Reading materials for learning TOEIC vocabulary based on corpus data

TANIMURA Midori
UTIYAMA Masao
National Institute of Information and Communications Technology

1. Introduction

It is agreed that presenting words in context rather than isolating them from context enhances learners' acquisition of vocabulary. For example, several recent studies have found positive evidence supporting the use of explicit vocabulary instruction in conjunction with reading (Paribakht & Wesche 1997). A great deal of vocabulary learning material which follows this idea has also been provided by publishers or on the Web, however, the selection of most words is mainly based on material writers' experience and intuition. Indeed, Kennedy (1987a, 1987b) points out that corpora are useful tools to look critically at existing language teaching materials about ways of expressing quantification and frequency of vocabulary in ESL (English as a second language) textbooks. Considering recent remarkable developments and the progress of computer technology, it is now possible to select vocabulary statistically, and create materials objectively based on vast amounts of corpus data.

The sorts of learners we principally have in mind are the ones who mastered senior high school level vocabulary and want to continue to increase their vocabulary. We attempt to bridge the gap between English vocabulary used in Japanese senior high school textbooks and those used in international communication through reading.

2. Literature review

What are the factors that should be taken into account when creating materials for learning vocabulary based on corpus data? It would be suggested that there are three major factors to be considered: (1) recycling and attention; (2) words in context; and (3) word selection by corpus data.

(1) Recycling and attention

Much research has been done about vocabulary learning through reading. In L1 study, Gardner (2004) points out that children (ages 5 to 12) learn 3000 new words every year and that 12th graders (ages 17 and 18) have approximately 40,000 words. "Word" can be defined as "a word family which consists of a base word, its inflected forms and common derivations (Laufer & Nation, 1995; Nation & Waring, 1997)."

The major source of acquiring vocabulary for them is reading in multiple and various contexts. However, in EFL situations, it is unlikely that learners meet the same words even though textbooks provide a variety of texts to learners, because learners are not exposed enough to acquire words. Nation (1990:44) for example points out that "It is important for teachers in countries where English is not used much outside the classroom to know whether they can rely on coursebook to provide enough repetition for vocabulary learning to be possible." Therefore, it is critically important in reading

that learners have the opportunity to keep meeting words that they have met before and these words need to be reinforced by another meeting in limited time (Paribakht & Wesche 1997). Indeed, there is interesting research done by Rott (1999) about the number of times a learner needs to be exposed to unknown words to acquire and retain them. The results indicate 6 times exposure produced significantly more vocabulary. Other researchers also have found similar results.

On the other hand, according to some psychologists (Craik & Lockhart, 1972; Craik & Tulving, 1975), repetition is not an important factor in vocabulary learning, but attention that is given to an item decides whether it will be remembered or not. Ellis (1995) and Robinson (1995) also indicate that vocabulary learning requires attention but it should be addressed to both meaning and form. Furthermore, in Fraser's study (1999), higher retention rates occur when words are noticed and potentially highly salient. She notes that this occurs with not only L1 cognates but also in L2 word association, and frequently encountered words.

(2) Words in context

Traditionally the way of learning vocabulary typically involves memorizing, making vocabulary cards and looking up dictionaries. Nation (1990) actually cites that one of the great attractions of learning lists of words is that large number of words can be learned in a very short time. He goes on to state that a lot of research has shown the effect of learning lists of words. However, Fanselow (1992) argues that learners can not store word meaning if they use word lists (Laufer, 1997; Nagy, 1997). A good example for this would be that learners often forget words just after the test was done. He further claims that learning words without the context is not only less effective in understanding meaning, but also it gives bad influence for reading.

On the other hand, comprehension of a text involves multiple exposures to the word in context. It requires deep processing of information about the words, leads to a more accurate understanding of word meaning, and fosters vocabulary acquisition (Ellis 1995). Another piece of research also has shown the fact that reading plus vocabulary instruction treatment led to a quantitatively and qualitatively superior gain (Qian 1996).

