Longitudinal Measurement of the Vocabulary Size of Japanese Junior High School Students: Developing a Vocabulary Size Test for Beginner Learners

Tsuyoshi Sato
*Hirosaki University Faculty of Education*

**Abstract**

As vocabulary plays an important part in L2 learning at the beginner level, investigation into vocabulary learning constitutes an essential part of second language acquisition research (Koda, 1989; Laufer, 1992; Meara, 1992). Assuming the case of Japanese, English education can provide a meaningful example. In addition, noting that precise measurements of the English vocabulary size of Japanese junior high school students have not been attained (Katagiri, 2000), the present study examines the growth of vocabulary size of Japanese junior high school students across their 3 years of English learning. For this purpose, a vocabulary size test (VST) was specially developed with a word list compiled from Japanese junior high school English textbooks based on the frequency and range of their vocabulary items (Sato, 2016). A total of 159 junior high school students participated in the research. The vocabulary size of the participants was measured five times over the period of 3 academic years at intervals of 6 months. Analysis of their vocabulary growth by frequency revealed a distinctive pattern for each level, yielding implications for more effective vocabulary teaching for learners at the beginner level in light of the characteristics of their vocabulary growth.

**Keywords:** vocabulary size, test development, Japanese junior high school students

1 **Introduction**

Extensive evidence suggests that vocabulary plays a vital role in L2 learning, showing a significant correlation with learner proficiency (Arnaud, 1992; Koda, 1989; Laufer, 1992; Meara, 1992). However ample research indicates that L2 learner vocabulary size is much smaller than those expected in the previous studies (Olmos, 2009; Pathan et al., 2019; Tang, 2007). Furthermore, most of these studies have focused on a “one-shot” measurement of vocabulary size at a given specific time, assessing participants’ attainment of certain vocabulary levels at the end of a course or academic year. In other words, there is still considerable ambiguity with regard to the pattern of vocabulary growth longitudinally from the beginning stage. Specifically, little research has evaluated how many English words learners know at the beginning of a certain course and how their vocabulary...
size develops throughout their academic career. Research on the process of learner vocabulary growth at the beginning stage likely holds significant implications for more effective vocabulary teaching at this important stage (Nation, 2006; van Zeeland & Schmitt, 2012).

In this regard, English education in Japan may become a suitable context which will lead to a test case in investigating beginner learners’ vocabulary knowledge. Japan is now undergoing a dramatic shift in English education from teaching the language mainly through grammar-translation methods to those intended to foster communicative competence. Various changes have been made, such as introducing English as a required subject into the elementary school curriculums since 2020, expecting teachers to teach English mainly using English. Among the changes, an increase in the amount of vocabulary to be taught is notable. According to the new Course of Study by the Ministry of Education, Culture, Sports, Science and Technology (MEXT, 2018), a total of 4,000 to 5,000 words should be taught over 10 years of English education from elementary school to senior high school. This denotes that 600 to 700 words must be introduced in elementary school, with an additional 1,600 to 1,800 words in junior high school, and with an additional 1,800 to 2,500 in high school. It is crucial to periodically assess students’ achievements in meeting these standards of vocabulary size in order to examine the validity of these standards and determine the necessary changes. However, little research has been conducted with Japanese junior high school students as beginner level English learners. Thus, assessment of the degree to which students have met the mandated standards in terms of vocabulary is mostly based on the experience, even intuition, of the teachers in charge of the lessons. Consequently, vocabulary size and growth of Japanese junior high school students have been overlooked and it remains a question to resolve.

2 Literature Review

It is generally accepted that vocabulary knowledge plays a crucial role in L2 learning. It is assumed that the more vocabulary a learner knows, the better his or her understanding and ease of communication becomes in the target language. There is also considerable evidence that the learner vocabulary size is strongly related to his or her proficiency in the language. Indeed, copious research has been conducted to measure vocabulary size, and has shown that it strongly correlates with several aspects of language proficiency, both in L1 and in L2 (Arnaud, 1992; Koda, 1989; Laufer, 1992; Meara, 1992; Pathan et al., 2019; Tang, 2007). Consequently, a variety of vocabulary size tests (VSTs) have been developed to measure learners’ vocabulary proficiency (Aizawa, 1998; Koizumi, 2003; Meara, 1992; Meara & Buxton, 1987; Mochizuki, 1998; Nation, 1990, 2001; Nation & Beglar, 2007; Sato, 2003; Schmitt et al., 2001). Using these tests, plentiful research on vocabulary size has been conducted.

