# Vocabulary in Japanese EFL Textbooks: A Bidirectional Coverage Analysis

Shusaku Nakayama, Meiji Gakuin University, Japan

**Abstract**

Textbooks developed based on the new Course of Study Guidelines have been in use since April 2021. One notable change can be seen in the increase of the variety of vocabulary words, from 1,200 words to between 1,600–1,800 words, that learners need to be familiar with before they graduate junior high school. In this study, vocabulary taught in these newly published textbooks designed for junior high school students in Japan will be assessed. For this purpose, a corpus of junior high school textbooks, which is sourced from one series of government-approved junior high school textbooks, is compared with the New General Service List (NGSL) that consists of 2,801 high-frequency words in general English. Results show that the textbook series is largely composed of vocabulary words contained in the NGSL with a greater than 95% coverage; however, it represents only a small part of the list with a smaller than 37% coverage. Additionally, breaking the NGSL down into 560-word frequency bands, the study investigates in which bands the textbook series focus and in which bands there is a deficiency. This distribution analysis indicates that words at higher frequency bands occur more frequently. The textbook series covers the first 560-word frequency band with a greater than 80% coverage, but the coverage levels sharply decrease after this band. Finally, pedagogical implications are suggested for textbook designers as well as for language learners.


Keywords: Coverage Analysis, Japanese EFL Textbooks, New General Service List, Textbook Corpus, Vocabulary

iafor

The International Academic Forum
www.iafor.org

**Introduction**

Last year, the previous Course of Study Guidelines for junior high schools came to an end after 11 years. From April 2021, the new Course of Study Guidelines has been implemented (Ministry of Education, Culture, Sports, Science and Technology [MEXT], 2017a). With the implementation, textbooks developed based on it have been made use of. One notable change can be seen in the increase of the variety of vocabulary words learners need to be familiar with before they graduate from junior high school. Unlike the previous Course of Study Guidelines, which declared that approximately 1,200 words must be learned throughout three years of English learning in junior high school (MEXT, 2009), the new Course of Study Guidelines (MEXT, 2017a) intends to familiarize students with 1,600–1,800 words. At first, this increase seems to be beneficial for learners as it is a widely accepted idea that an increase in vocabulary size has a positive effect on one's reading proficiency (Hu & Nation, 2000; Laufer & Ravenhorst-Kalovski, 2010); however, the increase could simultaneously cause some disadvantages for learners.

For example, an oversupply of new words within a single textbook possibly prevents language teachers from only giving a brief description of a word to learners because they may unintentionally spend a huge amount of time explaining the meaning of a new word. According to Nation (2013), three aspects are involved in vocabulary knowledge, namely, form, meaning, and use. Each has three components. Form refers to knowledge related to (a) pronunciation, (b) orthography, and (c) morphology. Meaning refers to knowledge regarding (a) the meaning of the word form and a word form that can express the meaning, (b) the concepts the word has and referents of the concept, and (c) other words semantically associated with the word. Use refers to knowledge related to (a) grammatical functions, (b) collocations, and (c) registers to use the word or the frequency of occurrence of the word. Consequently, memorizing only word meaning is not sufficient. However, the more frequently new words appear, the more likely it is that teachers will be required to spend more time explaining their meaning, possibly leading to teachers imparting superficial knowledge of those words to their students. To prevent this from occurring, textbooks should provide an appropriate number of various words that are worthwhile for students to learn. Furthermore, it is a reasonable learning goal for learners to reach vocabulary thresholds needed for comprehension of a text (Nation, 2006).

For these reasons, treating a wider variety of words does not necessarily benefit students. Therefore, textbook designers should consider how well their textbooks are able to increase learners' vocabulary size to satisfy vocabulary thresholds for comprehension, rather than the variety of words learners can learn through their textbooks. This study thus explored this issue. In addition, it examined how much of the vocabulary, taught in MEXT-approved textbooks intended for junior high school students, were high-frequency words in order to determine whether textbooks developed based on the new Course of Study Guidelines provided sufficient worthwhile words for students to learn in terms of frequency.

