研究論文

The Development of a Vocabulary Size Estimation Test for Japanese EFL Learners

Hitoshi Yashima

日本人英語学習者のための新しい語彙サイズ推定テストの開発

八　島　　等

Abstract

The study examines the Japanese EFL students' vocabulary size at the 1,000- to 4,000-word levels, using my newly developed vocabulary size estimation test (hereafter VSE test). The study also examines the rank correlation in the success rate in answering in my test. Data are analysed in terms of the response of knowing the meaning of a word in my previous checklist tests and the *JACET List of 8000 Basic Words* (hereafter JACET 8000), and according to participants' grades and schools. Results show that the average vocabulary size of students is much smaller than previously estimated by existing estimation tools. There is a statistically significant difference in the means in the vocabulary size according to grades and schools. There is a relatively strong rank correlation between the response of knowing the meaning of a word and the success rate in answering while there is a relatively weak rank correration between the rank in the JACET 8000 and the success rate in answering. Also there is a moderate to strong rank correlation between grades and between schools in the success rate in answering at each level with few exceptions. I conclude that educators should recognize the fact that there are many words difficult for students to acquire in common regardless of grades or schools and should change or devise the way of teaching these words.

1. Introduction

Little research has been done on the vocabulary size of Japanese EFL learners because it is very difficult to measure (Read, 2000). There are problems with terminology such as what constitutes a "word," and what it means to know a word? As there are various grammatical and morphological permutations of vocabulary (Schmitt, 2000: 2), unless we define what we count as a word, the vocabulary size measured can vary according to researchers. And as there are many degrees of knowing a word, unless we define which aspect to measure, the results also can vary according to researchers. We have to decide what kind of vocabulary test to administer and how many items the vocabulary test should have.

Nation and Waring (1997) indicate that learners need to know about 3,000 high-frequency words, so it is essential for educators to measure students' vocabulary size in order to help them develop their English skills. Yashima (2002) estimated the vocabulary size of Japanese EFL high school students, using Mochizuki's (1998) Vocabulary Size Test (hereafter VST) at the 1,000- to 6,000-word levels. The average vocabulary size for first-year students was 3,083 to 3,203 words while that of third-year students was 3,703 to 4,247. Mochizuki's test is a multiple-choice test and has 30 items at each level in order to test receptive knowledge. However, this may be too small to provide a valid indication of vocabulary size. Yashima's (2003) more exhaustive study yielded much lower estimated vocabulary size (hereafter EVS) across the board. This paper attempts to find a better estimation of the vocabulary size of Japanese EFL learners, using my newly developed VSE test, chiefly based on the results in Yashima (2003).

2. The study

2.1. Purpose

The purpose of this study is to clarify how many words Japanese EFL learners know at the 1,000- to 4,000-word levels receptively. This was based on the results in my previous checklist tests. I also examine the rank correlation between participants' ability to correctly discern word meanings and their prior responses as to whether or not they know the word. I also examine (1) the rank correlation between participants' success rates and the JACET 8000, and (2) the rank correlation between school grades and between schools.

2.2. Hypotheses

With the above points in mind, I have constructed the following hypotheses:

(1) The EVS of participants, regardless of grades or schools, will be much smaller than estimated by Mochizuki's (1998) VST in Yashima (2002).

(2) There is a strong rank correlation between the percentage of the response of knowing the meaning of a word in my previous checklist tests and the success rate in answering in this study.

(3) There is a strong rank correlation between the rank in the JACET 8000 and the success rate in answering in this study.

(4) There is a strong rank correlation between grades and between schools in the success rate in answering.

2.3. Participants

Participants in this study were 375 first-year students (15–16 years olds) from two Tokyo metropolitan senior high schools (Schools A and B), 201 second-year students (16–17 years olds) from two Tokyo metropolitan senior high schools and one private school (Schools A, C, and D), 144 third-year students (17–18 years olds) from one Tokyo metropolitan senior high school and one Saitama prefectural senior high school (Schools A and E), and 56 sophomores (19–43 years olds) from a science university. They had been learning English for 3 to 9 years.

