



P-ISSN 2355-2794  
E-ISSN 2461-0275

## Measuring the English Vocabulary Acquisition of Japanese Learners

Yoshie Ishikawa<sup>1</sup>  
Ayano Otaki<sup>\*2</sup>  
Hiromu Okamura<sup>3</sup>  
Tomohiko Shirahata<sup>2</sup>

<sup>1</sup>English Department, Faculty of Foreign Studies, Tokoha University, Shizuoka City 422-8581, JAPAN

<sup>2</sup>Department of English, Faculty of Education, Shizuoka University, Shizuoka City 422-8529, JAPAN

<sup>3</sup>Department of International Languages and Cultures, School of International Relation, University of Shizuoka, Shizuoka City 422-8529, JAPAN

### Abstract

*The measurable vocabulary knowledge of Japanese learners of English (JLEs) has yet to be fully investigated. This study, therefore, attempts a detailed investigation of JLEs' vocabulary knowledge to clarify the relationship between the frequency of words and their difficulty level and to identify the factors that affect the difficulty of vocabulary learning. We tested 192 university-level JLEs on 1,035 verbs, requesting that they choose a correct Japanese translation. The average number of correct answers for each verb was compared with its frequency ranking. Moreover, the characteristics of verbs that received high and low percentages of correct answers were examined. The three major findings were as follows. First, the average number of correct answers was 751.97 (72.65% of the verbs investigated). Second, the average number of correct answers decreased as the frequency of the verb decreased. However, this tendency diminished beyond the 4,000-word level. Finally, not a few low-frequency verbs garnered a high percentage of correct answers; these were verbs that included affixes or were English loanwords that were possibly known to JLEs. However, other low-frequency verbs received a*

---

\* Corresponding author, email: [otaki.ayano@shizuoka.ac.jp](mailto:otaki.ayano@shizuoka.ac.jp)

**Citation in APA style:** Ishikawa, Y., Otaki, A., Okamura, H., & Shirahata, T. (2024). Measuring the English vocabulary acquisition of Japanese learners. *Studies in English Language and Education*, 11(1), 59-75.

Received January 24, 2023; Revised June 17, 2023; Accepted June 21, 2023; Published Online January 31, 2024

<https://doi.org/10.24815/siele.v11i1.30460>

*low percentage of correct answers because they had abstract or multiple meanings and lacked clues that could facilitate an understanding of the word's meaning. We concluded that these were the main factors affecting the degree of difficulty in learning each vocabulary item. Regarding this study's pedagogical implications, efficient use should be made of English-based loanwords and affixes, and university-level JLEs should intentionally learn low-frequency words.*

**Keywords:** Vocabulary learning, vocabulary size, vocabulary tests, word lists, Japanese learners of English.

## 1. INTRODUCTION

Learning vocabulary is essential for acquiring language. Researchers believe that words are indispensable components of listening, speaking, reading, and writing (e.g., Webb & Nation, 2017). In Japan, vocabulary learning has been playing an increasingly important role in English education (Nakata, 2019). Regarding the number of words to be introduced into English classes, the latest edition of the Course of Study (Ministry of Education, Culture, Sports, Science and Technology, 2018) calls for 4,000 to 5,000 words to be taught from elementary to high school. This is an increase of 1,000 to 2,000 words over the preceding edition of the Course of Study. The Ministry emphasizes in the Course of Study that vocabulary knowledge is essential for the development of communication skills.

With this renewed awareness of the importance of vocabulary learning, there is a growing need to objectively measure the vocabulary knowledge of Japanese learners of English (JLEs) and gain an overview of their vocabulary learning (Masumi & Ishikawa, 2014). In university courses, learners' vocabulary size is used as an indicator to determine the difficulty level of planned materials. Some studies have attempted to measure the vocabulary size of university-level JLEs and examine the relationship between word frequency and the degree of difficulty (e.g., Aizawa & Iso, 2007; Brown, 2012; Okamoto, 2008). However, JLEs' vocabulary knowledge should be measured by employing more appropriate word lists that reflect the learning experience of JLEs, not that of native speakers (McLean et al., 2014). Moreover, the specific words that learners know and do not know should be investigated by testing them on all of the target words (Kitano & Chiba, 2018). Therefore, teachers can be made aware of the words their students should be taught, so they can make the content of their English language courses more appropriate.

To analyze the factors that affect the difficulty of vocabulary learning, this study aimed to measure the vocabulary knowledge of university-level JLEs. To that end, we made use of an existing vocabulary list, the New JACET List of 8000 Basic Words (New JACET 8000) (Mochizuki, 2016). We focused specifically on verbs on the list, which are significant as receptive and productive vocabulary. In line with the research objective, our research questions are listed below.

For the 1,000 or so verbs in the New JACET 8000,

1. How many verbs can university-level JLEs correctly identify the meanings of in Japanese?
2. Is the order of verb frequency related to the level of difficulty for university-level JLEs?

3. Regarding high-frequency verbs, what are the differences between the characteristics of verbs with a high percentage of correct responses and those with a low percentage?
4. Regarding low-frequency verbs, what are the differences between the characteristics of verbs with a high percentage of correct responses and those with a low percentage?

Through our analysis, we aimed to identify factors other than frequency that affect the ease or difficulty of individual words. Furthermore, we considered the implications of our findings for vocabulary learning at Japanese universities.

## 2. LITERATURE REVIEW

### 2.1 Tests of Vocabulary Knowledge for Second Language Learners

The Vocabulary Size Test (VST) (Nation & Beglar, 2007) aims to measure total receptive vocabulary knowledge and to present it as vocabulary size. Aizawa and Mochizuki (2010) developed a valid and reliable VST for JLEs, by presenting the meanings of English words in Japanese and improving the vocabulary list which the test relies on. It consists of seven levels ranging from the 1,000- to 7,000-word level. The format of the test involves selecting English words that correspond to the presented Japanese words, with 26 questions for each level.

