vocabulary size of between 5,900 and 6,300 lemmas. Chujo and Oghigian (2009) determined that 5,000 *word families* [consisting of a set of word forms that share a common meaning] (Read, 2000, p.19) would be necessary for the TOEFL (PBT). The difference in these figures could be attributed to the fact that the Chujo and Nishigaki (2003) study employed the High Frequency Word List (HFWL) to analyze the reading texts, while the Chujo and Oghigian (2009) study used Nation's 14K word list. Although both lists are based on the British National Corpus (BNC), the HFWL is arranged by lemma, while the Nation 14K list is arranged by word family. Chujo and Nishigaki (2003) also established that the ELT textbooks and materials that second-year college students in their study were using required only 4,050 lemmas (p.79) and concluded that the materials were insufficient preparation for TOEFL. Indeed, Chujo and Nishigaki stated that if these students wished to gain a high score on the TOEFL, they would "...need to make a determined and conscious effort" to expand their academic vocabulary (p.80). Using Nation's 14K list, Kaneko (2014) established that a vocabulary size of between 6,000 and 7,000 word families would be required to reach the 95% threshold on the TOEFL iBT reading section (p.47). There is considerable variance among these findings due to methodological differences and different versions of the TOEFL being investigated. Indeed Kaneko (2014, p. 3) suggests that earlier findings should be considered tentative.

2. Methodology

2.1 Study aim and research questions

The aim of this research is to illustrate how the lexical coverage construct can be employed to establish whether specific EFL reading textbooks provide sufficient vocabulary instruction for students to reach a specified adequate level of reading comprehension on the TOEFL (ITP). The extent to which the textbooks achieve this was established by calculating the size of vocabulary necessary to reach lexical coverage levels of 95% and 98% of TOEFL reading comprehension passages, and then comparing this data with similar data for the two textbooks. In order to achieve this, the following research questions were addressed:

- 1) What size of vocabulary is needed to achieve lexical coverages of 95% and 98% for the TOEFL (ITP) reading comprehension section?
- 2) What size of vocabulary is needed to achieve lexical coverages of 95% and 98% for the EFL textbooks?

2.2 The EFL textbooks and TOEFL reading texts

A total of twenty-seven texts were examined (nine TOEFL texts and nine each from the two textbooks). The nine TOEFL reading comprehension texts examined in this study were taken from three complete practice tests published by Educational Testing Services (ETS). The eighteen EFL textbook reading passages were taken from *Reading and Vocabulary Focus 1* and *Reading and*

Vocabulary Focus 2, published by National Geographic/Cengage Learning. At the time of the present study, the two books (hereafter referred to as *Focus 1* and *Focus 2*) were being used at the intermediate and upper-intermediate English levels respectively at APU. As described in the series introduction of each book, the texts are “...presented in level-appropriate language and used to build reading skills and to promote vocabulary learning” (McEntire, 2014; Gordon & Blass, 2014, p. xi). Each reading passage is preceded by an introduction to 10-12 topic-related vocabulary items, 6-8 academic words, and 6-8 multiword vocabulary items. A total of approximately 400-500 vocabulary items are explicitly introduced in each book. It can be said, therefore, that vocabulary instruction is a central component of both books.

Given the target TOEFL scores of the courses within which these textbooks are used, the suitability of these books, at least in lexical terms, could be established on the basis of how sufficient their vocabulary instruction is to meet these targets. Although neither *Focus 1* nor *Focus 2* were originally adopted into the curriculum in order to prepare students for the TOEFL (nor do the books' publishers claim that they provide suitable TOEFL preparation), considering their explicit focus on reading and vocabulary development, and the fact that the TOEFL has since become a central assessment tool in the courses in which these textbooks are being used, establishing to what extent a match has been successfully achieved between the textbooks and course assessment goals is a legitimate question to consider.

2.3 Text analysis method

The twenty-seven texts were analyzed using RANGE, a text-analysis software tool produced by Heatley, Nation and Coxhead (2002). The RANGE software employs Nation's 14K word lists and calculates the percentage of words of a given text that fall within each of the fourteen one-thousand word lists. By adding up the percentages, it is possible to establish what vocabulary size is required to reach a desired lexical coverage, in this case 95% or 98%. Each of the twenty-seven texts were typed into a Microsoft Word document and manually checked for spelling and other errors. Before each passage was analyzed, a number of modifications were made. The first modification concerns the handling of proper nouns. There appears to be no consensus in the literature about retaining or deleting proper nouns before an analysis of this type is conducted. Kaneko (2014, 2015) opted to retain them, arguing that removing them would affect lexical coverage by approximately 1% (2014, p. 46). Other researchers, however, including Chujo and Oghigian (2009) and Chujo (2004) have excluded them on the basis that “...they are of high frequency in particular texts but not in others” (Nation, 2001, as cited in Chujo, 2004, p. 233). My decision to delete proper nouns from the present samples was based on the fact that many *Focus 1* and *Focus 2* texts include a large number of non-English proper nouns (including locations and names such as *Kibera*, *Kiberan*, *Kiberans*, and *Pamoja*) which would not appear on any of the fourteen 1,000 word lists and would therefore skew the results. Indeed, an initial analysis indicated that as many as 8.67% of words of this type were being counted as “not on list”. For this reason, proper nouns were manually identified and deleted from all sample texts. Secondly, hyphens were removed from all hyphenated words, as