As Nation (2001) suggested, the number of words needed to use the language is utilitarian information to help decide how much vocabulary needs to be learned. In the case of second languages, it is generally accepted that 3,000
Sato: Developing a Vocabulary Size Test for Beginner Learners

Vocabulary Learning and Instruction, 10(1), 30–48.

to 4,000 word families are needed for intermediate learners (Nation, 2006; van Zeeland & Schmitt, 2012). This is based on the findings identifying this as the vocabulary size necessary for 95% coverage of various written texts such as novels, newspapers, and movies (Nation, 2006). Moreover, Nation suggested this level as the primary standard for the vocabulary size that English learners need in order to use English independently. van Zeeland and Schmitt (2012) investigated the relationship between lexical coverage and comprehension of passages in listening tasks for 36 native and 40 non-native speakers. Based on their findings, they concluded that 98% is a good coverage target. However, they suggested that for less stringent comprehension rates, 90 to 95% should be acceptable with spoken passages. Based on this, van Zeeland and Schmitt (2012) concluded that language users need to know from a low of 2,000 to 3,000 word families (for 90 and 95% coverage) to a high of 6,000 to 7,000 families for adequate to ideal listening comprehension (for 98% coverage). In conclusion, various research has suggested that for L2 learners, it is appropriate to set the vocabulary learning goal for beginner learners to increase the vocabulary size up to 2,000 to 3,000 word families.

However, an increasing number of studies which have measured L2 learners’ vocabulary size (Alonso, 2013; Gallego & Agustin Llach, 2009; Laufer, 1998; Qian, 2002), have concluded that the actual measure of their vocabulary size is generally much smaller than the requirements expected or proposed in those studies which induced the estimated vocabulary size based on the comprehension of the passages with different coverage rates. Pathan et al. (2019) investigated the receptive vocabulary growth of 123 first-year post-graduate university students in Pakistan. The vocabulary size of the students was measured at two different points and was estimated to be 3,500 word families and 4,090 word families at the beginning and end of the course, respectively. The VST developed by Nation and Beglar (2007) was adopted in this study. Pathan et al. concluded by expressing concern that the overall vocabulary size of the participants was much smaller than the proposed requirement of Nation’s (2006) 95% coverage. The research showed that even in Pakistan, where English is an official language and people are generally considered to have high English proficiency, learner vocabulary size is not satisfactory. Olmos (2009) measured the vocabulary size of final-year students in a high school in Murcia, Spain, using Nation’s (1990) Vocabulary Levels Test (VLT). A total of 49 students participated in Survey 1, measuring 1,000- and 2,000-word levels, and 38 students in Survey 2, measuring 3,000- and 5,000-word levels. The participants were considered to possess the vocabulary knowledge of a particular level if they scored above 88% of the established standard. The results showed that very few of the participants passed each level of the test: only three passed the 1,000-word level test, only one the 2,000-word level test, and not even a single participant passed either the 3,000- or 5,000-word level test. This suggests that the Spanish students, who spent 8–10 years studying English as a foreign language (EFL), were not able to reach the level necessary to be an efficient speaker of English. Tang (2007) investigated the vocabulary size of a total of 499 primary schools and secondary school EFL students in Hong Kong, also using VST, finding that the participants had a small vocabulary size and impoverished vocabulary knowledge. The primary school and secondary school students possessed a vocabulary size of about 925 and 2,891 word families, respectively. This suggests that Hong Kong
students still need help with basic English words even after learning English as a compulsory subject for 12,500 hours over 9 years.

Turning to Japanese learners, Kosuge (2003) measured the vocabulary size of Japanese junior high school students using a lemma-based VST for Japanese Learners of English (Mochizuki, 1998). The research concluded that at the end of their school year, the average vocabulary size of seventh-grade students was 1,185 lemmas and that of eighth-grade students was 1,859 lemmas. Katagiri (2000) researched the vocabulary size of 401 senior high school students using both written and listening versions of the same test. The results revealed that the vocabulary size of the participants ranged from 2,719 to 4,371 lemmas for the written version, and from 2,343 to 4,113 lemmas for the listening version. However, these findings need to be interpreted with caution because the VSTs adopted in these studies were based on lemma-based vocabulary lists, which are known to yield larger estimations than those based on word-family-based lists. According to Nation (2016), the approximate ratios for the 3,000 highest-frequency word families are 1:3:6 for word families, lemmas, and word types, respectively. Therefore, it is assumed that the estimated vocabulary size obtained from the studies above is around three times larger than would have been obtained by counting word families.

A serious limitation however with the above studies is that most of them adopt the multiple-choice format vocabulary size or levels tests such as VLT (Nation, 1990). As substantial research criticises, one of the major shortcomings of multiple-choice tests (MCT) is that they overestimate the learner vocabulary size because of the guessing effects (Gyllstad et al., 2015; Stewart, 2014; Stoeckel et al., 2020). Stoeckel et al. (2020) examined three studies which compared the score difference between meaning-recognition and meaning-recall tests, and found that the examinees answered correctly 13.9–62.9% of the meaning-recognition items that tested words unknown at the meaning-recall level of knowledge. Stewart (2014) revealed that the score overestimation for a four-choice vocabulary test could be as high as 25% for most ability levels of learners. Therefore, for precise measurement of the learner vocabulary size, much research suggests that it is preferable to avoid multiple-choice meaning-recognition tests and to adopt meaning-recall tests instead.