The VST has been employed in several studies measuring the vocabulary size of JLEs. McLean et al. (2014), for example, used the VST to measure the vocabulary size of 3,449 university-level JLEs; they reported a strong correlation between VST scores and university departments' standardized rank scores. Nonaka (2020), meanwhile, used the VST to measure increases in the vocabulary size of 98 junior college students. The results of his survey indicated that the student's vocabulary size increased by about 10% at the 2,000- to 4,000-word level and about 15% at the 5,000- and 6,000-word level. Sato et al. (2021) also used the VST to investigate the vocabulary size of 576 freshmen. They compared the test results of their students in April and those in December to examine the effects of the class organization on English proficiency. The results showed that the class division based on academic achievement was more effective for lower achievers.

Researchers generally agree that the VST can measure receptive vocabulary size in a relatively easy way (e.g., Stoeckel et al., 2019). However, the vocabulary size being measured is an "estimate": the VST divides a word list into different levels and uses words sampled from each level. Thus, the learner does not necessarily know the number of words indicated by the test result. Another limitation is that the test does not identify which words are known and which ones are unknown. For this study, therefore, we tested university-level JLES on all of the words in the vocabulary list to accurately determine which words they knew and did not know.

### 2.2 Lists of English Words Used for Vocabulary Size Measurement Tests

When examining learners' vocabulary size, it is necessary to determine the English words that need to be studied. McLean et al. (2014) used the VST proposed by Nation and Beglar (2007) to study JLEs' English vocabulary. The test was based on a word list compiled from the British National Corpus (BNC). McLean et al. (2014)

noted, however, that the BNC's frequency ranking is not necessarily appropriate for measuring JLEs' vocabulary size. One reason is that the order of frequency in the list does not match the vocabulary knowledge of JLEs. For example, the word "nil," which is unfamiliar to JLEs, is included with the high-frequency words at the 2,000-word level, whereas basic words like "quiz" are included with the relatively low-frequency words at the 5,000-word level.

Milton (2007) suggested that frequency ranking of words based on native speakers' language use might not be relevant to foreign language learners who mainly have textbooks to draw on, since the vocabulary in textbooks tends to be selected based on the contents of materials rather than frequency. In Japan, the contents of English language education up through high school are influenced by university entrance exams (Browne & Culligan, 2008). It is possible, then, that JLEs learn vocabulary items that native English speakers would consider low-frequency words. It also is possible that JLEs do not know some of the everyday vocabulary frequently used by native English speakers. Therefore, vocabulary test creators should recognize that the frequency order of a word list for JLEs does not necessarily match that of a word list for native English speakers.

### **2.3 Relationship between the Frequency and Difficulty of Words**

In planning vocabulary learning and teaching, it is essential to consider what makes it easy or difficult to learn individual vocabulary items. Webb and Nation (2017) suggested that some inherent aspects of a word, such as irregular spelling, difficult pronunciation, or unfamiliar word parts, could make learning more difficult. Aside from those features, word frequency is a major factor that determines the degree of difficulty of vocabulary. Studies have demonstrated that high-frequency words are recognized more quickly than low-frequency words, which applies to both L1 and L2 (e.g., Whitford & Titone, 2012). In addition, this effect of word frequency is greater for L2 speakers than for L1 speakers (Ishida, 2022). The word lists that language learners rely on indicate the frequency of the words listed and suggest that high-frequency words are easier to learn than low-frequency ones; the more frequently learners encounter a word, the easier it will be for them to learn it. Brown (2012) investigated whether the vocabulary knowledge of university-level JLEs followed the frequency model (i.e., the idea that words are learned in order of their frequency). He reported the results that high-frequency words tended to be well-known to the participants of his study, which confirmed the frequency model of vocabulary learning. Therefore, it is pedagogically important to examine the relationship between word frequency and word difficulty.

Aizawa and Iso (2007) investigated how frequency affects difficulty levels for JLEs. They administered two 80-question English vocabulary tests to 164 undergraduates. The participants answered 10 questions for each of the eight levels (in order of frequency) with 1,000 words in each level from the JACET List of 8000 Basic Words (JACET 8000) (JACET Committee of Basic Words Revision, 2003). For each question, three Japanese words were presented and participants were asked to select one English word from a group of six words that corresponded to each of the three Japanese words. The results showed a relationship between frequency and degree of difficulty up until the 4,000-word level of the JACET 8000. For words beyond the 4,000-word level, however, there was no tendency for the difficulty of learning to increase as the frequency decreased.

Okamoto (2008) also investigated the relationship between word frequency and level of difficulty in learning. She randomly selected 200 words from the Longman Dictionary of Contemporary English and the Collins Cobuild Learners' Dictionary, and then she asked 291 university students to choose the most appropriate Japanese translation from the five options or "no correct answer." After the test, 37 words containing inappropriate options were excluded, leaving 163 words for the analysis. These words were classified into 15 levels in order of frequency based on the JACET 8000 and the BNC, and the percentage of correct answers for each level was compared. The results showed that the relationship between frequency and difficulty began to weaken at the 6,000-word level but a correlation was observed up to the 8,000-word level.

Aizawa and Iso (2007) and Okamoto (2008) both suggested that, up to a certain level, difficulty increases as frequency decreases, but that beyond a particular level, increased difficulty and decreased frequency are not necessarily linked. There is no consensus, however, regarding the level at which the link begins to weaken. Therefore, additional research should be conducted to determine the threshold level beyond which difficulty levels are not linked to frequency and to scrutinize what factors influence the difficulty of learning words.

