

The Selection of 1,000 Basic Words Necessary for Japanese EFL Learners

Hitoshi Yashima

日本人英語学習者のための 基本1,000語の選定

八 島 等

Abstract

The study examines whether a word appearing in authorized textbooks for junior and senior high school students is a high-frequency word. The data are analyzed according to the frequency, appearance in authorized textbooks, and the importance of the word. The study attempts to select 1,000 basic words thought necessary for Japanese EFL learners to acquire. The study also examines the difficulty of words around the 1,000-word level in order to clarify whether or not senior high school students actually acquire these words. Results show that there are more than 30 very high-frequency words which do not appear in any authorized textbooks for junior high school students even among the 1,000 basic words selected. On the other hand, there are more than 10 lower-frequency words which appear in all the authorized textbooks. While first-year students acquire less than half of the words in the 1,000 prescribed words, third-year students acquire more than 80 percent of them. However, even third-year students have words they do not completely acquire. I conclude that educators should recognize this gap and should encourage their students to read and listen to as much additional material as possible to have them acquire these 1,000 words.

Introduction

Educators need to determine students' vocabulary size to help them develop

their English skills. Yashima (2004) estimated the vocabulary size of Japanese EFL high school and university students, using the Vocabulary Estimation Test (VET) at the 1,000- to 4,000-word levels taken from the *JACET List of 8000 Basic Words* (JACET 8000). The test was a multiple-choice test with 100 items at each level, an appropriate number to test receptive knowledge (Schmitt, 2000, p. 166). The mean estimated vocabulary size (EVS) for first-year students was 1,458.92 words while that for third-year students was 2,826.95. However, words at the 5,000- to 6,000-word levels were not measured, so Yashima (2005) estimated the vocabulary size of Japanese EFL learners, using a revised VET at 1,000- to 6,000-word levels taken from JACET 8000. The mean EVS for first-year students was 2,025.23 words while those of third-year students was 3,630.32. The study also showed that there were four words difficult even for sophomores to learn at the 1,000-word level, while there were four words whose success rates for first-years were over 70 percent at the 6,000-word level. This showed a need to reclassify all the target words and distractors into six word levels, according to students' success rates in answering. Yashima (2006, 2007b) examined the success rate of all these words. According to the results, Yashima (2007a) re-classified 1,800 words used in the VET at the 1,000- to 6,000-word levels into six word levels: elementary 1,000-, intermediate 1,000-, intermediate 2,000-, advanced 1,000-, advanced 2,000-, and advanced 3,000-word levels (E1,000, I1,000, I2,000, A1,000, A2,000, and A3,000 respectively) and examined participants' vocabulary. The results seemed to show that there were some words at the E1,000-word level difficult even for third-year students to acquire. The results show the necessity of re-defining basic words for students according to what they have studied, that is, authorized textbooks. Therefore, this paper attempts to identify the 1,000 basic words students need to acquire mainly based on their presentation in authorized textbooks for junior or senior high school students as well as frequency.

The study

Purpose

The study attempts to select the 1,000 basic words essential for Japanese EFL learners to acquire. In order to do that, I examine (1) words which do or do not appear in authorized textbooks for junior or senior high school students, and (2) how closely these words are connected with word frequency. I also examine words difficult for first- to third-year students to acquire among the 263 words around the 1,000-word level.

Hypotheses

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

- (1) There are some high-frequency words which do not appear in authorized textbooks for junior or senior high school students.
- (2) There are some low-frequency words which are essential for Japanese EFL learners to acquire.
- (3) There are some words which are difficult for the learners to acquire within the 1,000-word level while it is easy for them to acquire some words outside the 1,000-word level.

Participants

The participants in this study were 151 first-year students, 153 second-year students, and 218 third-year students from one Tokyo metropolitan senior high school. They had been learning English for 3 to 5 years.

Materials

As I would like the 1,000 basic words to have as wide coverage as possible, I chose materials many students use in school. Then, I chose to examine the most probably used authorized textbooks and dictionaries. Authorized textbooks for junior high school students were also examined because more than half of these basic words had already been learned in junior high schools. In the case of authorized textbooks for junior high school students (ATJ), books 1 to 3 of three ATJs were examined respectively because these three textbooks in all were used in more than two-thirds of all the junior high schools in Japan. Compulsory words designated in the two Courses of Study for junior high school students in the Heisei Period were also examined from the point of basic words. In the case of authorized textbooks for senior high school students (ATS), there are three levels: elementary (E), intermediate (I), and advanced (A). Three textbooks were examined at each level because all of them were used by more than thirty thousand students (See References).

Five probably known English-Japanese dictionaries were also examined. Two of them were for senior high school students: FAV3 and LH5 and the other three were for senior high school and university students: WIS2, G4, and OLEX. In relation to dictionaries, defining words in these two dictionaries were also examined: GBED and LDOCE because all the words were defined with 850 basic words in the former and all the definitions in the latter were written in the Longman Defining Vocabulary

around 2,000 common words (See References).

Word frequency was also examined because all the words that appear in authorized textbooks for junior or senior high school students were not high-frequency. Then, LEJD was examined because 3,000 most frequently used words in spoken and written English (LS and LW) in the English-speaking context. In addition, two word lists were also examined because they were compiled in terms of Japanese EFL learners as well as frequency: JACET 8000 and HUEVL (See References).