As discussed above, most studies of learner vocabulary size have generally focused on setting standards mainly for learners with a certain amount of learning experience using a cross-sectional study framework. This approach offers the possibility of substantial and valuable data. However, little research has sought to determine learner vocabulary size at the point when learners begin studying English and clarify the patterns of growth in their vocabulary size. Enhancing our understanding of how the vocabulary size of L2 learners at the beginning level develops to achieve the word level of 2,000 to 3,000 is significant as an initial step not only to improve effective curriculum development, but also to develop systematic vocabulary instructions for English learners.

2.1 Purposes of the Present Study

As stated above, the vocabulary size necessary for L2 learners to get 95% coverage of various written or spoken texts has been identified as 2,000 to 3,000
word families, constituting an appropriate first goal of vocabulary learning (Nation, 2001, 2013; Nation & Webb, 2011). However, neither the vocabulary levels of Japanese junior high school students at the beginning of each school year nor the pattern of development—how their vocabulary size increases—have been identified. In order to describe this longitudinal development, the present study set the following two objectives:

1. To determine the vocabulary size of Japanese junior high school students at the beginning and end of each school year.
2. To determine the longitudinal pattern of vocabulary size growth for Japanese junior high school students across 3 academic years.

3 Method

3.1 Participants

A total of 159 junior high school students in Japan participated in this study. The data from 10 participants who did not take part in one or more sessions were excluded from the analysis. The following analyses therefore were based on the remaining 149 participants. All participants are from a public junior high school attached to a national university in a rural area. It has five classes for each grade, thus it is regarded as a middle-size school. They had studied English from the first year of junior high school using MEXT authorised junior high school textbooks; at the time of the final measurement, they had studied English in Japan for 3 years as a required subject.

3.2 Materials

The experiment used the four-option multiple-choice VST, which was developed from the lemma-based vocabulary list compiled from all authorised junior high school English textbooks (seven publishers/textbooks for each grade) used in Japan (Sato, 2016). The vocabulary items are listed and ordered according to the indices attained from the following formula:

\[
\text{Ranking criterion of the list} = \frac{\text{range} \times \text{frequency}}{\text{grade level at which a word first appears in textbook}}
\]

The underlying concept of this formula is that in addition to frequency and range, learning period for any specific vocabulary item also has a strong effect on recognition of the word for Japanese learners. In other words, higher frequency and wider range words, those which students encounter earlier in their school years and have a longer period of exposure, are assumed to be learned better. Katagiri’s (2000) data suggested that less frequent words seem, not to be necessarily, more likely to impair comprehension for Japanese EFL learners. The results of previous research (Sato, 2016) revealed that the frequency or range alone does not correspond to Japanese learners’ vocabulary size. The
results demonstrate that some high-frequency words with wide ranges that are considered easier for English learners in general are in fact not so for Japanese junior high school students, and vice versa for low frequency and low range words. This is primarily because English in Japan is taught in the context of EFL, where students obtain most of their explicit vocabulary knowledge from school education. Based on this, Sato (2016) suggested that however high the frequency of a word is, it is rarely taken up by Japanese junior high school students until it is taught in English class. Therefore, Sato hypothesised that the learning period should be added as the third crucial factor along with frequency and range in developing word lists and VSTs for Japanese learners. For example, the word *language* first appears in the middle of the first year and appears a total of 817 times across all six authorised textbooks used in Japan. Its ranking criteria index is thus $817 \times 6 \div 1.5 = 3,264$. On the other hand, the word *energy* appears 840 times in five authorised textbooks, and first appears at the beginning of the second year. The ranking criteria index is thus $840 \times 5 \div 2 = 2,100$. Therefore, according to the index, *language* is listed higher than *energy*, even though its frequency is lower.

The VST used consists of four levels: Level 1—500-word level, Level 2—1,000-word level, Level 3—1,500-word level, and Level 4—2,000-word level. Each level comprises 30 test items; as each level includes 500 words, each test word represents about 15 words at that level, as can be seen in the Appendix. The VST adopts a Computer-Based Testing (CBT) system programmed to randomly extract the test items from the word list described above, which thus serves as an item bank. Using the CBT test format, it is expected to avoid memory effects from being used repeatedly with the same learners over a short span of time. The pilot test was conducted by Sato (2016) to confirm the equivalency in difficulties across the forms randomly generated by the computer programme. One-way AOVAs were conducted for three different forms, and results show that there were no significant differences in means of raw scores across the forms (see Sato, 2016, 2017, for more detail). Consequently, in the present study, the vocabulary size of the participants was measured five times using this VST, but each time they took the test, the test items were different, even at the same level.

### 3.3 Procedure

The study examines the longitudinal vocabulary growth of Japanese junior high school students across their 3-year junior high school English education. Using the VST described in section 3.2 above, vocabulary size of the students was measured five times over three academic years: in July and December in the first and second years and July in the third year. One level, which comprises 30 test items, was administered per session, allowing students sufficient time to finish the whole test (approximately 15 minutes per session). The VST consists of four levels, consequently, the participants had four test sessions within each test period. Considering the burden on the participants, each test session was administered at the beginning of every English class. Thus, 4 days were needed for each test period.
3.4 Data Analysis

To examine the longitudinal vocabulary size growth of junior high school students, the vocabulary size was measured five times over the period of three academic years at intervals of 6 months, in July and December. A two-way repeated-measures ANOVA was conducted, considering the test time (five times over the 3 years) and levels of vocabulary size (500-, 1,000-, 1,500-, and 2,000-word levels) as independent variables and the test score as a dependent variable.