2.4. Materials

2.4.1. Selection of target words and their distractors

In Yashima (2003), participants were tested on 2,087 words at the 1,000- to 3,000-word levels taken from the *Longman Dictionary of Contemporary English* (3rd ed.). All the words were re-classified into 1,000- to 4,000-word levels in accordance with the JACET 8000. An extra 469 words from the JACET 8000 at the 3,000- and 4,000-word levels were then added; as per Yashima (2003). The results were that the 1,000-word level had 551 words, the 2,000-word level had 797, the 3,000-word level had 544, and the 4,000-word level had 643. Each target word had two distractors at each level, based on Shizuka's (2003) study. Words which were almost the same in percentage terms at each level were selected as a target word and its two distractors, and words whose percentage was in the middle of the three words were selected as target words in principle. Only nouns, verbs, and adjectives were used.

2.4.2. Forms and arrangement of items and directions

Participants were required to select one of three English words that were closest in meaning to two Japanese words. These Japanese translations were adopted from the *Wisdom English-Japanese Dictionary*. 100 target words at each level were arranged in the order of the means of percentage of knowing their meanings, taking into consideration participants' motivation for the test and the effective administration of the test. They were also required to put a circle or a triangle round the number at the side of the item if they could answer it at a guess or because they knew the other two options.

2.5. Procedure

This study required one 50-minute session. Participants were given the test at the 1,000- to 4,000-word levels. Participants were told to answer as many questions as possible in order of numbers and levels, required to follow the directions mentioned in 2.4.2.

2.6. Scoring

If participants could answer 1 question at each level, they were regarded as having 10 words regardless of levels. However, there were two cases where participants could choose the correct answer if they didn't know it. Therefore, if participants put a circle round the number at the side of the item, they could answer it at a guess with a probability of one in three, so they were regarded as having 3.3 words. Moreover, if they put a triangle round it, they could answer it because they knew the other two options, so they were regarded as having 6.7 words.

3. Results

3.1. Estimated vocabulary size (EVS)

Table 1 shows the means and standard deviations in scores and EVS at the 1,000- to 4,000-word levels. The means vary according to grades. A one-way ANOVA revealed a significant main effect for grades in scores at the 1,000- to 4,000-word levels ($F(2, 772)=190.17$, $p<.001$, $F(2, 772)=194.08$, $p<.001$, $F(2, 772)=274.26$, $p<.001$, $F(2, 772)=272.50$, $p<.001$, respectively) and a significant main effect for grades in EVS ($F(2, 772)=299.67$, $p< .001$). A multiple comparison analysis using Tukey HSD's multiple range test revealed a significant difference between the means in scores and EVS for first- and second-years, first- and third-years, first-years and university sopho-

mores, second- and third-years, and second-years and university sophomores at $p<.001$ respectively at the 1,000- to 4,000-word levels. There was no significant difference between the means for third-years and university sophomores at any level or overall.

Table 1

Means and Standard Deviations in Scores and EVS for First- to Third-Year Senior High School Students and Sophomore University Students

Word level		*1st years* *N*=375	*2nd years* *N*=201	*3rd years* *N*=144	*Sophomores* *N*=56	*Total* *N*=776
1,000	*Mean*	698.63	814.63	941.39	903.37	788.50
	SD	119.29	121.34	88.06	91.77	148.58
2,000	*Mean*	473.86	652.08	823.63	788.21	607.61
	SD	169.55	165.13	154.62	143.35	216.74
3,000	*Mean*	220.48	425.92	629.82	602.04	377.19
	SD	139.50	165.74	203.37	174.94	233.01
4,000	*Mean*	65.95	239.60	432.11	422.96	204.64
	SD	96.90	147.00	220.33	182.18	210.01
EVS	*Mean*	1,458.92	2,132.23	2,826.95	2,716.60	1,977.94
	SD	446.09	548.36	622.69	543.02	758.93

