



Making the List: A Vocabulary Audit

Neil Millington and Joe Siegel

Abstract:

This paper reports on results of a pre- and post-quarter receptive vocabulary test that was conducted at Ritsumeikan Asia Pacific University during the second quarter of the Spring 2009 semester. The aim of the receptive vocabulary test was to ascertain levels of vocabulary attainment by learners in the Fundamental English II program in an effort to develop the current design and practice of the vocabulary component of this course. The paper begins by reporting on literature and research related to the fundamentals of vocabulary instruction. Then the research methods are described, after which the test results are reported. The paper reports results of the pre-quarter receptive vocabulary test that suggest more consideration to words on vocabulary lists maybe necessary. An increase in receptive lexical recognition scores on the post-quarter receptive vocabulary test allows the researchers to tentatively state that time devoted to in-class vocabulary coupled with self study outside of class leads to learner vocabulary gain. The paper concludes by suggesting periodic auditing of vocabulary lists would be beneficial in order to align student and teacher effort and time with vocabulary goals.

Key terms: Receptive vocabulary acquisition, lexical selection, facility value

Introduction

Vocabulary acquisition pertaining to English for Academic Purposes (EAP) continues to be a priority for a majority of university English students in Japan. The numerous TOEFL- and TOEIC-oriented vocabulary books on the EFL market as well as institution- and course-specific vocabulary lists attest to this fact. Vocabulary remains crucial to success on many EFL assessments, whether decontextualized terminology quizzes or more general tests of English ability. As Lewis, quoted in Sokmen (1997), states: “acquisition of vocabulary has assumed a more important role, and some would argue, the central role in learning a second language” (p. 237). Therefore, it is with good reason that teachers emphasize the importance of lexical development and that students invest valuable time and effort learning new words and preparing for corresponding evaluations.

Several complex issues converge in the design and implementation of a vocabulary teaching component. Among them are the initial selection and subsequent auditing of items taught and the pedagogical approaches used in classrooms. Many word lists are likely prepared before a course begins, a procedure that may limit any chance of diagnosing and analyzing students’ existing lexical knowledge. The risk of this approach is that words already known by a majority of learners will be introduced, thereby misappropriating time that may be better spent on new material, vocabulary or otherwise. Similarly, it is challenging for teachers to select activities that will effectively and efficiently produce desired outcomes of item acquisition and retention.

In an effort to build on current design and practice of the vocabulary component of the Fundamental English II (FE2) course at Ritsumeikan Asia Pacific University (APU), this paper reports on results of a pre- and post-quarter receptive vocabulary test (RVT). Under investigation is the hypothesis that scores on the 50 item pre- and post-quarter RVT will be higher at the end of the quarter than those at the beginning of the quarter. An increase would tentatively indicate that class time and in-class vocabulary practice are constructive towards course goals, which according to the Teacher's Handbook, include "developing [students'] vocabulary knowledge" (p. 3). Additionally, data from the pre-quarter RVT help to illustrate students' existing vocabulary knowledge. This information can affect future planning and instruction.

This paper begins with a brief review of literature relating to vocabulary selection and teaching, after which the design and administration of the RVT and the context of the project will be described. Next, pre- and post-quarter results as well as in-class instruction will be detailed. Data from the beginning and end of the term will be compared to determine what effect if any in-class instruction had on student receptive lexical knowledge. After discussing limitations of this study and future implications, this paper will close with a cautious yet encouraging conclusion that in-class vocabulary teaching does have a positive effect on word acquisition and retention in FE2 classes at APU.

Vocabulary instruction fundamentals

Literature and research related to the topic of this paper can be divided into two main themes: selection of words and the subsequent teaching thereof. The following review briefly highlights key matters pertaining to each theme beginning with selection.

The selection of certain lexical items for inclusion in a course or on a vocabulary list is most often the responsibility of a syllabus designer on a broad scale or of an individual teacher on a more localized scale. Concerned individuals must decide whether their selection process will be objective or subjective. In other words, item selection can be made using tools such as word lists and corpus studies or based on native speaker intuition or a combination thereof. Gairns & Redman (1986) recognize the responsibility educators have when determining lexical content. Assuming that students will invest time and effort attempting to acquire these lexical items, in addition to being evaluated on vocabulary tests, the creation of vocabulary lists should be taken as an important duty that can have profound implications.

