

琉球大学学術リポジトリ

Leaping a Lexical Gap : Does College Textbook Effectively Help Students Learn Less Frequent Vocabulary?

メタデータ	言語: 出版者: 琉球大学法文学部国際言語文化学科欧米系 公開日: 2017-02-20 キーワード (Ja): キーワード (En): 作成者: Goya, Hideki, 呉屋, 英樹 メールアドレス: 所属:
URL	http://hdl.handle.net/20.500.12000/36285

Leaping a Lexical Gap: Does College Textbook Effectively Help Students Learn Less Frequent Vocabulary?

Hideki Goya

Introduction

Textbooks used in Japanese grade schools (i.e., elementary school, junior high school, and senior high school) are written under the guidance (i.e., Course of Study) of the Ministry of Education, Science, and Culture (MEXT). As widely acknowledged, the current version of the Course of Study issued in 2008 announced that learning English as a foreign language activity was to become a part of the elementary school curriculum (2008). By the year 2020, a new Course of Study is scheduled to be in effect, which expects to emphasize that English learning should be viewed as a continuous act across all grade levels from elementary school to college. One can easily imagine that such a shift in English education, subsequently, will heavily influence what should be taught in class due to a hasty implementation of English learning at very young ages. The present study is a textbook analysis that allows us to evaluate whether a current textbook used in a college adequately provides sufficient amount of learning opportunities, especially word knowledge, under the prospective educational reform.

Literature Review

Word knowledge is indispensable for successful second language (L2) learning (Schmitt, 2008). This is because, as a number of studies agreed, word knowledge assists the L2 learners in reading comprehension (Grabe, 2009). In general, text comprehension is a cognitive product generated from lower processing skills (e.g., letter recognition, word meaning retrieval, morphosyntactic analysis, etc.) and higher counterparts (e.g., background

knowledge, cognitive reading strategies, and metacognitive strategies, etc.), in an interactive processing (Koda, 2005). In short, a deficiency of word knowledge may hinder the cognitive processes necessary for precise reading comprehension. Furthermore, the accumulated word knowledge eventually affects L2 learner' s academic works since some content learning is also done through reading textbooks (Nation, 2001). For the above reasons, providing effective word learning has become a focus of researchers and educators in second language teaching and learning (Nation & Meara, 2002).

One area of the research examined is text coverage or how much vocabulary per page a reader needs to know in order to have an adequate level of comprehension. Text coverage indicates a ratio of known words to the total number of words appearing in a text. A well-accepted consensus is that L2 learners need to have 95% text coverage in order to have sufficient comprehension (Laufer, 1989). What it means is that if learners find an unknown word every 20 word, they are likely to fail in comprehending what the text is about. A more recent study suggests that L2 learners need to know at least 98% of words on a page for written comprehension (Hu & Nation, 2000), meaning that if L2 learners find more than one unknown word every 50 word, they fail to comprehend the text.

With regards to text coverage, the further question to ask is how many words do students need to know in order to achieve this high coverage rate? Adolphs and Schmitt (2003) estimated that if only 90% coverage is required, the amount of words needed is around 800 word families, which is equal to 1,400 individual words. If 95% coverage is required, then the amount needed is over 2,000 word families, which may include many infrequent words. What is more, according to Nation' s (2006) calculation, 95% coverage requires knowledge of about 3,000 word families. Furthermore, the study indicates that 8,000 to 9,000 word families are necessary to read a range of authentic texts with 98% coverage. What this means is that if L2 learners wish to read

a wide variety of texts without unknown words being a problem, they need to know 8,000 to 9,000 word families. As a whole, what the current evidence suggests here is that achieving a sufficient level of reading comprehension requires at least learning the first two thousand most frequent words (i.e., K1 and K2) in English as well as many other infrequent words (i.e., above the K3 level).

Considering the importance of the amount of word knowledge to assure academic success to L2 learners, sources of such knowledge, especially textbooks, are worthwhile because, frankly speaking, L2 learners exclusively rely on them as an input provider in an EFL context. Given that the textbook is essential for EFL learners' input, in what way words are presented deserves pedagogical attention, especially how repeatedly (or frequently) words appear in the textbook. One consensus widely accepted regarding effective vocabulary learning is that the more frequently words appear (i.e., recycle), the more successfully learners will learn them. Early studies found that the repetition of unknown words in reading effectively facilitated word learning (Pitt, White, & Krashen, 1989; Saragi, Nation, & Meister, 1978). For example, Saragi et al. (1978) examined how repeated unknown words were learned by native speakers of English using 90 slang terms of Russian origin in English. Similarly, Pitt et al. (1989) replicated the study with non-native speakers of English. Although the impact of reading on word learning was diverse, significant findings still lent support for the effectiveness of reading for new word learning (Pitt et al., 1989).

A growing number of recent studies furthered this inquiry; namely, a threshold level study of repetition for unknown words to be successfully learned (Horst, Cobb, & Meara, 1998; Hulstijn, Hollander, & Greidanus, 1996) investigated the frequency effect of unfamiliar words and how many encounters were necessary while reading a graded reader. Consistent with previous studies (e.g., Saragi et al., 1978), the studies pointed out that

eight repeated encounters with unfamiliar words in extensive reading is a significant factor for successful vocabulary learning.

