

Vocabulary Size Targets for the TOEIC Test

KANEKO, Masaya
Yokohama City University

Abstract

Success on the Test of English for International Communication (TOEIC) is of value to many Japanese learners of English for academic and occupational purposes. Since this high-stakes test measures examinees' receptive knowledge, setting an accurate vocabulary-size target is crucial as research suggests that knowing 95–98% of running words in a text is necessary for adequate comprehension in listening and reading. The present study developed upon Chujo and Oghigian's (2009) study. A lexical frequency profile was created for twelve sets of listening and reading test papers derived from official TOEIC practice tests. Given that senior high school students in Japan are supposed to acquire vocabulary of 3,000 word families, findings of the present study argue that there is a lexical-size gap that test takers with this vocabulary size should fill. It is suggested that an additional 1,000 word families and 2,000–3,000 word families are necessary to yield 98% coverage on the listening and reading tests respectively. Pedagogical implications focusing on how to fill the gap are discussed in detail.

Keywords: TOEIC, vocabulary size, text coverage, BNC/COCA word lists

Introduction

For many Japanese learners of English, success on the Test of English for International Communication (TOEIC) is of value for academic and occupational purposes. For instance, Aoyama Gakuin University uses scores on the TOEIC or other high-stakes English tests as one of its admission criteria. Senior high school students with 730 points or more out of the maximum possible score of 990 are eligible to apply to the school through the admissions office (Murai, 2016, January 25). Waseda University administers the TOEIC Institutional Program test to all commerce-major freshmen (Institute for International Business Communication, 2015, June). High scores on the TOEIC test can also be beneficial for university graduates and business people as survey results released by the Institute for International Business Communication (2013) show. Of 228 enterprises listed on the First Section of the Tokyo Stock Exchange that responded to questionnaires, around 69% of the companies consider TOEIC scores in screening applicants. Also, about 16% of the companies use TOEIC scores in making promotional decisions.

Given various advantages expected from achieving high scores on the TOEIC test, one may contemplate how to prepare for the test. For TOEIC examinees, improving listening and reading proficiency is crucial because listening and reading comprehension questions alone constitute 74% of the total number of questions on the test (i.e., 148 out of 200 questions).

Research suggests that the percentage of known words in a text, or *text coverage*, is one significant factor determining success at listening (Bonk, 2000; Stæhr, 2009; van Zeeland & Schmitt, 2013) and reading (Hu & Nation, 2000; Laufer, 1989; Laufer & Ravenhorst-Kalovski, 2010; Schmitt, Jiang, & Grabe, 2011) for English learners. Overall, the above-mentioned empirical studies suggest that knowing 95% to 98% of words on a text allows listeners or readers to gain adequate comprehension. Setting 95% text coverage as a criterion, Chujo and Oghigian (2009) investigated how large a vocabulary it takes to reach this coverage level on the TOEIC test. Two sets of older and current versions of official TOEIC practice tests were examined using three different word lists, which were derived from the British National Corpus (BNC). Their results indicate that around 3,000 word families from Nation's (2006) BNC word-family lists are necessary to achieve 95% coverage on the current version of the TOEIC test.

