Vocabulary Levels Analysis:
Survey Results with University Students

Kenji ISHIHARA, Tae OKADA, and Shimpei MATSUI

Key words: vocabulary, level, recognition, production, survey

Introduction

In the report on an earlier vocabulary survey with some 350 university students (Ishihara, Okada, and Matsui, 1999), it was made evident that three categories of English words showed distinctive score distributions in vocabulary recognition tests (Figure 1), namely, (1) 92.2% of the students correctly recognized more than 80% of the test words that appeared in all of
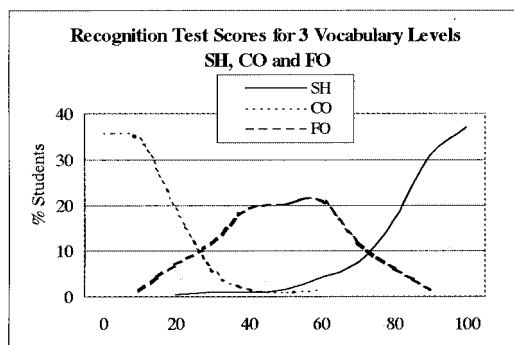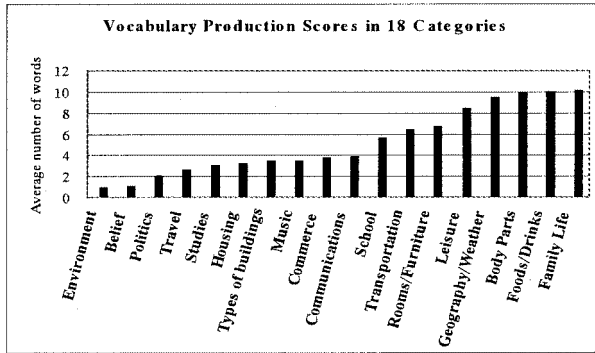
Figure 1

Figure 2

the ten major high school English textbooks (SH in Figure 1); (2) 60.9% correctly recognized 50-70% of the test words that appeared in some, but not all, of the textbooks (FO in Figure 1); (3) 70.5% correctly recognized less than 10% of the test words that appeared in none of the ten textbooks (CO in Figure 1). The three categories of recognition test words (SH, FO, and CO) were selected out of the 1500-item list compiled by Yoshioka and his colleagues (Yoshioka, 1997) which was designed for the first two years of English language instruction at the university level. Although it was assumed that SH, FO, and CO words were of increasingly higher levels in frequency scales, a more accurate measure of the students' word recognition abilities, preferably in reference to more commonly applied frequency lists, was considered to be of use for university instruction.

The same paper referred to above (Ishihara, Okada and Matsui, 1999) also reported that vocabulary production test results with the same group of some 350 students gave the authors the impression that production words in general were of lower levels than the words correctly recognized by the same participants. In a 5-minute vocabulary production test, the

participants produced approximately ten words in more personal, concrete semantic areas such as "family" and "food and drinks," while in more abstract semantic fields such as "politics" and "the environment," only one or two words were produced (Figure 2). A need was felt, therefore, for determining more accurately the levels, or discrepancies in levels, of university students' recognition and production vocabularies.

For the purpose of such levels analysis, three conditions are deemed essential. First, the recognition test words must be selected more directly on the basis of some widely recognized frequency list. In addition, the words produced by the participants in vocabulary production tests need to be analyzed as to their vocabulary levels. Finally, the recognition and production test results must be compared in reference to the scales of established vocabulary lists.

Two vocabulary lists used for levels analysis

To meet the first of the three conditions mentioned above, two word lists are used as points of reference in the current study: (1) JACET 4000 Basic Words (hereafter "JACET 4000") (JACET, 1993) as well as (2) Nation's First and Second Thousand Words (Nation, 1990) combined with Xue and Nation's University Word List (hereafter "Nation-Xue 3000") (Xue and Nation, 1984). Both JACET 4000 and Nation-Xue 3000 are largely based on frequencies of occurrence.

Table 1: Number of words listed in JACET 4000

Sub-lists	J1	J2	J3	J4	J5	Total
Number of words	531	508	972	871	1100	3982

JACET 4000 consists of 5 levels or sub-lists ranging from the most frequent and common, i. e., the lowest level (J1) with 531 words to the least

frequent, i. e., the highest level (J5) with 1100 vocabulary items. The specific numbers of words listed on the 5 sub-lists are as shown in Table 1, the total coming to 3982 words.

Nation-Xue 3000 consists of Nation's lists of First Thousand (N1) and Second Thousand (N2) words plus Xue and Nation's University Word List (hereafter UWL), the last of which is further divided into 11 sub-lists (X1, X2, etc.) according to frequencies of occurrence.

Table 2: Number of words listed in Nation-Xue 3000

Sub-lists	N1	N2	UWL	Total
Number of words	1008	956	799	2763

The numbers of words on these three lists are 1008, 956, and 799 respectively, as indicated in Table 2, with a total of 2763 words.