per the instructions provided with the text analysis software. The texts were then run through the RANGE software.

3. Text analysis results

Tables 1-3 display the lexical coverage results for the twenty-seven texts taken from *Focus 1*, *Focus 2*, and the TOEFL preparation tests. In Tables 1-3, “Word list” refers to each of the fourteen 1,000 word lists that constitute the Nation 14K word list. “Not on list” refers to the percentage of words in the texts which are not included in any of the word lists. For each text, the cumulative lexical coverage (%) achieved at each of the 1,000 word list levels is displayed. In this way, it is possible to show how large a vocabulary is required to reach a 95% and 98% lexical coverage. For *Focus 1*, 95% lexical coverage was achieved at the 3,000 word level, and 98% coverage was achieved at the 6,000 word level (Table 1). For *Focus 2*, a 95% lexical coverage was achieved between the 3,000 and 4,000 word lists, and 98% coverage was achieved at 6,000 words (Table 2). Finally, with regards to TOEFL texts, 95% and 98% lexical coverages were achieved at the 6,000 and 12,000 word list levels respectively (Table 3).

Table 1

Descriptive statistics and cumulative lexical coverage (%) for each Focus 1 text

Word list	Text 1	Text 2	Text 3	Text 4	Text 5	Text 6	Text 7	Text 8	Text 9	AVE
1K	87.9	83.9	88.2	85.6	80.1	85.7	82.8	76.9	90.4	84.6
2K	94.8	93.9	92.7	93.8	91.0	92.8	90.5	89.0	97.7	92.9
3K	96.5	95.3	96.5	96.0	94.1	95.1	93.1	91.1	98.5	95.1
4K	98.5	95.7	97.2	98.2	94.4	97.3	98.7	92.6	99.2	96.9
5K	98.5	95.9	97.2	100.0	96.0	97.8	98.8	93.2	99.4	97.4
6K	98.5	96.3	97.5	100.0	96.6	98.0	99.2	96.6	99.8	98.0
7K	98.5	96.5	97.5	100.0	99.7	98.0	99.2	96.6	99.8	98.4
8K	98.5	96.5	98.3	100.0	99.7	98.2	99.2	97.1	99.8	98.6
9K	98.5	96.7	98.3	100.0	100.0	98.4	99.2	97.1	99.8	98.7
10K	98.5	97.4	98.3	100.0	100.0	98.9	99.2	97.9	99.8	98.9
11K	98.5	97.4	100.0	100.0	100.0	98.9	99.2	97.9	99.8	99.1
12K	98.5	98.6	100.0	100.0	100.0	98.9	99.2	97.9	99.8	99.2
13K	98.5	98.6	100.0	100.0	100.0	99.1	99.5	97.9	99.8	99.3
14K	98.5	98.6	100.0	100.0	100.0	99.1	99.5	99.2	99.8	99.4
Not on list	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Table 2

Descriptive statistics and cumulative lexical coverage (%) for each Focus 2 text

Word list	Text 1	Text 2	Text 3	Text 4	Text 5	Text 6	Text 7	Text 8	Text 9	AVE
1K	80.43	86.25	83.71	85.62	84.53	78.77	89.42	81.91	84.56	83.9
2K	88.6	94.2	92.5	95.0	90.3	88.7	96.8	90.0	97.7	92.6
3K	90.1	96.0	96.2	95.9	92.3	89.6	97.2	96.9	98.5	94.7
4K	91.0	99.1	98.1	98.1	94.3	93.5	98.2	97.8	98.9	96.5
5K	97.0	99.1	98.8	98.2	94.4	94.0	98.6	98.2	99.0	97.5
6K	97.0	99.1	99.0	98.2	95.2	96.8	99.4	98.2	99.3	98.0
7K	97.2	99.1	99.3	98.2	96.2	97.1	99.4	99.1	99.3	98.3
8K	97.9	99.6	99.7	98.9	96.2	97.1	99.8	99.3	99.3	98.6
9K	98.1	99.6	99.7	99.3	96.2	97.1	99.8	99.3	99.5	98.7
10K	98.1	99.6	99.7	99.3	96.2	97.1	99.8	99.4	99.5	98.7
11K	98.1	99.6	99.7	99.3	96.2	99.7	99.8	99.4	99.5	99.0
12K	99.8	99.6	99.7	99.3	96.2	99.9	99.8	99.6	99.5	99.3
13K	99.8	99.8	99.7	99.3	96.6	99.9	100.0	99.6	99.5	99.3
14K	99.8	99.8	99.7	99.3	97.3	100.0	100.0	99.6	99.7	99.5
Not on list	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Table 3