(3) Word selection by corpus data

After a basic level of high frequency words is learned, all students need to develop knowledge of technical or topical vocabulary in order to grasp the idea of specialized texts which they will encounter in their own particular fields. To select words for specific purposes, there has been development of various word lists which attempt to count frequency in specific areas (O' Dell 1997:269). Nation (1990) summarizes several values which specialized word counts have. First, they provide a useful guide to which vocabulary learners should focus on in reading. Second, they can play the role of a checklist and a goal for learners. Through marking important words which they do not know, these words are brought to the learners' conscious attention and they are more likely to stay in their mind.

It should be kept in mind however that the results of frequency count often differ

depending on which corpus is used. As Nation and Newton (1997) mention, every field has its own technical vocabulary, which is used within a narrow range and normally used within a specialized field. In other words, in a particular field it may occur frequently, but in other fields it may scarcely occur. In fact, Grabe (1991) argues that in specialized academic settings, infrequent words may be the most important for L2 learners to know, because they may be closely connected to the topic being discussed.

The studies above reveal three points: (1) vocabulary should be recycled and it should obtain learners' attention; (2) vocabulary should be presented in contexts; and (3) vocabulary for specific areas should be selected by appropriate corpus data. However, there does not seem to be any traditional materials which meet all three criteria at the same time. In this paper, we propose new vocabulary learning materials for the TOEIC test in reading. In order to examine its usefulness and effectiveness, it was used in class for a year and it was compared with a vocabulary list book where new words and their meanings were isolated from contexts and which was basically based on material writers' experience.

To examine the validity of our reading texts, the following research questions were examined:

1. Which material is more effective and efficient to retain the meanings of vocabulary, texts or word lists?
2. Does general reading ability improve after learning words for the TOEIC test?

3. Features of our reading texts

The aim of producing our reading material is to help students effectively learn words for TOEIC in context, thus it is important to decide which new words learners should be exposed to and the degree to which those words are recycled. Our reading material consists of 642 words which occur most frequently in the TOEIC test according to research results by Chujo (2003) and Chujo et al. (2004). The bulk of the corpus analyzed in this study consists of parts of the Daily Yomiuri from 1989 to 2001 (about 25000 articles). The choice to use a newspaper is based on the assumption that it includes more words related to finance, business and economy covering a broad range of topics in the TOEIC test rather than a different genre such as the novel, and the Daily Yomiuri is the only Japanese newspaper which is available for non-commercial research purposes. Our reading material was constructed by the following steps (see Utiyama et al. 2004. for more details on how to create materials).

Step 1. Gather the texts (= articles) containing as many new and recycled words (that are shared with the 642 words) as possible.
Step 2. Sort the texts by size of how many new and recycled words are included.

In the end, 116 texts which contained 642 words were constructed as reading material.

Due to the algorithm, the texts selected at the earlier stage of this process had more new words than the ones selected at the later stage. That is, the number of new words gradually decreases as learners go on in their study. For example, in the first text, 43 new words were included which need to be learned for the TOEIC test while in the tenth text 16 new words were included. This means that most words in the TOEIC vocabulary were covered in the earlier texts. Thus, learners can learn vocabulary efficiently even if all reading texts can not be covered in the lesson. The reading material was also designed to contain as many recycled words out of the 642 words as possible so that learners meet the same words repeatedly through reading the texts. In this study, 60 texts out of 108 were used as reading texts through the academic year. The following statistics are thus based on the words in 60 texts.

The following is a sample of an English text. The third text is chosen here because the first and second texts do not include recycled words. In the original text of the English version, three different colors were used to distinguish new words and two kinds of recycled words. New words were colored in red, recycled words which occurred in previous texts were colored in blue and recycled words which occurred in the same text were colored in orange. In the following example, new words are underlined, recycled words in previous texts are in italics, and recycled words in the same text are boxed.