Although Chujo and Oghigian's (2009) study set an initial vocabulary-size target for the TOEIC test, there were several limitations in the procedure. First, the 3,000-word-family estimate was derived from a rather small corpus: two tests from the past and current versions of the TOEIC. Vocabulary loads may vary from test to test as Webb and Paribakht (2015) found when they profiled frequency levels of words appearing in the CanTEST, a standardized English test used in Canada. Considering possible deviations in frequency levels across tests, the sample size needs to be enlarged. Second, there was a mismatch between the test and the word lists used. Nation's (2006) word-family lists as well as other two lemma-based word lists were compiled from the BNC, which is largely written, British English as pointed out by Nation (2004). Given that the TOEIC test consists mainly of listening and reading comprehension questions and that both American and British English words are used, word lists derived from well-balanced samples of spoken and written English should provide more accurate results. Third, Chujo and Oghigian's (2009) vocabulary-size target might be underestimated because they mixed the listening and reading sections. Spoken and written English should be separated because spoken texts make greater use of high-frequency words than written English (e.g., Nation, 2006; Schannel et al., 1956; Webb & Paribakht, 2015). In fact, in measuring lexical requirements of standardized English tests, research has shown a strong tendency to investigate listening and reading tests separately (Taylor, 2014; Webb & Paribakht, 2015), to focus only on reading tests (Green, Unaldi, & Weir, 2010; Khalifa & Schmitt, 2010), or only on listening tests (Matthews & Chen, 2015; Stæhr, 2009). Fourth, the 95% coverage criterion may not be sufficient for some test takers. Laufer (1989) set 95% coverage as a cutoff point to separate good readers (defined as those who achieved 55% or more on a reading comprehension measure) from weak readers. However, not all TOEIC examinees are satisfied with such a modest comprehension level. They may wish for higher comprehension scores. As Schmitt et al. (2011) suggest, 98% of words in a written discourse should be known if readers desire high comprehension level (defined as 60% or more in reading comprehension scores). With regard to the ideal coverage in written texts, the consensus of opinion is 98% coverage (Hu & Nation, 2000; Khalifa & Schmitt, 2010; Laufer & Ravenhorst-Kalovski, 2010; Schmitt & Schmitt, 2014). Regarding listening, research has indicated that lower text coverage, say 90% (Bonk, 2000) or 95% coverage (van Zeeland & Schmitt, 2013), seems sufficient. However, repeated listening was allowed in these two

studies. Bonk (2000) allowed the participants to listen to the input text three times. Participants in van Zeeland and Schmitt's (2013) study listened to the target text twice. Since the listening of the TOEIC test is a one-off, 98% coverage seems more appropriate.

Given the above-mentioned limitations in Chujo and Oghigian's (2009) study, it seems reasonable to revisit a vocabulary-size target for the TOEIC test. The main aim of the present study is to explore how large a vocabulary is required to yield 98% coverage on the listening and reading sections of the TOEIC test respectively.

Method

Unit of Counting and Word Lists

In determining how large a vocabulary is required to provide 98% coverage on the listening and reading texts of the TOEIC test, *word families* were used as the unit of counting words in the present study. According to Schmitt and McCarthy (1997), a word family includes a base word (e.g., cover) with its inflections (covered, covers, covering) and derivatives (coverage, coverlet). Bauer and Nation (1993) categorized word families into seven different levels. In the present study, Level 6 was used. This level includes inflectional suffixes, and four types of affixes: frequent, regular, productive, and transparent derivational (Nation & Webb, 2011). Use of word families can be considered appropriate for high school graduates in Japan because it is in line with the current Course of Study Guidelines for upper secondary school. According to the guidelines compiled by the Japanese Ministry of Education, Culture, Sports, Science, and Technology (MEXT), words can be counted using lemmas or word families (MEXT, 2010, p. 44). Also, for takers of the TOEIC test, knowledge of word families is prerequisite because some questions in Part 5 (vocabulary measures) and 6 (cloze passages) directly test if they can identify the appropriate part of speech. Furthermore, corpus linguists generally agree on the use of word family for receptive purposes (e.g., Nation, 2006; Nation & Webb, 2011; Schmitt & Zimmerman, 2002). Therefore, word families can be considered appropriate in the present study since it aims to estimate vocabulary-size targets for the listening and reading sections of the TOEIC test.

In profiling frequency levels of words in the TOEIC test, Nation's (2012) word-family lists derived from the BNC and Corpus of Contemporary American English (COCA) were used. Nation's (2012) BNC/COCA lists would be more appropriate than his (2006) previous BNC lists because, as mentioned earlier, the TOEIC test involves both spoken and written English as well as British and American English.