**Literature review**

**What variety of words do learners need to know?**

Given that learners' goals are to reach vocabulary thresholds where sufficient comprehension can be achieved, what variety of words would learners need to know? This issue has been greatly controversial among researchers. In the 1990s, researchers suggested that 95% of the

running words in a text should be known for reading a text (Laufer, 1989; Liu & Nation, 1985) and for guessing from context (Liu & Nation, 1985). Subsequently, Hirsh and Nation (1992) argued that 97–98% coverage is needed to read a text for pleasure, which was empirically supported by Hu and Nation (2000). They gave the same text to four learner groups with a differing number of unknown words: 0%, 5%, 10%, and 20%. They conducted reading comprehension tests on the text and found a close relationship between the density of known words in a text and the likelihood of adequate comprehension of the text. None of the learners with 20% unknown words could comprehend the text adequately. With 90% and 95% of known words, some learners succeeded in comprehending the text; however, these thresholds could not ensure most learners' adequate comprehension of the text. They concluded 98% coverage was an optimal threshold for comprehension, which was echoed by Laufer and Ravenhorst-Kalovsk (2010); nevertheless, 95% coverage is also accepted as a minimum threshold for comprehension (Laufer & Ravenhorst-Kalovsk, 2010).

According to the previous studies reviewed here, two threshold levels can be suggested as reasonable targets for learners' goals, namely, 95% and 98% coverage levels. The most frequent 4,000–5,000 word families can provide 95% coverage (Laufer & Ravenhorst-Kalovsk, 2010); for 98% coverage, 8,000–9,000 of the most frequent word families are necessary (Nation, 2006).

**Research into vocabulary taught in Japanese EFL textbooks**

Most previous studies have investigated vocabulary in Japanese EFL textbooks primarily by measuring their vocabulary levels. For example, Chujo (2004) measured the vocabulary levels of combinations of junior and senior high school EFL textbooks by using the British National Corpus (BNC). The results showed that 95% coverage was achieved with the 3,200 most frequent lemmas of the list. In Wongsarnpigoon's (2018) study, it was found that 95% coverage of MEXT-approved textbooks designed for junior high school students was provided with the top 2,000 high-frequency words of the BNC.

Information on the vocabulary demands of textbooks is informative in evaluating the appropriateness of textbooks as teaching materials. If the vocabulary level of a textbook is too high to cover, for example, 95% coverage of it with basic words, the textbook might be inclined to teach vocabulary that learners will be less likely to encounter in the real world and would, therefore, need some modifications.

However, those previous studies might not be sufficient since textbooks are not reading materials but teaching materials. In this respect, the primary focus of research into vocabulary in textbooks should be on the investigation of how many types of high-frequency words students can potentially learn through textbooks. By doing so, textbook developers can ascertain how well their textbooks grow learners' vocabulary size to meet the necessary requirements for comprehension and would observe the potential deficiency in the variety of high-frequency words, which can then be used in revising textbooks. There are several studies on this issue conducted for coursebooks (e.g., Eldridge & Neufeld, 2009; O'Loughlin, 2012); all of these studies succeeded in insisting on the importance of supplemental input sources in addition to coursebooks. Despite this pedagogical importance, little research has adopted this perspective in Japan.

Please note that essentially, it would be almost impossible for junior high school students to meet all the necessary words for comprehension in textbooks throughout three years of

English learning as textbooks are too short to cover them. According to Waring (2009), 132,143 words are needed to meet each of the most frequent 5,000 words at least once, the requirement for achieving 95% coverage. However, the textbooks series analyzed in this study includes only 45,412 words (see Table 1). Nevertheless, this perspective can possibly provide valuable insights into the development of textbooks as observed in previous studies. Research questions addressed in the study were as below:

1. How much of the vocabulary in MEXT-approved textbooks consists of high-frequency words?
2. What variety of high-frequency words do learners meet in a set of three MEXT-approved textbooks?

As a follow-up study, I further investigated which levels of high-frequency words (e.g., the first 1,000-word level) MEXT-approved textbooks focused on and in which they had a deficiency in order to closely examine vocabulary in textbooks. The following question was used:

3. Which levels of high-frequency words do MEXT-approved textbooks focus on and in which levels is there a deficiency?

**Methodology**

**Textbook corpus**

In Japan, there are six series of MEXT-approved junior high school EFL textbooks, each consisting of three types of textbooks. Among them, the current study analyzed one series of textbooks, *Here We Go!* (Ota et al., 2021a,b,c), for the reason that unlike the other textbook series, it was possible to include all the English words that learners are supposed to encounter. Text data included those used in frontispieces to introduce each section, reading passages, language activities, example sentences to explain grammatical rules, and lists of new words. English words occurring in listening activities were transcribed by the author and included in the corpus. Significantly, to ensure representative sampling, the corpus includes both the reading and listening versions of the same passages, as the textbook series recommends that learners not only read passages, but also listen to them (see red circled part in Figure 1) and reading and listening passages often differ from each other. The information on the completed corpus is summarized in Table1.
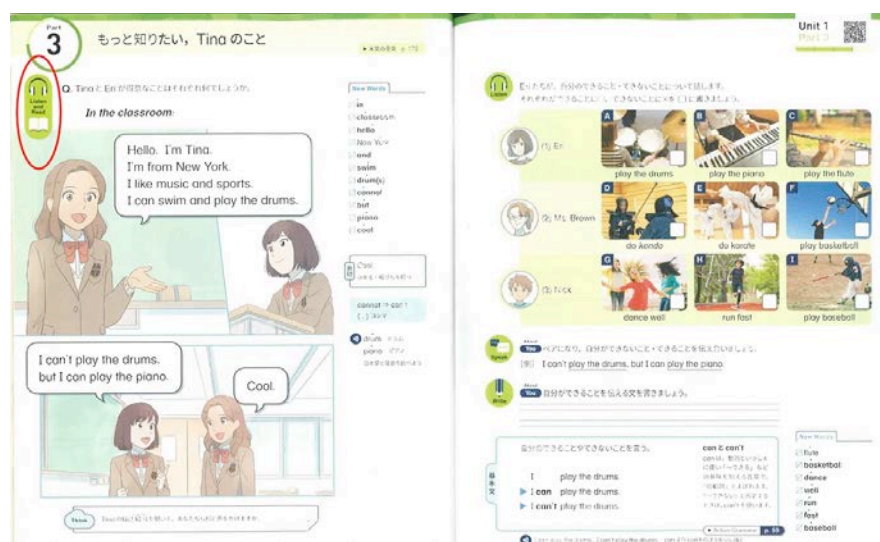

Figure 1: Screenshot of 'Here We Go!'

| Textbook | Type | Token |
|---|---|---|
| Here WE Go! 1 (Ota et al., 2021a) | 976 | 12,094 |
| Here WE Go! 2 (Ota et al., 2021b) | 1,319 | 15,835 |
| Here WE Go! 3 (Ota et al., 2021c) | 1,588 | 17,483 |
| Total | 2,434 | 45,412 |

Table 1: Information on the textbook corpus

**New General Service List**

Due to researchers' concerted efforts, there are numerous freely available word lists, for example, the General Service List (West, 1953) and the Academic Word List (Coxhead, 2000). In this study, the New General Service List (NGSL) ver. 1.01. (Browne et al., 2013), which is sourced from more than 273 million words within the Cambridge English Corpus and can cover approximately 92% of the words in a general text with 2,801 words (Browne, 2021), was used. Unlike those widely used word lists applying the word family count, where the base form of a word and its inflections and derived forms of the word are counted as one word, Browne et al. (2013) counted words with a modified lexeme approach that corresponds to the word counting unit of *flemmas* (McLean, 2017). Their modified lexeme and McLean's flemma word counting unit include the inflected forms of a word in different parts of speech in addition to those in the same part of speech.

**Procedure**

The procedure of the study largely followed that of Nakayama (2021) for the analysis of MEXT-approved textbooks intended for senior high school students. He adopted a unique approach for the study as currently available corpus-analysis tools compatible with the NGSL such as AntWordProfiler 1.5.1 (Anthony, 2021), the Online Graded Text Editor (Waring & Browne, n.d.), and VocabProfiler (Cobb, 2021a) are not geared toward analyzing the NGSL itself but only imported texts of interest. Another rationale for not using these tools were that they sometimes incorrectly identify the intended meaning of homographs or components of contracted forms (e.g., "I'd" used as "I would" or "I had").

Following Nakayama's (2021) study, this study also analyzed the textbooks with Text Lex Compare ver. 4.2. (Cobb, 2021b). This tool compares two different texts and informs us of the number of tokens, word types, word families, and word phrases used in, and unique to, each of the imported texts. To compare the textbooks with NGSL in a reasonable way, all the words in the textbooks were replaced with their headwords based on the modified lexeme approach by using the replacement function of Microsoft Word 2016. This was also used to separate contracted forms. Semantically opaque ones as in the example above and homographs were replaced with their headwords while considering the surrounding context.