Lastly, participants were tested on the 263 words around the 1,000-word level in order to examine whether the 990th word in rank for example, is really more difficult for Japanese EFL students to acquire than the 1,010th word. The receptive and productive knowledge of these words were examined on three scales: as for the receptive knowledge, 1 meant that they didn't know the meaning of the word, 2 meant that they didn't know the meaning exactly, and 3 meant that they knew the meaning exactly; as for the productive knowledge, 1 meant that they didn't know the spelling of the word, 2 meant that they didn't know the spelling exactly but could pronounce it, and 3 meant that they knew the correct spelling.

Procedure

I changed all the data into the numerical reading according to LDOCE and LEJD: 1, 2, and 3 showed that a word was one of the thousand most frequently used words, one of the next most frequent thousand, and one of the third thousand in the list of three thousand words most often used in English-speaking context. 4 showed that a word was outside the first 3,000-word levels in these two dictionaries. 5 showed that a word did not appear in them. If a word appeared in an authorized textbook for junior high school students regardless of its grade, it was given 1, and if not, it was given 1.9. If a word appeared for the first time in an authorized textbook for first-year senior high school students, it was given 2, and if not, it was given 4.9. As there appeared about 1,000 words in authorized textbooks from Books 1 to 3 for junior high school students, these words were given 1. The words which didn't appear in any of them were given 1.9 because some of them were regarded as the words which had already been learned in junior high school in the list of authorized textbooks for senior high school students and I would like to distinguish between these words and new words in authorized textbooks for senior high school students. Moreover, if a word appeared in the compulsory words designated in the two Courses of Study for junior high school students in the Heisei period (CS89 and CS99), it was given 1 respectively, and

if not, it was given 2 respectively.

If a word was marked as one of the most important words in a dictionary, it was given 1. If a word was marked as one of the second most important words, it was given 2. As the most important words in all the five dictionaries above varied from about 900 to about 1,250, the second important words could be referred to as the next thousand. However, as the number of the second most important words was about 1,800 or more in all the dictionaries except LH5, there were no words given 3 in these dictionaries except LH5. If a word was marked as one of the third most important in all the dictionaries except LH5, it was given 4. All the rest of the words appeared in these dictionaries were given 4 and the words which didn't appear in any of them were given 5. In addition, if a word appeared in the list of defining words in GBED and LDOCE, it was given 1, and if not, it was given 2 in the case of GBED and 3 in the case of LDOCE according to the number of the defining words in the two dictionaries.

As for word lists, in JACET 8,000, 1, 2, or 3 showed that a word was one of the 1,000-, 2,000-, or 3,000-word levels and 4 showed that a word was one of the 4,000- to 8,000-word levels. Moreover, in HUEVL, 1 or 2 showed that a word was one of the first- and second-word levels. However, as the second-word level has 1,778 words, a word which appeared in the third- to fifth-word levels was given 4. The words which didn't appear in either of the two word lists were given 5.

Lastly, participants were given the test of 263 words on both receptive and productive knowledge. They answered as many questions as possible. The frequency of the appearance of the answers 1 and 2 was examined according to and across school grades on receptive and productive knowledge respectively so that high-frequency words would be omitted even though they were within the 1,000-word level, because they were actually difficult for Japanese EFL students to acquire.

Scoring

All the words examined were given the numerical reading in terms of the above 25 items respectively and the average was calculated. Then they were arranged in order of small to large in the numerical reading: 1.00 to 3.91.

Supplementary vocabulary

When they were arranged in order of small to large, some of the words in the small numerical reading were derivatives. As I would like to select as many different words as possible, all the derivatives of the 1,000 basic words selected were put into

another vocabulary list, Supplementary Vocabulary (SV). The list had not only these derivatives but also proper nouns, informal words, abbreviations, and numbers which appeared in textbooks for both junior and senior high school students. It contained 300 words.

Conditions in selecting 1,000 basic words

In order to make a basic, essential word list for Japanese EFL learners, I, in principle, selected words which appeared in both ATJs and AJSs or which appeared in at least one AJS (I) and one AJS (A) respectively in the case of not appearing in ATJs. If a word appeared in an ATJ but did not in any AJSs, it might have appeared by chance in connection with topics though it was not a high-frequency word. Moreover, if a word appeared only in AJS (A), it might be a little difficult for most of them to function. However, words which regarded as important or necessary in activities in English classes were selected even though their numerical reading was a little high. Words which were the other of a pair and a set were also Selected even though their figure was a little high.

Classification of 1,000 basic words

I classified the 1,000 basic words into 14 categories: general things, animals, classroom English, colors, family, feelings, food and drink, occupations, parts of a body, seasons and time, shopping, sports, subjects, and travel (See Appendix). That's because I would like teachers as well as students to use them according to the new Course of Study.

