The Effect of Adapted Readings on High-frequency Vocabulary Development - A Pilot Test

ドールトン　フランク
Frank E. DAULTON

キーワード

［語彙］vocabulary　・［読書］reading

Introduction

The *high-frequency vocabulary* is the basic English for communication. University English teachers should reinforce the high-frequency vocabulary previously learned and introduce items that have been overlooked. One of my classes at Ryukoku University focussed on the most-frequent 2000 words of English in an adapted version of C. S. Lewis' (1950) *The Lion, the Witch and the Wardrobe.* In this paper, I summarize the importance of vocabulary and particularly high-frequency vocabulary. Next I discuss how teachers can adapt authentic materials for their students. Finally, in a pilot study I attempt to evaluate my students' initial knowledge of high-frequency words, and how successful my teaching materials had been. Although the results were inconclusive, they indicated that a second round of testing was justified.

Literature Review

A brief literature review will consist of three parts: the importance of vocabulary in EFL; the importance of high-frequency vocabulary in particular; and a few legal notes on adapting texts.

The Importance of Vocabulary in EFL

A solid vocabulary is necessary in every stage of learning, and communicative competence is heavily based on lexicon. As the maxim goes, without grammar, very little can be communicated, but without vocabulary, nothing can. Nation & Waring (1997) describe how lexis is a driving force in language acquisition:

> Vocabulary knowledge enables language use, language use enables the increase of vocabulary knowledge, knowledge of the world enables the increase of vocabulary knowledge and language use and so on (p. 7).

Conversely, learners indicate that lexical problems are their greatest single source of difficulty in the L2 (Meara, 1984), and native speakers faced with learner errors normally rate them as more disruptive and more serious than grammatical errors (Johannson, 1978).

For more on the crucial role of vocabulary in L2 acquisition, see: Coady & Huckin (1997); Huckin et al. (1993); Schreuder & Weltens (1993); Schmitt & McCarthy (1997). For more in the Ryukoku International Center Research Bulletin, see Daulton (2004).

The Importance of High-frequency Vocabulary in EFL

High-frequency vocabulary are the relatively small group of words that, because of how frequently they are used, cover a very large proportion of the running words in written texts and spoken English. With only a limited number of high-frequency words, it is possible to understand a large percent of words one will meet. Regarding high-frequency words, Nation (1990) writes:

> Any time spent learning them will be well repaid because they cover a lot of text and will be met often (p. 14).

For instance, knowledge of the most frequent 2000 word families (i.e. headword plus inflections and regular derivations) provides most of the vocabulary needed for basic, everyday oral communication (Schonell et al., 1956; cited in Schmitt et al., 2001 p. 55). The next 1000 word families provide additional material for spoken English and allow the learner to comfortably read authentic texts. Ideally, learners will understand 95 percent of what they read. At this proficiency, only one in twenty words are not comprehended, and the ratio of known to unknown words is sufficient to allow effective guessing (Nation, 1999 p. 120). In order to achieve such good comprehension of informal conversation and perhaps even some academic texts – as well as to have a well-rounded arsenal for production – Nation (1993) believes that at least 2000 but preferably 3000 word families are required. However, if one considers proper nouns in a text to be knowable to readers from context, then 2000 word families might be sufficient.

Especially for low-level learners, high-frequency words are important. West et al. (1934) write:

> In the earliest stage we need the maximum power of expression possible with the smallest number of words (p. 51).

They add that these words must be mastered to the point of ease and fluency (p. 52).

One approach to dealing with the high-frequency word families is the creation of word lists, as a careful choice of vocabulary can simplify teaching. There are many lists of the most frequently occurring words in English (see Nation & Waring, 1997). The most well-known and highly-esteemed was developed in the 1940s. It set a high standard for all lists to follow, and was so successfully compiled that it compares well to even the most modern of lists. It is West's (1953) *A General Service List of English Words* (GSL), which contains the approximately 2000

word families most frequently found in English text. It can be used as a resource for compiling simplified reading texts; older series of graded readers are based on this list. Moreover, the Vocabulary Levels Test, a computer-based text analysis tool, utilizes the GSL.