Considering the above finding, a good textbook should not only present learners with a variety of words, but it also should recycle them in order to facilitate successful learning subsequently. In light of efficacy of input provided by textbook, many such investigations have been conducted on the relationship between lexical richness and learner's uptake of the presented words. Vassiliu (2001) investigated EFL learners' vocabulary uptake among 47 participants learning English in Greece. He compiled vocabulary data from three different textbooks and divided vocabulary into eight frequency bands. Then the study examined whether 15 words from each band would be successfully remembered. Vassiliu concluded that the participants gained more vocabulary if they were presented with a fewer number of words.

In a Japanese EFL context, however, some studies showed a daunting finding. For instance, Sato (2007) analyzed the textbooks used in Japanese junior high schools and found that the recycling rate of words is about 15% to 25%, which indicates that two words out of 10 in a sentence are newly presented words. Given that words are better learned in recycling, such rate is far from sufficient for successful learning (Sato, 2007).

The recent educational reform in Japan worsened such aspect. Despite continuous decline of recycling rate of vocabulary (Ozasa & Erikawa, 2004), MEXT decided to increase the amount of vocabulary presented in grade schools (junior high and senior high schools) explained in the current version of Course of Study (2008). With regards to such aspect, some studies inquired lexical input provided by the textbooks at grade schools in Japan further. For instance, Nishigaki (2009) figured the number of words EFL learners in Japan should learn according to grade levels from elementary to graduate school after settling on 8,000 words as the targeted number for sufficient reading comprehension based on Nation's (2006) estimation. Nishigaki showed that

in elementary school, students should learn 500 words first. Once students enroll in English classes at junior high school, they should learn an additional 1,000 words; therefore they will have been exposed to 1,500 words total. By graduating senior high school, students have to learn 3,500 words in total in order to achieve an 8,000 word-level as adults. Table 1 illustrates how many words students at each level need to learn.

Table. 1

Targeted number of words Japanese EFL need to learn and its cumulative number

School Levels	Targeted # of words	Cumulative # of words
Elementary	500 words	500 words
Junior high	1,000 words	1,500 words
Senior high	2,000 words	3,500 words
College (General)	1,500 words	5,000 words
College (Specialized)	1,500 words	6,500 words
Graduate/adult	1,500 words	8,000 words

However, according to the guideline presented by Course of Study, students learning English as a foreign language in Japan will only learn 3,000 words by 12th grade, far fewer than the ideal figure presented in Nishigaki (2009). What is worse, the way MEXT counts words is unique, and the actual figure might be fewer. That is, according to the reference list of words included in an old version of the Course of Study, for instance, auxiliaries like “am,” “are,” “is,” “have,” and “has” are listed as individual entries in the list. Generally speaking, most lexicographers usually count them as a single entry (i.e., lemmatize or even word family). Moreover, as others (e.g., Chujo, Nishigaki, & Miyazaki, 2007) pointed out, many words overlap across

the textbooks used at elementary, junior high and senior high schools, which inevitably implies that EFL learners at Japanese grade schools have less opportunity to encounter various words.

Such concern drove researchers to investigate textbooks at each school level accordingly and compared the results. For instance, Hasegawa, Chujo, & Nishigaki (2008) analyzed English textbooks in order to investigate a possibility of missing link of vocabulary learning between junior high school and senior high school in Japan. They compared two sets of textbooks, one used in junior high schools and the other used in senior high schools. They found that the text coverage had declined for practical activities after the reform took place in 2008. They also acknowledged that such limited amount of vocabulary presented might cause Japanese EFL college students to struggle in using English for academic purposes (Hasegawa et al., 2008) in higher education such as in college.

To put it all together, what the above implies is that the amount of words the MEXT presented in the Course of Study may cause a severe shortage of word knowledge necessary for forthcoming higher education in college. In fact, the textbooks written for college English students published before the reformation help us easily see that the current textbook may be a stumbling block for a smooth transition of learners' English learning from high school to college. More specifically, the lexical gap between the two levels of school might eventually leads to poor language usage among EFL learners at the college level.

Given that college students may struggle with this lexical gap continuously widened via English learning at grade schools, an urgent response is needed to prevent many college students from failing in English learning. This issue is what drove the present study to investigate the lexical gap in textbooks. Thus, the purpose of the study was to investigate whether a possible learning gap can be identified between grade school and college

in terms of the lexical input presented in textbooks. As such, the following research questions were addressed:

1. What is the text coverage of frequent words (i.e., K1 and K2) in a textbook used in an EFL course at the college level?
2. How often are less frequent words (i.e., academic words) repeated in a textbook used in an EFL course at the college level?
3. Are there any lexical gap between textbooks used in grade schools and ones used in college?

Method

The purpose of the study was to see if there were any lexical gaps in the textbook used for freshmen in Japanese college English classes. In order to examine such gaps, the study described lexical input appearing in the textbook used for college level EFL English classes at the middle size Japanese university. The study compiled words appearing in the textbook and compared the data with the that from previous studies about the lexical description of grade school textbooks as well as with the current version of the Course of Study.