Results

High-frequency words which do or do not appear in authorized textbooks for junior or senior high school students

Table 1 shows the means and standard deviations in authorized textbooks, word lists, frequency, and English-Japanese dictionaries of 1,000 basic words. There is little difference across eight items except ATs, because words which did not appear in an ATs were given 4.9 and there were 245 words which did not appear in five or more out of nine ATs. The figures of almost all the items are less than 1.50, which means the 1,000 basic words are essential and important ones as well as high-frequency ones. Table 2 shows the distribution of total means of these 1,000 basic words. More than 600 words are 1.54 or lower and 942 words are 2.04 or lower. This is also a proof

that the 1,000 basic words are essential and important ones as well as high-frequency ones. The other 58 words are those that were supposed to be used by teachers as well as students in English classes. Table 3 shows the distribution of appearance of the 1,000 basic words in 3 ATJs. More than 40 percent of the 1,000 basic words appear in all the three ATJs. Less than 20 percent of them do not appear in any of these three ATJs. There are 34 words from 1.05 to 1.54, which do not appear in any of these three ATJs. These words are those which appear or are regarded as having already been learned in junior high school in from two, such as *president*, to seven, such as *draw* and *north*, out of these nine ATJs and most of them more than three.

Table 1

Means and standard deviations in authorized textbooks, word lists, frequency, and English-Japanese dictionaries of 1,000 basic words

	<i>ATJ</i>	<i>ATS</i>	<i>JACET</i>	<i>HUEVL</i>	<i>LS</i>	<i>LW</i>	<i>EJD</i>	<i>Total</i>
<i>Mean</i>	1.30	2.27	1.33	1.40	1.49	1.54	1.31	1.47
<i>SD</i>	.35	1.19	.53	.62	.76	.81	.41	.37

Table 2

Distribution of total means of 1,000 basic words

<i>~1.04</i>	<i>~1.54</i>	<i>~2.04</i>	<i>~2.54</i>	<i>~3.04</i>	<i>~3.54</i>	<i>~4.04</i>	<i>Total</i>
38	567	337	46	10	2	0	1,000

Table 3

Distribution of appearance of 1,000 basic words in 3 ATJs

	<i>~1.04</i>	<i>~1.54</i>	<i>~2.04</i>	<i>~2.54</i>	<i>~3.04</i>	<i>~3.54</i>	<i>~4.04</i>	<i>Total</i>
3 ATJs	37	355	37	12	3	0	0	444
2 ATJs	1	94	51	8	3	1	0	158
1 ATJ	0	84	114	19	2	1	0	220
0	0	34	135	7	2	0	0	178
Total	38	567	337	46	10	2	0	1,000

Tables 4-6 show the distribution of appearance of the 1,000 basic words in 3 ATJs(A)s, 3 ATJs(I), and 3 ATJs(E) including words which were regarded as having already been learned in junior high school respectively. Table 4 shows that more than

70 percent of the 1,000 basic words appear or are regarded as having already been learned in junior high school in all the three ATS(A)s. There are only 31 words which do not appear or are not regarded as such in any of these three ATS(A)s, more than 80 percent of which are from 1.55 to 2.04. Many of them are supposed to be used in English classes, such as *mark* and *print* or they are also words which are the other pair, such as *sugar* as opposed to *salt*. On the other hand, Tables 5 and 6 show that only half of these basic words appear or are regarded as having already been learned in junior high school in all the three ATS(I)s and ATS(E)s. Moreover, about 27 percent of these 1,000 basic words appear or are regarded as having already been learned in junior high school in these two ATS(I)s and there are only 58 words which do not appear or are not regarded as such in any of these three ATS(I)s. Eight words of them are those from 1.05 to 1.54, all of which appear or are regarded as having already been learned in junior high school in these two or three of ATS(A), such as *experience* or *power*. Only about 15 percent of these 1,000 basic words appear or are regarded as having already been learned in junior high school in the two ATS(E)s and there are as many as 252 words which do not appear or are not regarded as such in any of these three ATS(E)s. 56 words of them are those from 1.05 to 1.54, most of which appear or are regarded as having already been learned in junior high school in these two or three of ATS(A)s, or ATS(I)s, such as *list* or *seem*, *business* or *rise*. The distribution in Table 6 is similar to that in Table 3, which means that ATS(E)s are similar to ATJs in vocabulary in terms of these 1,000 basic words.

Table 4

Distribution of appearance of 1,000 basic words in 3 ATS(A)s, including words which were regarded as having already been learned in junior high school

	~1.04	~1.54	~2.04	~2.54	~3.04	~3.54	~4.04	Total
3 ATSs	38	524	142	12	3	2	0	721
2 ATSs	0	38	97	19	5	0	0	159
1 ATS	0	5	74	9	1	0	0	89
0	0	0	24	6	1	0	0	31
Total	38	567	337	46	10	2	0	1,000

Table 5

Distribution of appearance of 1,000 basic words in 3 ATS(I)s, including words which were regarded as having already been learned in junior high school

	~1.04	~1.54	~2.04	~2.54	~3.04	~3.54	~4.04	Total
3 ATSs	38	400	54	6	1	1	0	500
2 ATSs	0	134	118	12	3	1	0	268
1 ATS	0	27	121	21	5	0	0	174
0	0	6	44	7	1	0	0	58
Total	38	567	337	46	10	2	0	1,000