For more on high-frequency vocabulary appearing in *The Ryukoku Journal of Humanities and Sciences*, see Daulton (2003a).

Adapting Stories

Vocabulary (as well as grammar, etc.) can be presented in stories, which provide excellent stimulation for learners. In contrast to, for example, typical four-skills communication textbooks, stories include novelty, humor, conflict and surprise, which help students maintain interest and learn from context (Elley, 1989). As they encourage learners to focus on meaning rather than form, stories promote deep processing of input. Another benefit of stories is that they extend the learner's knowledge of their world.

Many publishers offer series of graded readers, for example, Oxford and the Oxford Bookworms series. Such materials are well suited for extensive reading – learners have access to a great range of topics and levels, and they are generally well designed and reinforce previously met vocabulary. However, using graded reader series requires more than a sufficient number of books for all students as well as space to store these books. For some classes and programs, this is not practical.

As an alternative to commercial materials, teachers can adapt authentic materials themselves to the level or pedagogical needs of students. For years, I have designed classes around materials of my own choosing, creating complementary videos, quizzes, discussion questions, short lectures, and so on. There are various advantages to adapting materials oneself. The teacher assumes full responsibility over students' progress. The teacher can make adjustments to teaching materials during the year. Moreover, while buying sets of graded readers can be quite expensive, adapting materials oneself is extremely economical, although it requires a great deal of time and labor.

The books I choose to adapt are the ones that have influenced my life so much that I want to share them with students. One is *The Lion, the Witch and the Wardrobe* by C. S. Lewis (1950). It is both an entertaining fantasy and a great literary work dealing with themes such as justice and love. {It is interesting to note that this story is scheduled to be released around Christmas in 2005 as a major motion picture by Disney, directed by Andrew Adamson (director of "Shrek" and "Shrek 2" and filmed in New Zealand (http://www.narnia.com).}

Naturally in adapting texts, one must keep in mind the intellectual property (I.P.) laws. In the case of Japan, the relevant laws are quite lax. According to the Copyright Research and Information Center, as a signatory to the Berne Convention since 1975, Japan protects foreign works up to 50 years from the death of the author. However, this is not retroactive, and foreign works published prior to 1975 are *in the public domain*. Therefore, teachers have a free hand in adapting and using older works. (Ironically, Japanese works are protected in other signatory countries by the much stricter copyright laws that prevail locally.)

The Study

Given the importance of vocabulary in general, and of high-frequency vocabulary in particular, one must identify effective means of developing learners' lexicons. I attempted to evaluate the efficacy of the C. S. Lewis reading materials I had adapted. This was done through: a vocabulary pre- and post-test over the course of a year; a comparison to the attendance trends in the previous year – when I was still using a commercial, four-skills textbook; and a questionnaire regarding students' compliance with homework.

The Students

The class in question was English lG/H, sometimes known as the "Output" class. Students were in the *Senshu-gata* sequence, where they had elected to study only English in there three foreign language classes each week. Their level was the lower of the two possible levels in the streaming system for *Senshu-gata*. There were 35 students in the class at approximately a 4:1 male to female ratio.

The level of motivation among the students varied widely. Indeed, some students had apparently chosen the option of three English classes -- as opposed to beginning another language – not because of their interest in English but because of their lack of interest in foreign languages in general. In response to this obstacle, I abandoned the commercial, four skills textbook I had previously used and began relying entirely on self-prepared materials.

The Reading Corpus

Each week, students read one chapter from the 17 chapters I had adapted from C. S. Lewis' (1950) *The Lion, the Witch and the Wardrobe.* I had shortened the original chapters to fit onto two A4 pages, and they averaged about 1000 words. I had replaced vocabulary that was beyond the first 2000 words of English as described by West's (1953) GSL. Most of the vocabulary was at the first-1000 level of the GSL with a spattering of words at the second-2000 level. The only exceptions were unavoidable words such as wardrobe, which were glossed with Japanese translations, and words corresponding to common English-based loanwords in Japanese such as balcony (*barukonii*), which were likely to have low learning burdens.