An important starting point for any lexical selection process is deciding whether the items need to be accessible to students in either a receptive or productive capacity or in some cases, both. Receptive vocabulary knowledge should be taken to mean that students understand the word or phrase when listening and reading. Meanwhile, productive ability entails being able to present the item through speaking and writing. Further issues in lexical selection are prescribed by White and include the following: frequency, coverage, range, availability, learnability, opportunity and centers of interest (cited in O'Dell, 1997).

Assembling a list of vocabulary words is by no means a simple task. In an effort to negate the often precarious subjectivity related to native speaker intuition, syllabus designers and teachers can rely on several word lists that have been created for just such a purpose. West's *General Service List* (1953) is often identified as the original word list and is praised as having had "by far the most lasting influence" (Carter, 1998, p. 206). More recently, several other word lists have begun to impact various EFL contexts:

- Van Ek et al. (1997) created a list for the Council of Europe
- Hindmarsh (1980) made a list based on Cambridge EFL assessments
- Xue & Nation (1984) compiled a *University Word List*

As the course under consideration for this paper is a university course, it would seem that Xue & Nation's list is the most relevant to inform selection for FE2.

Computerized corpora have become instrumental in creating word lists. Corpora, with their massive stores of linguistic data, can help inform those who create vocabulary components by providing valuable details of word frequency and usage. As O'Dell (1997) observes, corpora "allow more confident decisions on which frequent vocabulary to include" (p. 261). However, O'Dell (1997) points out one drawback of a corpus approach to vocabulary selection: the raw frequency data provided by corpora does not indicate which meaning of a word is used most regularly. Therefore, sole reliance on frequency data is insufficient and some aspect of subjectivity must be included in the process. It is likely that corpora can aid syllabus designers and teachers in narrowing a list of frequent words and phrases. It might then be left to the individuals to make decisions for their specific target audience and purposes.

Once words are chosen for inclusion, teachers need to address pedagogical concerns. It is important for educators to recognize that learning vocabulary involves multiple aspects. In fact, Richards (1976) and Nation (1990), cited in Schmidt & McCarthy (1997), identify eight characteristics involved in "knowing" a word: spelling, pronunciation, grammatical behavior, associations, collocations, meanings, frequency, and register. Likely only native or near-native speakers will be able to satisfy all eight of these characteristics in a timely fashion. Indeed, more modest expectations should be placed on university EFL students, who may only be fulfilling graduation requirements.

According to Schmitt & McCarthy (1997), Melka's (1997) description of vocabulary acquisition as an "incremental process...[and that] partial knowledge may be the norm" (p. 108) acknowledges that some of the facets mentioned above are more accessible to learners than others. As learners spend more time engaged with new words, they will probably acquire more knowledge about individual items. It has been suggested that seven or more meetings with new words will produce what McKeown et al., cited in Schmitt & McCarthy (1997), call "ownership" of that word. A charge for educators who teach a lexical component is how to make those encounters unique, stimulating and salient for learners.

Lexical acquisition theory proves a useful starting point from which to design and activate vocabulary work. One theory that may lead to robust vocabulary gain is the Depth of Processing Hypothesis, which states that: "the more cognitive energy a person expends when manipulating and thinking about a word, the more likely it is that they will be able to recall and use it later" (Craik & Lockhart, cited in Schmidt & McCarthy, 1997, p. 3). In other words, if students are encouraged to manipulate words by, for example, transforming word form or by applying words to their own experiences, they have a better chance of retention.

Another notion of how words are learned and retained is the Dual-Coding Theory. According Clark and

Paivio, quoted in Sokmen (1997), this theory states that: “the mind contains a network of verbal and imaginal representations for words” (p. 244). In this case, linking known words and/or images to new items helps to solidify words in memory. Broadly interpreted, these images may include concrete or abstract representations as well as actions and gestures.

From the theories described above, two points emerge as critical to a successful vocabulary component. One is that teachers need to design activities that force students to make cognitive investment in and attachment to selected words. A second point relates to the connectivity of the mental lexicon and to ways it should be focused on and developed in students via in-class proceedings. As Meara (1997) states: “[an] acquisition event consists of the building of a connection between a newly encountered word, and a word that already exists in the learner’s lexicon” (p. 118). Aitchison describes the mental lexicon as being “a network of associations, a web-like structure of interconnected links” (cited in Sokmen, 1997, p. 241). Sokmen (1997) continues: “If EFL students are to store vocabulary effectively, educators need to help them establish those links and build up those associations” (p. 241). Examples of how these theories are applied in practice in FE2 are mentioned in the section *In-class activities* below.