Table 6

Distribution of appearance of 1,000 basic words in 3 ATS(E)s, including words which were regarded as having already been learned in junior high school

	~1.04	~1.54	~2.04	~2.54	~3.04	~3.54	~4.04	Total
3 ATSs	35	364	32	3	2	0	0	436
2 ATSs	3	91	46	9	2	0	0	151
1 ATS	0	56	90	12	2	1	0	161
0	0	56	169	22	4	1	0	252
Total	38	567	337	46	10	2	0	1,000

Tables 7 and 8 show the distribution of appearance of high-frequency words from 1.05 to 2.04 in these 3 ATJs and in these 9 ATSs, except for 1,000 basic words respectively. As Table 7 shows, there are only six words from 1.05 to 1.54, which do not appear in the 3 ATJs, except for these 1,000 basic words, such as *condition* or *general*. As Table 8 shows, these words are those which appear or are regarded as having already been learned in junior high school in one or two ATS(A)s but in none of these ATS(I)s or ATS(E)s. Moreover, there are 224 words from 1.55 to 2.04 except for these 1,000 basic words, more than 80 percent of which do not appear in the 3 ATJs, such as *comfortable* or *ocean*. However, there are only 42 words of them, which do not appear or are not regarded as such in any of these nine ATSs, such as *attack* and *manage*.

Table 7

Distribution of appearance of high-frequency words from 1.05 to 2.04 in 3 ATJs, except for 1,000 basic words

	~1.54	~2.04	Total
3 ATJs	0(25)	0(34)	0(59)
2 ATJs	0(2)	2(5)	2(7)
1 ATJ	0(1)	34(2)	34(3)
0	6(2)	188(12)	194(14)
Total	6(30)	224(53)	230(83)

Table 8

Distribution of appearance of high-frequency words from 1.05 to 2.04 in 9 ATs, except for 1,000 basic words, including words which were regarded as having already been learned in junior high school

	~1.54	~2.04	Total
3 ATs or more	0(30)	50(50)	50(80)
2 ATs	3(0)	49(3)	52(3)
1 ATs	3(0)	83(0)	86(0)
0	0(0)	42(0)	42(0)
Total	6(30)	224(53)	230(83)

Note. The number in parentheses shows the number of SV.

Low-frequency words which are essential for Japanese EFL learners to acquire

As Table 2 shows, there are 58 words from 2.05 to 3.54, 12 words of which are from 2.55 to 3.54. These words are not high-frequency, but they are those that are supposed to be used by teachers as well as students in English classes, such as *classmate* and *textbook*. As Table 3 shows, there are nine words from 2.05 to 3.54, which do not appear in any ATJs. These words are classified into two groups: words which are supposed to be used in English classes, such as *dictionary* or *hobby*, and words which are the other pair, such as *senior* or *succeed* as opposed to *junior* or *fail* respectively.

Tables 4 and 5 show that there are only seven or eight words from 2.05 to 3.54, which do not appear or are not regarded as having already been learned in junior high school in any of these three ATS(A)s or ATS(I)s respectively. Most of them are supposed to be used in English classes, such as *interview* and *uniform* in these three ATS(A)s and *autumn* and *site* in ATS(I)s. On the other hand, Table 6 shows that there are 27 words from 2.05 to 3.54, which do not appear or are not regarded as such in any of these three ATS(E)s. Most of them are also supposed to be used in English classes, such as *grade* and *pardon*. Moreover, there are two words from 2.05 to 3.54, which do not appear or are not regarded as such in any of these nine ATs, that is, *absent* and *net*, which are also supposed to be used in English classes.

Degree of receptive and productive acquisition of 263 words that are around 1,000 in rank

Tables 9 and 10 show the means and standard deviations in appearance rates of 1 and 2 according to school grades on receptive and productive knowledge respectively. While first-year students acquire less than half of these 263 words on both knowledge, third-year students acquire more than 80 percent of them. Moreover, students acquire these words more receptively than productively regardless of school grades.

Table 9

Means and standard deviations in appearance rates of 1 and 2 according to school grades (Receptive)

		1	2	Total
1 st -years	Mean	38.38	16.98	55.37
n=151	SD	26.46	6.75	25.39
2 nd -years	Mean	15.39	10.56	25.95
n=153	SD	16.84	6.03	20.75
3 rd -years	Mean	3.30	11.13	14.44
n=218	SD	3.16	3.21	5.74
Total	Mean	19.03	12.89	31.92
n=522	SD	14.09	3.87	15.73

Table 10

Means and standard deviations in appearance rates of 1 and 2 according to school grades (Productive)

		1	2	Total
1 st -years	Mean	44.85	15.75	60.60
n=151	SD	26.40	6.92	23.61
2 nd -years	Mean	20.70	9.06	29.76
n=153	SD	18.74	4.49	20.45
3 rd -years	Mean	6.39	11.80	18.18
n=218	SD	4.04	3.00	6.33
Total	Mean	23.98	12.20	36.18
n=522	SD	14.97	3.40	15.19

Table 11

Distribution of appearance rate of 1 according to school grades (Receptive)