{To see my list of common loanwords corresponding to the first 3000 words in the British National Corpus, see Daulton (2003c). For how I collected that list appearing in *The Ryukoku Journal of Humanities and Sciences*, see Daulton (2003b)}.

The resulting reading contained 2773 word families, 3525 types (distinct words), and 17,961 tokens (occurrences of words, even repeated). The VocabProfile (VP) computer program, which performs lexical text analysis, was employed to examine this corpus by frequency level. The web version of the program (WebVP), which I utilized, is based on the original, off-line program (see Laufer & Nation, 1995), which is known as "Range" in its latest version. VP can divide any text into categories by frequency including: (1) the most frequent 1000 words of

English (K1); (2) the next thousand words of English (K2); (3) the 570-word Academic Word List (AWL; Coxhead, 1998); and (4) the remainder. The K1 and K2 vocabulary is based on West's (1953) GSL. The details of the reading corpus are summarized below in Table 1:

Table 1: Levels of Vocabulary, Number of Families, Number of Types, Number of Tokens, and Percent of Text of Reading Corpus

vocabulary level	types	tokens	percent
K1	1138	15,304	85.08%
K2	443	987	5.49%
AWL	13	35	0.19%
Off-List	462	1662	9.24%
total	2055	17,988	100%

The vast majority of words were at the K1 level -- 85.08 percent. A much smaller proportion was at the K2 level, merely 5.49 percent; these words were the target words for quizzes and tests conducted during the year. Off-list words were slightly more common at 9.24 percent. They included: frequently repeated proper nouns, such as character names (e.g. Peter and Aslan) and place names (Narnia and Cair Paravel); names of mythical creatures (e.g. fauns and centaurs); low-frequency words (e.g. witch and beaver); and unrecognized words (e.g. hilltop and ugh). The overall type-to token ratio was 0.11; on average, each word was used 8.75 times.

K1 words are not over-represented in the readings. Various research (e.g. Francis & Kucera, 1982; Hirsh & Nation, 1992) shows that the K1 words are by far the most common in English text. Indeed, the K1 words include virtually all of the function words. Incidentally, Daulton (2003d) found that in some Ryukoku university students' writing, 86.4 percent of words were at the K1 level and 3.7 percent at the K2 level.

Moreover, calculation shows that the type/token ratio is 0.16 at the K1 level and 0.65 at the K2 level. That is, the average K1 type was used about six times and the average K2 type was used about 1.6 times. In other words, while K1 tokens are about 15 times more common than K2 tokens, when considering individual words, K1 types are only 2.5 times more common than K2 types. K1 types are simply repeated more often than K2 types.

Within the first 2000 word families of English (e.g. K1 and K2), 962 word families were represented in the reading corpus. This data is summarized in Table 2:

Table 2: Levels of Vocabulary, Number of Families and Number of Types

vocabulary level	word families	types
K1	644	1138
K2	318	443
total	962	1581

On average, K1 word families were represented by 1.8 different words in the readings, and those at the K2 level by 1.4 words. In other words, average K1 word families were encountered 13.5 times in 1.8 forms (e.g. derivations and inflections) by students in the reading, and K2 families 2.2 times in 1.4 forms.

It can be noted that the Flesch Reading Ease, which is calculated using the average sentence length and the number of syllables per hundred words (Nisbet, J.D., 1960), was 88.

Various devices supported the homework reading assignments. An audio recording of the readings – in my voice – was provided to each student. Moreover, following each reading assignment, a crossword puzzle that focussed on ten of the K2 words of each chapter was administered. Four times over the year, blank-filling vocabulary tests with Japanese prompts were given, using 20 words drawn from the preceding 40 words that had appeared in puzzles. Because K2 words were plentiful but not often repeated in the texts, students had been given instructions to study the words with which they were unfamiliar. To further encourage reading, quizzes on the content of each chapter were administered. To dissuade students who had neglected a reading assignment from falling behind and dropping out, a video version of that week's reading assignment was shown to the class after all quizzing had finished. For students who had read the homework, the movie version – with English subtitles – reinforced their comprehension.