---

3. Consumer fraud incidents increasing

A growing number of people are falling victim to firms that ask for payment in advance but fail to deliver personal computers and other brand name goods *offered* at 40–70 percent discounts through newspaper inserts. As of Wednesday, consumer protection centers nationwide reported *receiving* more than 2,000 complaints from angry consumers claiming to have been victimized by such discount sales ploys. During the past year, *law* enforcement authorities have exposed four *organizations* nationwide said to have been behind such schemes. Victims order the products over the telephone––using various methods of payment––attracted by the notion of purchasing high quality products at low prices. Police and consumer protection centers are *advising* consumers to be on guard and consider the risks of paying for merchandise before delivery. According to the National Police Agency, *cases* of such discount sales ploys began to increase about two years ago. The agency has since stepped up efforts to crack down on such crimes. Japan Discount Service, a mail order firm based in Tokyo's Shinjuku Ward, claimed in a newspaper insert that those paying a 10,000 yen membership fee would be able to buy household appliances from Japan's top manufacturers at 60–70 percent off suggested retail prices. The *organization* attracted 40,000 members and was raided by Ibaraki Prefectural Police last September on suspicion of violating a telephone sales–related law. Investigators last March raided the offices of a marketing firm specializing in personal computer equipment sales, also based in Shinjuku Ward, for failing to deliver products to customers who had paid the *company* a *total* of 200 million yen.

It is assumed that those colored words work as markers to make learners conscious of them and get their attention.

Table 1 and Table 2 present how the target (TOEIC) words were recycled in the reading material.   Table 1 shows the high token frequency words.   For example, "company," the top ranking word, occurs 50 times throughout the texts.   This could occur in the same text more than once.   The word "according" in rank 4 is often used as "according to," but phrases are not considered in this study.

Table 1 High Token Frequency Words (the Top 20)

Ranking	Words	Number of Occurrences			
1	company	50	11	employee	18
2	official	45	11	director	18
3	bank	38	11	fiscal	18
4	according	27	14	price	17
5	tax	25	14	department	17
5	service	25	16	law	16
7	million	23	17	sale	15
7	number	23	17	president	15
9	agency	19	19	industry	14
9	expect	19	19	receive	14
			19	import	14
			19	loan	14

Table 2 shows the number of texts where each word occurs.   For example, "official," the top ranking word, occurs in 31 texts and learners most often meet this word in a different time through a year.

Table 2 Number of Texts Where Each Type Occurs (the Top 20)

Ranking	Words	Number of Occurrences			
1	official	31	9	offer	10
2	company	21	12	current	9
3	expect	14	12	provide	9
3	according	14	12	bank	9
5	number	12	15	firm	8
6	employee	11	15	concern	8
6	receive	11	15	agency	8
6	service	11	15	cost	8
9	million	10	15	organization	8
9	area	10	15	industry	8
			15	price	8
			15	fiscal	8

As shown in Table 1 and 2, target words are recycled and consequently learners are likely to be exposed enough to acquire target vocabulary efficiently.

4. Procedure

4.1. Participants

Ninety-nine Japanese English learners aged between 18 and 20 years old participated in this study.   There were three groups of participants; class A, B and C, who were placed according to the scores of their entrance exams.   They are first year students majoring in science, marketing, architecture and engineering in a university. In this university, first year students are required to take two English classes by the same Japanese instructors and one English conversation class by foreign teachers every week. A total of 60 lessons in both the first and the second semester taught by a Japanese teacher were targeted for this study.

4.2. Instruction

An instruction sheet was delivered to the learners.   To access the reading texts on the Web, learners were supplied with the URL, a username and a password.   Learners were told that the lessons using these reading texts were part of reading materials development for vocabulary learning.   Then learners were instructed on how to follow the Web page.

This is a part of the page where reading texts are listed.   Learners studied from 1. Streamlining to cost NTT over 1.4 tril. yen.   This text contains 296 words and was issued on November 9th 2001.   It contains 43 common words in the TOEIC vocabulary. All of the words were new and there were no recycled words.   Printed materials were also prepared for the learners who had difficulty accessing a computer.

TOEIC COURSE 1 (642 words, 116 articles)

１．Streamlining to cost NTT over 1.4 tril. yen (296 words - 2001/11/09)

43 common = 43 new + 0 old

new: account agreement allowance amount approve bank committee company compensation convention corporate cost current demand due earnings employee expect expense expensive

old:

6.7% (43/642) learned.296 words read.