### **3. METHODS**

#### **3.1 Participants**

The 192 JLEs who participated in this study included 132 first-year and 60 second-year students from two universities in Shizuoka Prefecture. They were enrolled in English classes that the authors of this study taught. The classes were held once a week and the vocabulary tests were given to the students in each class for 13 weeks. When administering the test, we explained the study purpose to the students, how the survey results would be used, and the fact that this research would be made public. Participants who took all 13 tests (192 students) were included in the study. Prior to enrolling in university, the participants had received English education in elementary and high school. The participants' majors included agriculture, engineering, informatics, sociology, and English. Despite the differences in majors, years in school, and which university they attended, the participants' overall English proficiency could be considered intermediate for Japanese university students. Most of them had scored in the 500s on the Test of English for International Communication (TOEIC) Listening and Reading Test.

#### **3.2 Materials**

We relied on the New JACET 8000, which aims to provide JLEs with a bank of vocabulary for daily life and for reading papers and giving presentations in English while pursuing academic studies at university. It was created using the procedures described below. First, the editors used the BNC and the Corpus of Contemporary American English (COCA) to create a base list that reflects native English language usage in five genres (spoken, fiction, magazine, newspaper, and academic). Subsequently, 2,188 basic words were selected from English textbooks authorized by the Ministry of Education, high school entrance examinations, university entrance

examinations, English proficiency tests, and English-English dictionaries. The words were arranged in order of frequency according to both the materials for Japanese learners and the base list. Next, other 3,553 words required for university disciplines were selected and their frequency was established by cross-referencing them with English newspapers and introductory academic books. Finally, 2,259 words were added from the base list in order of frequency as advanced-level vocabulary, totaling 8,000 words. In summary, the New JACET 8000 is based on the actual vocabulary used by native English speakers as well as the materials most frequently used by JLEs.

Nation (2013) estimated that the 1,000 most frequent word families<sup>†</sup> cover 81.94% of the text of the BNC and that the 3,000 most frequent word families cover 95.37%. Comparatively, educated adult native English speakers know about 20,000-word families. In Japan, a total of 4,000 to 5,000 words are introduced into English classes up to the high school level. We believe, therefore, that the New JACET 8000 could allow us to examine the actual retention of vocabulary learned in elementary and high school as well as the vocabulary size of university-level JLEs.

We limited our survey to verbs because the number of verbs in the New JACET 8000 (1,045 words) was appropriate for English vocabulary tests in 15 university classes during a semester. Verbs are important in both receptive and productive vocabulary because they play a significant role in expressing meaning and structuring a sentence. We believe that surveying all verbs in the New JACET 8000, rather than sampling a subset of verbs, would make it possible to scrutinize the factors affecting the degree of difficulty in learning each word.

### 3.3 Procedure

From the 1,045 verbs listed in New JACET 8000, we decided to exclude “be,” which is typically used in its modified form, and “fuck,” which is profanity, leaving 1,043 words for the survey. It became clear, however, that some of the incorrect options were confusingly similar to the correct options. Therefore, eight words (‘take’, ‘sell’, ‘pick’, ‘serve’, ‘finish’, ‘bear’, ‘perform’, and ‘cry’) for which there were very similar options were excluded, and 1,035 verbs were ultimately included in the study.

The test asked participants to identify the appropriate Japanese translation of an English verb from six options. The correct answer always corresponded to the most basic meaning of the verb in Japanese. We used multiple-choice so that the scorers’ subjectivity would not intervene in the scoring. In addition, by offering a relatively large number of choices (six), we attempted to minimize the guessing of correct answers as much as possible. The participants were allowed as much time as they needed to answer all questions; it ranged from 15 to 20 minutes for each session. The present authors administered the tests. They asked the participants to answer 80 questions on 80 words (83 in the last session) during a weekly English class held from April to July 2021 (13 weeks), beginning with the most frequently used words according to the New JACET 8000. Table 1 gives a sample of the test questions.

**Table 1.** Sample test questions.

1	do	A	Pull	b	watch	c	grow	d	feel	e	do	f	bring
2	say	A	Smile	b	know	c	cry	d	sing	e	hope	f	say
3	think	A	Answer	b	like	c	think	d	tell	e	worry	f	find

<sup>†</sup> Laufer (1992) suggested that 3,000 word families correspond to 5,000 words.

Table 1 continued...

4	get	A	Hit	b	finish	c	give	d	hold	e	put	f	get
5	see	A	move	b	see	c	stand	d	guess	e	make	f	lead

*Note.* The six choices were presented in Japanese.

### 3.4 Data Analysis

Participants' answers were scored by allotting one point for one correct answer and zero for an incorrect answer. The maximum number of correct answers was 1,035. First, to answer RQ1, we analyzed the total number of correct answers by calculating the mean score, standard deviation, and highest and lowest numbers.

Second, to address RQ2, we categorized the verbs based on the difficulty level (high school, university, and advanced) outlined in the New JACET 8000 and calculated the mean score, standard deviation, and correct percentage. Then, to compare the relationship between the number of correct answers and the frequency of verb use on the New JACET 8000, we categorized the 1,035 verbs into groups of 50 words each (i.e., 1-50, 51-100, etc.) according to their frequency of use. The verb that appeared as no. 1 on the New JACET 8000 was the most frequently used verb, with increasing numbers indicating a decreasing frequency of use. Linear regression analysis was performed to determine the relationship between the number of correct answers and the frequency of the word. A one-way repeated measures ANOVA (within subjects) was conducted to determine the statistically significant differences in the mean scores between each group of 50 words (e.g., between Group 1-50 and Group 51-100), followed by multiple comparison tests using the Bonferroni method.

Next, to further examine the relationship between correct responses and frequency of use, we divided the verbs into four groups and considered the characteristics of the verbs in each category. The four groups were as follows: high frequency and high percentage of correct responses (Group A), high frequency and low percentage of correct responses (Group B), low frequency and high percentage of correct responses (Group C), and low frequency and low percentage of correct responses (Group D). To answer RQ3, we examined the verbs in Group A and Group B.

To investigate RQ4, we first focused on the verbs in Group C and further subdivided them into three categories (Subgroup I: Words incorporated into Japanese as loanwords; Subgroup II: Words with affixes; Subgroup III: Others). Then, we examined the characteristics of Group D. For this analysis, we subdivided the verbs in Group D into three categories in the same way that we did for Group C.