Text data after these treatments were loaded into WordSmith 7 (Scott, 2016) to produce word lists; they were used to visually ascertain whether every word in the textbooks was successfully replaced with its headword. Information on the corpus after these treatments is summarized in Table 2.

| Textbook | Type | Token |
|---|---|---|
| Here WE Go! 1 (Ota et al., 2021a) | 608 | 11,227 |
| Here WE Go! 2 (Ota et al., 2021b) | 866 | 14,661 |
| Here WE Go! 3 (Ota et al., 2021c) | 984 | 16,576 |
| Total | 1,401 | 42,464 |

Table 2: Information on the corpus after the treatments

The NGSL comprises 52 supplemental words, including days of the week, months of the year, and numbers. These words were considered for calculating the lexical coverage of the textbooks to prevent their frequent recurrences in a single textbook from underestimating the lexical coverage of each book. In contrast, they were not considered for the calculation of the coverage of NGSL words, for the reason that not presenting all numerals or months of the year in a single textbook would be impossible. Using Text Lex Compare, the lexical coverage of the textbooks, i.e., (a) and the coverage of NGSL words, i.e., (b) were calculated with the following formulae:
  (a) (Number of words in the textbook – Number of words unique to the textbook)/Number of words in the textbook
  (b) (2,801 – Number of words unique to the NGSL)/2,801

Unlike other word lists that are often broken down into 1,000-word bands, the study used 560-word frequency bands just as researchers used the NGSL for the creation of a reliable and valid diagnostic and placement test of NGSL knowledge (Stoeckel & Bennett, 2015) or that of free flashcard learning apps (Browne & Waring, n.d.).

Japanese EFL learners begin to study English as a compulsory subject from the 5th grade of elementary school with 70 classes a year, and MEXT (2017b) intends to familiarize elementary school students with 600–700 words, implying that junior high school students have only started English learning and have not been exposed to most of the high-frequency words required for comprehension. In this respect, the skewness of high-frequency words within a frequency band is a crucial problem because they cannot encounter some higher frequency words. In order not to overlook the skewed distribution of words within a band as much as possible, the study analyzed the textbooks by a narrower size of word bands.

**Results**

**Lexical coverage of the textbooks by the NGSL**

According to Table 3, regardless of the number of tokens used in the textbooks, 95% of words were covered by the NGSL.

| Grade | Number of tokens | Number of tokens outside the NGSL | Coverage |
|---|---|---|---|
| JH1 | 11,227 | 551 | 95% |
| JH2 | 14,661 | 738 | 95% |
| JH3 | 16,576 | 727 | 96% |
| Total | 42,464 | 2,016 | 95% |

Table 3: Results of the lexical coverage analysis

**Coverage of the NGSL by the textbooks**

Table 4 illustrates the extent to which the textbooks covered NGSL words. Obviously, the textbooks intended for higher graders produced higher coverage for the NGSL. Users of this textbook series can potentially meet 37% of NGSL words.

The combined results of the two types of coverage analyses indicate that the textbooks were largely composed of NGSL words, but they covered only a very small part of the NGSL with 37% throughout three years of English learning in junior high school.

| Grade | NGSL words not covered by the textbook | Coverage |
|---|---|---|
| JH1 | 2,311 | 17% |
| JH2 | 2,113 | 25% |
| JH3 | 2,000 | 29% |
| Cumulative | 1,758 | 37% |

Table 4: Coverage of 2,801 NGSL words by the textbooks

**Distribution analysis**

Lastly, we examine which of the 560-word frequency bands the textbooks focused on and of which there was a shortage. In Table 5, the columns (a) stand for the percentage of tokens occurring at the frequency band; the columns (b) stand for the coverage of the frequency band by the textbook.

| Grade | NGSL 1 (1–560) | | NGSL 2 (561–1120) | | NGSL 3 (1121–1680) | | NGSL 4 (1681–2240) | | NGSL 5 (2241–2801) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (a) | (b) | (a) | (b) | (a) | (b) | (a) | (b) | (a) | (b) |
| JH1 | 81.2% | 48.9% | 6.4% | 16.4% | 3.3% | 11.1% | 2.0% | 5.2% | 1.4% | 5.7% |
| JH2 | 82.2% | 62.7% | 6.6% | 26.8% | 3.1% | 16.8% | 1.7% | 8.8% | 0.9% | 7.8% |
| JH3 | 81.4% | 72.9% | 6.2% | 31.6% | 3.9% | 20.5% | 1.8% | 10.0% | 1.1% | 8.0% |
| Cumulative | 81.6% | 80.5% | 6.4% | 42.5% | 3.5% | 30.2% | 1.8% | 16.8% | 1.1% | 16.2% |

Table 5: Distribution of words occurring in the textbook across five frequency bands

There are two common characteristics across the three levels of the textbooks. First, looking at the columns (a), NGSL words at higher frequency bands occurred more frequently in every textbook than those at lower frequency bands. Second, all the textbooks produced higher coverage for higher frequency bands.