Descriptive statistics and cumulative lexical coverages (%) for TOEFL reading passages

Word list	Text 1	Text 2	Text 3	Text 4	Text 5	Text 6	Text 7	Text 8	Text 9	AVE
1K	74.7	75.4	68.8	72.8	77.1	82.1	68.3	77.1	69.7	74.0
2K	84.0	88.5	77.7	83.7	87.6	89.7	77.9	87.0	83.2	84.4
3K	86.5	90.0	87.3	88.5	89.2	93.5	86.9	91.0	89.2	89.1
4K	89.7	93.5	90.1	91.8	91.7	97.9	89.5	92.6	93.4	92.2
5K	90.0	94.2	92.1	93.4	97.0	98.5	93.5	96.3	94.3	94.4
6K	93.9	94.6	93.8	93.4	97.9	99.1	94.0	96.9	94.9	95.4
7K	94.7	95.8	94.2	94.3	98.4	99.4	94.5	98.5	94.9	96.1
8K	94.7	97.7	95.2	94.6	98.4	99.4	95.5	98.8	97.6	96.9
9K	95.4	98.1	95.9	95.5	98.7	99.4	96.5	99.4	97.6	97.4
10K	95.4	98.4	96.2	95.8	98.7	99.4	97.0	99.7	97.9	97.6
11K	96.1	98.8	96.2	96.1	98.7	99.4	97.5	99.7	98.2	97.8

12K	96.1	98.8	96.6	97.0	98.9	99.4	98.0	99.7	98.2	98.1
13K	96.1	98.8	96.6	97.9	99.2	99.7	98.5	99.7	98.5	98.3
14K	96.1	99.6	98.3	98.2	99.2	99.7	99.0	99.7	99.7	98.8
Not on list	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

With regard to Research Question 1, the vocabulary sizes needed to achieve a lexical coverage of 95% and 98% for the TOEFL reading section are approximately 6,000 word families and 12,000 word families respectively, based on the above data. With regards to Research Question 2, the vocabulary sizes needed to achieve a lexical coverage of 95% and 98% for *Focus 1* were 3,000 and 6,000 words families respectively. Likewise, 95% and 98% coverages were reached at approximately 3,500 word and 6,000 word levels for *Focus 2*.

4. Discussion and implications

The results of previous studies suggest that a minimum vocabulary size of between 6,150-6,300 words (lemma) is required to achieve 95% lexical coverage on the TOEFL (PBT) reading section (Chujo and Nishigaki, 2003; Chujo, 2004). Similarly, Kaneko (2014) calculated a figure of between 6,000 and 7,000 word families for the TOEFL (iBT). The current results are consistent with these findings. In order to achieve a 95% lexical coverage for *Focus 1* and *Focus 2*, however, a learner would only need to know between approximately 3,000 and 3,500 words. There is, therefore, a gap of approximately 2,500-4,000 words. As described above, successful reading comprehension is largely predicated on achieving an adequate lexical coverage of any given text. The implication here is that the lexical coverage that readers must achieve in order to achieve adequate comprehension of *Focus 1* and *Focus 2* is considerably lower than that of TOEFL reading texts. In order for students to achieve a suitable lexical coverage of TOEFL texts, learners would therefore need significant opportunities for supplementary vocabulary instruction; indeed, this reflects the conclusion reached by Chujo and Nishigaki (2003). The gap becomes even wider at a 98% lexical coverage which, although hypothesized to be necessary for comprehension scores slightly beyond the target TOEFL score, points to a shortfall of approximately 6,000 words – double the vocabulary size required for either *Focus 1* or *Focus 2*. Put simply, neither *Focus 1* nor *Focus 2* provides a range of vocabulary sufficient to meet the lexical demands of the target TOEFL scores. It should also be noted that there appears to be only a modest difference in the lexical demands of *Focus 1* and *Focus 2* (approximately 500 words at the 95% threshold), which is a potential concern for curriculum developers who aim to expose their students to progressively more lexically demanding content.