Materials

The textbook chosen for the analysis is entitled English for Global Citizen (*Ishihara et al., 2012*). This textbook is used in a requisite English class for freshmen at College English at a local university in Japan. Every instructor teaching College English in the institution is required to use this textbook whose goals are to develop not only English proficiency, but also cultivate content knowledge as well as other cognitive skills such as critical thinking and information literacy.

The textbook contains six units consisting of activities aiming to develop four integrated language-skills (listening, reading, writing, and

speaking) with a focus on vocabulary learning. The units are built upon particular topics such as Language Diversity, Ecology and Economy, DNA and Technology, Nutrition Education, Humanity Today, and Global Community. Each topic provides learners with a reading passage in the reading section. Words are presented mostly in context by means of monologues.

Apparatus

In order to compile the data for analyses, the present study employed one of the computer programs widely used for vocabulary description studies, VocabProfile, based on Laufer and Nation's Lexical Frequency Profiler available in Complete Lexical tutor version 8.0 in the Internet (<http://www.lextutor.ca>). VocabProfile generates various lexical descriptions such as word frequency lists, text coverage with the first and the second thousand most frequent words (K1 and K2), type-token ratio, tokens per type, and lexical density as well as the number of tokens and types of words. The present study ran the program on a desktop computer to type in the lexical information extracted from the textbook after converting word documents to text documents.

Procedure

The lexical data was excerpted only from the reading section of the six units and given to the researcher as a data set in word documents. The researcher first checked the data by unit to see if any words included in the raw data were unidentifiable by the system. Once those items were identified, the researcher manually went back to the text and spelled them out so that they could be read by the system adequately. In doing this, the number of "Unknown" words (i.e., unknown to the system) can be reduced.

Data Analysis

Data analyses were done descriptively through the VocabProfile program. This online program estimates frequency-based lexical information along with other information such as tokens, types, and families by entering texts.

According to Nation (2001), token can be calculated by counting every word form in a text, so each appearance of a word is counted. Type is another counting unit and can be calculated by counting any different form of a word occurring in a text. What this means is that words such as *have* and *has* are counted as two different words in token; however, *have* is counted once even if it appears twice in type. Lastly, a word family is another way to count words. It consists of a headword, its inflected forms, and its closely related derived forms; therefore, words *have*, *has*, *had*, and *having* are categorized as the same word and counted only once in word-family. As reviewed earlier, the way MEXT counts words seemingly fall into the type; however, the present study adopted word family as the counting unit when necessary for later comparisons.

Other indices of lexical richness were also employed in the analyses; they were Tokens per Type ratio (hereafter, TpT) and Lexical Density (hereafter, LD). The TpT score was calculated by dividing the number of tokens by the number of types, and the LD was calculated by dividing the number of content word' s tokens by the number of tokens. The TpT score indicates specificity of the text per se because certain words are repeated frequently. On the other hand, the LD score indicates how dense the text is in terms of the ratio of content words to the total number of tokens. That is, a higher LD is considered dense since content words primarily convey information (Laufer & Nation, 1995). Note that some studies claimed that TpT was invalid as a reliable measurement. The present study acknowledges this fact and understands that TpT may not be able to adequately indicate lexical

richness (Laufer & Nation, 1995; Vermeer, 2004). This is true, especially when comparing different sets of text or learner-produced writings. Since the present study subjected only a single set of data to subsequent analyses, expected interpretations do not suffer from such invalidity.

According to the coding description available on the VocabProfile site, it says the program classifies texts into four categories by frequency; that is, the first and second thousand most frequent words, academic words, and offlist words (e.g., remainder words). The first list contains the first most frequent thousand words (i.e., K1) and the second list contains the second most frequent thousand words of English (i.e., K2). The third list provides the academic words of English (hereafter, AWL). The AWL word list contains 550 words that frequently appear in academic contexts such as college textbooks of any subjects (Coxhead, 2000). The last list contains the remaining words that are not found on the other lists, so it is called the offlist words. Words in the offlist, in other words, are less frequent words that are not found in the first and second most frequent words as well as in the AWL list.

Result

After running the VocabProfile program, several lists of lexical descriptions were obtained, such as the overall lexical description (i.e., tokens, types, TpT, and LD), text coverage of three different categories (i.e., K1, K2, AWL, and Offlist words), and a detailed description of AWL and Offlist words' TpT scores.

Overall Description of Lexical Input of EfGC

First of all, the total token, the number of words used for the six passages, was 4,638. For more detailed information, the number of tokens in each unit was 774, 778, 756, 755, 842, and 733 for Unit 1 to Unit 6, respectively. As for the total number of type, it was 1,387 words, and the

number of type in each unit was 316, 268, 343, 368, 358, and 339 for Unit 1 to Unit 6, respectively. A total tokens per type (TpT) score was 3.34, and each unit showed 2.45, 2.90, 2.20, 2.05, 2.35, and 2.16 accordingly. The lexical density (LD) for the combined passages was .55, and each passage' s LD was .57, .54, .53, .53, .58, and .56 for Unit 1 to Unit 6, respectively. Table 2 illustrates the above numbers.