２．Council recommends measures to more effectively cope with (255 words – 1995/09/12)

4.3. Treatment

The treatment in this study was carried out from May 2004 to January 2005. Two kinds of materials were used as a supplement for learning vocabulary for the TOEIC test, one was our reading texts and the other a commercially available vocabulary list book which learners were required to purchase and study in each lesson. In this vocabulary list, each word with a pronunciation guide followed by its meanings is presented on the left-hand side, and some specific words are accompanied by typical usage with the Japanese translation on the right-hand side. This vocabulary list book, however, does not present example sentences of all the words.

Learners were required to study both vocabulary learning materials as homework, and they were tested in each lesson. During the summer holidays, the learners were also required to study both sets of vocabulary. Through the year, 60 out of the 116 reading texts were covered, while all words in the vocabulary list book were covered. The reason all the words in reading texts were not covered was the logistics of the lesson schedule, but 559 words out of 642 were covered; which amounts to 87% of the important words.

4.4. Measure

Two kinds of tests were conducted: (a) vocabulary retention test; and (b) reading comprehension tests following the TOEIC format which consisted of 40 multiple choice tests. The retention test (see 4.5 test designs) was done to examine which vocabulary learning materials were more effective to recall. Learners were asked to write Japanese translations of the words which were listed in the test sheet. Since many words have multiple senses and are different parts of speech (e.g. nouns, adjectives and verbs), learners were instructed to write as many meanings as possible. If one of the meanings which learners listed was compatible with the one in reading texts or vocabulary list book, a score was given. The reading part of the TOEIC test was used in order to assess learners' initial and final reading levels on the first and the last day of the semester. At the end of the semester a questionnaire was conducted to examine what the students thought of the reading texts and the Web system they actually used for self-study.

4.5. Test designs
4.5.1. Word selection

To select the words for the retention test, electric word lists needed to first be prepared. Since there was no electric version in vocabulary list book, every word was

typed and stored as a text file.   Next, each word in the vocabulary list book and in the reading texts was given a ranking in JACET.   It is necessary to note that the words in JACET 8000 are categorized as parts of speech and thus the same words occur several times in different rankings.   As table 3 shows, it turned out that only the words "upward" and "forward" occurred more than once.

Table 3 Frequency of "upward" and "forward"

Word	Part of Speech	Ranking
upward	Adv	3902
upward	A	6089
forward	A/Adv	642
forward	Adv	4187

In this study, the higher ranking is adopted; 3902 for "upward" and 642 for "forward". The original ranking was used for the other words, since they occur only once.   All the words in the reading texts and the vocabulary list book were then combined, and finally three types of combined word rank lists were made: (1) words which occur in the reading texts but not in the vocabulary list book; (2) words which do not occur in the reading texts but occur in the vocabulary list book; and (3) words which occur in both materials.   The left column of Table 4 shows the total number of words which occur in each type.   The right column shows the number of the words which do not occur in JACET.   In order to select the words for retention test, those words which do not occur in JACET were excluded.

Table 4 Distribution of Word Occurrence in Three Types

	Total Number of Words	JACET✕
(1) Texts ◯ – Lists ✕	177	20
(2) Texts ✕ – Lists ◯	2206	221
(3) Texts ◯ – Lists ◯	370	6

Twenty words in each type (1), (2) and (3) and thus a total of sixty words were selected for the retention test.   To conform to the level of difficulty, the ranking distribution in JACET 8000 was examined as in Figure 1-3.
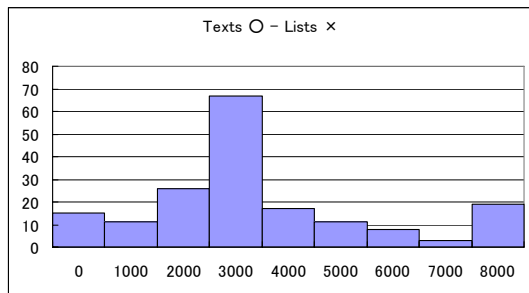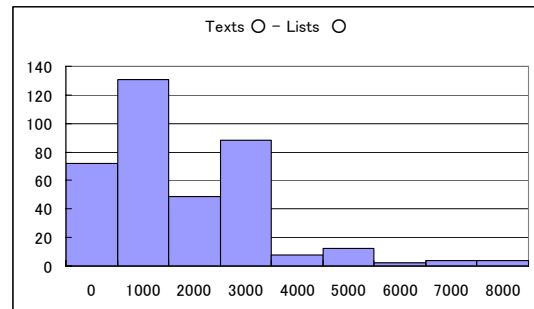
Figure 1 Texts ◯ - Lists ✕
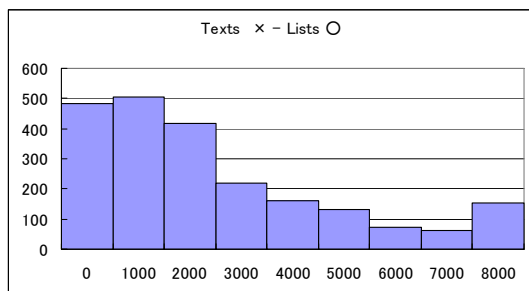