As a way of comparing JACET 4000 with Nation-Xue 3000, percentages of the JACET words included in Nation-Xue 3000 are shown in Table 3: 88.5% and 2.6% of J1 words, for instance, are included in Nation's First Thousand (N1) and Second Thousand (N2) words respectively; 65.4% and 14.8% of J2 words are likewise included in N1 and N2 respectively, etc. The percentage of JACET words that do not appear in Nation-Xue 3000 increases steadily from 8.5% for J1 to 45.6% for J5. In brief, the higher the levels, the greater the differences are between JACET 4000 and N-X 3000 in actual words listed. JACET 4000 claims to list "basic" words, and this

Table 3: Percentage of JACET 4000 words listed in Nation-Xue 3000

% words	J1	J2	J3	J4	J5	Not in J1-J5	Total
N1	88.5	65.4	35.1	18.9	13.9	2.6	36.7
N2	2.6	14.8	30.5	30.1	24.3	15.1	23.0
UWL	0.4	6.7	13.4	17.9	16.2	82.3	12.6
Not in NX	8.5	13.2	21.1	33.1	45.6	/	27.8
Total	100	100	100	100	100	100	100

claim is reflected statistically in the fact that 82.3% of UWL-words are not listed there, although the other 17.7% are scattered among J1-J5; notably 17.9% and 16.2% of J4- and J5- words respectively are the same vocabulary items as those in the University Word List.

Given such similarities and differences between the two lists, using both these lists as points of reference is judged preferable, to assure a more balanced view than when employing a single list.

Vocabulary recognition test for levels analysis

For the recognition test in this study, Nation's Levels Test (Nation, 1990) was adopted. A practical adjustment was necessary, however, to the limited amount of time allocated for the recognition test, so that Nation's original version (36 test words for each of the five vocabulary levels, i. e., 2000-, 3000-, 5000-, university- and 10,000-word levels) was divided into three equal parts with 12 test words for every level. The three shortened versions were given to three different groups of approximately 100 second-year university students each, all non-English majors, totaling 344 students (Cf. Appendix 1). In consequence, survey data were collected on all 180 test items of Nation's original version. The participants are fairly representative of the student body in that they are from all six faculties and are in eight different English classes to fulfill part of the foreign language requirement.

When the 36 test words of the 2000-word level are compared with the vocabulary items listed in Nation-Xue 3000 (Table 4), all but one word or 97.2% are included either in Nation's First or Second Thousand words; thus, the match between the test words and Nation-Xue 3000 is, for practical purposes, almost perfect. In the 3000- and 5000-word levels, on

Table 4: Levels Test words compared with Nation-Xue 3000 (Figures in %)

Word levels		2000	3000	5000	UWL	10,000	Total
Listed	N1	42	3	0	0	0	8.9
	N2	56	0	3	0	0	11.7
	UWL	0	19	19	97	6	28.3
Not listed		3	78	78	3	94	51.1
Subtotal of Listed		97.2	22.2	22.2	97.2	5.6	48.9
Total %		100.0	100.0	100.0	100.0	100.0	100.0

Table 5: Levels Test words compared with JACET 4000 (Figures in %)

Word levels		2000	3000	5000	UWL	10,000	Total
Listed	J1	0	3	0	0	0	0.6
	J2	19	6	0	0	0	5.0
	J3	50	17	3	0	0	13.9
	J4	25	44	11	11	0	18.3
	J5	6	8	28	25	0	13.3
Not listed		0.0	22.2	58.3	63.9	100.0	48.9
Subtotal of Listed		100.0	77.8	41.7	36.1	0.0	51.1
Total %		100.0	100.0	100.0	100.0	100.0	100.0

the other hand, 78% of the test words are not found in Nation-Xue 3000, which is not surprising, since Nation-Xue 3000 is not expected to list words on these levels. At the university-word level as in the case of the 2000-word level, all but one word (97.2%) are included in UWL.

A comparison of Nation's Levels Test words with JACET 4000 (Table 5) also reveals a relatively close match for the lowest levels or high-frequency categories of words: 100% for the 2000-word level, and 77.8% for the 3000-word level. As the levels go higher, however, the numbers of Levels Test words included in JACET 4000 decrease, until, in the highest 10,000-word level, none of the 36 Levels Test words appear, understandably, in JACET 4000. Although aiming at listing "basic" words, JACET 4000 lists 41.7% and 36.1% of Nation's test words at the 5000-word and UWL levels

respectively. Viewed differently, what is considered basic in JACET 4000 includes vocabulary items beyond the 4000-word level in terms of Nation's Levels Test.

This set of comparative data might well be useful in considering or interpreting the recognition test results reported below.

Recognition test results and levels analysis

The mean scores of the 344 survey participants (Appendix 1) for the five levels of the vocabulary recognition test are as shown in Table 6 and Figure 3. As seen in the mean scores, the average score is lowest in the highest or 10,000-word level with 23.7%, going increasingly higher toward the lowest

Table 6: Scores of the 344 students for the five levels of vocabulary recognition test

Vocabulary Levels	10,000	UWL	5000	3000	2000	Total
Full Scores	6	6	6	6	6	30
Mean Scores	1.42	2.31	2.83	4.22	5.02	15.80
Standard Deviations	1.14	1.47	1.60	1.47	1.13	4.16
Mean Scores in %	23.7	38.5	47.2	70.3	83.7	52.7
Correlation w/ Total	0.51	0.65	0.69	0.61	0.55	/

Figure 3

Levels Test Scores
(Means for 344 students)
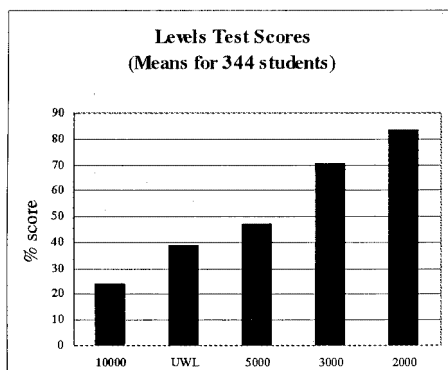
Table 7: Number of participants in 12 score brackets of recognition test