Table 2 also shows the means and standard deviations in scores and EVS at the 1,000- to 4,000-word levels. The means vary according to schools. A *t* test revealed a significant difference between the means in scores at the 1,000- to 4,000-word levels for first-years ($t(373)=9.98$, $p<.001$, $t(373)=10.21$, $p<.001$, $t(373)=7.17$, $p<.001$, $t(373)=5.60$, $p<.001$, respectively) and for third-years ($t(142)=3.40$, $p<.01$, $t(142)=4.25$, $p<.001$, $t(142)=4.21$, $p<.001$, $t(142)=4.11$, $p<.01$, respectively). Also there was a significant difference between the means in EVS levels for first- and third-years ($t(373)=10.28$, $p<.001$, $t(142)=4.11$, $p<.001$ respectively). The one-way ANOVA revealed a significant main effect for schools in scores at the 1,000- to 4,000-word levels for second-years ($F(2, 198)=43.77$, $p<.001$, $F(2, 198)=65.31$, $p<.001$, $F(2, 198)=24.76$, $p<.001$, $F(2, 198)=45.91$, $p<.001$, respectively) and a significant main effect for schools in EVS for second-years ($F(2, 198)=43.77$, $p<.001$). The multiple comparison analysis revealed a significant difference between the means in scores and EVS for Schools A and D and Schools C and D at $p<.001$ respectively at the 1,000- to 4,000-word levels. There were also a significant difference between the means in scores for Schools A and C at $p<.001$ at the 2,000-word level and at $p<.05$ at the 3,000-

word level. There was a significant difference between the means in EVS for Schools A and C at $p<.05$. There was no significant difference between the means for Schools A and C at the 1,000- and 4,000-word levels.

Table 2

Means and Standard Deviations in Scores and EVS for First- to Third-Year Senior High School Students according to Schools

Word level		*1st years*		*2nd years*			*3rd years*	
		School A	*School B*	*School A*	*School C*	*School D*	*School A*	*School E*
		N=140	*N*=235	*N*=121	*N*=48	*N*=32	*N*=106	*N*=38
1,000	*Mean*	769.47	656.44	785.87	784.46	968.63	926.99	981.58
	SD	94.98	112.22	112.80	102.01	28.92	97.47	27.45
2,000	*Mean*	576.40	412.77	632.00	551.89	878.27	792.68	909.98
	SD	130.14	160.76	151.64	154.54	79.41	159.30	99.91
3,000	*Mean*	283.31	183.05	412.68	352.60	585.98	589.43	742.51
	SD	143.95	122.57	147.47	145.30	159.91	203.09	158.31
4,000	*Mean*	100.86	45.15	225.99	177.93	383.53	396.85	530.45
	SD	113.88	78.35	130.07	110.53	165.55	220.07	191.53
EVS	*Mean*	1,730.04	1,297.40	2,056.55	1,866.88	2,816.41	2,705.94	3,164.52
	SD	407.79	386.13	470.93	448.88	401.80	635.24	440.94

3.2. Rank correlation between the percentage of the response of knowing the meaning of a word and the success rate in answering

Table 3 shows the rank correlation between the percentage of the response of knowing the meaning of a word and the success rate in answering. The correlation coefficients were calculated using Kendall's tau and Spearman's rho. They are relatively or rather high: Kendall's tau was .61 to .70 at the 1,000- and 2,000-word levels regardless of grades and overall and it was .62 to 68 at the 3,000- and 4,000-word levels for first-years and overall. It was .54 to .60 at the 3,000- and 4,000-word levels for second- and third-years and university sophomores. Spearman's rho was .79 to .88 at the 1,000- and 2,000-word levels regardless of grades and overall and it was .80 to 88 at the 3,000- and 4,000-word levels for first-years and overall. It was .72 to .78 at the 3,000- and 4,000-word levels for second- and third-years and university sophomores.

3.3. Rank correlation between the rank in the JACET 8000 and the success rate in answering

Table 4 shows the rank correlation between the rank in the JACET 8000 and the

success rate in answering, using Kendall's tau and Spearman's rho. The correlation coefficients are relatively or rather low: Kendall's tau was .20 to .29 at the 1,000-, 2,000-, and 4,000-word levels regardless of grades and overall. It was .14 to .18 at the 3,000-word level regardless of grades and overall. Spearman's rho was .32 to .44 at the 1,000-, 2,000-, and 4,000-word levels regardless of grades and overall except for university sophomores at the 1,000-word level (.28). It was .21 to 28 at the 3,000-word level regardless of grades and overall.