Materials

Twelve sets of the listening and reading sections from official TOEIC practice tests (Educational Testing Service, 2005, 2007, 2008, 2009, 2012, 2014) were examined using a lexical frequency profiler called RANGE (Heatley, Nation, & Coxhead, 2002). More specifically, transcripts of the CD recordings from Part 1 to 4 (excluding directions) constituted listening texts for analyses. Words appearing in reading passages from Part 7 comprised reading texts. Written multiple-choice items and comprehension questions were excluded from analyses (spoken multiple-choice items from Parts 1 and 2 were included) because this decision allows researchers to compare results of the present study with those of

studies on other high-stakes tests. For instance, Green, Unaldi, and Weir (2010) investigated vocabulary appearing in reading passages of the International English Language Testing System (IELTS). In examining the listening section of the IELTS, Matthews and Chen (2015) focused exclusively on words appearing in listening passages. In determining vocabulary-size targets for the listening and reading (excluding graph-reading tasks) sections of the Test of English for Academic Purposes (TEAP), Taylor (2014) profiled frequency levels of words in the listening and reading passages respectively.

Before analyses, several modifications were made to the original texts. First, proper nouns and numerals were removed. One rationale was that proper nouns that are regular words affect coverage figures. For example, the proper noun *John Foster* affects frequency figures because the RANGE program classifies *Foster* as the regular verb *foster*, which occurs at the 3,000-word frequency level based on Nation's (2012) BNC/COCA lists. Other names such as *Bill*, *Brown*, and *Burns*, all appearing in one or more of the official TOEIC practice tests, would also affect results. The other reason behind deletion of proper nouns was that proper nouns are likely to be identified easily from contexts and do not require previous learning (Khalifa & Schmitt, 2010; Nation & Webb, 2011; Taylor, 2014; Webb & Paribakht, 2015). Therefore, proper nouns were removed from the text. Numerals were also deleted from the text because these non-words do not affect reading/listening comprehension. For the same reason, email addresses, telephone numbers, home addresses, symbols, and letters were excluded from analyses in the present study. AM and PM were also deleted because RANGE cannot identify homographs and homonyms. AM is categorized as an inflected form of the verb *to be* instead of AM, which represents the time of day. Some abbreviations were changed back to full forms because RANGE cannot identify them appropriately. These include *CEO*, *PhD*, *Pty*, *g*, *pg.*, *ma'am*. In the modified texts, they appear as *chief executive officer*, *doctor of philosophy*, *proprietary*, *gram*, *page*, *madam*. Similarly, the word *biz* was changed back to its original form *business*. Regarding hyphenated words, spaces were added before and after hyphens because RANGE cannot treat multiword units. One exception was the word *e-mail*. For this word, hyphens were simply removed because Nation's (2012) BNC/COCA word lists contain the word *email*.

After making the above-mentioned modifications to all the listening and reading passages, frequency levels of words in the texts were determined based on Nation's (2012) BNC/COCA word lists.

Results

Table 1 illustrates cumulative text-coverage figures on each listening test. In determining frequency levels of words in the listening section, coverage figures for exclamations such as *Ah* and *Oh*, commonly appearing in listening tests, were included in the coverage of known words at all levels. The most frequent 1,000 word families plus exclamations provided an average of 87.48% coverage. The most frequent 2,000 word families yielded a mean coverage of 94.74%. Average coverage of the most frequent 3,000 word families was 97.57%. In order to reach the crucial 98% coverage, vocabulary size between 3,000 and 4,000 word families was required.

Table 1

Vocabulary Size and Cumulative Text Coverage for the Listening Tests of the TOEIC

	Volume 1		Volume 2		Volume 3		Volume 4		Volume 5		Volume 6	
	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
Word families												
Exclamations	0.2	0.18	0.15	0.23	0.21	0.14	0.36	0.41	0.23	0.17	0.27	0.17
1,000	88.57	88.93	87.02	88.36	88.37	88.06	88.38	86.17	86.61	85.69	87.28	86.34
2,000	94.78	94.56	94.98	96.03	94.9	95.42	95.18	94.76	94.04	93	95.13	94.04
3,000	97.28	97.27	97.96	98.24	97.63	97.78	97.75	97.7	97.46	96.94	97.12	97.69
4,000	98.1	98.07	98.7	99	98.54	98.35	98.73	98.75	98.37	97.95	97.91	98.53
5,000	98.61	98.59	99.03	99.53	99.03	98.89	99.32	99.22	98.73	98.59	98.53	98.83
Tokens	3525	3252	2689	2619	2862	2799	3073	2957	3068	2969	2917	2986

Note. Coverage figures reaching 98% are in boldface.