Ranks	Scores	Number of students	Percentage of students
1	5-6	5	1.5
2	7-8	8	2.3
3	9-10	24	7.0
4	11-12	36	10.5
5	13-14	62	18.0
6	14-16	62	18.0
7	17-18	64	18.6
8	19-20	38	11.0
9	21-22	27	7.8
10	23-24	11	3.2
11	25-26	5	1.5
12	27-28	2	0.6
	Total	344	100

Figure 4
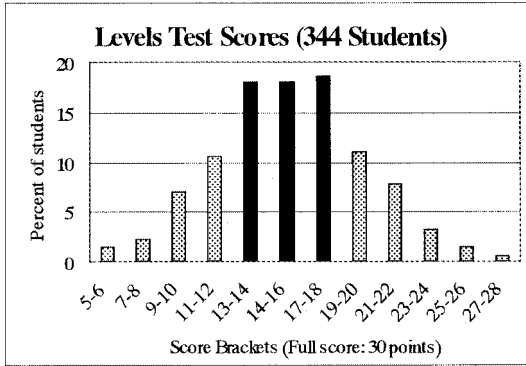
2000-word level, where the mean score is 83.7%. Among the five levels, the standard deviation is highest in the 5000-word level, suggesting that the vocabulary ability differs most among the participants in this category. Discrepancies among the standard deviation values, however, are relatively small in general, which might be assumed to reflect the fact that the individuals in the entire group vary relatively little in their vocabulary profile. In fact, as shown in Table 7 as well as in Figure 4, a majority of the

Table 8: Percentage of students for each score on 5 vocabulary levels

Scores	10000	UWL	5000	3000	2000
0	23.3	9.9	9.3	2.3	0.3
1	34.3	21.8	12.2	2.9	1.5
2	26.5	28.2	20.1	7.3	2.6
3	10.8	18.6	23.5	15.4	4.7
4	4.1	12.8	20.9	23.3	14.2
5	1.2	6.7	9.9	27.3	36.0
6	0.0	2.0	4.1	21.5	40.7
Total	100	100	100	100	100

Figure 5
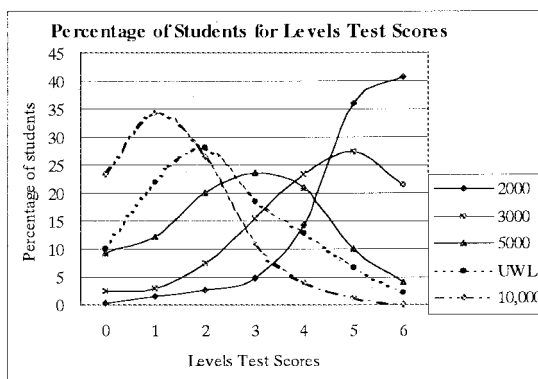
344 students (54.6%) (Appendix 1) scored between 13 and 18 points out of the full score of 30 points. If the bracket on either side of this middle core group is added, the result would be 76.1%, implying further that the participants under investigation form a relatively large homogeneous core group in terms of their vocabulary recognition abilities, with a minority of less than a quarter distributed on either end of the spectrum.

Beyond this overall profile of the vocabulary recognition abilities of the participants under survey, the recognition test results offer more specific evidence on the levels of correctly recognized vocabulary items. A closer look at the data (Table 8 and Figure 5) reveals that 1.2% of the participants

correctly recognize as many as five of the six text words at the 10,000-word level, and another 4.1% four out of six words. Assuming that these students are also able to correctly recognize meanings of the words, or at least a majority of the words, on all four lower levels, more than 5% or 2 students in an average class of 50, for instance, might be safely presumed to know close to 10,000 words. A majority or 57.6% of the students, however, know none or only one word at the 10,000-word level.

The University Word List (UWL) is apparently above 5000-word level and below 10,000-word level. At this level, 8.7% of the participants recognize five or all six test words correctly; when the 12.8% who recognize four words are added, 21.5% (10-11 students in a class of 50, for instance) presumably know more than half of the words at the UWL level. Below the UWL level, a near-standard distribution curve is obtained for the 5000-word level, with 34.9% of the participants recognizing a majority of the test words, and 64.5% forming the middle group who recognize two, three or four items out of the six.

The distribution curves for the two lowest vocabulary levels, i. e., 3000- and 2000-word levels, practically form mirror images of those for the two highest levels. At the 3000-word level, 48.8% of the participants correctly recognize the meanings of five or all six test items; no less than 72.1% know four to six words at this level. At the lowest level of 2000 words, 40.7% of the participants recognize all six test words, and 76.7%, five or all six.

In summary, some 80% of the participants could be assumed to know most of the words listed in Nation's First and Second Thousand words. In addition, 72.1% of the students presumably know close to 3000 words, and some 34.9%, more than half of the words at the 5000-word level, for either

of which no concrete word list is available in Nation-Xue 3000. At the UWL level, approximately 59.9% of the students recognize only a very small number of items or none. Finally, 5.3% of the students recognize 4 or five out of the six test items at the 10000-word level, while 84.1% recognize two words or fewer (Table 8 and Figure 5).