Table. 2

Overall description of lexical input of EfgC

	Tokens	Types	TpT	LD
Unit 1	774	316	2.45	.57
Unit 2	778	268	2.90	.54
Unit 3	756	343	2.20	.53
Unit 4	755	368	2.05	.53
Unit 5	842	358	2.35	.58
Unit 6	733	339	2.16	.56
Overall	4638	1387	3.34	.55

Text Coverage

The program also reported the text coverage. As for the overall text coverage, it consisted of the K1 and K2 words (i.e., first two thousand most frequent words) for 80.60%, AWL (i.e., academic) words for 8.90%, and the Offlist (i.e., less frequent words without AWL words) words for 10.50%. More specific description of each unit was also reported. That is, Unit 1 consisted of the K1 and K2 words for 83.47%, the AWL words for 9.95%, and the Offlist words for 6.59%, and Unit 2 contained the K1 and K2 words for 79.18%, the AWL words for 6.17%, and the Offlist words for 14.65%. Unit 3 contained the K1 and K2 words for 83.60%, the AWL words for 3.70%, and the Offlist

words for 12.70%, and Unit 4 contained the K1 and K2 words for 85.16%, the AWL words for 5.83%, and the Offlist words for 9.01%. Unit 5 had the K1 and K2 words for 73.87%, the AWL words for 12.83%, and the Offlist words for 13.30%, and lastly, Unit 6 was covered with the K1 and K2 words for 79.18%, the AWL words for 6.17%, and the Offlist words for 14.65%. Figure 1 illustrates the above numbers.

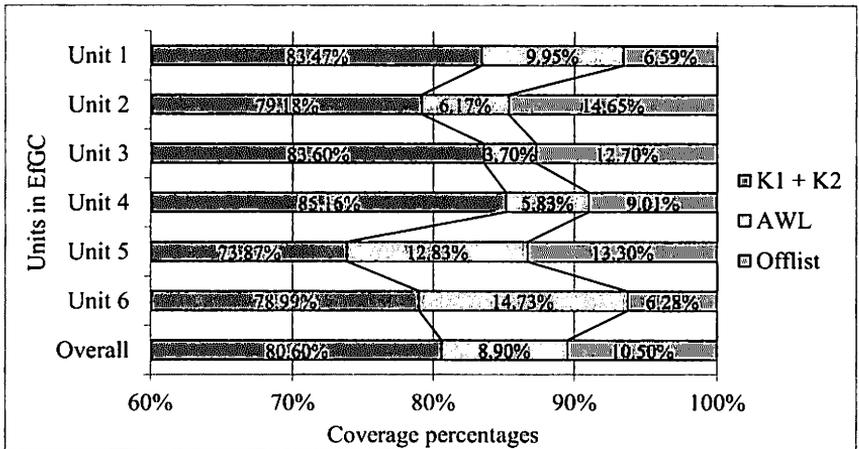


Figure 1. Text coverage of words in each unit and combined units.

Detailed Description of Academic Words (AWL) and Offlist Words

In order to compare the AWL words and the Offlist words, the tokens and types of the two categories were reported. First, there were 77 tokens and 53 types of AWL words in Unit 1, 48 tokens and 29 types in Unit 2, 28 tokens and 25 types in Unit 3, 44 tokens and 36 types in Unit 4, 108 tokens and 76 types in Unit 5, and lastly 108 tokens and 71 types in Unit 6. Figure 2 shows the above numbers.

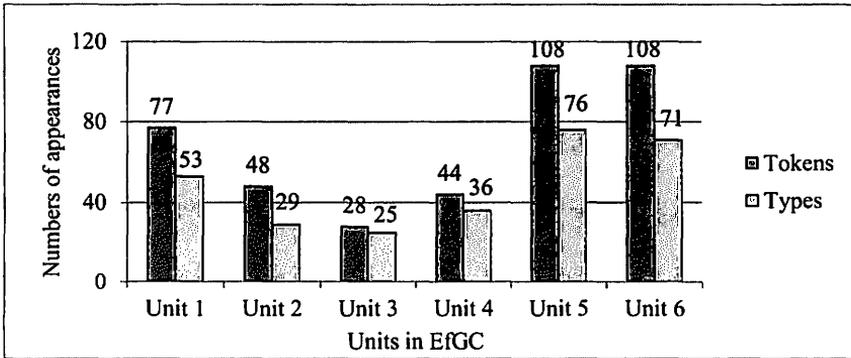


Figure 2. Tokens and types of AWL words in each unit.

As for the Offlist words description, there were 51 tokens and 35 types of the Offlist words in Unit 1, 114 tokens and 31 types in Unit 2, 96 tokens and 65 types in Unit 3, 68 tokens and 57 types in Unit 4, 112 tokens and 61 types in Unit 5, and lastly 46 tokens and 32 types in Unit 6. Figure 3 shows the above numbers.