## 4. RESULTS AND DISCUSSION

### 4.1 Overall Results

As shown in Table 2, of the 1,035 English verbs presented, the average number of words for which 192 participants correctly selected the Japanese translation was found to be 751.97; the correct response rate was 72.65%. The highest number of correct answers for a single participant was 980 (94.69%) and the lowest was 558 (53.91%), a difference of 422 words (40.78%). From these results, we can presume that the average vocabulary size of the participants was approximately 5,600 words

(72.65% of 8,000 words). This aligns with Aotani (2012), who estimated that the average vocabulary size of university-level JLEs who had learned English vocabulary using word lists to prepare for entrance exams was between 4,800 and 6,400 words. Thus, regarding RQ1 (i.e., How many verbs can university-level JLEs correctly identify the meanings of in Japanese?), the average number of correct responses was 751.97 (72.65%). Thus, university-level JLEs with TOEIC scores in the 500s were found to know the Japanese translations of more than 70% of the verbs in the New JACET 8000.

**Table 2.** Test results.

Total number	Average number of correct answers (%)	Standard deviation	Highest number of correct answers (%)	Lowest number of correct answers (%)
1,035	751.97 (72.65%)	76.64	980 (94.69%)	558 (53.91%)

## 4.2 Frequency of Use and Difficulty in Learning

The next analysis considered the number of verbs at each level in the New JACET 8000 and the average percentage of correct answers (see Table 3). The 367 verbs at the high school level were regarded as basic English verbs that everyone learns; for these, there was a high percentage of correct answers (93.64%). This indicates that Japanese translations of English verbs learned in high school are mostly established for university-level JLEs. The next 553 “university-level” verbs are those that appear frequently in English newspapers and in introductory academic materials that university students need to know for their respective majors. The correct response rate for these verbs was 64.76%, which is not high. This could be partly because the participants were either freshmen or first-semester sophomores. Consequently, they were largely in the early stages of study in their major. The remaining 115 words were “advanced” vocabulary, which was selected in order of frequency from the base list based on the BNC and the COCA as words necessary for the study and the daily life of college students. The correct response rate was even lower, 43.94%, showing that university-level JLEs are not very familiar with the low-frequency words selected from the corpora of native English speakers.

**Table 3.** Number of verbs tested by level and average percentage of correct answers.

	Number of verbs	Numbers in the test list	Average number of correct answers (%)	Standard deviation
High school	367	1-367	343.66 (93.64%)	13.47
University	553	368-920	357.77 (64.70%)	57.73
Advanced	115	921-1,035	50.54 (43.95%)	10.99

Next, as shown in Table 4<sup>‡</sup> and Figure 1, the mean score for each 50-word category based on the frequency ranking in the New JACET 8000 represents the overall tendency between word frequency and the number of correct responses. These data show how the average number of correct responses changed according to the frequency ranking in the New JACET 8000. The overall trend was that, as the frequency decreased, the average score decreased. The  $R^2$  value was 0.93, which

<sup>‡</sup> The results are indicated for up to 1,000 (in order of frequency) of the 1,035 words surveyed.



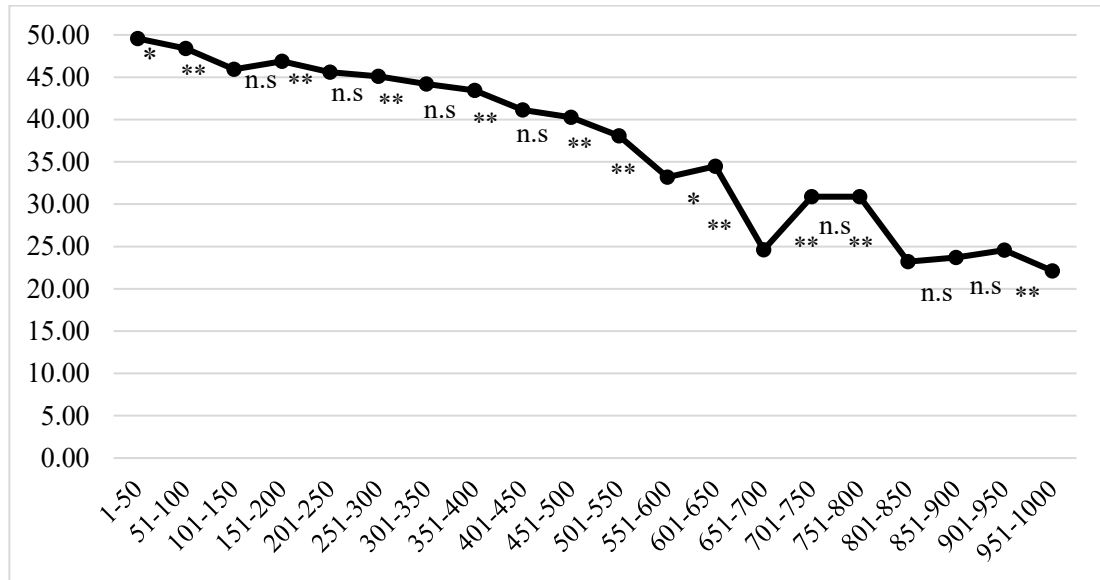
means that the linear regression line could explain 93% of the observed variables (average score). We can safely say, therefore, that the frequency ranking in the New JACET 8000 is related to the difficulty ranking of verbs for JLEs. However, the average scores from 651 to 1000 did not reflect this overall tendency. The  $R^2$  (from 651 to 1000) was 0.33, indicating that the linear regression line could only explain approximately 33% of the observed variables. By contrast, the linear regression line from 1 to 650 could explain approximately 92% of the observed variables ( $R^2 = 0.92$ ).