4. Results and Discussion

4.1 Vocabulary Size of Japanese Junior High School Students

The descriptive statistics of the test results of each word level are shown in Table 1. The vocabulary size was estimated by multiplying the ratio of correct answers by 2,000. Therefore, the estimated mean vocabulary size of the participants was: 1,295 in July of the first year; 1,420 in December of the first year; 1,570 in July of the second year; 1,655 in December of the second year; and 1,721 in July of the third year. Thus, the participants’ vocabulary grew by an average of 425 words across the 3 academic years. It should be noted that the dispersion of the participants’ vocabulary size remained relatively stable over the 3 years, while the mean vocabulary sizes increased.

However, as discussed above, the vocabulary sizes obtained from the present study can be overestimated compared with the actual vocabulary knowledge of the participants because of the MCT formats (Gyllstad et al., 2015; Stewart, 2014;

<table>
<thead>
<tr>
<th></th>
<th>500 WL</th>
<th>1,000 WL</th>
<th>1,500 WL</th>
<th>2,000 WL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st-J Mean</td>
<td>23.62</td>
<td>21.77</td>
<td>18.94</td>
<td>13.36</td>
</tr>
<tr>
<td>SD</td>
<td>4.42</td>
<td>4.75</td>
<td>4.83</td>
<td>4.21</td>
</tr>
<tr>
<td>95% CI</td>
<td>[22.91, 24.34]</td>
<td>[21.00, 22.53]</td>
<td>[18.15, 19.72]</td>
<td>[12.68, 14.04]</td>
</tr>
<tr>
<td>1st-D Mean</td>
<td>26.09</td>
<td>21.96</td>
<td>20.64</td>
<td>16.54</td>
</tr>
<tr>
<td>SD</td>
<td>4.90</td>
<td>4.59</td>
<td>3.75</td>
<td>4.66</td>
</tr>
<tr>
<td>95% CI</td>
<td>[25.29, 26.88]</td>
<td>[21.21, 22.70]</td>
<td>[20.03, 21.25]</td>
<td>[15.78, 17.29]</td>
</tr>
<tr>
<td>2nd-J Mean</td>
<td>28.88</td>
<td>25.15</td>
<td>21.73</td>
<td>18.42</td>
</tr>
<tr>
<td>SD</td>
<td>2.59</td>
<td>4.06</td>
<td>4.02</td>
<td>4.35</td>
</tr>
<tr>
<td>95% CI</td>
<td>[28.46, 29.30]</td>
<td>[24.50, 25.81]</td>
<td>[21.08, 22.38]</td>
<td>[17.71, 19.13]</td>
</tr>
<tr>
<td>2nd-D Mean</td>
<td>29.23</td>
<td>26.15</td>
<td>23.58</td>
<td>20.32</td>
</tr>
<tr>
<td>SD</td>
<td>2.35</td>
<td>3.60</td>
<td>4.23</td>
<td>3.78</td>
</tr>
<tr>
<td>95% CI</td>
<td>[28.85, 29.61]</td>
<td>[25.57, 26.74]</td>
<td>[22.90, 24.27]</td>
<td>[19.71, 20.93]</td>
</tr>
<tr>
<td>3rd-J Mean</td>
<td>29.62</td>
<td>28.11</td>
<td>24.72</td>
<td>20.79</td>
</tr>
<tr>
<td>SD</td>
<td>1.34</td>
<td>2.25</td>
<td>3.98</td>
<td>4.13</td>
</tr>
<tr>
<td>95% CI</td>
<td>[29.40, 29.84]</td>
<td>[27.74, 28.47]</td>
<td>[24.08, 25.37]</td>
<td>[20.12, 21.46]</td>
</tr>
</tbody>
</table>

Note. The maximum score for each level of the VST (Sato, 2016) was 30. WL = world level, SD = standard deviation, CI = confidence interval, 1st-J = July in the first year, 1st-D = December in the first year, 2nd-J = July in the second year, 2nd-D = December in the second year, 3rd-J = July in the third year.

Vocabulary Learning and Instruction, 10(1), 30–48.
Stoeckel et al., 2020). One possible way to eliminate the effect of overestimation as a result of guessing is adopting “correction-for-guessing formula” (Frary, 1988; Rowley & Traub, 1977). This formula estimates adjusted scores based on the number of questions which students got wrong on the test. The formula is: corrected score = # correct – (#incorrect/#options–1). For example, suppose a participant took the four-option multiple-choice VST adopted in this study and happened to get 105 correct and 15 incorrect items, then, the participant’s score would be decreased from 105 to 100: 105 – (15/3) = 100. Consequently, it is plausible to estimate the mean vocabulary size of the participants somewhere between 1,060 and 1,295 in July of the first year, 1,227 and 1,420 in December of the first year, 1,426 and 1,570 in July of the second year, 1,540 and 1,655 in December of the second year, and 1,628 and 1,721 in July of the third year as described in Figure 1.