Table 3

Rank Correlation between the Percentage of the Response of Knowing the Meaning of a Word and the Success Rate in Answering

Word level		*1st years* *N*=375	*2nd years* *N*=201	*3rd years* *N*=144	*Sophomores* *N*=56	*Total* *N*=776
1,000	Kendall	.66**	.63**	.66**	.70**	.69**
	Spearman	.85**	.83**	.84**	.87**	.88**
2,000	Kendall	.69**	.66**	.62**	.61**	.70**
	Spearman	.87**	.84**	.80**	.79**	.86**
3,000	Kendall	.68**	.56**	.54**	.60**	.63**
	Spearman	.86**	.75**	.72**	.77**	.81**
4,000	Kendall	.68**	.59**	.54**	.54**	.62**
	Spearman	.88**	.78**	.74**	.72**	.80**

Note. $^{**}p<.01$.

Table 4

Rank Correlation between the Rank in JACET 8000 and the Success Rate in Answering

Word level		*1st years* *N*=375	*2nd years* *N*=201	*3rd years* *N*=144	*Sophomores* *N*=56	*Total* *N*=776
1,000	Kendall	.23**	.22**	.24**	.20**	.24**
	Spearman	.33**	.32**	.35**	.28**	.34**
2,000	Kendall	.29**	.28**	.25**	.25**	.29**
	Spearman	.42**	.41**	.37**	.37**	.41**
3,000	Kendall	.16*	.18**	.18**	.14*	.18**
	Spearman	.26*	.26**	.28**	.21*	.27**
4,000	Kendall	.23**	.25**	.29**	.25**	.28**
	Spearman	.34**	.38**	.44**	.36**	.41**

Note. $^{*}p<.05$. $^{**}p<.01$.

3.4. Rank correlation between schools in the success rate in answering

Table 5 shows the rank correlation between schools in the success rate in answering, using Kendall's tau and Spearman's rho. These are relatively high:

Kendall's tau was .67 to .80 at every level between Schools A and B and between Schools A and C except at the 4,000-word level between Schools A and B (.54). It was .53 to .54 at the 1,000- and 3,000-word levels while it was .60 to .67 at the 2,000- and 4,000-word levels between Schools A and E. It was .48 to .54 at every level between Schools A and D and between Schools C and D except at the 2,000-word level between Schools A and D (.64) and between Schools C and D (.61). Spearman's rho was .85 to .93 at every level between Schools A and B and between Schools A and C except at the 4,000-word level between Schools A and B (.72). It was .71 to .72 at the 1,000-, and 3,000-word levels while it was .79 to .85 at the 2,000, and 4,000-word levels between Schools A and E. It was .60 to .72 at every level between Schools A and D and between Schools C and D except at the 2,000-word level between Schools A and D (.81) and between Schools C and D (.80).

Table 5

Rank Correlation between Schools in the Success Rate in Answering

Word level		*1st years*	*2nd years*			*3rd years*
		S A vs *S B* (140) (235)	*S A* vs *S C* (121) (48)	*S A* vs *S D* (121) (32)	*S C* vs *S D* (48) (32)	*S A* vs *S E* (106) (38)
1,000	Kendall	.73**	.72**	.50**	.48**	.54**
	Spearman	.90**	.89**	.64**	.60**	.72**
2,000	Kendall	.70**	.72**	.64**	.61**	.67**
	Spearman	.87**	.89**	.81**	.80**	.85**
3,000	Kendall	.80**	.75**	.54**	.49**	.53**
	Spearman	.93**	.91**	.72**	.65**	.71**
4,000	Kendall	.54**	.67**	.50**	.50**	.60**
	Spearman	.72**	.85**	.70**	.69**	.79**

Note. $^{**}p<.01$.

3.5. Rank correlation between grades in the success rate in answering

Table 6 shows the rank correlation between grades in the success rate in answering, using Kendall's tau and Spearman's rho. These are relatively high: Kendall's tau was .62 to .78 at every level between any schools regardless of levels except at the 4,000-word level between first- and third years and between first-years and university sophomores (.56 and .58 respectively). Spearman's rho was .80 to .93 at every level between any schools regardless of levels except at the 4,000-word level between first- and third-years and between first-years and university sophomores (.75 and .76 respectively).