Figure 3 Texts ◯ - Lists ◯

Figure 2 Texts ✕ - Lists ◯

In JACET 8000, ranking 1 means the most frequent word and ranking 8000 means the least.   While figures 1 and 2 show that more than 10 words occur at the 5000 word level, figure 3 shows that most words are distributed from 0 to 3000 word level and not many words occur above the 4000 word level.   Therefore five words were selected from the 0 word level (from 1 to 999) to the 3000 word level (from 3000 to 3999) for each type (see appendix).

It was predicted that the words in type (3) were most retained, those in type (1) were less retained and those in (2) were least retained, because in type (3) words were most recycled and in type (2) words were least recycled.

4.5.2. Data analysis

To examine the efficiency and effect of the reading texts and the vocabulary list book, a retention test was conducted and the score gained for each of the three types of words were calculated.   The gain score and three types were then run as the dependent variable and the independent variable in the statistical tests.   One way ANOVA was run and subsequent analysis by using the LSD post hoc test was then performed.

In order to evaluate general reading ability, t score was used to compare the TOEIC test scores before and after the research.

5.   Results

5.1.   Results of research questions

Research question 1: Which material is more effective and efficient to retain the meanings of vocabulary, texts or word lists?

Table 6 presents means and standard deviation for retention test scores according to the three types.   The average score of (1) is the highest and that of (2) is the lowest.

Table 6 Retention Test Scores

	Mean	SD
(1) Texts ○ – Lists ×	14.9	2.2
(2) Texts × – Lists ○	10.8	3.2
(3) Texts ○ – Lists ○	13.7	3.7

Table 7 shows the result of one way ANOVA and it shows that the three types are significantly different.   Table 8 is a result of LSD post hoc test.

Table 7 One Way ANOVA

	Ssq	Df	Ms	F-ratio	Significance
Between Groups	875.3603	2	437.6801	46.68327	8.05E-14
Within Groups	2756.404	294	9.375524		
Total	3631.764	296			

Table 8 LSD Post Hoc Test

LSD		Mean Diff. (I-J)	Std. Error	Sig.	95% Confidence Interval for Mean	
(I) SUBJECT	(J) SUBJECT				Lower Bound	Upper Bound
(1)Texts ○ - Lists ×	(2)Texts × - Lists ○	*4.091	0.435	0	3.234	4.947
	(3)Texts ○ - Lists ○	*1.202	0.435	0.006	0.345	2.058
(2)Texts × - Lists ○	(1)Texts ○ - Lists ×	*-4.091	0.435	0	-4.947	-3.234
	(3)Texts ○ - Lists ○	*-2.889	0.435	1.53E-10	-3.744	-2.032
(3)Texts ○ - Lists ○	(1)Texts ○ - Lists ×	*-1.202	0.435	0.006	-2.058	-0.345
	(2)Texts × - Lists ○	*2.889	0.435	1.53E-10	2.032	3.745
*		a statistically significant (p < 0.05) difference in score				

The result of table 8 can be summarized as follows: type (1) > type (2); type (1) > type (3); and type (3) > type (2).   As might be expected, learners' scores in type (1) were higher than in type (2) and the scores in type (3) were higher than in type (2). Words in type (1) and (3) are recycled while the words in type (2) occur only.   Thus these results indicate that learning recycled words in reading texts is a more effective way to remember vocabulary than learning from a vocabulary list book.   In other words, the reading texts would make it possible for the learners to pay more attention to the target words which they need to learn and the contexts help the learners to understand the meaning more easily.