Production test results

For the present survey of vocabulary production abilities, three versions of vocabulary production tests were devised with four semantic areas each (Appendix 2). The semantic areas were selected on the basis of the earlier survey results of spring 1998 (Ishihara, Okada and Matsui, 1999) so that they might range from the most concrete and personal (such as "family life" and "body parts") to the more abstract and general (such as "politics" and "technology"). Since the test time was limited to five minutes, the maximum number of words allowed under each semantic area was restricted to 20 words. The three sets of 4 semantic areas were presented in Japanese to three different groups of about 100 participants each, who wrote, in a brain-storming fashion, as many English words as possible in each semantic area. (For more details on the test and scoring procedures, see Ishihara *et al*., 1999.) In consequence, survey data were obtained regarding 12 semantic areas in total.

For the 340 participants, the mean number of words written in five minutes is 25.9 with a standard deviation of 8.5 words (Table 9). When

Table 9: Production test scores

Number of participants	340
Mean	25.9
Standard Deviation	8.5
Words by all participants	8821

Table 10: Participants in score brackets

Brackets	Number of students	Percentage of students
7-11	12	3.5
12-16	33	9.7
17-21	57	16.8
22-26	77	22.6
27-31	69	20.3
32-36	49	14.4
37-41	32	9.4
42-46	7	2.1
47-51	3	0.9
52-57	1	0.3
Total	340	100

Figure 6

Production Test Scores in Brackets
(340 students)
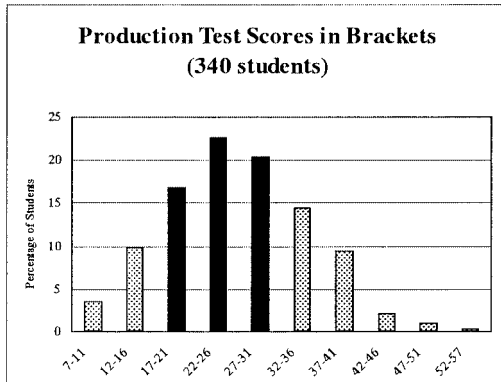
divided into ten score brackets ranging from 7 through 57, 59.7% of the
participants are in the three middle brackets between 17 and 31 words; if the
participants in the bracket of 32-36 words are added to this middle bracket
group, 74.1% of the participants form the middle core in the entire statistical
profile (Table 10 and Figure 6).

When the number of words produced in five minutes is counted
separately for each of the 12 semantic fields (Table 11 and Figure 7), it is
evident, exactly as in the previous survey results of 1999 (Figure 2), that

Table 11: Vocabulary production scores for 12 semantic fields

	Semantic Fields	Mean
1	Politics	1.74
2	Buildings	2.83
3	Science/Technology	2.95
4	Environment	3.55
5	Communications	5.24
6	Commerce	6.03
7	Foods and Drinks	6.05
8	Traffic/Travel	8.55
9	Rooms/Furniture	9.30
10	Family Life	9.52
11	Body Parts	10.49
12	Geography/Weather	10.90

Figure 7
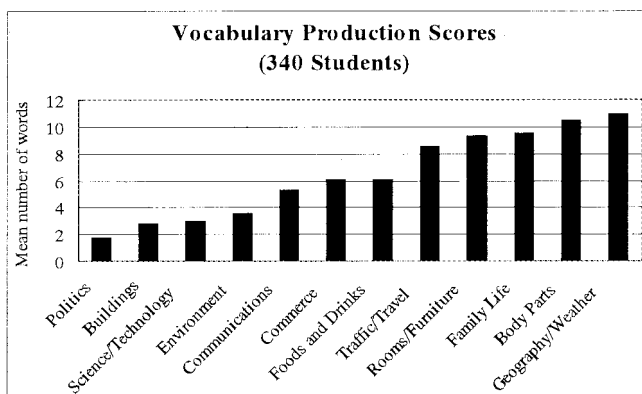
more concrete or personal semantic areas tend to invite more vocabulary production than the less concrete or personal. Thus, in the semantic areas of "geography and weather" and "body parts," the average numbers of words produced are 10.90 and 10.49 respectively, while on the other side of the spectrum, in the semantic category of "politics," 1.74 words are produced on the average, and for "types of buildings" and "science and technology," the mean numbers are close to each other, i. e. 2.83 and 2.95 respectively.

Levels analysis of sample production words

    For the purpose of determining the levels of words produced by the participants in reference to JACET 4000 and Nation-Xue 3000, sampling is needed to obtain a manageable size of corpus for analysis. Therefore, 20 participants with the largest numbers of words produced were selected for each of the twelve semantic areas; the words produced by these participants total 2731 (Table 12). Each word was counted only once, regardless of how many times it occurred in the corpus, resulting in 857 different words as indicated in Table 12. Notice, in passing, the general tendency that the larger the total number of words produced, the smaller the percentage of different words: for the semantic areas of "body parts" and "rooms and furniture," the percentages of different words are 19.1% and 28.5% respectively, while on the other polarity, "politics" and "science and technology" have the percentages of 43.3% and 45.4% respectively.