Context

This project was conducted at APU during the second quarter of the Spring 2009 semester. The students involved were all enrolled in the FE2 course, which has entry requirements including a TOEFL score of 400. A majority of the students were Japanese; however, several students from other Asian countries such as China, Korea and Mongolia also studied in FE2. These classes met four times a week for 95 minutes.

The FE2 course includes Computer Aided Language Learning (CALL) and a series of video lectures aimed at improving students’ abilities to comprehend extended exposures to English. Weekly vocabulary lists comprised of 40 words each were also a central focus. Quiz and test scores based on these vocabulary lists equaled 30% of grades. These vocabulary lists and the pedagogical approaches to them form the main motivation for the project described in this paper. The following section outlines the methods used to investigate students’ initial vocabulary knowledge as well as the vocabulary gain over the course of the quarter.

RVT design & administration

Items on the RVT were selected from five weekly vocabulary lists that were to be covered during the second quarter. Eight words were randomly selected from each list, totaling 40 words. Students were asked to indicate by answering “yes” or “no” to the question “Do you know this word?” Students were not asked to demonstrate knowledge of or ability with the presented words.

In an attempt to neutralize any students who might lose motivation and/or fabricate or overestimate their existing knowledge, two countermeasures were inserted in the RVT. First, five “assumed known” words were added: apartment, homework, restaurant, family, and classroom. The designers believed that by adding these words, all students would be able to answer “yes” to at least some of the items, thereby maintaining their interest in the test. Second, five “nonsense words” were added which were meant to offset any students who might have chosen to rush through the test indiscriminately, answering “yes” to all questions. Thus, the 40 authentic vocabulary words, five “assumed knowns” and five “nonsense words” combined to make the 50 item RVT, which was administered via the Blackboard online teaching platform at APU. The same items were used on the

pre- and post-quarter tests.

The first instrument was completed by 253 students, while 254 completed the end of quarter RVT. Data was summarily collected using the same internet-based program. Collected data revealed compelling information both about students' initial vocabulary ability as well as the vocabulary gains they made over the quarter. These results are discussed in the following sections.

Lexical selection process - Vocabulary lists

The vocabulary lists are based on a particular theme and are comprised of 40 words. The lists are broken down into three sections. The first section comprises 10 theme specific words deemed to be common TOEFL words. These words were selected from a TOEFL vocabulary book and were cross-referenced with other words on the list to ensure that there was no overlap. The second section consists of 15 words taken from a content-based video lecture. The final section for each vocabulary list has 15 words selected from in-house interviews taken on campus. The videos that were selected for use in-class were selected based on the range of vocabulary within them. The aim was to have a selection of low and high frequency words. The lists were not cross referenced with any particular wordlist.

Pre-RTV results

The Pre-receptive vocabulary test (Pre-RTV) was conducted on June 15th 2009 with 253 university students studying at the FE2 level. Before taking the test students were given the following instructions:

This test is NOT graded.

Please try to be honest in your answers.

You have 15 minutes to complete this test.