In contrast, the score in type (1) is higher than type (3).   There is a significant difference between type (1) and type (3), which was not expected.   Looking at the average score in Table 7, however, the difference between the two types was 1.2 and thus it is not as big as the difference between type (1) and (2) or (2) and (3).   Even though the level of difficulty was carefully conformed when selecting words for the test, learners may have known some words.   Overall, this result indicates that texts are more effective and efficient to retain meanings of vocabulary than word lists are.

Research question 2: Does general reading ability improve after learning words for the TOEIC test?

Table 9 presents the mean and standard deviation for the pre- and post-tests of a TOEIC reading section.   T-test shows that pre- and post-tests are significantly different ($t = -2.37$, $df = 76$, $p = 0.02$).

Table 9 Pre- and Post-test Scores of a TOEIC Reading Section

	Mean	**SD**
Pre-test	20.2	5.9
Post-test	21.5	7

Whereas there is a significant difference between pre- and post tests, the difference between the scores is not much.   Compared to the development of vocabulary skills, the range of reading skills shows low rates of increase.   Thus it would be difficult to conclude that development of vocabulary skills correlate with reading skills although the result indicates that vocabulary learning would have a positive effect on general reading ability.   It should be noted that there are still many words left the students need to know for understanding the reading texts.   Moreover reading involves different types of knowledge, cognitive and linguistic skills and strategies.   Vocabulary is

important to reading comprehension but would play a part of role in reading development.

5.2. Results of Questionnaire

A questionnaire was also conducted at the end of the semester. It was designed to be completed in a short time. Most of the questions involved simple choices of option or judgment presented on the familiar semantic differential scale, and expressed by a circle. The last question invited comment and opinions in which we hoped possible issues may be raised.

Learners were asked to assess reading texts on a scale of 1 to 5 for each of the following.

Table 10 Results of Scaling in Questionnaire

Questions	Score 5	-	1	Average
(1) Do you think it was good to have a vocabulary check test in each class?	very good	-	not good	4.1
(2) Do you think your vocabulary size for TOEIC increased compared with in April?	increased	-	not increased	3.4
(3) Is reading texts more fun than previous learning styles?	very fun	-	not fun	3
(4) Is reading texts useful?	very useful	-	not useful	3.7
(5) Do you think reading texts is easy and appropriate?	appropriate	-	not appropriate	3.4
(6) Do you want to use reading texts next year?	very much	-	not very much	3.5
(7) Do you think there are better materials than reading texts?	many	-	not many	3.5

The average scores of the questions were between 3 and 4. The score of question (1) had the highest score of 4.1 and a substantial majority felt that learning vocabulary was unavoidable and that they needed to be tested to remember vocabulary. The second highest score was 3.7 in question (4). Learners felt that these reading texts were fairly useful but at the same time, the score of question (3) was marked the lowest, suggested that their learning might not be enjoyable with reading these texts. In terms of the necessity of increasing vocabulary size, there is no doubt that learners felt they needed to learn vocabulary and the materials had the capacity to arouse their motivation. But in terms of reading for pleasure, it is not surprising that the learners found the reading difficult, because the texts used here were from the newspaper which included TOEIC vocabulary in them, and their topics are not very familiar. The result of question (6) and (7) showed some discrepancy. The figure suggests that whereas learners wanted to use these texts next year, they also felt that there must be better

materials.

The learners were also asked to answer the following questions from (8) to (10) and give their reasons and comments.

Table 11 Results of Fill-in Type Questions

Questions	Answers
(8) About how many pages do you really study hard?	Mean: 30.6 out of 60 texts
(9) About how many minutes do you spend studying this material?	Mean: 12.9 minutes
(10) Which do you think is better for learning words, (a) vocabulary list book, (b) reading texts or (c) difficult to answer.	The number of learners (a) 12, (b) 35, (c) 38

The figures in (8) and (9) showed that the learners were studying seriously about 30 out of 60 texts, and spending about 13 minutes reading each text.   This result indicated that they specially focused on the words which were colored and not on all the sentences, because it would be difficult to understand the passage in such a short time. The answers in (10) presented that three times as many learners chose (b) reading texts is than chose (a) vocabulary list book, but more than half the learners chose (c) difficult to answer.   The main comments by the learners who chose reading texts and those who chose vocabulary list book are shown in Table 13.