Table 6

Rank Correlation between Grades in the Success Rate in Answering

Word level		*1st* vs *2nd* (375) (201)	*1st* vs *3rd* (375) (144)	*1st* vs *Soph* (375) (56)	*2nd* vs *3rd* (201) (144)	*2nd* vs *Soph* (201) (56)	*3rd* vs *Soph* (144) (56)
1,000	Kendall	.74**	.67**	.66**	.75**	.68**	.71**
	Spearman	.91**	.85**	.84**	.90**	.85**	.86**
2,000	Kendall	.78**	.70**	.73**	.77**	.74**	.72**
	Spearman	.93**	.86**	.90**	.93**	.91**	.90**
3,000	Kendall	.67**	.65**	.77**	.71**	.71**	.73**
	Spearman	.85**	.84**	.88**	.88**	.88**	.91**
4,000	Kendall	.65**	.56**	.58**	.69**	.64**	.62**
	Spearman	.84**	.75**	.76**	.87**	.82**	.80**

Note. $^{**}p<.01$.

4. Discussion

As can be seen in Tables 1 and 2, there is a clear difference in scores and EVS among grades and schools at each level and overall with few exceptions. The results of the ANOVA confirm that there was a statistically significant difference among grades at each level and overall except the case between third-year students and university sophomores and the case between Schools A and C at the 1,000- and 4,000-word levels. The tables also show that as grades advance in senior high school, participants acquire about 700 words in EVS each year on average. The reason why the difference in EVS between first- and second-years in School A are only about 300 is that all the first-year participants were in the advanced English class. If the data of the other 100 students are added, the EVS is much smaller.

In comparison with Mochizuki's (1998) VST in Yashima (2002), the means in EVS estimated by the current test are much smaller. In the case of first-years, the means are 1297.40 to 1730.04 in this study while they were 2961.61 to 3082.42 in Yashima, and in the case of third-years, 2705.94 to 3164.52 while 3703.36 to 3853.03. Japanese EFL high school students really know few words at the 5,000- and 6,000-word levels, taking into consideration the Course of Study. Therefore, even though the participants were given a test at these two levels in the same way as in this study, the means in EVS would not increase so much. My first hypothesis could be accepted.

Table 3 shows that there is a relatively strong rank correlation in both Kendall's tau and Spearman's rho between the percentage of the response of knowing the meaning of a word in my previous checklist tests and the success rate in answering in this

study, especially at the 1,000- and 2,000-word levels regardless of grades and overall. The results from one public senior high school and one private university were relatively true of other senior high schools. A much stronger rank correlation was expected, but there were 14, 15, 18, and 29 words whose difference in rank between the two means are twenty or more at the 1,000- to 4,000-word levels respectively. Participants mistook, for example, *firm*, *stuff*, *weak*, *ceiling*, *deserve*, *flesh*, *conservation*, *county*, and *yield* at the 2,000- to 4,000-word levels for *farm*, *staff*, *week*, *sailing*, *reserve*, *fresh*, *conversation*, *country*, and *field* respectively in my previous checklist tests. Moreover, there are so many loan words from English in Japanese and they didn't know the other meanings of their original English words, for instance, *concrete* and *draft* at the 3,000- and 4,000-word levels. Therefore, there was a big difference in rank between the percentage of the response of knowing the meaning of these words in my previous checklist tests and the success rate in answering in this study. My second hypothesis could also be accepted.

Table 4 shows that there is a relatively or rather weak rank correlation in both Kendall's tau and Spearman's rho between the rank in the JACET 8000 and the success rate in answering, especially at the 3,000-word level regardless of grades and overall. Although the JACET 8000 takes into consideration senior high school textbooks, as the number of the rank of a word in the JACET 8000 advances, it doesn't necessarily mean that it is more difficult for learners to acquire. One main reason for that is that for most of Japanese EFL learners, authorized textbooks are a major source of study and that as only the total number of words learned in junior and senior high schools are shown in the Course of Study, low-frequency words sometimes appear in these textbooks, depending on topics. My third hypothesis was rejected.