Mean	0	10	20	30	40	50	60	70	80	90	100	Total
1 st -years	41	58	30	20	20	28	24	17	22	3	0	38.38
2 nd -years	141	52	27	16	10	8	4	5	0	0	0	15.39
3 rd -years	251	11	1	0	0	0	0	0	0	0	0	3.30
Total	98	49	65	24	18	9	0	0	0	0	0	19.03

Tables 11 to 13 show the distribution of appearance rate of 1 or 2 or the sum of 1 and 2 according to school grades. While there are only 18 words whose appearance rates of the sum of 1 and 2 is less than 20 percent for first-years, there are more than 80 percent of all the words tested for the third-years. There are two words whose

appearance rate of the sum of 1 and 2 is 70 percent or more for the total participants. These two words, *bill* and *indeed*, are those from 1.55 to 2.04, but they do not appear in any ATJs and appear only in one ATS(I) and three ATS(A)s respectively, which caused them to be excluded from the 1,000 basic words. Moreover, there are 40 words whose appearance rate of the sum of 1 and 2 is 50 percent or more. Although some of them, such as *conversation* and *certainly*, meet the condition of being included in these 1,000 basic words, 50 percent or more of the participants do not acquire them. That's why these words are excluded. On the other hand, although such words as *chocolate* and *driver* are outside 1,000 in rank, these two words are less than 10 percent in appearance rates of the sum of 1 and 2 for the total participants, and they are also essential for them. Therefore, they are included in these basic 1,000 words.

Table 12

Distribution of appearance rate of 2 according to school grades (Receptive)

<i>Mean</i>	<i>0</i>	<i>10</i>	<i>20</i>	<i>30</i>	<i>40</i>	<i>50</i>	<i>60</i>	<i>70</i>	<i>80</i>	<i>90</i>	<i>100</i>	<i>Total</i>
1 st -years	42	141	72	7	1	0	0	0	0	0	0	16.98
2 nd -years	128	117	17	1	0	0	0	0	0	0	0	10.56
3 rd -years	114	143	6	0	0	0	0	0	0	0	0	11.13
Total	60	189	14	0	0	0	0	0	0	0	0	12.89

Table 13

Distribution of appearance rate of the sum of 1 and 2 according to school grades (Receptive)

<i>Mean</i>	<i>0</i>	<i>10</i>	<i>20</i>	<i>30</i>	<i>40</i>	<i>50</i>	<i>60</i>	<i>70</i>	<i>80</i>	<i>90</i>	<i>100</i>	<i>Total</i>
1 st -years	0	18	40	37	25	27	25	30	33	28	0	55.37
2 nd -years	71	61	46	32	10	13	18	7	4	1	0	25.95
3 rd -years	41	187	29	6	0	0	0	0	0	0	0	4.44
Total	7	70	55	56	35	26	12	2	0	0	0	31.92

Discussion

Table 3 shows that there are 34 words from 1.05 to 1.54, which do not appear in any ATJs. Words from 1.05 to 1.54 mean that they are high-frequency ones, such as *draw* and *step*. On the contrary, there are two words from 3.05 to 3.54, which appear

in one or two ATJs. Words from 3.05 to 3.54 mean that they are not so high-frequency ones, that is, *classmate* and *snowy*. These two words are definitely used in English classes by teachers as well as students, while they are not frequently used in an everyday situation. The findings show that it is not safe that we regard all the words that appear in ATJs as easy or high-frequency ones. It is true of ATSs. Although there are no words from 1.05 to 1.54, which do not appear or are not regarded as having already been learned in junior high school in any ATS(A)s, there are six or 56 words from 1.05 to 1.54, which do not appear or are not regarded as such in any ATS(I)s or ATS(E)s, such as *set* and *wide* or *business* and *report* respectively. *Experience*, *matter*, *price* are those which do not appear or are not regarded as such in any ATS(I)s and ATS(E)s in common. Contrary to that, *classmate* and *snowy* appear seven and six out of these nine ATSs respectively. Moreover, as Tables 7 and 8 show, there are six words from 1.05 to 1.54, which do not appear in any ATJs, except for these 1,000 basic words, that is, *condition*, *general*, *mile*, *rather*, *sense*, and *sort*. Three of them appear or are regarded as such in two out of three ATS(A), and the other three appear only in one ATS(A)s. These six words do not appear in any of these ATS(I)s and ATS(E)s. It can be safe to say that there are some high-frequency words which do not appear in ATJs or ATSs and that there are also some low-frequency words which appear in ATJs and ATS(E) from the basis of this study, especially Tables 3 and 6. My first hypothesis could be accepted.

Table 2 shows that there are 58 words from 2.05 to 3.54, 12 words of which are from 2.55 to 3.54. These words are lower-frequency, but they are those that are supposed to be used by teachers as well as students in English classes. Students can use *excited*, *surprising*, and *surprised* as an example of their impressions and *email*, *hobby*, and *ski* as their introductions. Teachers can use *absent*, *homework*, and *textbook* as their instructions and *cloudy*, *rainy*, and *snowy* as their statements of the weather. These 58 words are classified into three groups: words which are supposed to be used in English classes, such as *grade* or *notebook*, words which are the other pair, such as *appear* or *succeed* as opposed to *disappear* or *fail* respectively, and words which are the other set or sets, such as *autumn* and *snowy*. While *spring*, *summer*, and *winter* are within the 1,000-word level, *autumn* is 2.24. If *autumn* is included, four seasons are completed, though *fall* is within the 1,000-word level. Moreover, as for the weather, only *fine* is within the 1,000-word level, and the others, that is, *cloudy*, *rainy*, and *snowy* are outside the 1,000-word level. As these four words are one set, the other three are included in the 1,000 basic words, though *snowy* is 3.3. It can be said that there are

some low-frequency words which are essential for students to acquire. Therefore, the second hypothesis could be accepted.