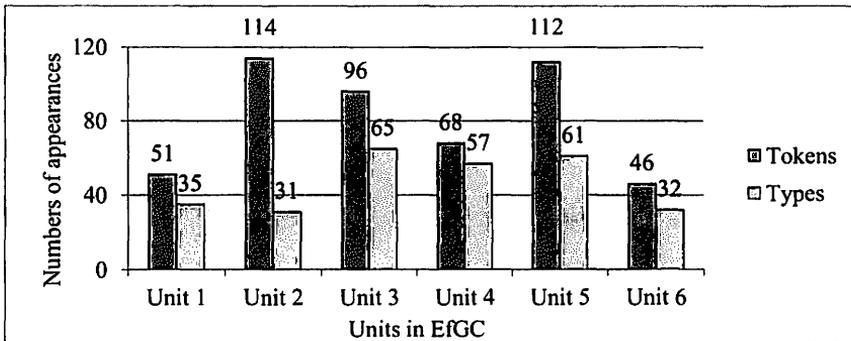


Figure 3. Tokens and types of Offlist words in each unit.

For a more specific comparison, the present study looked into a ratio

of AWL and Offlist words, namely, the tokens per type ratio (TpT). The program reported the TpTs as 1.46, 3.68, 1.48, 1.19, 1.84, and 1.44 for AWL words in each unit respectively, and 1.45, 1.66, 1.12, 1.22, 1.42, and 1.52 for Offlist words accordingly. Its overall TpT for all texts combined were 1.69 for the AWL words and 1.84 for the Offlist words. Figure 4 shows the above numbers.

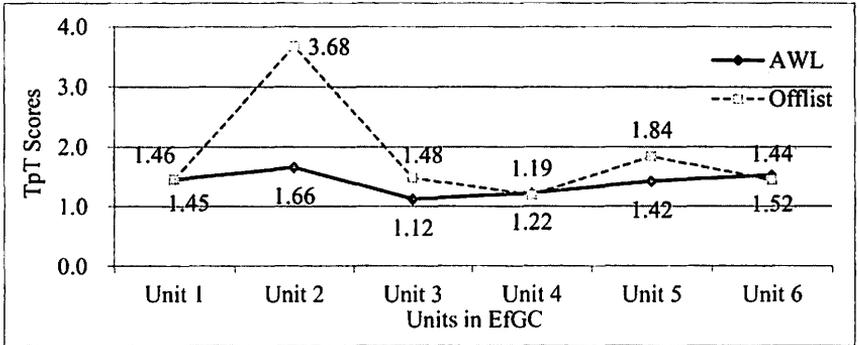


Figure 4. Tokens per type (TpT) ratios for both AWL and Offlist words.

Discussion

The main purpose of the present study was to examine whether any gap would exist between grade schools and college in terms of lexical input quantitatively in a textbook used in a university level EFL class. More specifically, the study ran several corpus analyses to seek clues to shed more light on the following aspects: (1) the text coverage of frequent words (i.e., K1 and K2) in the textbook; (2) the repetition rate of less frequent words (i.e., academic words) in the textbook; and (3) the probable lexical gap between a vocabulary size presented in grade schools and the lexical input presented in the college textbook. Let's look at the points by referring to our data.

As for the text coverage of various lexical types, its overall rate was

definitely worthwhile for consideration. Figure 1 in the earlier section showed that words in “English for Global Citizen” (EfGC) consisted of 80.60% of K1 and K2 words, 8.90% academic words from AWL, and 10.59% of words from the Offlist. These numbers clearly suggest that EFL learners with EfGC under the current investigation were apt to be exposed to many less frequent words (i.e., approximately 20% of all words).

With regards to reading comprehension, however, a question naturally arose from the above description: how much vocabulary did college students need to know in order for sufficient comprehension to be achieved when reading EfGC? The study further looked into the inquiry by running the subsequent analyses with a function available in the VocabProfile program: lexical distribution by frequency bands. This frequency-based analysis in accordance with the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA).

The program counted 3,245 tokens as the number of K1 level words. Subsequently, there were 606 tokens, 358 tokens, 96 tokens, 38 tokens, 40 tokens, 14 tokens, 13 tokens, 1 token, and 1 token according to each frequency band from K2 level to K10 level. The accumulated number of tokens from K11 to K25 was 17, and 115 tokens of the Offlist words were automatically

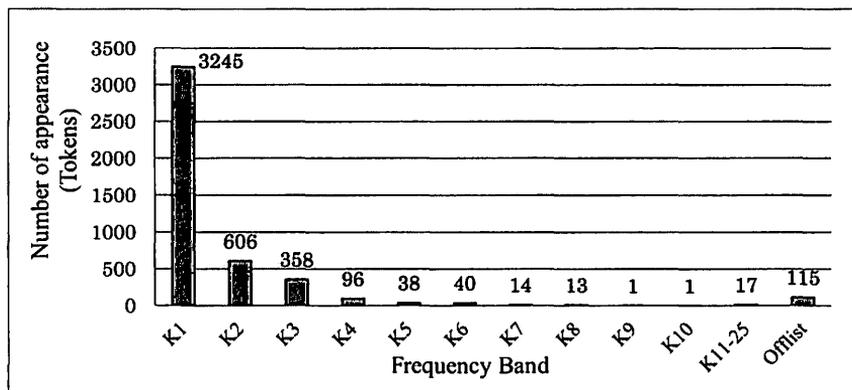


Figure 5. The number of tokens according to frequency bands

excluded from the analysis as a default function. Figure 5 illustrates the above number.