According to the Course of Study (MEXT, 2008), Japanese junior high school students are supposed to learn 1,200 words. Compared to this standard, the estimated vocabulary sizes of the participants of the present study were quite large, even in July of the first year at the beginning of their formal study of English. It is notable that the average size of their receptive vocabulary was beyond the 1,200 word level in December of the first year even when strict estimation considering the effect of overestimation because of guessing was used. Therefore, it is probable that the participants understand almost all the 1,200 words, that is, by reading, even at the very beginning of their English study. From the viewpoint of vocabulary knowledge, one of the greatest concerns of some English teachers in Japan is that students may have difficulty reading passages without Japanese verbatim translation, as is suggested by the Course of Study (MEXT, 2018). The findings contribute to minimising the concern in this regard and teachers should be encouraged to try to conduct reading lessons without word-for-word translation.

Figure 1. The process of vocabulary growth.

Vocabulary Learning and Instruction, 10(1), 30–48.
These unexpected results—that the participants’ vocabulary size is as high as, if not higher than, the levels set by MEXT—may be partially explained by the abundance of English loan words experienced in students’ daily lives. For students exposed to English loan words, it is not so difficult to either understand or infer the meaning of English words in receptive ways (Nation and Webb, 2011). Daulton (2008) investigated which and how many of the first 3,000 words of the British National Corpus list occur in Japanese and reported that over half are used in Japanese as loan words in one form or other. Under the current Course of Study (MEXT, 2018), Japanese students start studying English as a required subject from junior high school, and consequently, most of the English words introduced at this stage are high-frequency basic words. Indeed, Morinaga et al. (1978, 1979) suggested that Japanese authorised English textbooks include a substantial amount of loan words. As Nation and Webb (2011) suggested, learners’ knowledge of loan words is a “tremendous resource for learning English” (p. 56). Therefore, many teachers assume that their students do know something of English for a variety of reasons. For example, English activities in elementary schools in Japan are prevalent, some students learn English from a young age at cram schools. This raises the need for further research investigating what and how many English loan words found in Japanese appear in the authorised English textbooks.

The results further suggest that it is necessary to revise the vocabulary test to measure larger vocabulary sizes (2,000 words) because of the revision of the Course of Study (MEXT, 2018). The number of vocabulary items to be taught in junior high school is to be increased up to 1,600 to 1,800 words, in addition to the 600 to 700 words to be taught in elementary school in the new English curriculum implemented from 2020 (MEXT, 2018). Therefore, students under the current Course of Study are expected to learn 2,200 to 2,500 words in total by the end of junior high school. Thus, it would be valuable to develop a VST capable of measuring vocabulary sizes up to 3,000 or 4,000 words.

Lastly, further studies to measure the productive vocabulary size of Japanese junior high school students would help teachers to make their vocabulary teaching more effective. It is problematic that the English Course of Study provided by the Japanese government fails to indicate whether the 1,600 to 1,800 words specified are receptive, productive, or both. It states only that “1,600 to 1,800 words should be taught in junior high school English lessons (MEXT, 2018, p. 33).” This seems not to be specific enough to show the goal of vocabulary teaching for teachers. As teachers might interpret this as requiring them to teach all these words as productive vocabulary, some may be overwhelmed by the large number of vocabulary to teach, while others might assign much harder work than necessary. Developing separate receptive and productive vocabulary lists, specifying the words to be learned by junior high school students based on empirical research rather than intuition, will contribute substantially to more effective vocabulary teaching and learning.

4.2. The Patterns of Vocabulary Growth by the Word Level

Table 2 shows the results of a two-way ANOVA, comparing the mean scores and standard deviations (SD) between the test times and the four levels of the
Table 2. Results of the Two-Way ANOVA for the Vocabulary Size Test

<table>
<thead>
<tr>
<th>Sources</th>
<th>SS</th>
<th>Df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Levels</td>
<td>37319.58</td>
<td>3</td>
<td>12439.86</td>
<td>796.04</td>
<td>&lt;0.001</td>
<td>0.81</td>
</tr>
<tr>
<td>Test Time</td>
<td>16204.01</td>
<td>4</td>
<td>4051.00</td>
<td>259.23</td>
<td>&lt;0.001</td>
<td>0.32</td>
</tr>
<tr>
<td>Interaction</td>
<td>868.20</td>
<td>12</td>
<td>72.35</td>
<td>4.63</td>
<td>&lt;0.001</td>
<td>0.02</td>
</tr>
<tr>
<td>Error</td>
<td>46256.27</td>
<td>2,960</td>
<td>15.62</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100648.06</td>
<td>2,979</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. SS = sum of squares, df = degrees of freedom, MS = mean sum of square, F = F ratio.