Cumulative coverage figures on each reading test are shown in Table 2. The most frequent 1,000 word families provided an average of 74.63% coverage. With vocabulary size of the most frequent 3,000 word families, which is a learning requirement for senior high school students in Japan, a mean of 95% coverage was achieved. In order to reach 98% coverage, vocabulary size of the most frequent 5,000–6,000 word families was necessary. Practice Test 2 from the *New Official TOEIC Practice Test Volume 6* (Educational Testing Service, 2014) required vocabulary of the most frequent 7,000–8,000 word families. The major reason is the occurrence of a large number of compound words such as *feedback*, *breathhtaking*, and *skyscraper*. There are 23 compound words on this test, which account for 1.18% of the running words. If these compounds are known to test takers, then the most frequent 5,000–6,000 word families would be sufficient.

Table 2

Vocabulary Size and Cumulative Text Coverage for the Reading Tests of the TOEIC

	Volume 1		Volume 2		Volume 3		Volume 4		Volume 5		Volume 6	
	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
1,000	74.16	75.29	72.55	75.46	76.99	74.67	75.98	77.12	75.58	72.54	73.21	71.97
2,000	88.08	89.82	86.62	88.49	90.19	88.71	90.25	89.81	88.63	86.74	87.47	86.39
3,000	95.38	95.66	94.93	95.21	95.22	94.66	96.19	95.95	94.36	94.69	94.56	93.24
4,000	96.85	97.49	96.59	97.66	96.86	96.58	97.64	97.75	96.26	97.2	96.71	95.49
5,000	97.6	98.76	97.4	98.34	97.96	97.56	98.82	98.39	96.98	98.1	97.47	96.46
6,000	98.17	98.98	98	98.54	98.41	98.2	99.21	98.61	98.27	98.62	98.78	97.38
7,000	98.38	99.2	98.64	98.62	98.56	98.39	99.41	98.91	98.38	98.71	99.16	97.89
8,000	98.63	99.27	98.9	98.78	98.71	98.54	99.57	99.02	98.68	98.85	99.29	98.25
Tokens	2794	2671	2346	2486	2008	2657	2544	2671	2637	2112	2370	1955

Note. Coverage figures reaching 98% are in boldface.

Discussion

If high school graduates acquire the most frequent 3,000 word families as stated in the current Course of Study Guidelines for upper secondary school (MEXT, 2010), then an average of 95% coverage can be expected both in the listening and reading sections of the TOEIC test. This finding is in line with that of Chujo and Oghigian's (2009) study. However, in order to reach 98% coverage, an additional 1,000 word families was necessary for the listening test, and 2,000–3,000 word families for the reading test. These findings pose two important pedagogical issues.

First of all, classroom practitioners should ensure that students can identify the most frequent 3,000 words both phonologically and orthographically. As some researchers (e.g., Milton & Hopkins, 2006; van Zeeland & Schmitt, 2013) note, being able to recognize written words does not necessarily guarantee that these learners can perceive the same words in a spoken form. In fact, Milton and Hopkins (2006) found that there was a significant difference between written and aural vocabulary size among 126 EFL participants (38 Arabic and 88 Greek speakers). Although no Japanese learners of English were involved in the study, their findings indicate that some EFL learners may have an imbalance between their written and spoken vocabulary size. This difference can be explored using written vocabulary measures such as the New Vocabulary Levels Test (McLean & Kramer, 2015) and spoken vocabulary tests such as the Listening Vocabulary Levels Test (McLean, Kramer, & Beglar, 2015).

In addition to students' vocabulary size, depth of their lexical knowledge should be examined. As Nation (2013) states, knowing vocabulary involves various aspects of words. More specifically, teachers should examine students' various lexical knowledge such as associations, grammatical functions, collocations, register, and word-family members of the most frequent 3,000 words.

After ensuring that students can recognize the first 3,000 words in forms of written and spoken words and that they are well familiar with various facets of the most frequent 3,000 words, teachers should help them acquire additional 2–3,000 words beyond the most frequent 3,000 words. But how can this lexical-size gap be filled?