This tendency was also observed when comparing the average scores between each category of the 50 words. A one-factor ANOVA was performed ( $F(19, 3629) = 1027.77$ ,  $p < .001$ , partial  $\eta^2 = .84$ ), followed by multiple comparisons using the Bonferroni method. When looking at each 50-word category, the trend was particularly pronounced from 1 to 650 words; there was a tendency for the average number of correct responses to decrease with the decrease in frequency. However, the average number of correct responses increased significantly from 651 to 700. After the 800th word, although the frequency of use decreased, the average number of correct answers did not significantly decrease.

We can say, therefore, that the correct responses from 1 to 650 were highly influenced by frequency; that is, the average correct responses tended to decrease as frequency decreased. However, this trend began to weaken after the 601st verb, and it collapsed after the 800th. From 651 to 1000, correct responses might have been affected by factors other than frequency.

**Table 4.** Results for each 50-word category from the 50 most frequently used verbs ( $n = 192$ ).

Frequency ranking	Mean	SD	Percentage of correct answers
1-50	49.58	0.63	99.16%
51-100	48.40	4.10	96.80%
101-150	45.92	8.45	91.84%
151-200	46.86	2.18	93.73%
201-250	45.61	2.88	91.22%
251-300	45.10	4.14	90.21%
301-350	44.18	3.67	88.36%
351-400	43.43	4.94	86.85%
401-450	41.14	6.07	82.27%
451-500	40.27	6.59	80.54%
501-550	38.06	7.06	76.13%
551-600	33.19	6.93	66.39%
601-650	34.49	6.29	68.98%
651-700	24.59	7.71	49.19%
701-750	30.88	5.58	61.75%
751-800	30.88	5.58	61.75%
801-850	23.20	5.65	46.40%
851-900	23.71	5.15	47.43%
901-950	24.56	4.73	49.13%
951-1000	22.11	5.27	44.23%



Note. \*:  $p < .05$ , \*\*:  $p < .01$ , n.s.: not statistically significant

**Figure 1.** Test results for each of the 50 New JACET 8000 list rankings.

The 601st most frequent verb was the 3,717th word on the New JACET 8000 list. This is consistent with [Aizawa and Iso's \(2007\)](#) finding that frequency was not linked with difficulty at the 4,000-word level. This could be because, as [Okamoto \(2008\)](#) suggested, the actual differences between the frequencies of each word become less significant as frequency decreases. Moreover, [Aizawa and Iso \(2007\)](#) suggested that frequency is linked to difficulty with high-frequency words because the participants had learned them from textbooks or other materials, although this does not apply to low-frequency words. If this is the case, the threshold beyond which difficulty level is not linked to frequency might be affected by learners' proficiency since more proficient learners will be acquainted with more words. For example, [Okamoto \(2008\)](#) found that this threshold was at the 6,000-word level, while [Aizawa and Iso \(2007\)](#) suggested that it was at the 4,000-word level. [Okamoto \(2008\)](#) attributed this discrepancy to differences in the participants' proficiency levels. Thus, regarding RQ2 (Is the order of verb frequency related to the level of difficulty for university-level JLEs?), the general trend was that the level of difficulty increased as frequency decreased. However, the relationship between frequency and difficulty tended to weaken as frequency decreased below the 601st word in the tests conducted in this study.

### 4.3 Relationship between the Frequency and Percentage of Correct Answers

For the next analysis, we divided the verbs into four groups. The classification was made by considering verbs with a frequency ranking of 1-260 as high frequency and those with a frequency ranking of 776-1,035 as low frequency. Correct responses of 70% or more were considered high, while those of less than 30% were considered low. There were 250 verbs in Group A, 0 in Group B, 52 in Group C, and 77 in Group D (see Table 5).

**Table 5.** Groupings of verbs based on their frequency rank and percentage of correct answers.

Percentage of correct answers	Frequency ranking in the New JACET 8000	
	1-260	776-1035
70% to 100%	Group A (250 words) High frequency and high correct response rate	Group C (52 words) Low frequency and high correct response rate
0% to less than 30%	Group B (0 words) High frequency and low correct response rate	Group D (77 words) Low frequency and low correct response rate

The results for Group A showed that 250 of the 260 high-frequency words (96.15%) had high correct response rates of 70% or higher. Among the 10 verbs that did not have high correct response rates, five were included in the 101-200 category ('apply', 'fix', 'notice', 'preserve', 'stick'), and five were in the 201-260 category ('appreciate', 'describe', 'dine', 'represent', 'suppose'). Regarding Group B, no verb belonged to this group. These results indicate that high-frequency verbs did not result in low percentages of correct responses for university-level JLEs. Therefore, most of the verbs identified as high frequency in the New JACET 8000 were familiar to them. Thus, with regard to RQ3 (Regarding high-frequency verbs, what are the differences between the characteristics of verbs with a high percentage of correct responses and those with a low percentage?), no high-frequency verb received a low percentage of correct responses from the participants. We can infer, therefore, that frequency level is a factor that affects the degree of difficulty in learning high-frequency verbs.

We then analyzed the characteristics of the 52 verbs in Group C. Table 6 shows the results for the three subgroups in Group C. The breakdown is as follows: 18 words in Subgroup I (words incorporated into Japanese as loanwords), 32 in Subgroup II (words with affixes), and two in Subgroup III (not applicable to Subgroup I or II). Verbs belonging to Subgroup I are listed as loanwords in the *Daijisen* Digital Edition (Matsumura, 2022), an online Japanese-language dictionary. Verbs for which a derivative was listed in the dictionary were also included in Subgroup I because JLEs would be able to recognize them by association (e.g., 'circulator' for 'circulate'). It is possible that the participants unconsciously used their knowledge of loanwords to identify the Japanese translations of English verbs. Nation (2003) suggested that using loanwords incorporated into one's native language as a cue for vocabulary learning is an effective way to increase the vocabulary size of L2 learners. Rogers et al. (2015) experimented with JLEs and reported the results that English-based loanwords could decrease the difficulty of learning English words for JLEs.