Table 3. Results of the Post Hoc Comparison (Level and Test Time)

<table>
<thead>
<tr>
<th>Level</th>
<th>Test Time</th>
<th>Mean Difference</th>
<th>p</th>
<th>Cohen's d</th>
</tr>
</thead>
<tbody>
<tr>
<td>500 WL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st-J</td>
<td>1st-D</td>
<td>2.46</td>
<td>&lt;0.001</td>
<td>0.20</td>
</tr>
<tr>
<td>1st-D</td>
<td>2nd-J</td>
<td>2.79</td>
<td>&lt;0.001</td>
<td>0.22</td>
</tr>
<tr>
<td>2nd-J</td>
<td>2nd-D</td>
<td>0.35</td>
<td>0.94</td>
<td>0.03</td>
</tr>
<tr>
<td>2nd-D</td>
<td>3rd-D</td>
<td>0.40</td>
<td>0.91</td>
<td>0.03</td>
</tr>
<tr>
<td>1,000 WL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st-J</td>
<td>1st-D</td>
<td>0.19</td>
<td>0.99</td>
<td>0.02</td>
</tr>
<tr>
<td>1st-D</td>
<td>2nd-J</td>
<td>3.19</td>
<td>&lt;0.001</td>
<td>0.26</td>
</tr>
<tr>
<td>2nd-J</td>
<td>2nd-D</td>
<td>1.00</td>
<td>0.19</td>
<td>0.08</td>
</tr>
<tr>
<td>2nd-D</td>
<td>3rd-D</td>
<td>1.95</td>
<td>&lt;0.001</td>
<td>0.16</td>
</tr>
<tr>
<td>1,500 WL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st-J</td>
<td>1st-D</td>
<td>1.70</td>
<td>0.03</td>
<td>0.14</td>
</tr>
<tr>
<td>1st-D</td>
<td>2nd-J</td>
<td>1.09</td>
<td>0.68</td>
<td>0.09</td>
</tr>
<tr>
<td>2nd-J</td>
<td>2nd-D</td>
<td>1.85</td>
<td>&lt;0.001</td>
<td>0.15</td>
</tr>
<tr>
<td>2nd-D</td>
<td>3rd-D</td>
<td>1.14</td>
<td>0.09</td>
<td>0.09</td>
</tr>
<tr>
<td>2,000 WL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st-J</td>
<td>1st-D</td>
<td>3.17</td>
<td>&lt;0.001</td>
<td>0.25</td>
</tr>
<tr>
<td>1st-D</td>
<td>2nd-J</td>
<td>1.89</td>
<td>&lt;0.001</td>
<td>0.30</td>
</tr>
<tr>
<td>2nd-J</td>
<td>2nd-D</td>
<td>1.90</td>
<td>&lt;0.001</td>
<td>0.15</td>
</tr>
<tr>
<td>2nd-D</td>
<td>3rd-D</td>
<td>0.47</td>
<td>0.84</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Note. WL = word level, 1st-J = July in the first year, 1st-D = December in the first year, 2nd-J = July in the second year, 2nd-D = December in the second year, 3rd-D = December in the third year.

VST considering the test time (five times over the 3 years) and levels of vocabulary size (500-, 1,000-, 1,500-, and 2,000-word levels) as independent variables and the test score as a dependent variable. The results reveal a significant interaction between the two factors of the level and test time ($F(12, 2960) = 4.63, p < 0.001$, $\eta^2 = 0.02$ [small effect size, Field, 2005]). In addition, there were significant differences between the main effect at the four levels ($F(3, 2960) = 796.04, p < 0.001$, $\eta^2 = 0.81$ [large effect size]) and the test times ($F(4, 2960) = 259.23, p < 0.001$, $\eta^2 = 0.32$ [large effect size]).

To examine the pattern of vocabulary size growth in detail, the main effects were analysed using post hoc analysis. The results of Tukey’s multiple comparison indicate that there were significant differences between the four levels (500-word level, > 1,000-word level, > 1,500-word level, > 2,000-word level), showing that each level measures the participants’ vocabulary knowledge distinctively. This confirms the validity of the four-level construction of the VST (Sato, 2016).

The results for the simple main effect of the test time reveal the distinctive patterns of vocabulary growth for each level, as shown in Table 3 and Figure 2.
The vocabulary size did not always increase linearly. Instead, a distinctive pattern is revealed for each level. For the 500-word level, the participants showed a significant increase until July of the second year and reached a plateau after that, indicating a ceiling effect. These results can be interpreted to suggest that the participants had acquired the vocabulary of the 500-word level by July of the second year. For the 1,000-word level, on the other hand, there were significant increases observed from December of the first year through July of the second year, and from December of the second year through July of the third year, the 7th to 8th grade levels. For the 1,500-word level, a significant increase was observed from July through December of the second year. Finally, for the 2,000-word level, the participants’ vocabulary grew at each point until December of the second year.