One way to address this issue is to directly teach 2,000 to 3,000 words beyond the most frequent 3,000 word families, and help students strengthen their knowledge through extensive reading or listening. Test takers can gain a partial knowledge of mid-frequency vocabulary using Nation's (2012) BNC/COCA word-family lists. However, deepening knowledge on various aspects of vocabulary beyond the most frequent 3,000 word families through the use of commercial graded readers is not plausible, as Schmitt and Schmitt (2014) point out. As these researchers note, most published graded readers finish at around the 3000-word frequency level. For instance, there are seven different levels available in Longman's *Penguin Readers*. However, even Level 6, the most advanced level, deals with only 3,000 head words (essentially word families). *Oxford Bookworms* also has seven grades, but even the most advanced stage employs only 2,500 head words. Compared to these two graded readers, Level 6 of *Cambridge English Readers* is promising as it deals with 3,800 head words, indicating that aural input provided from the accompanying CDs can be used to prepare for the listening section of the TOEIC test. However, none of the three commercial graded readers sets fully prepare test takers for the reading section. One possible solution is to use Nation's (n.d.) mid-frequency readers, which are freely downloadable from his website. There are three different grades available: 4,000, 6,000, and 8,000 word-family levels. Use of mid-frequency readers can augment learners' knowledge beyond the most frequent 3,000 word families. If learners have difficulty with a particular level, say the 6,000-word frequency level, classroom practitioners can reduce the vocabulary burden using the RANGE program. Some words from the 6,000-word frequency level can be replaced with more common synonyms from a higher frequency level (i.e., 5,000- or 4,000-word frequency level).

The other way to help test takers with a 3,000-word vocabulary to prepare for the TOEIC test is to create word lists specifically aimed for the listening or the reading section of the TOEIC test. In creating word lists, frequency and range are crucial criteria (Nation & Webb, 2011). However, making such word lists is beyond the scope of the present study and not plausible due to the limited sample size. That being said, there seem to be useful words that occur across tests. Of words beyond the most frequent 3,000 word families, there are 68 word families that appear in a quarter of tests (i.e., range 3 or more out of 12 tests). Learning these words would provide additional 1.78% coverage on combined 12 reading tests, although learning the most frequent 3,000 word families plus these additional 68 word families would still not reach 98% coverage on the reading tests. As more official practice and actual tests become available in the future, we may be able to create word lists which provide 98% coverage on the reading section.

Conclusion

Before the findings of the present study are considered conclusive, we need to think about its limitations. First, test takers and classroom practitioners need to pay close attention to homographs and homonyms because RANGE (Heatley et al., 2002) cannot deal with them

(Webb & Nation, 2011). For instance, the word family *sincere* appears in 11 out of 12 reading test papers. In teaching this frequently occurring word, attention should be paid to its multiple meanings. In most cases, *sincerely* is used as a closing. However, *sincerely* has another sense, which expresses a true or honest feeling. Test takers need to keep this distinction in mind. Similarly, the word *regard* appears across 10 different reading tests. In teaching this word, a caution is necessary because RANGE cannot separate *regarding* (as an inflectional form of the verb *regard*) from the preposition *regarding*.

Another limitation lies in the fact that coverage figures for the most frequent 1,000 word families include those of phrasal verbs. For instance, both *go* and *over* are at the 1,000-word frequency level. As Garnier and Schmitt (2015) point out, some learners may try to derive the meaning from each individual word. However, knowing each word or having a vocabulary size of the most frequent 1,000 word families does not ensure that the test taker knows the meaning of the phrasal verb *go over*. Classroom practitioners should test learners' knowledge of phrasal verbs. Creating tests based on Liu's (2011) phrasal-verb list is highly recommended since the list was compiled from the BNC/COCA.

The other limitation is that test takers need to know abbreviated forms and their corresponding full forms. As described in the Method section, abbreviations were changed back to their full forms because RANGE (Heatley et al., 2002) cannot recognize them. However, there are only seven words that required this attention (i.e., *CEO, PhD, Pty, g, pg., ma'am, biz*). Classroom practitioners can address this issue in a single lesson.