**Table 6.** Detailed results for Group C.

Percentage of correct answers	Sub-group	Frequency ranking in the New JACET 8000		
		776-900	901-1000	1001-1035
100% to 90%	I	circulate<circulator hug	chop, nominate, peel	
	II	deepen, generalize, repay, restructure, rethink, rewrite, soften, unlock	darken, memorize, rearrange, socialize	

Table 6 continued...

90% to 80%	I	crawl, elevate<elevator, smash	bind, knit, sprinkle<sprinkler	
	II	differentiate, diversify, disconnect, empower, mislead, prolong, tighten	imprison, reheat, straighten, terrify	
80% to 70%	I	shave<shaver subscribe<subscription	drip, kneel<knee, sniff<sniffer, surrender	Flip
	II	disapprove, heighten, suppress, underlie	lengthen, exclaim	horrify, loosen, proclaim
	III	comprehend	murmur	

Subgroup II consisted of verbs composed of base forms with affixes. They accounted for 32 (61.54%) of the 52 verbs in Group C. Table 7 summarizes these 32 verbs by prefixes and suffixes. The 11 different prefixes, which have the function of changing the meaning of the base verb, are included in 17 of the 32 verbs. Among these, ‘re-’ was the most commonly used, appearing in five words, followed by ‘dis-’ and ‘pro-’ with two tokens each. Suffixes, meanwhile, change the part of speech of the base verb, and in all, four different suffixes were attached to 15 words. The most commonly used was ‘-en’, appearing in eight words, followed by ‘-fy’ and ‘-ize’ with three tokens each.

Researchers have proposed learning strategies for increasing vocabulary knowledge based on knowledge of prefixes and suffixes. Nation (2013) suggested that knowledge of affixes can be effective for inferring the meaning of unknown words and recognizing their relationship with other parts of speech belonging to the same word family. Mochizuki and Aizawa (2000) investigated JLEs’ knowledge of affixes to determine their order of difficulty. They reported that the most commonly acquired affixes were ‘re-’, ‘un-’, and ‘pre-’. These prefixes were found in some of the verbs in Group C. We can infer, therefore, that the participants’ knowledge of affixes might have helped them understand the meanings of those verbs.

**Table 7. Breakdown of Subgroup II in Group C.**

Prefix	Verb	Suffix	Verb
re-	rearrange, repay, restructure, rethink, rewrite	-en	darken, deepen, heighten, lengthen, loosen, soften, straighten, tighten
dis-	disapprove, disconnect	-fy	diversify, horrify, terrify
pro-	prolong, proclaim	-ize	generalize, memorize, socialize
em-	empower	-ate	differentiate
ex-	exclaim		
im-	imprison		
mis-	mislead		
pre-	preheat		
sup-	suppress		
un-	unlock		
under-	underlie		

The two verbs classified into Subgroup III (‘comprehend’ and ‘murmur’) are those for which knowledge of loanwords and affixes does not necessarily aid in understanding their meanings. To consider the possible reasons for the high percentage

of correct responses, we examined whether these verbs were included in three of the most commonly used English vocabulary books in Japanese high schools: English Vocabulary Target 1900 (Usami & Urata, 2020), System English Vocabulary (Shimo & Tone, 2019), and Speed Reading English Vocabulary (Kazahaya, 2015). Both verbs were found in two of these books, and ‘comprehension’, a derivative of ‘comprehend’, was found in all three (see Table 8). This suggests that the participants might have learned the two words before they started college.

**Table 8.** “Comprehend” and “murmur” in English vocabulary books.

	English Vocabulary Target 1900	System English Vocabulary	Speed Reading English Vocabulary
“comprehend”	○	○	×
“comprehension”	○	○	○
“murmur”	×	○	○

Next, we considered the results of Group D (see Table 9). The 77 words were divided into three subgroups in the same way that the verbs in Group C were divided. The breakdown was as follows: seven words for Subgroup I (words incorporated into Japanese as loanwords), 28 for Subgroup II (words with affixes), and 42 for Subgroup III (not applicable to Subgroup I or II). First, of the 77 verbs in this group, seven of them (‘deploy’, ‘dispatch’, ‘divert’, ‘flicker’, ‘offset’, ‘drape’, and ‘tailor’) are English loanwords in Japanese. However, this presumably did not help the participants understand their meanings since these words are not commonly used in Japanese and thus might not have been familiar to the participants.

**Table 9.** Detailed results for Group D.

Percentage of correct answers	Sub-group	Frequency ranking in the New JACET 8000		
		776–900	901–1000	1001–1035
less than 30%	I	divert, offset		Drape
	II	alleviate, contaminate, contemplate, deduct, deter, disguise, disperse, entail, fasten, recede, reside	condemn, detain, fabricate, testify	
	III	dilute, flatter, ignite, mourn, spawn, stagger, stalk, wrestle	allege, doom, grin, jerk, shudder,	amend, dwell, linger, lurk, preach, rattle, resent, sob, startle,
less than 20%	I	dispatch, tailor	deploy	flicker
	II	encompass, evacuate, exacerbate, mitigate, outweigh, recreate, saturate	aggravate, retort, stipulate	
	III	augment, forge, refute, shatter, shove, taint	chuck, curb, fling, perch, render, slam, solicit	clad, giggle, grate, grunt, stroll
less than 10%	I			
	II	underscore	replenish	
	III		poise, simmer, tumble	

Table 10 summarizes the 27 verbs in Subgroup II (three of them include both a prefix and a suffix). They are characterized by (i) abstractness of the base verb (e.g.,

‘replenish’, ‘exacerbate’, ‘aggravate’, ‘saturate’) or (ii) polysemy of the base verb (e.g., ‘outweigh’, ‘underscore’, ‘fasten’, ‘testify’); for example, ‘fast’, which is the base of ‘fasten’, has multiple meanings since it can mean “adhered to” in addition to the more common meaning of “moving at high speed.” As Laufer (1997) noted, abstractness and polysemy are semantic features that make it more difficult to learn the meanings of words. De Groot and Keijzer (2000) conducted an experiment and reported that abstract words were more difficult to learn and more susceptible to forgetting than concrete words. It might have been difficult, therefore, for our participants to recognize the meanings of the 27 verbs, even though they included affixes, because of the meanings of the base verbs.