The Instrument

The Vocabulary Levels Test -- also known as the "Levels Test" -- was used to measure the students pre- and post-ability on both the K1 and K2 levels of English vocabulary. The Levels Test is divided into a number of distinct frequency levels. As the 2000 level describes the basic and essential vocabulary for aural communication, the corresponding test was used in this study. The Levels Tests sample words at a given level in order to create a test that is quick to take, easy to mark and easy to interpret. In contrast to subsequent levels of the Levels Test, which contain one level of vocabulary, the 2000 level contains two types of words – both K1 and K2 vocabulary, sampled at a 1:2 ratio.

As a diagnostic test, the Levels Test's main purpose is to allow teachers to quickly

ascertain whether their learners need to be working on high-frequency words, or can move on to the various types of low-frequency words -- such as academic vocabulary and technical vocabulary. As a multiple choice test, it gives credit for partial knowledge of a word. Target items are given in clusters of three, with six choices (three correct answers and three distractors) for each cluster. The original Levels Test was created by Paul Nation (Victoria University of Wellington). It first appeared in 1983 and was later republished in Nation's (1990) book. Since then, various versions have been employed around the world in both assessment and research (e.g. Cobb, 1997; Schmitt and Meara, 1997; Laufer and Paribakht, 1998). Schmitt, et al. (2001) found that the Levels Test provides accurate estimates of the vocabulary size of students at the targeted frequency levels. For a fuller description of the test, see Nation (1990).

During the second administration of the Levels Test, students were asked how regularly they had done the reading homework. They were to write a circle if they had completed 90 percent of assignments or better, a triangle if they had completed 60 but less than 90 percent, and an "X" if they had done less. The results of students who had not faithfully complied with the homework (i.e. 90 percent or better) would have to be excluded from the pre-test / post-test comparison. Moreover, if large numbers of students had not faithfully done the homework, this would be considered in future curriculum planning.

The Data

The pre-test was given to 29 students (six students were absent that day) on the second class of the year. "The Revised 2,000 Word Level Test Form A," which contains 27 items in nine three-item sets was used. Students were given 10 minutes to complete the test. Their performance is summarized in Table 3:

Table 3: Pre-test Data -- Average, Median, Low, High Scores by Number Correct and Percentage

	# Correct out of 27	Percent
Average	22.8	84.6%
Median	24	88.9%
High	27	100.0%
Low	14	51.9%

Students' average accuracy was 22.8 out of 27 (84.6 percent), with the median being 24 (88.9 percent). This indicates that most students already had a good command of the basic English vocabulary, albeit this competence may be "passive" rather than "active" that is, students

may recognize high-frequency words when encountered but not necessarily be able to produce them. The high score was a perfect score and the low score was 14 (51.9 percent). The largest cluster of scores was at 26 out of 27 correct (96.3%). Seven of 29 students had this score. As the particular item these seven students had incorrectly answered varied, there was no faulty test item.

During the school year students read one chapter each week, eight chapters in the first semester and eight in the second semester. One chapter remained after the post-test; the post-test was administered early due to publishing deadlines.

Towards the end of the second semester, when 16 of the 17 chapters had been completed, the class was given a comparable post-test – "The Revised 2,000 Word Level Test Form B," which contains different items. Of the 29 students who sat for the original test, 22 were present. Unfortunately, of these students, only 8 reported that they had completed more than 90 percent of the reading assignments. (Three others reported completing more than 60 but less than 90 percent.) The performance of the 8 students (6 boys and 2 girls) who had both sat for the pre-test and had completed 90 percent or more of the reading assignments is presented below alongside these individual's average pre-test results:

Table 4: Post-test Data -- Average, High and Low Scores for Eight Students' Pre- and Post-tests

	Pre-test # Correct out of 27	Pre-test Percent	Post-test # Correct out of 27	Post-test Percent
Average	23.4	84.6%	24.9	92.2%
High	26	96.3%	26	96.3%
Low	15	55.6%	23	85.2%

We can see an increase in the average score, from 23.4 (85.7 percent) correct in the pre-test to 24.9 (92.2 percent) in the post-test. The high score remained the same while one student's (low) score rose from 15 (55.6 percent) to 24 (88.9) – no longer the low score. If this student's possibly anomalous scores were to be removed, overall progress among the remaining seven students would be even harder to detect. Indeed, the combination of absences on the day of the test and particularly that so few students faithfully completed the homework resulted in a very small sample for this comparison.