With regard to hypothesis 4, Tables 5 and 6 show that there is a relatively or rather strong rank correlation between grades and between schools in the success rate in answering at each level, especially at the 2,000-word level except between Schools C and D. This suggests that although, as mentioned in 3.1, there is a statistically significant difference in scores and EVS among grades and among schools, there are many words easy or difficult for Japanese EFL learners to acquire in common regardless of grades or schools at each level, especially at the 2,000-word level. These results are not so surprising because many words at the 2,000-word level are first learned in the 10th year. So there are many words even first-years acquire. As for second-years, as School D is a private school and the means in EVS was almost the same as the means for third-years, the rank correlation with public schools was weak. My fourth hypothesis could be accepted.

5. Conclusion and implications

This study shows that the vocabulary size of Japanese EFL learners is much smaller than previously estimated by Mochizuki's (1998). The means in EVS for first- and second-year students, on average, were 1458.92 and 2132.23 respectively. The results suggest that although it seems to Japanese teachers of English that this reflects their real vocabulary size much more, this is not good enough to perform functionally in English. The means in EVS for third-years and university sophomores, on average, were 2826.95 and 2716.60, which were the minimum vocabulary size to perform functionally. This study also shows that there is a relatively strong rank correlation between the percentage of the response of knowing the meaning of a word in my previous checklist tests and the success rate in answering in this study regardless of grades or levels and overall. Kendall's tau is .60 to .70 in almost all cases although there are 14, 15, 18, and 29 words whose difference in rank between the two means are twenty or more at the 1,000- to 4,000-word levels respectively. The results indicate that at each level, there are words easy or difficult for learners regardless of schools.

The findings on words that are difficult to acquire indicate that there are still some words difficult to learn for third-years or sophomores, even at the 1000-word level. Moreover, there are also some words difficult to learn although their ranks in the JACET 8000 are beginning at each level. As authorized textbooks are not written just in the order of word frequency, there is a chance that these words don't or rarely appear in them. So these words are difficult to acquire. Therefore, these findings suggest that educators should encourage their students to read and listen to additional materials, taking word frequency into consideration.

The findings on the rank correlation among schools and grades indicate that there are many words easy or difficult for Japanese EFL learners to acquire in common regardless of grades or schools at each level. Kendall's tau is .60 to .80 in most cases. These findings suggest that educators should recognize the fact and that they should change or devise the teaching way to teach words difficult for learners to acquire in common.

The results show that the current VSE test can be administed by other schools. However, the order of the arrangement of some target words or these target words themselves should be changed, and some other target words should also be changed because they are the words whose meanings in Japanese are the same as another

word at another level. After changing these points, I would like to measure learners' vocabulary size again by the revised version.

Notes

1. This paper is based on the contents of a presentation titled "The Development of a Vocabulary Size Estimation Test for Japanese EFL Learners" given at the 30th JASELE annual convention in Nagano. Some of the data has been updated since the presentation.
2. All participants gave the author their consent to use any information gathered in this study.

References

JACET Basic Words Revision Committee. (Ed.). (2003). *JACET list of 8000 basic words*. Japan Association of College English Teachers.

Longman dictionary of contemporary English (3rd. ed.). (1995). Longman.

Mochizuki, M. (1998). Nihonjin eigo gakushusha no tame no goi size test [The vocabulary size test for Japanese EFL learners]. *The IRLT Bulletin*, *12*, 45–52.

Ministry of Education, Science and Culture. (1989). *The course of study for upper secondary schools*. Ministry of Education, Science and Culture.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

Nation, I. S. P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6–19). Cambridge University Press.

Read, J. (2000). *Assessing vocabulary*. Cambridge University Press.

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge University Press.

Shizuka, T. (2003, August 9–10). *Can testers predict the worst-functioning distractors? Implications for multiple-choice item writing*. Paper presented at the 29th JASELE annual convention in Miyagi.

Yashima, H. (2002, August 7–8). *The vocabulary size of Japanese EFL senior high school students (2)*. Paper presented at the 26th KATE annual convention in Chiba.

Yashima, H. (2003, August 9–10). *The vocabulary size of Japanese EFL learners*. Paper presented at the 29th JASELE annual convention in Miyagi.

Wisdom English-Japanese dictionary. (2003). Sanseido.