**Table 10.** Breakdown of Subgroup II in Group D.

Prefix	Verb	Suffix	Verb
re-	recede, recreate, replenish, reside, retort	-ate	aggravate, alleviate, contaminate, contemplate, evacuate, exacerbate, fabricate, recreate, saturate, stipulate
dis-	disguise, disperse	-en	Fasten
en-	encompass, entail	-fy	Testify
con-	condemn, contaminate, contemplate		
de-	deduct, detain, deter		
ex-	exacerbate		
out-	outweigh		
under-	underscore		

Lastly, regarding the 43 words in Subgroup III, which are neither loanwords nor words including affixes, we assume that they do not present any clues that could help JLEs understand their meaning. Therefore, with regard to RQ4 (Regarding low-frequency verbs, what are the differences between the characteristics of verbs with a high percentage of correct responses and those with a low percentage?), verbs with high percentages were incorporated into Japanese as loanwords or had affixes. Thus, knowledge of loanwords and affixes might have helped the participants understand their meanings. This explains the fact that low-frequency verbs with a high percentage of correct responses in this study. Meanwhile, the low-frequency verbs with low percentages of correct responses can be categorized into three groups: words including the affix and base with an abstract meaning or multiple meanings, English loanwords that are not commonly used in Japanese, and words that do not provide learners with clues to recognize their meaning. Webb and Nation (2017) suggested that the degree of difficulty in learning a word is determined by how well the word can be adapted into the learner’s existing systems of knowledge. Thus, these factors might explain the difficulties in vocabulary learning among university-level JLEs. These results could indicate that as for low-frequency verbs, certain factors, aside from frequency level, such as affixes, English loanwords, abstractness, and polysemy, affect the degree of difficulty for JLEs.

## 5. CONCLUSION AND IMPLICATIONS

This study examined the vocabulary size of university-level JLEs and investigated the relationship between frequency ranking and the order of word difficulty. There were two major findings. First, the average number of correct responses was 751.97 (out of 1,035), and the correct response rate was 72.65%. This

is consistent with the average vocabulary size of university-level JLEs estimated by Aotani (2012). Second, difficulty tended to increase with decreasing frequency up to a certain frequency rate; however, frequency and difficulty were not regularly linked beyond that rate. This finding is also consistent with previous studies (Aizawa & Iso, 2007; Okamoto, 2008).

Additionally, we examined the characteristics of high-frequency verbs and compared those with a high percentage of correct responses to those with a low percentage. The results showed that none of the high-frequency verbs were classified as having a low correct response rate, which means that there will be no factors other than the frequency level that affect the difficulty of learning high-frequency words. Furthermore, we examined the characteristics of low-frequency words by comparing those with a high percentage of correct responses with those with a low percentage. It turned out that approximately 40% of the low-frequency verbs were classified as having a high correct response rate. Through a close examination of these verbs, they were found to be either incorporated into Japanese language as loanwords or composed of bases and affixes. Such characteristics could have helped the participants understand the meaning of the verbs and could be the reason for the high correct response rate, as well as the break in the link between frequency and degree of difficulty. Meanwhile, the other 60% of the low-frequency verbs that received a low percentage of correct responses had abstract or multiple meanings or did not present any clues about understanding their meanings. These factors, along with the frequency level, could have increased the degree of difficulty.

Our findings have the following pedagogical implications for vocabulary learning at a Japanese university level. First, since many low-frequency verbs had higher percentages of correct answers when knowledge of affixes and loanwords could be applied, learners should be made aware that affixes and loanwords are effective tools for increasing vocabulary knowledge. At least 80% of loanwords in Japanese derive from English (Ishino, 1983). Therefore, as Daulton (2008) suggested, JLEs should be encouraged to take advantage of these English-based loanwords as potential cognates, even though their first language has few similarities with their target language. Second, we found that more than half of the low-frequency verbs were unknown to the participants or difficult for them to understand, even though they were English loanwords or included affixes. Therefore, explicit teaching and intentional learning should focus on these words.

Regarding the scope for future research, first, it is necessary to examine whether participants' proficiency is a factor that determines the point at which the relationship between frequency and degree of difficulty begins to weaken, as suggested by Aizawa and Iso (2007) and Okamoto (2008). Second, it is necessary to examine whether there are other factors affecting the difficulty level in learning vocabulary items aside from frequency, affixes, loanwords, abstract meaning, and polysemy. Finally, we focused on verbs from a vocabulary list. Words belonging to other parts of speech should be investigated to get a better grasp of the vocabulary knowledge of university-level JLEs.

## REFERENCES

- Aizawa, K., & Iso, T. (2007). Estimating word difficulty: The divergence from frequency levels. *Annual Review of English Language Education in Japan*, 18, 111-120. [https://doi.org/10.20581/arele.18.0\\_111](https://doi.org/10.20581/arele.18.0_111)