Further Analysis

In addition to the data regarding the development of high-frequency English, it was possible to examine two additional factors of importance.

Attendance

Possibly an indication of the success or failure of my new curriculum could be found in a comparison of the attendance tendencies of the subjects' class and the 2003 class, which had used the commercial, four-skills textbook. Unfortunately, this data provided mixed results. This data is summarized in Table 5:

Table 5: Attendance Tendencies -- Year and Semester, Percent of Dropouts, Percent with Perfect Attendance

year and semester	percent of dropouts	percent perfect attend
2003 1st	8.1%	37.8%
2003 2nd	14.7%	29.4%
2004 1st	8.9%	38.2%
2004 2nd	18.8%	35.5%

"Dropouts" are students who had attended classes at the beginning of the semester but subsequently dropped out without explanation or exceeded the allowable number of absences (one third of classes) and had been disqualified from completing the class. We see that in both years, dropouts were more common in the second semester, and we see that this percentage was higher in 2004 than in 2003.

"Perfect attendance" is for students who had attended every class during the semester, although they may have been tardy. These figures are better in the first semester than the second, and the 2004 year's results are better than those of 2003. That is, while fewer students dropped out in 2003, more students had perfect attendance in 2004. This may indicate that the homework, which was likely more challenging than in the previous year, had a polarizing effect – tending to overwhelm unmotivated students while engaging motivated ones.

Compliance with Homework Assignments

As mentioned above, it appears that only about a third of the students faithfully complied with the reading assignments. As this was a major component of the class, this result is troubling. However, although no data is available on how well students complied with the textbook

homework assignments in 2003, a similar proportion is likely to be found.

Conclusions

Okamoto (forthcoming) has reported the dramatic decline in vocabulary knowledge of first and second year university students over a school year in groups of English non-majors. With such a regrettable state as a possibility, while my students had not made uncontestable progress during the year, neither had their vocabulary apparently deteriorated. Of course, the effects of my students' other English classes and outside exposure can not be overlooked.

The average scores of my students' on the two Levels Tests were relatively good. This indicates that most students would likely benefit from fluency work and are ready to focus on the less common vocabulary, such as academic and technical vocabulary, and to develop other skills such as guessing from context for low frequency words.

There is room for improvement in the reading materials. Only about half of the GSL word families were to be found in the homework readings. K2 words were particularly few. It is likely that the reading homework would have been more effective if somehow all of the families had been used, however it is not clear whether this is a practical possibility. And while the repetition of K1 word families was rather good at an average of 14 times – likely to be sufficient for incidental learning (e.g. Tinkham, 1993) – the repetition for K2 families was relatively low at about 2 times.

Thus more valuable data may have been obtained had an examination of only the K1 words of The Revised 2,000 Word Level Tests been included. Although this would limit the number of test items, Okamoto (forthcoming) has shown that the Levels Tests can be shortened even further without affecting their accuracy. The drawback of this procedure is that students who had perfect or near perfect scores in K1 vocabulary on the pre-test would have no room for improvement; yet examining only the weaker students (based on pre-test data) might be difficult as these students are the least likely to complete the semesters and homework assignments.

In conclusion, it was not possible to conclusively evaluate the success of the adapted reading homework and the curriculum designed around it. Improvement could neither be demonstrably shown in vocabulary size nor classroom attendance. In particular, students must be better encouraged to comply with homework, thus creating a larger group for test-result comparison. It is warranted to conduct a second study in the 2005 school year.

References

Coady, J. and Huckin, T. (eds.; 1997). Second Language Vocabulary Acquisition. Cambridge University Press, New York.

Cobb, T. (1997). Is there any measurable learning from hands-on concordancing? System 25: 301-15.

Copyright Research and Information Center.
http://www.cric.or.jp/cric_e/

Coxhead, A. (1998). An Academic Word List. Occasional Publication Number 18, LALS, Victoria University of Wellington, New Zealand.