Given the scale of the vocabulary size needed to achieve ideal coverage of the TOEFL, a secondary implication of these results relates to the way in which vocabulary instruction is provided. In light of the gap which has been highlighted here, teachers need to consider how best to provide students with vocabulary instruction, supplementary or otherwise, to assist them

in reaching such targets. While a discussion of specific curriculum development and teaching approaches is beyond the scope of the present study, it remains a pressing issue, given that the TOEFL will remain part of the APU English program assessment framework for the foreseeable future.

5. Conclusion

Before concluding this paper, there are a number of methodological limitations which should be mentioned. Firstly, the relationship between lexical coverage and reading comprehension (described above) is based on studies that did not use TOEFL reading texts. Given that studies have shown that TOEFL takers employ test-taking strategies and techniques which can impact overall test performance, the results of the present study should be regarded as tentative. Further studies are necessary to identify the relationship between lexical coverage, vocabulary size and actual TOEFL reading comprehension performance. The second limitation is the reliance on the Philips (2003) TOEFL score calculation formula with which the target reading comprehension scores (%) for intermediate and upper-intermediate English were established. Although it is impossible to accurately establish how TOEFL scores can be converted into a percentage without authentic test papers and authentic test scores, this aspect must be acknowledged as a limitation of the present study. A final limitation relates to the processing ability of the RANGE analysis software. As described above, *Focus 1* and *Focus 2* emphasize multiword vocabulary items, such as “*to hope for the best*”. However, the RANGE software cannot recognize multiword units, therefore each of these items is counted individually as a total of five words (which all fall within the first 1,000 word list) rather than as a complete, and more lexically demanding, unit. Given the emphasis on multiword units in both *Focus 1* and *Focus 2*, the RANGE results may have been affected by this technical limitation.

The aim of this research was to employ the lexical coverage construct to establish whether certain EFL reading textbooks provide sufficient vocabulary instruction within a specific context and given specific TOEFL target scores. The results indicate that a significant gap exists between the levels of vocabulary that *Focus 1* and *Focus 2* expose students to and the level of vocabulary necessary to reach a lexical coverage required to achieve adequate comprehension of TOEFL reading texts. In other words, there appears to be a discrepancy between teaching materials and assessment goals. Notwithstanding the aforementioned limitations to the study, the results illustrated that lexical coverage can be employed to statistically establish the suitability of textbooks given specific achievement targets and how it could, therefore, help to inform the selection of classroom textbooks and instructional materials.

Endnote:

¹ This refers to the reading and vocabulary development focused “B” courses of the APU Standard English Track curriculum.

References

- Chujo, K. (2004). Measuring vocabulary levels of English textbooks and tests using a BNC lemmatised high frequency word list. *Language and Computers*, 51, 231-249.
- Chujo, K., & Nishigaki, C. (2003). Bridging the vocabulary gap: From EGP to EAP. *JACET Bulletin*, 37, 73-84.
- Chujo, K., & Oghigian, K. (2009). How many words do you need to know to understand TOEIC, TOEFL & EIKEN? An examination of text coverage and high frequency vocabulary, *The Journal of Asia TEFL*, 6(2), 121-148.
- Educational Testing Service (1999). *TOEFL Practice Tests Volume 2*.
- Educational Testing Service (2010). Find out why it's the English language test more institutions trust, retrieved from: <https://www.ets.org/c/mrm/ets00068/>
- Gordon, D., & Blass, L. (2014). *Reading and vocabulary focus 2*. Boston, National Geographic Learning.
- Hsueh-Chao, M. H., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a foreign language*, 13(1), 403-430.
- Kaneko, M. (2014). Is the vocabulary level of the reading section of the TOEFL internet-based test beyond the lexical level of Japanese senior high school students? *Vocabulary Learning and Instruction*, 3(1), 44-50.
- Kaneko, M. (2015). Vocabulary size required for the TOEFL iBT listening section. *The Language Teacher*, 39(1), 9-14.
- Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans to thinking machines*, 316-323. Clevedon, England: Multilingual Matters.
- Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. *Second language vocabulary acquisition: A rationale for pedagogy*, 1, 20-34.
- 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.
- McEntire, J. (2014). *Reading and vocabulary focus 1*. Boston, National Geographic Learning.
- Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I. S. P., & Heatley, A. (1994). Range: A program for the analysis of vocabulary in texts [software]. Retrieved from <http://www.victoria.ac.nz/lals/staff/paul-nation/nation.aspx>
- Phillips, D. (2003). *Longman Complete Course for the TOEFL Test: Preparation Course for the TOEFL Test*. NY: Longman.
- Read, J. (2000). *Assessing vocabulary*. 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, 26-43.

Spolsky, B. (1990). The prehistory of TOEFL. *Language Testing*, 7(1), 98-118.