Tables 9 and 10 show that while first-year students acquire less than half of the 263 words on both knowledge, third-year students acquire more than 80 percent of them, as mentioned in 3.3. Moreover, Table 13 shows that while there are only 18 words whose appearance rates of the sum of 1 and 2 is less than 20 percent for first-years, there are 228 words for third-years. The findings show that there are more than 30 words even third-year students do not acquire quite well receptively though they are around the 1,000-word level. As Table 13 shows, there are 40 words whose appearance rates of the sum of 1 and 2 is 50 percent or more for the total participants, 15 words of which are from 1.55 to 2.04, such as *above* and *below*. As 50 percent or more of the participants do not acquire them, they are excluded from the 1,000 basic words though some of them meet the conditions of being included in the 1,000 basic words, such as, *ahead* and *certain*. On the contrary, as Table 2 shows, there are 46 words from 2.05 to 2.55, which are outside the 1,000-word level. As mentioned in 3.1, these words are those which are supposed to be used by teachers as well as students in English classes. In addition to that, as Table 13 shows, there are 77 words whose appearance rates of the sum of 1 and 2 is less than 20 percent for the total participants, and 7 words of them are those which are from 2.05 to 2.55 but are included in the 1,000 basic words, such as *actor* and *tape*. These words are included in the 1,000 basic words because more than 80 percent of all the participants acquire them receptively. Therefore, my third hypothesis could be accepted.

Conclusion and implications

This study shows that even in the 1,000 basic words, there are 34 words from 1.05 to 1.54, which do not appear in any of these three ATJs. These words are high-frequency, such as *low* and *pay*. Moreover, there are also 178 words at the same range of means in it, which appear in one or two of these three ATJs. Some of them are *dark* and *warm*. As these words are basic, important words as well as high-frequency ones, they should be learned in junior high school. However, many of junior high school students will be high school students without seeing at all. On the other hand, the study also shows that there are three words from 2.55 to 3.04 in it, which appear in all these three ATJs. These words were *cloudy*, *e(-)mail*, and *homework*. They are lower-frequency, but they are definitely supposed to be used in English classes in junior high school. It can be safe to say that that is the reason why they appear in all

the ATJs. There are also some words at the same range of means even in it, such as *judge* and *print*, which appear in one of these three ATJs, though they do not appear in any of these nine ATJs. They are also definitely supposed to be used in English classes in junior high school. The results indicate that all the words that appear in an ATJ are not basic, important, and high-frequency, that there are some words which do not appear in it though they are such high-frequency words, and that some words appear in ATJs though they do not appear in any ATJs. These findings suggest educators in junior high school should recognize they should devise ways to have their students acquire words which do not appear in their ATJs but are basic and important. They also suggest that educators in high school should recognize the situations in junior high school and should not overestimate their students' vocabulary and that they should also devise ways to have their students acquire such words.

This study also shows that in the 1,000 basic words, there are some words from 2.55 to 3.54, such as *absent* and *textbook*, which are supposed to be used by teachers as well as students in English classes as mentioned above. These words are lower-frequency because they do not so often in daily lives. It also shows that if basic words are selected only according to word frequency, the other pair or set(s) of a word may be excluded from these words. While *fine*, whose means is 1.19, appears in all these three ATJs and it is also regarded as having already been learned in junior high school in all these nine ATJs, the other words concerning weather, *cloudy*, *rainy*, and *snowy*, are from 2.64 to 3.30, and appear in two of these three ATJs but do not appear in three or four ATJs. If only words less than 2.05 are selected, these words will be omitted. However, although they are not so high-frequency words, they are regarded as having already been learned in junior high school in two-thirds of ATJs. It may be because the authors of ATJs regarded them as important for students. Therefore, these findings suggest that educators should recognize the fact that the other pair or set(s) of a basic, important word may not appear in their ATJs or ATJs and that they should teach them as a pair or set(s) in order to have their students express themselves.

The findings on students' receptive and productive acquisition of words which are around the 1,000-word level indicate that while first-year students acquire less than half of these words, third-year students acquire more than 80 percent of them. The results indicate that even third-year students have words they do not completely acquire, though they are around the 1,000-word level. The results also show that there are some words which are difficult for students to acquire such as *above* and

below, though they are within the 1,000-word level, while there are some words which are easy for students to acquire, such as *chocolate* and *tape*, though they are outside the 1,000-word level. Therefore, these findings suggest that educators should recognize this gap and should encourage their students to read and listen to as many additional materials as possible to have them acquire these 1,000 words. They also suggest that educators should judge the difficulty for their students to acquire on the basis of word frequency in the context of teaching as well as in daily lives.