Overall, the textbooks for higher graders better covered the NGSL. Another notable fact is that the coverage of the NGSL decreased not gradually but sharply after the first frequency band. Students who use this textbook series cannot be exposed to over 50% of NGSL words through the second to third frequency bands.

**Discussion**

The lexical coverage analysis showed that every textbook of interest was largely composed of NGSL words with a greater than 95% coverage. This result reasonably matched Wongsarnpigoon's (2018) findings that 95% coverage of MEXT-approved junior high school textbooks was provided with the top 2,000 high-frequency words in the BNC. Therefore, the vocabulary words the students who use this textbook series meet is worth learning in terms of

frequency. The distribution analysis identified that the textbooks covered a wider variety of words at higher frequency bands with higher frequency. Additionally, the textbooks for higher graders produced higher coverage for the NGSL as well as for each frequency band. In conclusion, learners can possibly learn core words in descending order of frequency, and the textbook series is designed so that users can widen their vocabulary size of high-frequency words as they go to higher grades.

In contrast to these advantages, the coverage analysis of the NGSL exposed the shortage of high-frequency words in the textbooks in terms of variety. Users of the textbook series cannot meet 63% of NGSL words (i.e., 1,043 NGSL words) throughout three years of English learning. Moreover, the distribution analysis identified that students cannot be exposed to most high-frequency words other than those at the first frequency band. Therefore, having learners use supplemental input sources would be helpful for them.

Here, it could be worth mentioning that the textbooks analyzed in this study better covered the NGSL than most of the MEXT-approved senior high school textbooks analyzed in Nakayama's (2021) study, not only at an overall coverage level but also at each frequency band level. He analyzed three MEXT-approved senior high school textbooks intended for 1st, 2nd, and 3rd graders respectively. Among them, all the 1st graders' books did not cover the NGSL as much as the 1st graders' book analyzed in this study. In addition, the 1st graders' book produced higher coverage for each of the five frequency bands than four types of senior high school textbooks. For the 3rd graders' junior high school textbook, it produced higher coverage, both for the NGSL and for every frequency band, than six types of senior high school textbooks. This superiority of junior high school textbooks is possibly due to differences in the design of corpora. Compared to the current study, limiting the scope of language data to reading passages produced lower coverage than that observed in this study. Nevertheless, junior high school textbooks may prove more useful for learning high-frequency words.

For junior high school textbooks to include a wider variety of high-frequency words, a specific suggestion to textbook designers is to decrease the number of recurrences given to items recurring many times and use the space allocated to them in order to present high-frequency words not taught in the book. Repeated exposure to an item is an important factor in vocabulary learning (Waring & Takaki, 2003; Webb, 2007). For this reason, researchers disputed the insufficient recycling of words within a textbook (Waring, 2011; Wongsarnpigoon, 2018). Nevertheless, textbooks are not designed for those who learn vocabulary in an incidental manner. Textbook users would study hard for term tests and can receive a pedagogical intervention, so they can learn vocabulary through a variety of ways such as language activities, teachers' explanations, peers' utterances, homework, and so on. Thus, it is less likely that students, at least Japanese learners, are exposed to every word appearing in textbooks only once. In other words, students may need fewer encounters to gain knowledge of one word than those who get the input of vocabulary only from reading books. Textbook designers need to consider the distribution of the word within the textbook, rather than the number of repetitions of the item. Specifically, textbook designers should develop textbooks so that high-frequency words appear periodically to increase the chance that learners can become familiar with them and not forget them easily. This kind of repetition is necessary for learners from the perspective of the spacing effect, the phenomenon that people can better retain what they have learnt when their repeated exposure to items have time intervals than when their repeated exposure to them are massed (Nakata, 2015; Sobel et al., 2011). Therefore, if textbook designers make concerted efforts to allow

high-frequency words to appear periodically within a textbook, it could be possible for textbooks to expose learners to a wider variety of high-frequency words and to acquire them even with a relatively limited number of recurrences.