    When the lists of different words produced for the twelve semantic areas are compared, a fair number of items are found in more than one semantic

Table 12: Sample production words used for levels analysis

Semantic fields	All sample words	Different words	% different/all
Politics	104	45	43.3
Science/Technology	119	54	45.4
Environment	131	55	42.0
Buildings	139	52	37.4
Commerce	202	92	45.5
Communications	213	77	36.2
Traffic/Travel	269	101	37.5
Foods/Drinks	279	78	28.0
Family Life	292	67	22.9
Geography/Weather	304	75	24.7
Rooms/Furniture	333	95	28.5
Body parts	346	66	19.1
Total	2731	857	31.4

Table 13: Percentage of words listed in JACET 4000 found among sample production words
by 20 of the survey participants

JACET 4000	J1 + J2	J3	J4	J5	Not listed
Number of content words	738	921	862	1066	3587
Number of sample words	302	188	111	94	77
% Prod/Content words	40.9	20.4	12.9	8.8	2.1

Table 14: Percentage of words listed in Nation-Xue 3000 found among sample production
words by 20 of the survey participants

Nation-Xue 3000	N 1	N 2	UWL	Not listed
Number of content words	861	946	807	2614
Number of sample words	308	183	66	215
% Prod/Content words	35.8	19.3	8.2	8.2

area; for instance, some of the words in "science and technology" are also among those in "communications" or "the environment," so that the 857 sample words are further processed to eliminate overlapping. Eventually, the remaining 772 different sample words are submitted to levels analysis. It must be noted that the 772 sample words are all "content words" and do not include any "function words"; both JACET 4000 and Nation-Xue 3000 are scanned to delete such items as prepositions and pronouns. The resulting numbers of content words in the separate levels for both of these word lists are indicated in the second row of Tables 13 and 14 respectively.

As shown in the bottom row of Table 13 as well as in Figure 8, a comparison of the 772 sample words with the content words in JACET 4000 shows that 40.9% of the 738 content words in J1 and J2 and 20.4% of the 921 content words in J3 is among those produced by the survey participants. On the other hand, 12.9% and 8.8% of the content words in J4 and J5 respectively are found in the corpus of the sample words.

Similarly, a comparison between the 772 sample words with the content words listed in Nation-Xue 3000 (Table 14) reveals that 35.8% of the 861
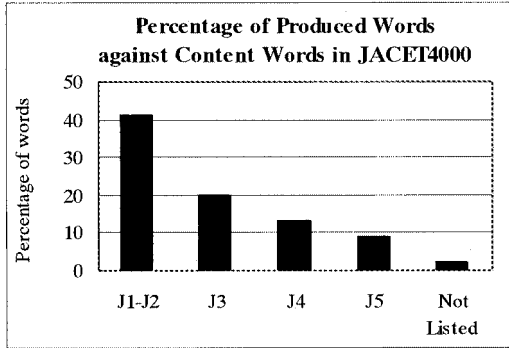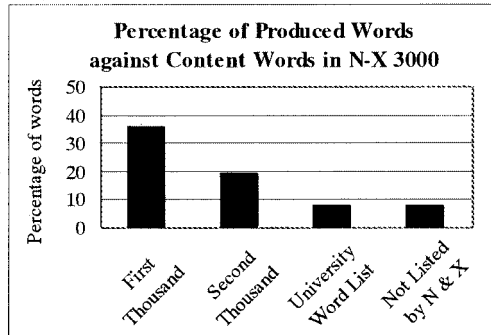
Figure 8

**Percentage of Produced Words
against Content Words in JACET4000**

Figure 9

**Percentage of Produced Words
against Content Words in N-X 3000**

content words in the First Thousand is produced by the 20 sample
participants. However, for the Second Thousand level, the percentage
declines relatively sharply to 19.3% of its 946 content words. Considering
the relatively high vocabulary level of Xue-Nation's UWL, 8.2% of its
content words appearing among the sample words is impressive; all the
more so when compared with a fairly close value of 8.8% for J5. This
could imply that the highest levels of production vocabulary range from the
4000- through 7000-word levels, with a relatively small decline toward the

Table 15: Percentage of sample words in JACET 4000

JACET 4000	J1-J2	J3	J4	J5	Not listed	Total
words	302	188	111	94	77	772
%	39.1	24.4	14.4	12.2	10.0	100

Table 16: Percentage of sample production words in Nation-Xue 3000

N-X 3000	1st thousand	2nd thousand	Univ wd list	Not listed	Total
words	308	183	66	215	772
%	39.9	23.7	8.5	27.8	100

higher end.

As for the proportions within the list of 772 sample words, 39.1% are among those in J1 and J2; similarly, 39.9% are among the first thousand words in Nation-Xue 3000 (Tables 15 and 16). Slightly less than a quarter, namely, 24.4% and 23.7%, are listed in J3 and Nation-Xue's Second Thousand words respectively. Consequently, in terms of both JACET 4000 and Nation-Xue 3000, a majority, i. e. 63.5% and 63.6% respectively, of the sample words produced are within the 2000-word level.

As indicated in Table 14, 14.4% and 12.2% of the sample words are in J4 (or 3000-word level) and J5 (or 4000-word level) respectively, leaving the other 10.0% among those not listed in JACET 4000. Some of the 215 sample words (or 27.8%) that are not found in Nation-Xue 3000 are presumably beyond the 2000-word level and below the UWL level.

Also among the sample words not listed either in JACET 4000 or Nation-Xue 3000 are some highly contemporary vocabulary items, especially those related to electronics and communication technology as well as relatively specialized terms regarding musical genres and instruments.