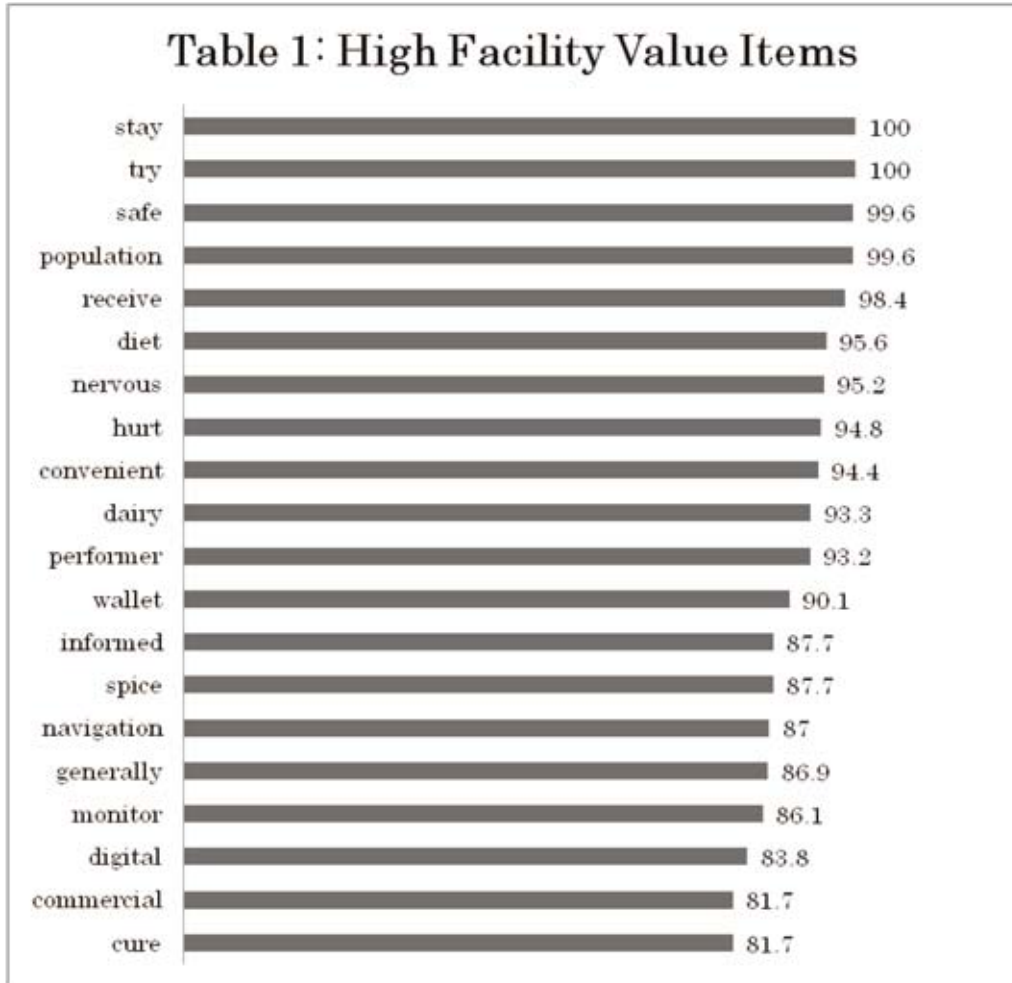
Please do not use your dictionaries.

Do you know the meaning of these words?

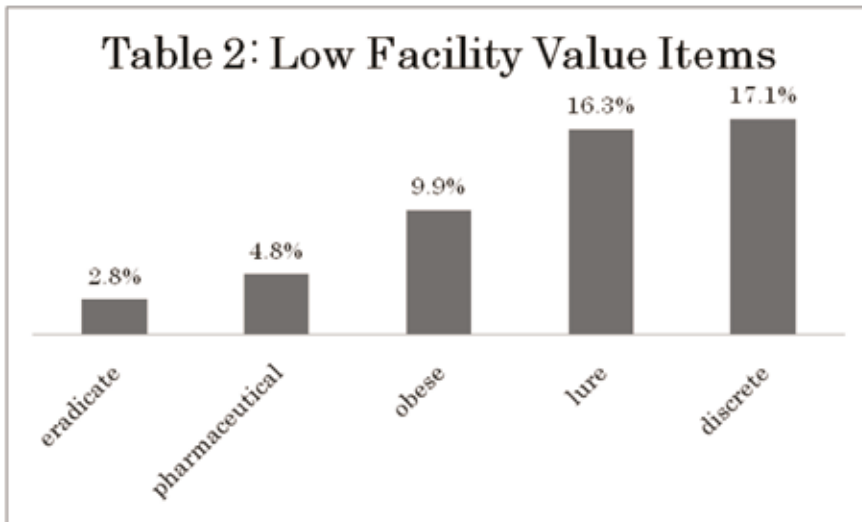
A first glance at the Pre-RTV test results shows that there are a substantial number of items with a high facility value. The facility value measures the difficulty of individual items and shows "the percentage of students to answer the item correctly" (Alderson et al., 1995, p. 80). Usually written as a proportion, this "simple measure immediately gives item writers some idea of how easy the item is" (Alderson et al., 1995, p. 81). The data from the Pre-RVT show that twenty of the forty items had a facility value above 0.8. This demonstrates that over 80% of the students believed they knew these twenty vocabulary words prior to any instruction in the FE2 course (see Table 1). A closer analysis of these twenty items shows that over 90% of students believed they knew 10 of these words, and all 253 students indicated that they knew the words *try* and *stay*.