Notes

1. This paper is based on the contents of a presentation titled “The Selection of 1,000 Basic Words Necessary for Japanese EFL Learners” given at the 36th JASELE annual convention in Osaka. 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

- All Aboard! English I* (AA) (E). (2008). Tokyo Shoseki.
- BIG DIPPER English Course I* (BD) (I). (2009). Suken Shuppan.
- Course of Study* (CS89). (1989). Ministry of Education, Science and Culture.
- Course of Study* (CS99). (1999). Ministry of Education, Science and Culture.
- CROWN English Series I* (CR) New Edition (A). (2008). Sanseido Publishing.
- FAVORITE English-Japanese Dictionary* (FAV3) (3rd ed.). (2005). Tokyo Shoseki.
- GENIUS English-Japanese Dictionary* (G4) (4th ed.). (2006). Taishukan Publishing.
- Hokkaido University English Vocabulary List* (HUEVL). (1999). Hokkaido University.
- JACET Basic Words Revision Committee. (Ed.). (2003). *JACET List of 8000 Basic Words* (JACET 8000). Japan Association of College English Teachers.
- LIGHTHOUSE English-Japanese Dictionary* (LH5) (5th ed.). (2007). Kenkyusha.
- Longman Dictionary of Contemporary English* (LDOCE) (3rd ed.). (1995). Addison Wesley Longman Ltd.
- Longman English-Japanese Dictionary* (LEJD). (2006). Pearson Kirihara.
- NEW CROWN ENGLISH SERIES* (NC) New Edition. (2006). Sanseido Publishing.
- NEW HORIZON English Course* (NH). (2006). Tokyo Shoseki.
- OLEX English-Japanese Dictionary* (OLEX). (2008). Obunsha.
- Power On English I* (PO) (I). (2008). Tokyo Shoseki.

- PRO-VISION ENGLISH COURSE I* (PR) New Edition (A). (2008). Kirihara Shoten.
- 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.
- SUNSHINE ENGLISH COURSE* (SS). (2006). Kairyudo Publishing.
- The General Basic English Dictionary* (GBED). (1960). Hokuseido Press.
- Yashima, H. (2004, August 14–15). *The development of a vocabulary size estimation test for Japanese EFL learners*. Paper presented at the 30th JASELE annual convention in Nagano.
- Yashima, H. (2005, August 6–7). *Identifying words difficult for Japanese EFL learners to acquire by a revised vocabulary estimation test*. Paper presented at the 31st JASELE annual convention in Hokkaido.
- Yashima, H. (2006, August 5–6). *Words difficult for Japanese EFL learners to acquire*. Paper presented at the 32nd JASELE annual convention in Kochi.
- Yashima, H. (2007a, August 4–5). *The development of a vocabulary achievement test for Japanese EFL learners*. Paper presented at the 33rd JASELE annual convention in Oita.
- Yashima, H. (2007b, August 17–18). *Words easy or difficult for Japanese EFL learners to acquire*. Paper presented at the 31st KATE annual convention in Chiba.
- Yashima, H. (2008, August 5–6). *The influence of basic words on Japanese EFL learners' estimated vocabulary size*. Paper presented at the 34th JASELE annual convention in Tokyo.
- UNICORN ENGLISH COURSE I* (UN) (A). (2008). Bun-eido Publishing.
- VISTA English Series I* (VI) New Edition (E). (2008). Sanseido Publishing.
- Viva English! I* (Vi) NEW EDITION. (2008) (E). Daiichi Gakushusha.
- Vivid English Course I* (VV) NEW EDITION (I). (2008). Daiichi Gakushusha.
- Wisdom English-Japanese Dictionary* (WIS2) (2nd ed.). (2007). Sanseido Publishing.

Appendix

1,000 Basic Words (829 out of 1,000 words) (* shows that the word is from 2.05 to 3.91)

1. General things (570 words)

a, the, it, one; ability; able; about; accept; accident; across, cross; act; action; activity; add; address; adult, child; advice; again; against; age; agree; air; alive; all, almost, half, nothing; allow; along; already still, yet; also, too; although; always, ever, never, often,