- Aizawa, K., & Mochizuki, M. (2010). *Eigo goishidou no jissen aidiashu* [Collection of practical idea for English vocabulary teaching]. Taishukan.
- Aotani, M. (2012). *Eigo gakushuron* [A theory of English language learning]. Asakura Shoten.
- Brown, D. (2012). The frequency model of vocabulary learning and Japanese learners. *Vocabulary Learning and Instruction*, 1(1), 20-28. <http://dx.doi.org/10.7820/vli.v01.1.brown>
- Browne, C., & Culligan, B. (2008). Combining technology and IRT testing to build student knowledge of high frequency vocabulary. *The JALT CALL Journal*, 4(2), 3-16. <https://doi.org/10.29140/jaltcall.v4n2.59>
- Daulton, F. (2008). *Japan's built-in lexicon of English-based loanwords*. Multilingual Matters, Ltd.
- De Groot, A. M. B., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning*, 50(1), 1-56. <https://doi.org/10.1111/0023-8333.00110>
- Ishida, T. (2022). Nihonjin eigogakushusha to eigobogowasha no tangohindo kouka hikaku [Comparison of word frequency effects between Japanese learners of English and native English speakers]. *Journal of Nihon Fukushi University Education Center*, 10, 23-30.
- Ishino, H. (1983). *Gendai gairai gokou* [Thoughts on current loanwords]. Taishukan.
- JACET Committee of Basic Words Revision. (Ed.). (2003). *Daigaku eigo kyouiku gakkai kihongo risuto JACET 8000* [JACET list of 8000 basic words]. Japan Association of College English Teachers.
- Kazahaya, H. (2015). *Sokudoku eitango 2 joukyuuhen kaitei dai 4 han* [Rapid reading English vocabulary 2 Advanced level, 4<sup>th</sup> ed.]. Z-kai.
- Kitano, M., & Chiba, K. (2018). Goi chishiki no ichisokumen: Nihonjin daigakusei no rikai ga husokushiteiru koshiyohindogoi ni kansuru kousatsu [A profile of vocabulary knowledge: Identifying high frequency English words unknown to Japanese university students]. *Journal of Department of International Studies, Bunkyo University*, 29(1), 19-29. <http://id.nii.ac.jp/1351/00006829/>
- Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 126-132). Macmillan. [https://doi.org/10.1007/978-1-349-12396-4\\_12](https://doi.org/10.1007/978-1-349-12396-4_12)
- Laufer, B. (1997). What's in a word that makes it hard or easy: Some intralexical factors that affect the learning of words. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 140-155). Cambridge University Press.
- Masumi, A., & Ishikawa, S. (2014). Koukousei no eigoryoku o suiteisuru goitesuto no kentou: Saizu tesuto to sokudo tesuto no hikaku [Examining a vocabulary test to estimate high school students' English proficiency: A comparison of size test and speed test]. *Journal of International Communication Center, Kobe University*, 11, 14-30. [http://www.lib.kobe-u.ac.jp/handle\\_kernel/81008801](http://www.lib.kobe-u.ac.jp/handle_kernel/81008801)
- Matsumura, A. (2022). *Daijisen dezitaru ban* [Daijisen, Digital ed.]. Shogakukan.
- McLean, S., Hogg, N., & Kramer, B. (2014). Estimations of Japanese university learners' English vocabulary sizes using the vocabulary size test. *Vocabulary Learning and Instruction*, 3(2), 47-55. <http://dx.doi.org/10.7820/vli.v03.2.mclean.et.al>



- Milton, J. (2007). Lexical profiles, learning styles and the construct validity of lexical size test. In H. Daller, J. Milton, & J. Treffers-Daller (Eds.), *Modelling and assessing vocabulary knowledge* (pp. 47-58). Cambridge University Press.
- Ministry of Education, Culture, Sports, Science and Technology. (2018). *Kotogakko gakushu shido yoryo* [A course of study for high school]. Kairyudo.
- Mochizuki, M. (2016). *JACET 8000: The new JACET list of 8000 basic words*. Kiriara.
- Mochizuki, M., & Aizawa, K. (2000). An affix acquisition order for EFL learners: An exploratory study. *System*, 28(2), 291-304. [https://doi.org/10.1016/S0346-251X\(00\)00013-0](https://doi.org/10.1016/S0346-251X(00)00013-0)
- Nakata, T. (2019). *Eitango gakushu no kagaku* [Science of English vocabulary learning]. Kenkyusha.
- Nation, I. S. P. (2003). The role of the first language in foreign language learning. *Asian EFL Journal*, 15, 1-8. <https://doi.org/10.26686/wgtn.12560333.v1>
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2<sup>nd</sup> ed.). Cambridge University Press.
- Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9-13. <https://doi.org/10.26686/wgtn.12552197.v1>
- Nonaka, T. (2020). On development of English vocabulary size of Japanese junior college students. *Nigata Seiryō Junior College Research Bulletin*, 50, 181-190.
- Okamoto, M. (2008). Assessing lexical knowledge of English at low-frequency levels: Implications for vocabulary teaching. *Annual Review of English Language Education in Japan*, 19, 141-150. [https://doi.org/10.20581/arele.19.0\\_141](https://doi.org/10.20581/arele.19.0_141)
- Rogers, J., Webb, S., & Nakata, T. (2015). Do the cognacy characteristics of loanwords make them more easily learned than noncognates? *Language Teaching Research*, 19(1), 9-27. <https://doi.org/10.1177/1362168814541752>
- Sato, K., Terauchi, H., Matsuya, A., Funaki, T., Yamada, H., & Cannegieter, D. (2021). Future directions of English language education at Takachiho University: Based on the vocabulary knowledge of its first-year students. *Takachiho Ronso*, 55(4), 1-17.
- Shimo, K., & Tone, M. (2019). *Shisutemu eitango 5teiban* [System English vocabulary (5<sup>th</sup> ed.)]. Sundaibunko.
- Stoeckel, T., Stewart, J., McLean, S., Ishii, T., Kramer, B., & Matsumoto, Y. (2019). The relationship of four variants of the Vocabulary Size Test to a criterion measure of meaning recall vocabulary knowledge. *System*, 87, 102161. <https://doi.org/10.1016/j.system.2019.102161>
- Usami, M., & Urata, F. (2020). *Eitango tagetto 1900 6teiban* [English vocabulary target 1900 (6<sup>th</sup> ed.)]. Obunsha.
- Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.
- Whitford, V., & Titone, D. (2012). Second-language experience modulates first- and second-language word frequency effects: Evidence from eye movement measures of natural paragraph reading. *Psychonomic Bulletin & Review*, 19, 73-80. <https://doi.org/10.3758/s13423-011-0179-5>