These results offer a significant insight into FE2 students' existing vocabulary knowledge. With over 80% of students claiming to know 50% percent of the words tested in the Pre-RVT, it would appear that instructors and course designers could offer more consideration to the words on the existing vocabulary lists. One possible option would be to remove these items from the vocabulary lists and to replace them with words that more students are less likely to know. Of course it would be problematic to establish a minimum percentage any certain word must have in order to be included in the course but it is clear that a high percentage of learners feel they know these words prior to being given any instruction. A second option might be to reduce the amount of

time devoted to creating and administering vocabulary activities that focus on these words and allow learners more time to become familiar with other items from the list.



Of the remaining twenty items from the Pre-RVT, there were three items that had a facility value between 0.6 and 0.8, five items that were between 0.4 and 0.6 and a further seven between 0.2 and 0.4. In addition, there were five items that had a particularly low facility value of less than 0.2, indicating that they were very difficult for the group. Less than 20% of the participants thought they knew the words *discrete*, *lure*, *eradicate*, *pharmaceutical* and *obese* (see Table 2).



These results also offer instructors and course designers valuable information regarding the existing vocabulary knowledge of learners on the FE2 course, and such information could allow them to make informed decisions regarding vocabulary instruction. It was reported above that as learners spend more time engaged with new words, they have the opportunity to acquire more knowledge about individual items. It is clear from the results that there are several items that are “new words” for the majority of learners. Having this information prior to instruction could allow course designers to develop activities that would give learners more exposure to and practice with these words. Reported below are the various methods of instruction that teachers used during the second quarter of the Spring 2009 FE2 course.

Methods of instruction

As reported above, vocabulary lists based on a particular theme and comprising of 40 words each were given to students each week. Along with each word, there is a definition and a selection of these words has accompanying synonyms. Students are expected to remember the vocabulary, definitions and synonyms and are tested the following week. According to the FE2 syllabus, quiz and test scores based on these vocabulary lists equaled 30% of quarter grades. Below is a brief description of the series of standardized Vocabulary Activities used in FE2. Each activity was designed taking into consideration the vocabulary acquisition theory reported in the *Vocabulary instruction fundamentals* section above.

Online Interactive Activities

Interactive online activities allow learners to practice vocabulary during and outside of class. There are three games that focus on Definitions (Flashcards, Matching and Crossword) and three other games that focus on using the vocabulary in context (Pick a Letter, Matching and Fill in the Blank). Learners can also access definitions to vocabulary and are given the opportunity to listen to the pronunciation of each word through an MP3 Audio.

In-Class Activities

Vocabulary Conversations

Learners are given opportunities in-class to practice a conversation in pairs. The conversation has been written to include a selection of words from the vocabulary list in context. Gradually certain words are removed from the conversation, encouraging students to recall the vocabulary through memory and context.

Picture Activities

Students are shown a series of pictures and are asked to discuss with their partner which word from their vocabulary list is closest in meaning to each picture. The students are also encouraged to justify their answer by providing an explanation of why they feel that vocabulary word is suitable.

Find Someone Who

A list of different scenarios is given to the students, each sentence containing at least one vocabulary word from the list. Students are encouraged to interact with their classmates to discover whether they have encountered that particular situation.

Information Grids

As above, different scenarios are given to students. Students are encouraged to interact with their classmates to discover whether they have encountered a particular situation. Learners then write their answers and later post them in online discussion forms.

TOEFL Style Readings

There are a series of TOEFL style readings that incorporate vocabulary words for learners to complete each week. Each reading has been written to match materials selection criteria regarding length, context, and grammatical and lexical complexity. Each reading is between 200-250 words with a familiar context for learners. Lexical and grammatical complexity was checked using JACET 8000 and Flesch-Kincaid. The only words outside the ranges suggested by the selection criteria were words from the vocabulary list. The accompanying questions also include vocabulary and synonyms from the vocabulary list.