In conclusion, given that senior high school students in Japan are expected to acquire a vocabulary of 3,000 word families (MEXT, 2010), university freshmen or high school graduates are recommended to learn an additional 2,000–3,000 word families if they do not want to deal with more than 2% unknown vocabulary (i.e., one unknown word in every 50 words) while engaging in the reading section of the TOEIC test. Results of the present study show that the listening section of the TOEIC test requires a much smaller vocabulary size than the reading section. Even so, test takers with vocabulary size of the most frequent 3,000 word families would have to deal with 2.47% unknown vocabulary (one unknown word in every 40 words) while listening. Test takers should strive to fill this lexical-size gap if 98% coverage is desired. That being said, it should be noted here that 98% coverage is a rather tentative target figure for the TOEIC test because there has been no research investigating the extent to which text coverage affects listening and reading comprehension scores on the TOEIC test. Until empirical research is conducted, results of the present study should be considered tentative. One important pedagogical implication is that commercial graded readers may not be sufficient for TOEIC preparation. Use of mid-frequency readers (Nation, n.d.) or word lists aimed for the TOEIC is strongly recommended for TOEIC test takers.

Acknowledgements

I am grateful to three anonymous reviewers for their insightful input on an earlier version of this paper.

References

Bauer, L., & Nation, I. S. P. (1993). Word families. *International Journal of Lexicography*, 6(4),

253–279.