These distinctive patterns of learners’ vocabulary growth are explained by the characteristics of the words in each level and the contents taught at each period. For example, most basic and frequently used words, such as *do, have, like, can, what, how, good, many, time, Japan,* and *English* are included in the 500-word level. These words are used when discussing general topics rather than being highly content or topic specific. It is important to note that the vocabulary knowledge needed to express all tenses—past, present, and future—are included in the 500-word level, such as the base forms of important basic verbs and their past tense forms (*did, went, had, and came*), and auxiliary verbs (*will, may, etc.*). This is consistent with the structure of the syllabus of Japanese English textbooks,
Sato: Developing a Vocabulary Size Test for Beginner Learners

where the verb tenses are introduced in the order of present, past, and future by July of the second year. In other words, Japanese junior high school students are expected to acquire basic sentence patterns and tenses of English using high-frequency words included in the 500-word level, which are used commonly in most interactions. Therefore, it is crucial for Japanese students to be able to understand the vocabulary included in the 500-word level by the end of the second year, which provides a fundamental base for their future English learning. This substantiates the suggestion made by Kosuge (2003) that loss of motivation and onset of difficulties with studying English begins to appear in the middle of the second year. Indeed, Japanese teachers of English often refer to the second year as “the crossroads of English learning,” marking a significant period for students. Therefore, measuring student vocabulary growth at this point as a formative assessment may contribute to appropriate adjustments to classes and provision of remedial instruction if necessary.

The vocabulary of the 1,000 and 1,500-word levels, on the other hand, consists of words related to more specific topics or contents, for example, Japanese culture or tradition (shrine, temple, cherry blossom, pray, and ancient), or social issues (environment, fossil fuel, volcano, wheelchair, and volunteer). Once students learn the fundamental structures of English together with the vocabulary at the 500-word level, they are exposed to less frequently used words to express more specific contents. Vocabulary growth at the 1,000 and 1,500-word levels beginning after December of the first year reflects this delayed exposure to these words in the textbooks.

Finally, words of 2,000-word level are characterised as “topic-specific vocabulary reflective passage content,” that is, vocabulary directly related to the topics dealt with in the textbook chapters. These words are neither high frequency words in general English use, nor is this vocabulary borrowed into Japanese as loan words. These words are necessary for the specific topics or stories of the textbook, thus they appear only in the relevant passages in English textbooks. For example, words such as lullaby, sugarcane, and thatch appear in a story about Okinawa in World War II, and polar bear, aurora, and iceberg appear in a biography of a photographer working in the Arctic. Knowledge of these words grows continuously from the first year to the second year because learners rarely have knowledge of the vocabulary of this level from the beginning. The mean score for this level is 13.36 out of 30, indicating room for development. However, the present study failed to account for the reason why the vocabulary knowledge of this level develops at each time point, even though these words appear only in limited lessons or one particular story. Further research is needed to seek possible explanations.

While intuition leads English teachers in Japan to generally assume that vocabulary size increases linearly from high-frequency to low-frequency words, the present study reveals that this is not the case. The results show a distinctive pattern according to each level and/or the characteristics of the words. Although future studies on this topic are needed to verify or generalise these findings considering that this pattern would change if a different set of textbooks were used at this school, the present research offers invaluable findings not only for vocabulary teaching but also for preparing appropriate reading materials for Japanese

Vocabulary Learning and Instruction, 10(1), 30–48.
students by controlling the vocabulary level to match the target learners based on
the data obtained by measuring their vocabulary size.

5. Conclusion

The present study primarily aimed to investigate the pattern of vocabulary
growth of beginning English learners. The findings suggest that the vocabulary
size of junior high school students in this context increases differently across word
frequency levels. This implies the following regarding the growth in vocabulary
size of junior high school students.

It is clearly evident that the learners already have a fairly large vocabulary
even in July (fourth month) of the first year at the beginning of their formal study
of English, estimated at around 1,060 and 1,295 words. Learner vocabulary there-
after develops, showing a distinctive pattern for each level. The 500-word level
increases significantly until July of the second year, at which point it plateaus.
These are words to express general topics or contents that students should surely
acquire as soon as possible. The 1,000-word level vocabulary size increases from
December (ninth month) of the first year to July of the second year, and from
December of the second year to July of the third year. For the 1,500-word level, a
significant increase occurs from July through December of the second year. These
words mainly relate to limited or specific topics. The 2,000-word level, including
textbook topic-related words, develops consistently until December of the second
year. The observed patterns of vocabulary development can be explained by the
relationship between the characteristics of vocabulary items included in each level
and the topics or contents of the textbooks at a time.

This research provides a new and effective approach for vocabulary instruc-
tion at the beginning level for Japanese English learners. The findings of the pres-
ent study also constitute a valuable resource both for material development and
setting guidelines for material selections appropriate to the levels of learners. They
should guide instructors in controlling the vocabulary during the preparation of
reading materials that will be helpful to students based on the data obtained by
measuring their vocabulary size. These results should be confirmed by further
research with junior high school students in different contexts in a case study
approach considering the possibility that vocabulary of every learner does not
develop precisely as outlined above.