Post-RVT results

It was reported above that two countermeasures were inserted in the RVT to discover whether students were approaching the test in the manner expected. First, five “assumed known” words were added with the aim of maintaining student interest in the test. Second, five “nonsense words” were added to offset any students who might choose to rush through the test indiscriminately. Henning (1987) advises, “If the examinees do not approach the testing situation in the expected manner, the results may prove to be invalid” (p. 92). The five assumed known words all had a facility value of over 0.9 on both the Pre- and Post RVT. In addition, the five nonsense words all had very low facility values of less than 0.14 on both tests. This shows that the students were indeed approaching the test in the manner expected and the test does have a degree of response validity.

The post receptive vocabulary test (Post-RTV) was conducted at the end of the Spring 2009 semester and 254 students participated. The test was administered under the same conditions as the Pre-RVT and students were given the same instructions.

An analysis of the overall test scores shows increases in receptive lexical recognition. The overall mean score

for the 254 students was 78.93% indicating that the students thought they knew almost 80% of the forty items. This compares with the Pre-RVT data when the overall mean score was only 65.03%. Results also show that students indicated that their receptive knowledge had increased for the majority of items. When the pre- and post-quarter data are compared, results show that more students indicated that they knew 34 of the 40 items in comparison to the Pre-RVT, while three items remained the same and only three items showed a lower score on the Post-RVT.

The number of items that had a considerably high facility value of over 0.8 increased from 20 to 25 and of these 25 words 19 had a facility value of over 0.9. Perhaps more interesting is the fact that none of the forty words had facility value of less than 0.2, and only 4 words had a facility value of between 0.2 and 0.4. In other words there were real increases in receptive lexical recognition in comparison with the Pre-RVT when 12 items had a facility value of less than 0.4. These increases in receptive lexical recognition suggest that the time devoted to in-class vocabulary practice, along with student self-study outside of class, leads to vocabulary gain. A closer analysis of some the individual items follows and provides additional explanation of these overall results.

Comparison between Pre- and Post-RVT

Due to an outbreak of H1N1 influenza strain, classes during the spring semester were suspended for one week. During this week, students were given no in-class instruction and had no access to the in-class materials described above. Students were informed that there would be no weekly quiz on this particular vocabulary list (Media week) but the words from this would be included in their final vocabulary test. This seemingly unfortunate incident provides interesting data to analyze. The Pre-RVT and Post-RVT results are shown in Table 3 below.

Table 3: Media Week Vocabulary Data

Vocabulary	Pre-RVT Scores	Post-RVT Scores	Increase/Decrease
commercial	81.7	82.2	0.5
performer	93.2	91.3	-1.9
vague	34.5	36.2	1.7
dealer	32.5	44.5	12
mutual	28.1	29.2	1.1
wallet	90.1	90.6	0.5
informed	87.7	87.7	0
terrorism	79	78.7	-0.3
Mean Scores	65.85	67.55	

The results show the Pre-RVT mean score for this group of words taken from the Media vocabulary list was 65.85%, similar to the overall mean score of 65.03% on the same test. What is interesting though, is the mean score for the Post-RVT, which was 67.55% in comparison to the overall mean score for this test of 78.93% with only one item showing above a 2% increase. Furthermore, two of the eight items, *performer* and *terrorism* showed a lower score on the Post-RVT. This would suggest that the time devoted to in-class vocabulary practice, along with student self-study outside of class, leads to a greater vocabulary gain in comparison with only self-study.

Items showing an increase

Six items, reported in Table 4 below, showed an increase of over 30% from Pre-RVT to Post-RVT. On the Pre-RVT less than 10% of students indicated that they knew *eradicate*, *obese* or *pharmaceutical*, yet on the Post-RVT these figures had increased by over 40%. In addition, *cellular*, *enclosure* and *lure* all showed increases of over 30%. Each of the words below was used in at least one of the in-class vocabulary activities reported above. One tentative conclusion that could be drawn from this data is that the time devoted to in-class vocabulary practice of these particular words produced some vocabulary gain.