For a more detailed picture of text coverage from the K1 to K25 frequency level of words, the program reported accumulated percentages as the frequency level increased. As partially reported in the previous section, the text coverage of K1 words was 71.41% and that of K2 was 13.34%. The cumulative coverage of K1 and K2 was 84.75%. What should be noted here was that the program reported lexical distributions of overall tokens without the Offlist words according to the frequency band from the K1 level to the K25 level; therefore, numbers in the generated results appeared slightly different from ones shown in earlier tables and figures. Such exclusion of the Offlist words made the coverage appear seemingly higher. In fact, the exclusion increased the coverage up to 93.63% with K3 words added, up to 94.74% with K4 words added, and up to 95.58% with K5 words added. The coverage end with 97.10% with K10 words added, and the final coverage resulted in 97.46% when all words from the K1 to K25 levels were included. Figure 6 shows the cumulative text coverage according to the frequency bands below.

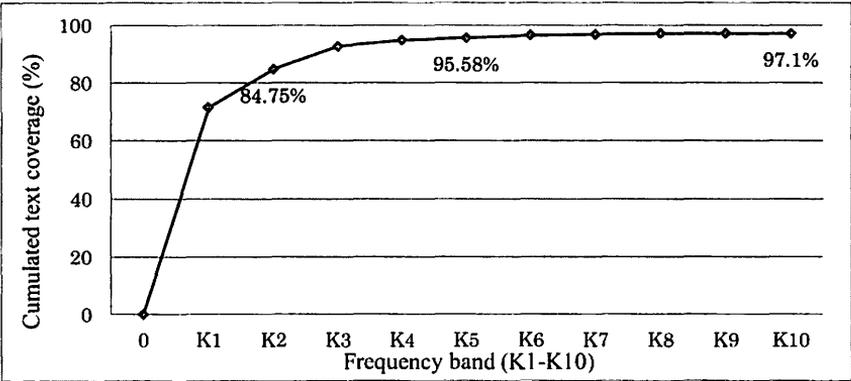


Figure 6. Cumulative text coverage according to frequency band

As seen in the cumulative text coverage presented in the above figure, the text coverage of K1 and K2 words in our data covered only 84.75%, and even if all the words were added up, its overall coverage only resulted in 97.1%. In other words, without the Offlist words, the coverage never reached beyond the 98% level. The present finding adequately showed that even if learners know 3,000 of the most frequent words (K1 to K3 words), they would most likely fail in reading comprehension as the previous studies reviewed in the literature suggested that learners would have to know at least 90% of the words on a page in a lenient estimation (Adolphs & Schmitt, 2003). For a more strict estimation, 95% or even higher is required for sufficient understanding of the text (Nation, 2006). Apparently, none of the above ideal coverage suggested in the literatures was achieved in our data. Thus, learners with EfGC may struggle with reading comprehension of the passages given in each unit.

The above is of importance as far as incidental word learning benefits is concerned: the research showed the pedagogical value of repeated encounters to learn less frequent words. That is, in order for successful incidental word learning to occur as a byproduct of reading comprehension, providing an input rich environment is essential. In the research paradigm, many researchers claimed that the threshold level of known word coverage per page (or text coverage) for adequate comprehension must be 95% or above in order to utilize the contextual information to guess any unknown words (e.g., Milton, 2009). Collectively, given that EfGC contains only 84.75% basic vocabulary (i.e., K1 and K2 words), adequate reading comprehension may be difficult to be achieved. The learners' limited reading comprehension in turn interferes with successful word guessing and, as a consequence, prevents learners from learning unknown words incidentally.

Thus far, our data corroborates that EfGC may fail to provide lexical learning opportunity by itself due to an excessive number of less frequent

vocabulary. Contrary to such view, however, the textbook can be also recognized as a good source of vocabulary learning because of such diverse lexical variation at the same time. Considering this paradox of lexical richness in the textbook, another worthy discussion deserves careful attention. That is, just because the textbook presented a lot of less frequent words, it does not necessarily mean that learners can learn the words successfully. Despite its inconsistency reviewed in the previous studies of how many times words need to be repeated for assuring successful word learning, the consensus reported was that the repetition of unknown words in reading effectively facilitates word learning (e.g., Horst, Cobb, & Meara, 1998). In this light, in order to examine whether EfGC is a good source of words learning or not, it is necessary to determine how words were repeated.

Unfortunately, our data did not directly report the repetition rate of specific items. Instead, its general rate of repetition for the less frequent words (i.e., AWL and Offlist words) can be interpreted through the token per types ratio (TpT) calculation in Figure 4 provided in the earlier section. Simply put, the TpTs indicate lexical variation by dividing the total number of tokens by the total number of types. This formula partially provides the repetition rate of overall vocabulary. If the obtained TpT is 1.0, none of the words has been repeated twice. If the TpT is larger than 1.0, this indicates that some of the words have repeated twice or more.