Comparison between vocabulary recognition (R)
and vocabulary production (P) test scores

The simplest way to compare vocabulary recognition test scores (R score) with vocabulary production test scores (P score) is, of course, by means of the correlation figure for the entire survey group, which, in the present case, is 0.35 for 335 participants (Appendix 1) where p<0.148 at 1% level of confidence. For further comparison from a different angle, the participants were divided into five P-score brackets from those with the largest numbers of P words to those with the fewest: 50-59, 40-49, 30-39, 20-29 and 0-19. The mean R scores were then calculated for each of the five P-score groups on all five vocabulary levels as shown in Table 17 and Figure 10. The five P-score groups are represented in Figure 10 by five different lines, with their mean R scores plotted horizontally. Thus, the top group of participants with the P scores 50-59, for instance, has the mean R score of 3.0 words for the 10,000-word level, 4.5 words both for the university- and 5000-word levels, 5.5 words on the 3000-word level, and full 6.0 words for the 2000-word level. Conversely, the participants who recognized, on the average, 4.7 words on the 2000-word level produced fewer words, i. e. 0-19 words, constituting the lowest P-score group. The three other P-score groups in between, i. e. 40-49, 30-39 and 20-29, are in the middle between the highest

Table 17: Percentage of students for Recognition and Production Test score brackets

Production Test Scores	Recognition Vocabulary Levels				
	10000	UWL	5000	3000	2000
50-59	3.0	4.5	4.5	5.5	6.0
40-49	1.0	3.1	3.3	5.0	5.4
30-39	1.7	2.5	3.3	4.4	5.2
20-29	1.4	2.3	2.7	4.3	5.1
0-19	1.1	1.9	2.4	3.7	4.7

Figure 10
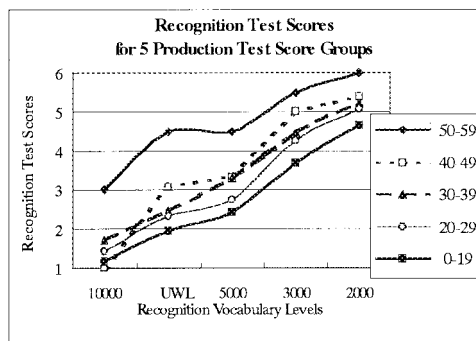
and the lowest in R scores as well, forming almost perfect parallel patterns as seen in Figure 10. This means that, in general, the participants who produced larger numbers of words correctly recognized notably larger numbers of words than those who produced fewer words, and vice versa.

In addition, for every one of the five P score groups, there is a steady increase in R scores from the 10,000-word level toward the 2000-word level; the rate of increase is remarkably similar for all five groups, as illustrated by relatively parallel lines in Figure 10. Regarding the top two P score groups, the R scores for the UWL and 5000-word levels are either perfectly identical, namely 4.5, or nearly so, 3.1 and 3.3 respectively. A somewhat similar pattern, although not as evident, is also seen for the P score group of 20-29 (2.3 and 2.7 respectively). This might well reflect special emphasis or efforts made in vocabulary studies at these two relatively high levels in English language instruction prior to the current survey.

One other method of comparison is via C-test scores (C scores). The C-test is assumed to reflect relatively comprehensive L2 abilities including reading and writing proficiencies as well as vocabulary knowledge (Ishihara

et al., 1999; Writing Research Group, 1999; Laufer, 1998; Tsuchiya, 1998). For the entire survey group, the correlation figures of C scores with R and P scores are 0.32 and 0.43 respectively where p<0.148 at 1%.

In order to compare the R and P scores on the basis of C scores, the survey participants are first grouped into six C-score brackets ranging from 10-24 to 85-100; then the average R and P scores are calculated for each of the 6 C-score groups. As shown in Table 18 and Figure 11, the higher the C scores are, the higher are the R and P scores in general: thus for the C-score group of 10-24, the R and P scores are 42.2 and 21.9 respectively in percentile, while for the C-score group of 85-100, they are 60.4 and 42.7, with the four middle groups falling between these two extremes.

In Figure 11, where R (solid line) and P (broken line) scores are plotted horizontally for the six C-score groups, the R and P scores form remarkably parallel lines, although in the top C-score group of 85-100, the R-score does not differ much from that for the second highest C-score group of 70-84, 60.4 for the top group and 60.6 for the second highest. Thus, the difference between the R and P scores are notably constant for the six C-score groups. Although the test procedures may well be responsible to some extent for these score differences, they could also be presumed, in some degree at least, to reflect the gap between the R and P abilities of the entire survey group. If so, it should be noted that the differences between R and P

Table 18: Recognition (R) and Production (P) test scores for 6 C-test (C) score brackets

C scores	% R scores	% P scores
10-24	42.2	21.9
25-39	46.6	27.0
40-54	50.2	29.4
55-69	54.3	33.8
70-84	60.6	38.1
85-100	60.4	42.7

Figure 11

Levels Test (R) and Production Test (P) Scores for 6 C-Test Score Groups
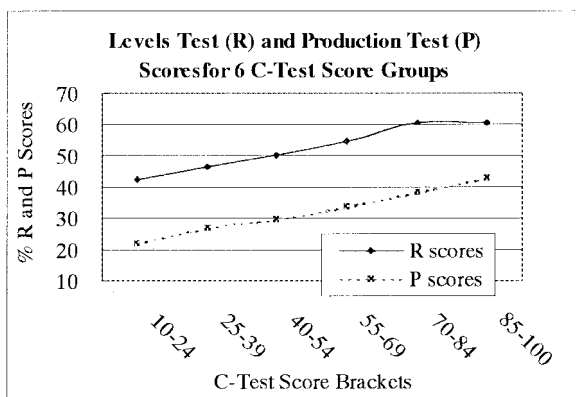
abilities are generally constant for all the C-score groups, with virtually no narrowing or widening for any of the six groups.