Table 4: Pre- and Post-RVT scores for six items

Vocabulary	Pre-RVT	Post-RVT	Increase/Decrease
eradicate	2.8	52.4	49.6
obese	9.9	57.5	47.6
pharmaceutical	4.8	48.8	44
cellular	36.8	80.3	43.5
enclosure	30.6	64.6	34
lure	16.3	49.6	33.3

Although these results are encouraging, this data also raises some concerns. With the exception of *cellular* each of these words still has a relatively low facility value (less than 0.65), showing that many students believed they did not know these words despite being given instruction. One possible solution to this problem would be to devote more in-class time to the study of these words. It was reported above that a high percentage of learners felt they knew 50% of the words from the Pre-RVT prior to being given any instruction. Perhaps devoting more time to creating and administering vocabulary activities that focus on these more difficult words might lead to an overall greater increase in vocabulary knowledge.

Limitations

During the analysis of the data collected, several shortcomings became apparent. These limitations relate to student responses and the percentages generated from those responses. First, the wording on the RVT asks students if they “know” the designated word. However, as noted above, there are several aspects of “knowing” a word, including orthographic, semantic and syntactic (in Schmidt & McCarthy, 1997). It is unclear to which aspect of the word students were referring. A second limitation relates to the percentages used when categorizing the selected vocabulary items. The use of 20% increments was used at the discretion of the researchers and was arbitrarily selected for convenience. Finally, while the base percentages referred to above provide useful insights into existing vocabulary knowledge, it would be problematic to establish a minimum percentage any certain word must have in order to be included in the course.

Future implications

This study used a RVT in order to ascertain students’ receptive vocabulary knowledge upon entry into FE2 courses. The pre- and post-quarter survey was also used to monitor any receptive vocabulary gain. The data gathered and subsequent reflection should help inform future selection of items to be included on vocabulary lists as well as pedagogical considerations in this course. With regards to vocabulary selection, it is clear that

a general and ongoing audit of course vocabulary lists would help to identify words students already have receptive knowledge of. Eliminating words such as *receive* (98.4%) would make more space and time available for words students have legitimate reasons to study. Moreover, extra pedagogical attention could be paid to those less salient words such as *pharmaceutical* (4.8%). These are words to which students would benefit from additional exposure. Furthermore, consistent use of a pre-quarter or post-quarter RVT could help teachers and course designers to modify vocabulary lists to maximize the vocabulary gain of each specific group of learners.

Conclusions

This paper has reported the findings of a pre- and post-quarter RVT in an attempt to determine levels of vocabulary attainment over a one quarter period at the FE2 level in APU's English program. In addition, the pre-quarter test results offered some insight into students' existing vocabulary knowledge prior to the quarter. Based on the pre-quarter data, it seems more consideration to the words on the existing vocabulary lists is necessary because high percentages of students claimed to know several words prior to their introduction. This fact suggests that other perhaps more advanced items merit inclusion. When the pre- and post-quarter data are compared, it is clear that for many items, receptive vocabulary knowledge did develop, as expected. These increases in receptive lexical recognition suggest that the time devoted to in-class vocabulary practice, along with student self-study outside of class, leads to vocabulary gain. This vocabulary growth will likely transfer to improved receptive abilities and may be recognized in listening and reading skills and test scores.

From a broader perspective, this study strengthens the notion that periodic auditing of current vocabulary lists may be necessary in order to pinpoint items that may or may not warrant inclusion. When such informed decisions are made regarding which lexical items to introduce and teach, class time as well as student and teacher efforts will be better aligned with vocabulary goals.

References

- Alderson, C. J., Clapham, C & Wall, D. (1995). *Language Test Construction and Evaluation*. Cambridge University Press:Cambridge.
- Carter, R. (1998). *Vocabulary: Applied Linguistic Perspectives* (2nd ed.). New York: Routledge.
- Gairns, R. & Redman, S. (1986). *Working with words*. Cambridge University Press: Cambridge.
- Henning, G. (1987). *A Guide to Language Testing*. Cambridge University Press: Cambridge.
- O'Dell, F. (1997). Incorporating vocabulary into the syllabus. In Schmitt, N. & McCarthy, M. (eds.) *Vocabulary: Description, Acquisition & Pedagogy*. Cambridge: Cambridge University Press. Pp. 258-278.
- Schmitt, N. & McCarthy, M. (eds.). (1997). *Vocabulary: Description, Acquisition & Pedagogy*. Cambridge: Cambridge University Press.
- Sokmen, A. (1997). Current trends in teaching second language vocabulary. In Schmitt, N. & McCarthy, M. (eds.) *Vocabulary: Description, Acquisition & Pedagogy*. Cambridge: Cambridge University Press. Pp. 237-257.