Our finding turned out to be definitely withering from the word learning perspective. Our data showed that although each of the TpTs of the less frequent words (e.g., AWL words and Offlist words) varied according to the units, the TpT of all texts combined was 1.69, ranging from once to 19 times (e.g., *global*) for the AWL words and 1.84 ranging from once to 14 times (e.g., *carbon*) for the Offlist words. In short, although a few of words were repeated sufficiently, almost all of the less frequent words were repeated less than twice. Considering the presented consensus as well as the reported

numbers here, less than twice is insufficient for successful word learning. That is, the repetition rates of AWL words and that of the Offlist words in EfGC do not assure successful learning of new words. As a whole, it may be safe to conclude that the examined textbook is not helpful for learners to acquire new words incidentally due to the poor lexical environment in terms of text coverage of known words as well as repetition rate of less frequent words.

The final question was about a possible lexical gap between vocabulary learned in grade schools and that presented in the college textbooks. Taking the earlier findings into consideration, a drawn conclusion is that the textbook of EfGC may provide a rich variety of lexical input ranging from the K1 level to the K25 level. However, the textbooks seemingly fail to provide sufficient learning opportunities with regards to the repetition rate of those infrequent words in context. Another key fact that helps to deepen the discussion is to ask about the expected amount of vocabulary learned while in the grade schools. As illustrated in the literature review section, an estimation of vocabulary size expected to learn at each grade school in Japan (Nishigaki, 2009) suggests that the students are supposed to have learned 3,500 words before college enrollment. As for college students, they would need to know cumulatively 5,000 words during general education classes and 6,500 words in learning their majors.

According to the guideline provided by the MEXT, the Course of Study suggests EFL students at Japanese grade schools are guided to learn 3,000 words by 12th grade (2008). Given that every English class is supposed to be structured based on the strict guideline, the amount of words that learners have learned by the time they enroll in a college English course under the current investigation is likely to leave them with a severe vocabulary shortage. More specifically, as indicated in our data, college students in the first semester will be exposed to the five thousand most

frequent words (i.e., K5 level of words) or above in English class with EfGC. What is more, those less frequent words are only presented in written texts less than twice, which is much lower than the threshold level for successful word learning as discussed in the previous studies. Such lexical dispersion may not help college freshmen in neither reading comprehension nor word learning because they are in a interdependent relationship (e.g., Koda, 2005); rich reading comprehension assists learners to guess any unknown words based on available clues gained, while too many unknown words in reading prevent learners from successful comprehension. Taken together, the present study can safely draw a conclusion that there exists a lexical gap in transition from grade school to college in terms of lexical input presented in the textbook.

Although identifying a possible lexical gap is definitely essential to practice textbook-based pedagogy more effectively, we should bear several caveats in mind as limitations of the study. One limitation of the study was its reliance on written input for the investigation. That is, the present study performed lexical input analyses exclusively on written texts; therefore, a future study needs to consider oral input given by recorded audio and instructors.

How to count words also plays a crucial role in the direct comparison between the lexical descriptions that is another limitation of the study. The present study exclusively analyzed the data descriptively. Thus, in order to generalize our finding, a future study needs to employ a various textbooks used in college and analyze them with inferential statistics. As relevance to generalization of our finding, a future study is necessary to utilize lexical input gained from grade school textbooks so that a direct comparison can be available. Furthermore, it is necessary to investigate what specific words are repeated more and less in relation to its uptake among learners so that a probable pedagogical influence can be discussed adequately. The probable

methodological flaw in the study could be a reliance on TpT as the unit to interpret the repetition rate, although it only indicates the averaged repetition rate of all words combined.

Nevertheless, the present findings can draw some pedagogical implications. One bright side to the study was that we saw that EfGC provides rich lexical input. To maximize the efficacy of EfGC for the purpose of word learning, we need to recycle infrequent words more often for effective word learning. First, words presented in EfGC should be additionally repeated as part of the oral input provided by an instructor. The study found that the textbook provides a lot of infrequent words which can be viewed as a rich environment for word learning. However, what is lacking in the textbook is a sufficient amount of exposure to the words or recycling of the words. Considering the fact that a textbook only provides written input, an instructor should make use of different input modalities. Thus, it is necessary to recycle the words orally such as through the instructor's output, teaching devices, and learners' interactions. Another implication similar to the first point is to utilize written worksheets. That is, instructors can provide additional learning opportunities by recycling the infrequent words in additional worksheets so that learners can tackle these words frequently, both receptively and productively. Lastly, instructors may need to encourage learners to create an individual word list in which learners keep unfamiliar words encountered during their academic work so that they can conduct mnemonic learning by themselves. In following the instructions, learners will be exposed to the lexically rich environment repeatedly, which in turn allows successful word learning take place in and out of class.