Table 13 Comments about Reading Texts and Vocabulary List Book

Reading texts	Vocabulary list book
"Important words for the TOEIC test are recycled."	"It is easy to look up words."
"I can learn the words in context."	"It contains many words."
"Words in color are the key to understanding the contents of information,"	"It has direct translation and easy to know the meanings while reading texts do not have direct translation."
"I can learn not only important words but also I can gain knowledge on a variety of topics."	"It is easy to use because I do not have to access a computer."
"This is good for reading as well as for vocabulary."	"Several meanings are presented for each word."
"I think it is more practical for learning vocabulary for the TOEIC test."	
"It gives me an opportunity to work on myself and this helped me with improving my proficiency."	

In a practical sense there may be pros and cons in both materials. Comments indicate that learners who chose (a) vocabulary list book tend to use it as a dictionary. On the other hand, learners who chose (b) reading texts tend to use them for vocabulary learning aids and for improving general English proficiency. Learners who chose (3) difficult to answer seem to flexibly use the materials depending on their purposes. One of the learners who chose (c) difficult to answer commented that both materials are helpful to learn vocabulary.

Prospective readers of these reading materials were people who have learned a basic level of high frequency words and already acquired reading skills. For the learners who have just started learning English, word lists may be a useful tool to learn new words.

6. Conclusion

This research investigated the effect of corpus-based reading texts in learners' vocabulary development for the TOEIC test. The approach we used was to recycle and reinforce frequent vocabulary in context to get learners' attention. We showed the syllabus designs of how we used reading texts in and outside of the classroom. We also presented the test design of how we selected the words for the vocabulary retention test. The result of the pre/post reading tests showed that vocabulary learning enhanced general reading comprehension. The result of the retention test showed that reading texts is more efficient and effective to remember vocabulary compared with a vocabulary list book. The results of the questionnaires showed that reading texts promotes a positive attitude toward reading even though it is not very easy. The challenges that teachers and learners may encounter when using the reading texts and how they are dealt with were also addressed.

This way of creating our vocabulary learning materials and presenting it on the Web in this study can be applied for any kind of specialized fields, such as medicine, engineering, architectures and chemical, and it would contribute to fostering learners' vocabulary learning more efficiently. Considering the burden that learners feel when learning English, this way of creating materials based on statistical and objective corpus data is absolutely necessary and it is the role of teachers and material writers to select vocabulary and reduce the learning burden which many learners have.

Our vocabulary learning Web site, VOCABRIDGE, is now available on the internet.

References

Chujo, K. (2003). Eigo shokyuushamuke TOEIC goi 1 & 2 no sentei to sono kouka [Selecting "TOEIC vocabulary 1 & 2" for beginning level students and measuring its effect on a sample TOEIC test]. *Journal of the College of Industrial Technology, Nihon University*, *36,* 27-42.

Chujo, K., Ushida, A., Yamazaki, M., Genung, A., Uchibori, A., & Nishigaki, C. (2004). Bijuaru beishikku niyoru TOEIC-yoo goiryoku yoosei sofutowuea no shisaku (3) [The development of English CD-ROM material to teach vocabulary for the TOEIC test (utilizing Visual Basic): Part 3]. *Journal of the College of Industrial Technology, Nihon University*, *37*, 29-43.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology*, *104*, 268-284.

Ellis, R. (1995). Modified oral input and the acquisition of word meanings. *Applied Linguistics, 16*, 409-441.

Fanselow, J. (1992). *Try the opposite*. Tokyo: Simul Press.

Fraser, C. A. (1999). Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition, 21*, 225-241.

Gardner, D. (2004). Vocabulary input through extensive reading: A comparison of words found in children's narrative and expository reading materials. *Applied Linguistics, 25*, 1-37.

Grabe, W. (1991). Current developments in second language reading research. *TESOL Quarterly, 25*, 375-406.

Kennedy, G. (1987a). Expressing temporal frequency in academic English. *TESOL Quarterly, 21*, 69-86.