sometimes; among, between; amount; and, but; another, else, other; answer, ask; any, some; anyone, everybody, everyone, nobody, somebody, someone; anything, everything, something; anywhere, everywhere, somewhere; appear, *disappear; area; around; arrive, leave, reach, start; as; at; away; back; bad, good, nice, terrible, well, wonderful; bank; basic; be, exist; beach; sea; beat; beautiful; because; become; bed; begin, end; beginning; behind, front; believe; belong; big, chief, main, small; bit; block; blow; board; boat; book; borrow, lend; both; either; boy, girl; break, destroy, keep; bridge; bring, take; build, building; burn; business; busy; by, until; call, telephone; can, may, must, shall, should, will, would; care, careful, carefully; carry; case; catch; cause, effect; center, middle; chair, desk, table; character; chance, opportunity; change; check, judge, *grade, level, mark; chief; choose; church; city, town; clean; clear; climb; clock, watch; close, open, shut; *cloudy, fine, *rainy, *snowy, cold, cool, hot, warm, cloud, rain, snow, wind, blow, sun, *umbrella, weather; club; collect; college, university; come, go; common; communication; community; complete; computer, *net, *site; concert sing, song; continue; control; copy; corner; correct, exactly, perfect, wrong; country, foreign; couple; course, road, street; cover; create; culture; cut; dance; danger, dangerous, safe; decide; decision; deep; degree; depend; describe; design; develop; die, dead, death, life, live; difference, different, same, similar; difficult, easy, easily, hard, simple; direction, distance, left, right; discover; disease, fine, health, healthy, hospital, ill, sick, discussion; do; door; down, up; draw; dream; drive; drop, rise; dry, wet; during, while; each, every; earth, environment, ground, moon, world; education; effort; *e(-)mail; encourage; energy; enough, full; enter; event; even; example, form, model, style, type, way; except; excuse; exercise, *practice; expect; experience; explain; fact, real, really, true, truth; fail, mistake, pass, *succeed, success,; famous; *fan; far, near; few, half, little, no; some; *festival; final, finally; first; finish, stop; floor; follow; forest, tree, wood(s); forget, remember; friend; future; garden, yard, *gate; get; give; goal; good(-)by(e); ground; group; grow; guess, suppose, think, wonder; sure; guide; have, let, make; hard, soft; have, hold; hear, listen; he, she, they; hello; help; high, school, *junior, *senior, low; *hobby; home, house; hope, want, wish; hurry; how, what, when, where, whether, which, who, why; I, we, you; idea, thought; imagine; important; in, out, within; information; inside, outside; international, man, people, person, personal, woman; nation, national; introduce; join; key; kick; kind; king, *queen; knock; know, knowledge; language, word, sign; large, small; laugh, smile; leader; learn, study; library; light; listen, say, speak, talk, tell, speech; long, short; look, meet, see; lose, win; lot, many, much; loud, quiet, silent, voice; matter;

mean, meaning; memory; message; mix; mountain, river; move; movie, theater; name; necessary; need; new, old, young; no, yes; object; office; on, off; once, *twice; original, over, under; page; *pair; paper; part; party; pen, pencil; period; *piano; picture; plan; play, please; point; poor, rich; possible; *post; print; problem, question; program; project; promise; quick, quickly, slow, slowly; radio, television; raise; read, write; ready; reason; record; repeat; report; research; rest; result; return; ring; *roll; rule; run, walk; *screen; secret; shake; share; show; side; sit; size; sleep, wake; speed; stage; stand; state; step; story; strong, weak; student, study, subject, teacher; system; team; test; thank; to; top; trouble; *uniform; very; video; village; *volunteer; wait; walk; welcome; window; with, without; work

2. Animals (7 words)

animal; bird; cat, dog, horse, *pet; *rabbit

3. Classroom English (13 words)

*absent, present; class, lesson, *classmate; classroom; *dictionary; *homework; *interview; *notebook; *pardon; *tape; *textbook

4. Colors (11 words)

black; blue; bright, col(u)or; dark; brown; gold; green; red; white; yellow

5. Family (15 words)

aunt, uncle; baby; brother, sister; daughter, son; family; father, mother; *grandfather, *grandmother; husband, wife; parent

6. Feelings (40 words)

afraid; alone; angry; cry; dear; doubt; enjoy; *excited; *exciting; favo(u)rite; fear; feel; feeling(s); fun; funny; glad, great, happy, sad; God; hurt; interest; interested; interesting; like, love; lucky; natural; nice; pretty; serious; sorry; surprise; *surprised; *surprising; tear; tired; want, wish; worry

7. Food and drink (38 words)

apple; bottle; box; bread, rice; breakfast, dinner, lunch; cake; *chocolate; coffee; cook; cup; dish; drink, food; eat; egg; fruit, *vegetable; glass; hungry; ice; kitchen; meal; meat; milk; orange; plate; *salt, sugar; service; smell; sweet; taste; tea; wash; water

8. Occupations (14 words)

actor; artist; company; doctor, farm, farmer, *nurse; *driver; job; player; scientist;
*singer; *worker; *writer

9. Parts of a body (23 words)

arm; blood; body; brain; cell; ear; eye; face; finger; foot; hair; hand; head; heart; leg;
*lip; mind; mouth; neck; nose; shoulder; skin; tooth

10. Seasons and time (39 words)

*autumn; fall; spring; summer; winter; afternoon, evening, *noon; morning, night;
after, before, ago; today, tomorrow, tonight, yesterday; century; day, month, week,
year; holiday, vacation, weekend; birthday; date; hour, minute, moment; period, sea-
son; time; current, modern, past; early, late, soon

11. Shopping (24 words)

bag; buy; camera; *cap; card; charge; cheap, dollar, money; expensive; clothes; *coat;
dress, shirt; gift; hat; market; pay; sale; sell; shoe(s); shop; store; *supermarket

12. Sports (13 words)

ball; base; baseball; *basketball; football; game, race; shoot; *ski; *soccer; sport; swim;
tennis

13. Subjects (4 words)

art; history; music; science

14. Travel (18 words)

*abroad; airport, station; bus, train; camp; car; east, north, south, west; hotel; plane,
ship; *traffic; travel; trip; visit