- Bonk, W. (2000). Second language lexical knowledge and listening comprehension. *International Journal of Listening*, 14(1), 14–31. doi: 10.1080/10904018.2000.10499033
- Chujo, K., & Oghigian, W. (2009). How many words do you need to know to understand TOEIC, TOEFL and EIKEN? An examination of text coverage and high frequency vocabulary. *The Journal of Asia TEFL*, 6(2), 121–148.
- Educational Testing Service. (2005). *TOEIC test shin koshiki mondaishu* [new official TOEIC practice test]. Tokyo, Japan: The Institute for International Business Communication.
- Educational Testing Service. (2007). *TOEIC test shin koshiki mondaishu Volume 2* [new official TOEIC practice test] (Vol. 2). Tokyo, Japan: The Institute for International Business Communication.
- Educational Testing Service. (2008). *TOEIC test shin koshiki mondaishu Volume 3* [new official TOEIC practice test] (Vol. 3). Tokyo, Japan: The Institute for International Business Communication.
- Educational Testing Service. (2009). *TOEIC test shin koshiki mondaishu Volume 4* [new official TOEIC practice test] (Vol. 4). Tokyo, Japan: The Institute for International Business Communication.
- Educational Testing Service. (2012). *TOEIC test shin koshiki mondaishu Volume 5* [new official TOEIC practice test] (Vol. 5). Tokyo, Japan: The Institute for International Business Communication.
- Educational Testing Service. (2014). *TOEIC test shin koshiki mondaishu Volume 6* [new official TOEIC practice test] (Vol. 6). Tokyo, Japan: The Institute for International Business Communication.
- Garnier, M., & Schmitt, N. (2015). The PHaVE List: A pedagogical list of phrasal verbs and their most frequent meaning senses. *Language Teaching Research*, 19(6), 645–666. doi: 10.1177/1362168814559798
- Green, A., Unaldi, A., & Weir, C. (2010). Empiricism versus connoisseurship: Establishing the appropriacy of texts in tests of academic reading. *Language Testing*, 27(2), 191–211. doi: 10.1177/0265532209349471
- Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). RANGE and FREQUENCY programs. Retrieved from <http://www.victoria.ac.nz/lals/about/staff/paul-nation>
- Hu, M., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403–430.
- Institute for International Business Communication. (2013). *Jojokigyo ni okeru eigo katsuyo jittai chosa hokokusho* [Survey on how English is used at enterprises listed on the First Section of the Tokyo Stock Exchange]. Retrieved from http://www.toEIC.or.jp/toEIC/about/data/katsuyo_2013.html
- Institute for International Business Communication. (2015, June). *TOEIC newsletter* (Vol. 125). Retrieved from http://www.toEIC.or.jp/toEIC/about/data/doc_newsletter.html
- Khalifa, H., & Schmitt, N. (2010). A mixed-method approach towards investigating lexical progression in Main Suite Reading test papers. *Cambridge ESOL: Research notes*, 41, 19–25.
- 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., & 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.
- Liu, D. (2011). The most frequently used English phrasal verbs in American and British English: A multicorpus examination. *TESOL Quarterly*, 45(4), 661–688. doi: 10.5054/tq.2011.247707
- Matthews, J., & Chen, J. (2015). Recognition of high frequency words from speech as a predictor of L2 listening comprehension, *System*, 52, 1–13. doi: 10.1016/j.system.2015.04.015
- McLean, S., & Kramer, B. (2015). The creation of a New Vocabulary Levels Test. *Shiken*, 19(2), 1–9.
- McLean, S., Kramer, B., & Beglar, D. (2015). The creation and validation of a listening vocabulary levels test. *Language Teaching Research*, 19(6), 741–760. doi:10.1177/1362168814567889
- Milton, J., & Hopkins, N. (2006). Comparing phonological and orthographic vocabulary size: Do vocabulary tests underestimate the knowledge of some learners. *The Canadian Modern Language Review*, 63(1), 127–147. doi: 10.3138/cmlr.63.1.127
- Ministry of Education, Culture, Sports, Science, and Technology. (2010). *Kotogakkou gakushu shido yoryo kaisetsu: gaikokugo hen · eigo hen* [Description of the course of study guidelines for upper secondary school: Foreign languages education, English]. Tokyo, Japan: Kairyudo.
- Murai, S. (2016, January 25). Changes in store for TOEIC, but test still not total gauge of fluency: experts. *The Japan Times*. Retrieved from <http://www.japantimes.co.jp/news/2016/01/25/reference/changes-store-toeic-test-still-not-total-gauge-fluency-experts/#.VyM2qMtJn4g>
- Nation, I. S. P. (n.d.). *About mid-frequency readers*. Retrieved from <http://www.victoria.ac.nz/lals/about/staff/paul-nation>
- Nation, I. S. P. (2004). A study of the most frequent word families in the British National Corpus. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (pp. 3–13). Amsterdam, Netherlands: John Benjamins.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, 63(1), 59–82. doi: 10.3138/cmlr.63.1.59
- Nation, I. S. P. (2012). *The BNC/COCA word family lists*. Retrieved from <http://www.victoria.ac.nz/lals/about/staff/paul-nation>
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). New York, NY: Cambridge University Press.
- Nation, I. S. P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Boston, MA: Heinle.
- 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. doi: 10.1111/j.1540-4781.2011.01146.x
- Schmitt, N., & McCarthy, M. (1997). *Vocabulary: Description, acquisition and pedagogy*.

Cambridge, UK: Cambridge University Press.

- Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary testing. *Language Teaching*, 47(4), 484–503. doi: 10.1017/s0261444812000018
- Schmitt, N., & Zimmerman, C. B. (2002). Derivative word forms: What do learners know? *TESOL Quarterly*, 36(2), 145–171.
- Schonell, F. J., Meddleton, I. G., Shaw, B. A., Routh, M. R., Popham, D. F., Gill, J. D., . . . Stephens, C. (1956). *A study of the oral vocabulary of adults* (Report No. 1). Brisbane, Australia: The University of Queensland Press.
- Stæhr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition*, 31(4), 577–607. doi: 10.1017/s0272263109990039
- Taylor, L. (2014). A report on the review of test specifications for the reading and listening papers of the test of English for academic purposes (TEAP) for Japanese university entrants. *Eiken Report*, 1–43. Tokyo, Japan: Eiken Foundation of Japan. Retrieved from <http://www.eiken.or.jp/teap/group/report.html>
- Van Zeeland, H., & Schmitt, N. (2013). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics*, 34(4), 457–479. doi: 10.1093/applin/ams074
- Webb, S., & Paribakht, T. S. (2015). What is the relationship between the lexical profile of test items and performance on a standardized English proficiency test? *English for Specific Purposes*, 38, 34–43, doi: 10.1016/j.esp.2014.11.001