The limitations of this study should also be noted. First, the vocabulary sizes
of the third-year students are probably larger than the coverage of the VST adopted
in the present study (2,000 words). In addition, for the fact that the vocabulary size
to be taught is increased in the new Course of Study (MEXT, 2018), the need is to
revise the vocabulary test to measure larger vocabulary sizes. Second, the cur-
rent study failed to examine either individual test takers or test items. Therefore,
the degree to which test takers and words conform to or deviate from these pat-
terns remains unclear. Research comparing item measure and person measure
indices and investigating how well the difficulty levels actually correlate with the
adjusted frequency values utilising Rasch Analysis is already underway. Third, as
the participants are from a junior high school attached to a national university,
they likely have larger vocabulary size than students from public schools. In that sense, application of this research method targeting students with various backgrounds should be fruitful in extending its implications. Finally, it is probable that the participants’ vocabulary size was overestimated for adopting multiple-choice VSTs because of the guessing effects (Gyllstad et al., 2015; Stewart, 2014; Stoeckel et al., 2020). Therefore, it is urgent to develop alternative test formats which estimate learner vocabulary size precisely with more confidence. Stewart (2014) and Stoeckel et al. (2020) suggested that it can be achieved with adapting item response theory (IRT) or computer adaptive tests. Fourth, as the present study only focused on learners’ receptive vocabulary size, a further study should examine the pattern of their vocabulary growth from a different aspect, such as productive vocabulary learning or depth of vocabulary knowledge.

Acknowledgements

I would like to thank the English teachers of the junior high school for cooperating in the 3-year longitudinal research, without whose help this would never have been possible. I would also like to thank Editage (www.editage.com) for English language editing.

References


**Appendix**

*Sample of the Vocabulary Size Test for the 500-Word Level*

**中学生のための語彙サイズテスト (500 word Level)**

1. **now**
   a. 一緒に～しましょう、～しよう
   b. 今
   c. ～の後に
   d. ～です。～である。～にある、～にいる

2. **your**
   a. あなた（たち）の
   b. そこへ（で・に）
   c. ～を食べる
   d. ～について、およそ、約

3. **our**
   a. ～を使う
   b. 私の
   c. 私たちの
   d. 水

4. **park**
   a. 公園
   b. しばしば
   c. ～を試す、試み
   d. 魅力的、興味深い

5. **word**
   a. 一緒に
   b. (直前の) を選った、～を買った、選れていた
   c. 冬
   d. ことば、語

6. **speak**
   a. 長い、長く
   b. 外へ
   c. 幸せな
   d. 話す
Sample of the Vocabulary Size Test for the 1,000-Word Level

中学生のための語彙サイズテスト (1000 word Level)

ID 名前
1〜30の単語の意味として最もふさわしいものを①〜④の中から選び記号を○で囲みなさい。

1. enough
   a. ～しましょうか、～しませんか
   b. 行った、催した
   c. 十分な（に）
   d. ピザ

2. die
   a. 忙しい
   b. 声
   c. 転、ブラン
   d. 死ぬ

3. such
   a. 地球
   b. 他に（の）
   c. そのような
   d. すべて

4. news
   a. 外で、屋外で
   b. 隠かな
   c. 知らせ、ニュース
   d. 音楽家

5. bridge
   a. 事務所、会社
   b. 橋（はし）
   c. 木
   d. おばあちゃん

6. cup
   a. ～の間で（に）
   b. 城
   c. 茶碗、カップ
   d. 地震

Sample of the Vocabulary Size Test for the 1,500-Word Level

中学生のための語彙サイズテスト (1500 word Level)

ID 名前
1〜30の単語の意味として最もふさわしいものを①〜④の中から選び記号を○で囲みなさい。

1. professional
   a. 大学
   b. ～を修がらせる、おびえす
   c. 目の不自由な
   d. 職業（係）の、木造（の）

2. shock
   a. 事故（はし）、世話
   b. ～を作り出す、生み出す
   c. 黒板
   d. 衝撃、振動

3. address
   a. （うちけしの文で）もう（～ない）
   b. （洋服の）ボタン
   c. 住所
   d. お母さん、ママ

4. square
   a. ある方
   b. 四角の、正方形
   c. 難い
   d. 植物、～を植える

5. act
   a. 行進、行進する
   b. 行動する
   c. 祖先、先祖
   d. とても、本当に

6. medal
   a. 神
   b. 島（どく）
   c. ～の上／の／で
   d. 齢
Sample of the Vocabulary Size Test for the 2,000-Word Level

中学生のための語彙サイズテスト (2000 word Level)

ID 名前
1〜30の単語の意味として最もふさわしいものを a〜dの中から選び記号を○で囲みなさい。

1. import
   a. 目が覚めた
   b. 半分
   c. 〜を移す (シフトする)
   d. 離れる

2. relay
   a. ばっと放す (落下する)
   b. 〜を稼ぐ (かせぐ), 〜を得る
   c. ひざ
   d. リレー戦争

3. profession
   a. 苦い
   b. 〜を主張 (しょう張る)
   c. 職業, 専門教
   d. ナス

4. final
   a. 郵便箱, ポスト
   b. 色, 風景
   c. (ほうきやブラシなどで) 掃く (はく)
   d. 決勝戦

5. tug-of-war
   a. 機、場所
   b. 地下鉄
   c. 價引き
   d. 残念なこと

6. acid rain
   a. わらぶき
   b. 鍵 (なわ), 鍵 (つな)
   c. 酸性物
   d. うなずく

Vocabulary Learning and Instruction, 10(1), 30–48.