Designing textbooks in this way can result in a shortcoming that learners may be exposed to each item in a way that differs from how it is used in the real world and cannot learn multiple aspects involved in one word as described by Nation (2013). To compensate for this, engaging students in extensive reading can be helpful. As learners are exposed to known words repeatedly, they can deepen their existing knowledge of them (Waring & Takaki, 2003; Webb, 2007). However, in this way of learning there are some drawbacks to consider. First, it would take a huge amount of time to learn core words not appearing in textbooks as no one is able to anticipate when learners can meet unknown words; therefore, it may result in wasting time (Waring, 2009). According to Cobb (2008), reading over 375,000 words using graded readers is still not sufficient for learners to gain knowledge of even the 3,000 most frequent words. Second, for extensive reading to work well, it is indispensable that "only a small proportion of the language they need to use is not familiar to them" (Nation, 2007, p. 2). It is possible that such beginners as junior high school students in Japan cannot read even simplified text in the first place and, therefore, they cannot enjoy the benefits of extensive reading.

These disadvantages of extensive reading can be compensated for with deliberate learning (Nation, 2015). To teach high-frequency words that students cannot encounter in their textbook series, teachers can ask their students to learn those words from word cards (Nation, 2013). However, if teachers did not have the word list of the textbook they use in the classroom, a coverage analysis of a word list comprised of high-frequency words (e.g., NGSL) by the textbook, as done in this study, would be indispensable, validating the usefulness of this analytical approach, which has been less adopted in Japan. This deliberate learning would allow learners to work on learning worthwhile words without wasting time. At the same time, learners can probably acquire a fundamental capability for doing extensive reading.

A reasonable approach is to incorporate an extensive reading program into a regular English course as suggested by Waring (2009), together with deliberate learning (Nation, 2015). By doing so, it is possible for the two language sources to compensate for each other's limitations and for learners to take advantage of both.

**Conclusion**

This study analyzed vocabulary taught in one series of MEXT-approved textbooks intended for junior high school students. The two types of lexical coverage analyses suggested that what users of the textbook series encounter is highly likely to be high-frequency words with a greater than 95% coverage of the NGSL, but they cannot be exposed to a large proportion of the list, with a smaller than 37% coverage, which answered the first and second research questions. For the third research question, the distribution analysis indicated that the vocabulary words that learners often meet in the textbooks were those at the first 560-word frequency band; throughout three years of English learning using this textbook series, users can be exposed to over 80% of NGSL words at this band. However, the analysis indicated a shortage of NGSL words through the second to third frequency bands, with a smaller than 50% coverage of each band. Lastly, the importance of learning sources other than textbooks

was discussed as pedagogical implications. Specifically, the combined approach of extensive reading and deliberate learning using word cards can be a suitable option.

This study indicated that MEXT-approved textbooks for junior high school students do not provide a sufficient variety of words needed for comprehension, and this is also the case for senior high school textbooks (Nakayama, 2021), indicating the possibility that Japanese EFL learners have not mastered core words before they graduate from senior high school. Therefore, language teachers at all school levels including the junior high school, senior high school, and college cannot presuppose that their students have already mastered at least core words and disregard teaching them. In fact, they might need to put a primary focus on core words in anticipation of learners' insufficient knowledge of those words, or else, some Japanese EFL learners will never be able to read any authentic text without help from other sources such as a dictionary and teachers.

**References**

Anthony, L. (2021). *AntWordProfiler* (Version 1.5.1) [Computer software]. Waseda University. https://www.laurenceanthony.net/software

Browne, C. (2013). The new general service list: Celebrating 60 years of vocabulary learning. *The Language Teacher*, *7*(34), 13–16.

Browne, C. (2021). The NGSL project: Building wordlists and resources to help EFL learners (and teachers) to succeed. In E. Forsythe (Ed.), *Teaching with Technology 2020 Selected papers from the JALTCALL2020 Conference* (pp. 1-18). The JALT CALL SIG.

Browne, C., & Waring, R. (n.d.) *Word-learning app*. http://www.newgeneralservicelist.org/ngslnawl-iphone-apps

Browne, C., Culligan, B., & Phillips, J. (2013). The new general service list. http://www.newgeneralservicelist.org.