Conclusion

The present investigation adequately allowed us to evaluate whether a current textbook used in college adequately provides a sufficient amount

of learning opportunities, especially word knowledge under the current educational reform. The main findings of the study showed that EfGC substantially supplied the number of vocabulary as lexical input while its repetition of infrequent words was not at an adequate level. At the same time, the reported estimation of vocabulary size according to grade levels seemed to have deviated from the actual lexical input suggested by the Course of Study. Undoubtedly, textbooks are essential materials in EFL from the content learning as well as foreign language learning perspective. Given that vocabulary is important as a basic component of language, we as instructors should not take the lexical input presented in the textbook for granted. We must pay closer attention to how well learners are exposed to what words and how frequently that exposure occurs. Such thoughtfully planned lexical recycling in a class will fill in any possible lexical gap existing in a transition from English learning at grade schools to one at college.

References

- Adolphs, S. & Schmitt, N. (2003). Lexical coverage of spoken discourse. *Applied Linguistics*, 24, 425-438.
- Cobb, T. *Range for texts v.3* [computer program]. Accessed 15 March 2016 at <http://www.lex tutor.ca/cgi-bin/range/texts/index.pl>.
- Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, 34, 213-238.
- Fewell, N., Ishihara, M., Ishikawa, R., Kanemoto, M., Kelly, T., O' Callaghan, K., Shibata, M., Toya, M., & Yoshimoto, Y. (2013). *English for Global Citizens*. M. Ishihara (Ed.). Tokyo, Japan: Eihosha.
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge University Press.
- Horst, M., Cobb, T., and Meara, P. (1998). Beyond A Clockwork Orange: Acquiring second language vocabulary through reading. *Reading in a*

Foreign Language, 11, 207-223.

- Hu, M., & Nation, I.S.P. (2000). Vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403-430.
- Hulstijn, J., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal*, 80, 327-339.
- Koda, K. (2005). *Insights into second language reading*. Cambridge University Press.
- Laufer, B. (1989). What percentage of text lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316-323). Clevedon: Multilingual Matters.
- Laufer, B., & Nation, I. S. P. (1995). Lexical richness in L2 written production: Can it be measured? *Applied Linguistics*, 16, 307-322.
- Milton, J (2009). *Measuring Second Language Vocabulary Acquisition*. (2009). New York: Multilingual Matters.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, 63, 59-82.
- Nation, I.S.P. & Meara, P. (2002). Vocabulary, An introduction to applied linguistics, 35-54.
- Pitts, M., White, H., & Krashen, S. D. (1989). Acquiring second language vocabulary through reading: A replication of the *Clockwork Orange* study using second language acquirers. *Reading in a Foreign Language*, 5, 271-276.
- Saragi, T., Nation, I. S. P., & Meister, G. F. (1978). Vocabulary learning and reading. *System*, 6, 72-78.

- Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, 12, 329-363.
- Vassiliu, P. (2001). Lexical input and uptake in the Greek low level EFL classroom. PhD. University of Wales, Swansea
- Vermeer, A. (2004). The relation between lexical richness and vocabulary size in Dutch L1 and L2 children. In P. Bogaards, & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition and testing* (pp. 173-189). Amsterdam and Philadelphia, PA: John Benjamins.
- 小篠敏明, 江利川春雄 (2004) 「英語教科書の歴史的研究」東京: 辞游社
- 佐藤剛 (2007) 「中学生の語彙指導に関する実証的研究: コーパスを利用して」『教科研究 TOTAL English』, 12, pp. 16-17.
- 西垣知佳子 (2009) 「小・中・高・大の語彙指導の連携を考える」『GOD 英語通信』, 46, pp.14-17.
- 中條清美, 西垣知佳子, 宮崎海理 (2009) 「小学校 5・6 年生「英語ノート」の語彙一覧」『日本大学生産工学部研究報告 B』, 42, pp. 99-115.
- 長谷川修治, 中條清美, 西垣知佳子 (2008) 「中・高英語検定教科書語彙の実用性の検証」『日本大学生産工学部研究報告 B』, 41, pp. 49-56.
- 文部科学省 (2008). 「小学校学習指導要領」 Accessed 15 March 2016 at http://www.mext.go.jp/a_menu/shotou/new-cs/youryou/syo/gai.htm.
- 文部科学省 (2008). 「中学校学習指導要領」 Accessed 15 March 2016 at http://www.mext.go.jp/a_menu/shotou/new-cs/youryou/chu/gai.htm.

Leaping a Lexical Gap: Does College Textbook Effectively Help Students Learn Less Vocabulary?

Hideki Goya

論文要旨

2020年までに完全実施が予定されている新学習指導要領では、初等教育から高等教育までの英語教育により一貫性が求められるであろう。外国語としての位置付けにある日本の英語教育において、教科書は言語材料の主なインプット源であり、早期の英語教育導入によって教科書の質的な在り方が問われであろう。本研究では、大学の英語教育で使用される教科書の効果的な語彙学習得の機会について調べた。分析の結果、使用されている教科書は、多くの低頻出語を提示しているが、その繰り返し頻度は低く、付随的な語彙学習得には十分でない事が分かった。同時に、大学入学までの語彙学習得目標と、大学で使用されている教科書の語彙レベルに差があることが示された。以上の結果から、本研究は学習者の大学英語教育への順応が困難となる可能性を指摘した上で、授業改善の工夫として、様々なタスクや口頭での精微化を通じ、低頻出語の反復使用（リサイクル）を心掛ける教育的工夫の必要性を示唆した。