Discussion and concluding remarks

For a long time, vocabulary acquisition was a largely neglected aspect of EFL research and teaching (Meara, 1982). In his widely read book *Teaching English in Difficult Circumstances* (1960), Michael West, the author of *A General Service List of English Words* (1953), apparently felt it necessary to defend his position on vocabulary teaching against the then prevalent trend, stating that:

> I have been accused of devoting too much attention to content words. In a reading vocabulary Content words are the major problems; but in a Speech vocabulary they cannot be totally neglected, as in the present violent swing to structuralism they tend

to be (West, 1960: 38-39).

Theories of structural linguistics exerted a strong influence over the theory and practice of teaching English, with their tenet that "in learning a second language . . . vocabulary is comparatively easy, in spite of the fact that it is vocabulary that students fear most. The harder part is mastering new structures in both content and expression" (Gleason, 1961: 7). Giving some priority to structure learning over vocabulary learning may be good for initial stages of second or foreign language acquisition. For later stages, however, instructional focus should be shifted more onto vocabulary learning so that students might learn to communicate in increasingly meaningful ways in the target language. There is some truth, especially from the educational point of view, in calling structural grammar a "skeleton grammar," as did some traditional grammarians. It must be "fleshed out" with content words.

Since the 1980s, as communicative competence was cogently argued for, research for vocabulary instruction has become activated. Krashen (1982: 80) states that:

> Another way teachers help students understand messages containing structures that are "beyond" them is by emphasizing vocabulary. While knowledge of vocabulary may not be sufficient for understanding all messages, there is little doubt that an increased vocabulary helps the acquirer understand more of what is heard or read. Thus, more vocabulary should mean more comprehension of input, and more acquisition of grammar. This "new view" is quite different from earlier positions. Language teachers had been told to

restrict introduction of new vocabulary in order to focus on syntax. Now we are saying that vocabulary learning will actually contribute to the acquisition of syntax.

Moreover, it may very well be assumed that when native speakers read, they focus chiefly on content words, utilizing the grammar knowledge as used in newspaper headlines or telegraphic sentences, and employing the function words for clarifying a dubious interpretation (Bowen *et al.*, 1985). In honing this skill to be fluent readers, students of English, when they have mastered the basic syntax used in newspaper headlines and telegraphic sentences, must try to enrich their vocabulary. Just as syntax is systematically taught, vocabulary must also be presented in a structured way.

Echoing these theoretical views on the importance of vocabulary instruction, students also seem to recognize the necessity of word study. Table 19 and Figure 12 summarize the results of a questionnaire on learning English (Appendix 3) with 257 university students. The questionnaire asks to what extent the students feel the difficulty, importance and need for further study in four language study areas, *i. e.* vocabulary, thought or content, socio-cultural aspects, and grammar. The participants respond by choosing one of three alternatives: (1) very little, (2) a little, and (3) very

Table 19: Questionnaire results on Difficulty, Importance and Need for Further Study

Percentage of Students (n: 257 students)			
Study areas	Difficult	Important	Necessary
Vocabulary	26.1	47.5	63.4
Thoughts	23.7	56.4	42.4
Socio-Culture	21.8	43.6	45.1
Grammar	17.9	32.3	27.6

Kenji Ishihara, Tae Okada, and Shimpei Matsui

Figure 12

Questionnaire Results (257 students)
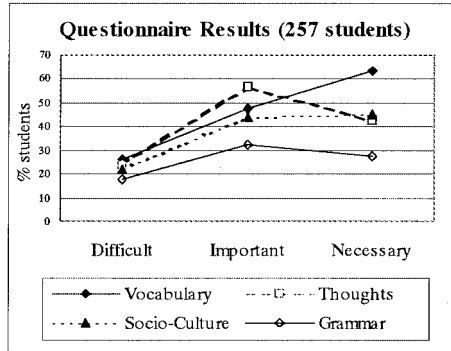
much. The figures in Table 18 represent the percentages of participants whose choice was the highest degree "very much" in the four study areas. Thus, 63.4% of the respondents feel the greatest need for learning vocabulary, followed by 45.1% and 42.4% for socio-cultural aspects and thought patterns or content respectively. Grammar, from the students' point of view, is the lowest in difficulty and importance, as well as in need for further study. For reading comprehension, thought patterns or content came at the top with 56.4% of the respondents, followed by vocabulary with 47.5%. This might perhaps reflect one or both of two situations: (1) the students may have experienced a greater difficulty in comprehending the overall content of English text rather than individual words; (2) understanding thoughts or content is more strongly emphasized in the university English classroom than in earlier language instruction. Vocabulary was considered most difficult by 26.1%, or more than one out of four students.

                    *   *   *

  The data presented earlier in this paper seem to point toward the

vocabulary levels relevant for university instruction. First, considering the fact that nearly 40% of the sample production vocabulary is at the 1000-word level (39.1% and 39.9% in reference to JACET 4000 and N-X 3000 respectively) and around a quarter (24.4% and 23.7% respectively in reference to the same two lists), at the 2000-word level, it is desirable that production of a larger number of words at the 2000-word level be encouraged along with the 3000-word level (14.4% of the present sample in reference to JACET 4000). Vocabulary production exercises need to be designed on this basis.