Chujo, K. (2004). Measuring vocabulary levels of English textbooks and tests using a BNC lemmatised high frequency word list. In J. Nakamura, N. Inoue, & T. Tabata (Eds.), *English corpora under Japanese eyes* (pp. 231-249). Amsterdam: Rodopi.

Cobb, T. (2008). Commentary: Response to McQuillan and Krashen. *Language Learning & Technology*, *12*(1), 109–114.

Cobb, T. (2021a). Text Lex Compare v.4.3 [Computer software]. https://www.lextutor.ca/cgi-bin/tl_compare/

Cobb, T. (2021b). Web VocabProfile [Computer software]. https://www.lextutor.ca/vp/comp/

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*(2), 213-238.

Eldridge, J., & Neufeld, S. (2009). The graded reader is dead, long live the electronic reader. *The Reading Matrix*, *9*(2), 224-244.

Hirsh, D., & Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, *8*(2), 689– 696.

Hu, M., & Nation, I. S. P. (2000). Unknown vocabulary and reading comprehension. *Reading in a Foreign Language*, *13*(1), 403-430.

Laufer, B. (1989). What percentage of text lexis is necessary for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316–323). Bristol: 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, N., & Nation, I. S. P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal*, *16*(33), 33-42.

McLean, S. (2017). Evidence for the adoption of flemma as an appropriate word counting unit. *Applied Linguistics*, *39*(6), 823-845.

Ministry of Education, Culture, Sports, Science and Technology. (2017a). Chugakkou gakushu shidou youryou kaisetsu gaikokugo hen [*Expository comments on the course of study guideline for foreign languages in junior high schools*]. Tokyo: Kairyudo.

Ministry of Education, Culture, Sports, Science and Technology. (2017b). Shogakkou gakushu shidou youryou kaisetsu gaikokugo katsudou gaikokugo hen [*Expository comments on the course of study guidelines for foreign languages in elementary schools*]. Tokyo: Kairyudo.

Nakata, T. (2015). Effects of expanding and equal spacing on second language vocabulary learning. *Studies in Second Language Acquisition*, *37*, 677-711.

Nakayama, S. (2021, November 12–15). *A close examination of vocabulary in Japanese EFL textbooks* [Conference presentation]. Paper presented at JALT2021: Reflections and New Perspectives, Online.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review/ La Revue canadienne des langues vivantes*, *63*(1), 59-81.

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge: Cambridge University Press.

Nation, I. S. P. (2015). Principles guiding vocabulary learning through extensive reading. *Reading in a Foreign Language*, *27*(1), 136-145.

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

Ota, H et al. (2021a). *Here we go! English course 1*. Tokyo: Mitsumura Tosho

Ota, H et al. (2021b). *Here we go! English course 2*. Tokyo: Mitsumura Tosho

Ota, H et al. (2021c). *Here we go! English course 3*. Tokyo: Mitsumura Tosho

Scott, M. (2016) *WordSmith tools version 7* [Computer software]. Stroud: Lexical Analysis Software. https://www.lexically.net/wordsmith/

Sobel, H. S., Cepeda, N. J., & Kapler, I. V. (2011). Spacing effects in real-world classroom vocabulary learning. *Applied Cognitive Psychology*, *25*, 763-767.

Stoeckel, T., & Bennett, P. (2015). A test of the new general service list. *Vocabulary Learning and Instruction*, *4*(1), 1-8.

Waring, R. (2009). The inescapable case for extensive reading. In A. Cirocki (Ed.), *Extensive reading in English language teaching* (pp. 93-112). München: Lincom Europa.

Waring, R. (2011). Extensive Reading in English Teaching. In H. Widodo, & A. Cirocki (Eds.), *Innovation and Creativity in ELT methodology* (pp. 69-80). Hauppauge: Nova Publishers.

Waring, R. & Browne, C. (n.d.). Online graded text editor [Computer software]. https://www.er-central.com/ogte/

Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, *15*(2), 130-163.

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, *28*(1), 46-65.

West, M. (1953). *A general service list of English words*. London: Longman, Green & Co.

Wongsarnpigoon, I. (2018). Vocabulary in junior high school textbooks and exams. In P. Clements, A. Krause, & P. Bennett (Eds.), *Language teaching in a global age: Shaping the classroom, shaping the world* (pp. 154-160). Tokyo: JALT.

**Contact email:** shusaku-nakayama@outlook.jp