Daulton, F. E. (2003a). The high-frequency vocabulary of English. The Ryukoku Journal of Humanities and Sciences 24, 2: 19-42.

Daulton, F. E. (2003b). Common gairaigo cognates for the high-frequency vocabulary of English. The Ryukoku Journal of Humanities and Sciences 25, 1: 51-78.

Daulton, F. E. (2003c). List of high-frequency baseword vocabulary for Japanese EFL students #2. The Internet TESL Journal 9, 3.
<http://iteslj.org/lists/Daulton-BasewordVocab2.html>

Daulton, F. E. (2003d). The effect of Japanese loanwords on written English production – a pilot study. JALT Hokkaido Journal 7: 4-14.
<http://journal.jalthokkaido.org/index.htm>

Daulton, F. E. (2004). The importance of the high-frequency 'built-in lexicon' for EFL in Japan. Ryukoku International Center Research Bulletin 13: 69-80.

Elley, W. B. (1989). Vocabulary acquisition from listening to stories. Reading Research Quarterly 24: 27-43.

Francis, W. N. and Kucera, H. (1982) Frequency Analysis of English Usage. Houghton Mifflin Company, Boston.

Hirsh, D. and Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? Reading in a Foreign Language 8, 2: 689-696.

Huckin, T., Haynes M., and Coady, J. (eds.) (1993). Second Language Reading and Vocabulary. Ablex Publishing Corporation, Norwood, NJ.

Johannson, S. (1978). Some Aspects of the Vocabulary Learned and Scientific English. Gothenburg Studies in English, Acta Universitatis Gothoburgensis: 42.

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

Laufer, B. and Paribakht, T.S. (1998). The relationship between passive and active vocabularies: effect of language learning context. Language Learning 48: 365-91.

Meara, P. (1984). The study of lexis in interlanguage. In A. Davies, C. Criper, and A. Howatt (eds.), Interlanguage Edinburgh University Press, Edinburgh: 225-235.

Nation, I.S.P. (1990). Teaching and Learning Vocabulary. Heinle & Heinle Publishers, Boston.

Nation, I. S. P. (1993). Vocabulary size, growth and use. In R. Schreuder and B. Weltens

(eds.). The Bilingual Lexicon. John Benjamins, Amsterdam/ Philadelphia: 115-134.
Nation, I. S. P. (1999). Learning Vocabulary in Another Language. English Language Institute Occasional Publication No. 19. Wellington.
Nation, I. S. P. (2001). Learning Vocabulary in Another Language. Cambridge University Press, Cambridge.
Nation, P. and Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt and M. McCarthy (eds.). Vocabulary: Description, Acquisition and Pedagogy. Cambridge University Press, Cambridge: 6-19.
Nisbet, J. D. (1960). Frequency counts and their uses. Educational Research 3: 51-64.
Okamoto, M. (forthcoming). University Students' Lexical Acquisition and Attrition in English as a Foreign Language.
Schreuder, R. and Weltens, B. (eds.; 1993). The Bilingual Lexicon. John Benjamin Publishing Company, Amsterdam/ Philadelphia.
Schmitt, N. and Meara, P. (1997). Researching vocabulary through a word knowledge framework: word associations and verbal suffixes. Studies in Second Language Acquisition 9: 17-36.
Schmitt, N., Schmitt, D., and Clapham, C. (2001). Developing and exploring behaviour of two versions of the Vocabulary Levels Test. Language Testing 18, 1: 55-88.
Schonell, F., Meddleton, I., Shaw, B., Routh, M., Popham, D., Gill, G., Mackrell, G., and Stephens, C. (1956). A Study of the Oral Vocabulary of Adults. University of Queensland Press / University of London Press, Brisband and London.
Tinkham, T. (1993). The effect of semantic clustering on the learning of second language vocabulary. System 21, 3: 371-380.
West, M. (1953). A General Service List of English Words. Longmans, Green and Co., London.
West, M. P., Swenson, E. et al. (1934). A critical examination of basic English. Bulletin of Department of Educational Research, University of Toronto: 2.