For recognition, key levels for university students seem to be the 5000-word level and higher. Approximately half the words at the 5000-word level need to be reinforced or consolidated in varied contexts, while the other half need to be newly acquired. Through reading with a focus on content, vocabulary at levels higher than the 5000-word level might also be introduced in relation to specific topics. The notable similarity between the test score distribution of 1998 for FO-words based on the Yoshioka List (Figure 1) and that for the 5000-word level of the present survey (Figure 5) suggests that the FO words in the Yoshioka List might legitimately be assumed as relevant for university instruction. Other lists (such as Barnard, 1971) should also be consulted in further research for specific items at this level. The minimum goal for recognition words might be set for the acquisition of sufficient vocabulary for using monolingual English dictionaries (such as the Cambridge defining words). Further research is needed to make comparisons of specific vocabulary items of this type.

The comparative data on R and P scores reported above suggests that more instructional efforts should be made to improve production abilities, both in terms of the number of words produced per unit of time as well as in

narrowing the gap between R and P vocabulary levels, since, as indicated in the present set of data, R is consistently at higher levels than P.

Vocabulary instruction needs to be structured just as grammar instruction must be systematic. Vocabulary lists based on frequency of occurrence are certainly useful for organizing instruction, but other aspects must also be taken into pedagogical consideration. A morphological or word family approach as well as notional categorization based on content or semantic areas must be incorporated in good balance. This latter might further be conceived in two different ways: a more linguistic approach in which references are made to synonyms, antonyms, paraphrases, derived forms, etc., and a situational approach based on communication situations. Further elaboration is left for future research.

In any case, vocabulary instruction cannot be dealt with in isolation. It must be a constant concern throughout L2 instruction whether the focus happens to be on reading, writing, listening or speaking. Also, since the role of self-instruction is particularly important in vocabulary acquisition, the teacher's responsibility for effective proposals, encouragement, and monitoring is of primary importance. Further research is needed before more specific proposals can be made, particularly those regarding the relevant vocabulary items for acquisition.

References

Barnard, H. (1971). *Advanced English vocabulary: Workbook 1*. Rowley, Mass.: Newbury House.

Barnard, H. (1972). *Advanced English vocabulary: Workbook 2*. Rowley, Mass.: Newbury House.

Bowen, J. D., Madsen, H. and Hilferty, A. (1985). *TESOL techniques and procedures*. Rowley, Mass.: Newbury House.

Gleason, H. A. (1961). *An introduction to descriptive linguistics (Revised Edition)*. New York: Holt Reinhart Winston.

Ishihara, K, Okada, T. and Matsui, S. (1999). English vocabulary recognition and production: A preliminary survey report. *Doshisha Studies in Language and Culture, 2, 1,* 143-175.

JACET. (1993). *JACET kihongo 4000* (JACET 4000 basic words). Tokyo: The Japan Association of College English Teachers.

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon Institute of English.

Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics, 19*, 255-271.

Meara, P. (1982). Vocabulary acquisition: A neglected aspect of language learning. In V. Kinsella (Ed.), *Surveys 1: Eight state-of-the-art articles on key areas in language teaching* (pp. 100-126). Cambridge: Cambridge University Press.

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Boston: Heinle and Heinle.

Tsuchiya, T. (1998). What makes reading a difficult task? The implications of a metrical analysis of college-level reading materials. *JACET Bulletin, No. 29*, 193-206.

West, M. (1953). *A general service list of English words*. London: Longman, Green and Company.

West, M. (1960). *Teaching English in difficult circumstances*. London: Longman, Green and Company.

Writing Research Group, JACET Kansai Chapter. (Ed.). (1999). *Teaching writing in colleges and universities: Practical reports, 3.*

Xue, G. and Nation, I. S. P. (1984). A university word list. *Language Learning and Communication*, *3*, 215-229.

Yoshioka, K., Matsui, S., Okada, T. and Ochi, R. (1997). *Kokusaika-jidai ni daigakusei ga shuutoku subeki goi no kenkyuu* (A study on the vocabulary for college study in the period of internationalization).

Appendices

Appendix 1: Number of participants 1999 (all non-English majors in eight of the English classes required of the second-year students)

Faculties	Divinity	Letters	Law	Economics	Commerce	Engineering	Total
Participants	1	47	65	95	77	74	359

Of the 359 students, 344 participated in the Levels test, 340 in the Production test, and 335 in both.

Appendix 2

   -

(1)   (2)

(3)   (4)

   -

(1)   (2)

(3)   (4)

   -

(1)   (2)

(3)   (4)

Appendix 3

Questionnaire 1999-2000:

(1)   (2)   (3)

1.

A.	(1)	(2)	(3)
B.	(1)	(2)	(3)
C.	(1)	(2)	(3)
D.	(1)	(2)	(3)

2.

A.	(1)	(2)	(3)
B.	(1)	(2)	(3)
C.	(1)	(2)	(3)
D.	(1)	(2)	(3)

A.	(1)	(2)	(3)
B.	(1)	(2)	(3)
C.	(1)	(2)	(3)
D.	(1)	(2)	(3)

Nation

JACET 4000

Nation 2000 Xue

2000 80

3000 50 5000 33

30 10000 1

64 2000

C

5000 3000