Kennedy, G. (1987b). Quantification and the use of English: A case study of one aspect of the learner's task. *Applied Linguistics, 8*, 264-286.

Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know, words you can't guess. In J. Coady and T. Huckin (Eds.), *Second language vocabulary acquisition* (pp.20-34). Cambridge: Cambridge University Press.

Laufer, B., & Nation, I. S. P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics, 16*, 307-322.

Nagy, W. (1997). On the role of context in first- and second-language vocabulary learning. In N. Schmitt and M. McCarthy (Eds.), *Vocabulary: Description, acquisition*

and pedagogy (pp. 64-83). Cambridge: Cambridge University Press.

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York: Newbury House Publishers.

Nation, I. S. P., & Newton, J. (1997). Teaching vocabulary. In J. Coady and T. Huckin (Eds.), *Second language vocabulary acquisition* (pp.238-254). Cambridge: Cambridge University Press.

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

O' Dell, F. (1997). Incorporating vocabulary into the syllabus. In N. Schmitt and M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp.258-278). Cambridge: Cambridge University Press.

Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady and T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy vocabulary* (pp.174-199). New York: Cambridge University Press.

Qian, D. D. (1996). *ESL vocabulary acquisition*: Contextualization and Decontextualization. *The Canadian Modern Language Review*, *53*, 120-142.

Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning, 45*, 283-331.

Rott, S. (1999). The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition through reading. *Studies in Second Language Acquisition*, *21*, 589-619.

Utiyama, M., Tanimura, M., & Isahara, H. (2004). Constructing English reading courseware. *PACLIC-18*, 173-179.
from http://www2.nict.go.jp/jt/a132/members/mutiyama/pdf/reading-paclic18.pdf

Sources for World Wide Web

VOCABRIDGE
http://www.kotonoba.net/~mutiyama/vocabridge/

The Daily Yomiuri database purchase
http://ndk.co.jp/yomiuri/kijideta/guidance/index.html

Appendix

JACET Rank	Words	Lists	Texts
0	DEAL	×	○
0	AHEAD	×	○
0	TOTAL	×	○
0	DIRECTOR	×	○
0	GUIDE	×	○
1000	SECURITY	×	○
1000	QUARTER	×	○
1000	NEARBY	×	○
1000	CHAIN	×	○
1000	AIRPORT	×	○
2000	SPECIALIST	×	○
2000	USER	×	○
2000	PRIORITY	×	○
2000	CONSULTANT	×	○
2000	WEEKLY	×	○
3000	REQUIREMENT	×	○
3000	CONVERT	×	○
3000	ALLOWANCE	×	○
3000	INADEQUATE	×	○
3000	ANALYST	×	○
0	PUBLIC	○	×
0	DESIGN	○	×
0	COMMUNITY	○	×
0	SPREAD	○	×
0	FREEDOM	○	×
1000	BREATH	○	×
1000	DESPITE	○	×
1000	IMPRESSION	○	×
1000	PERSUADE	○	×

JACET Rank	Words	Lists	Texts
1000	DISTANT	○	×
2000	AFTERWARD	○	×
2000	COMPOSE	○	×
2000	EMPHASIZE	○	×
2000	LITERALLY	○	×
2000	AWKWARD	○	×
3000	ABUSE	○	×
3000	EXPENDITURE	○	×
3000	ACCUSE	○	×
3000	GENUINE	○	×
3000	COMPROMISE	○	×
0	FIGURE	○	○
0	INDIVIDUAL	○	○
0	EXPRESS	○	○
0	CROWD	○	○
0	OPERATION	○	○
1000	CHARGE	○	○
1000	RESPONSIBLE	○	○
1000	EXPORT	○	○
1000	ORGANIZE	○	○
1000	OBTAIN	○	○
2000	TRANSFER	○	○
2000	APPLICATION	○	○
2000	URBAN	○	○
2000	APPROVE	○	○
2000	PARTICIPATE	○	○
3000	PROCEDURE	○	○
3000	REASONABLE	○	○
3000	WITHDRAW	○	○
3000	APPROXIMATELY	○	○
3000	ANTICIPATE	○	○