Measuring the Effectiveness of a Vocabulary Programme

Jeannette Watts
Jeannette Watts is the Foundations Chair at Dubai Women’s College, Higher Colleges of Technology, UAE. She has previously taught in New Zealand, Ireland, Thailand and the Solomon Islands, and holds a Master’s degree in Applied Linguistics from Victoria University of Wellington. Her academic interests include vocabulary teaching and learning, and assessment.
Reason for the Study

As with most courses preparing students for academic study, the previously existing Higher Foundations English course at Dubai Women’s College included the key component of vocabulary. This is not surprising as vocabulary learning is essential for development in both oral and written language and as such is an essential part of any language programme. Further, high levels of vocabulary are associated with successful reading, being both “predictive and reflective of high reading achievement” (Pikulski & Templeton, 2004, p.1).

It is especially important for Foundations students to have a course focusing on vocabulary because evidence indicates that upon entry they have an inadequate level of vocabulary to meet the demands of academic study. It was decided therefore in 2003 to introduce a vocabulary programme which was based on sound principles of vocabulary teaching and learning, and which directly taught the high frequency words in a systematic manner. As Nation (2005) writes, “principled planning of vocabulary learning is more important than particular [vocabulary teaching] techniques” (p. 5).

It seemed timely after five years of implementation to reflect on the effectiveness of our vocabulary teaching, which is the purpose of this article. It begins with a review of the literature surrounding these principles by examining the five issues of: incidental or deliberate learning, decontextualised or contextualized instruction, the selection and presentation of vocabulary, learning as a cumulative process, and the role of testing in learning. This is followed by an evaluation of the programme in terms of these principles by reporting on a pre- and post-test, and concludes with recommendations.

Literature Review

Incidental or Deliberate Learning?

While many teachers believe learners will simply ‘pick up’ vocabulary by being exposed to it, this is not supported in the literature, at least for learners at foundations level of acquisition. In a review of the literature Nielson observes that “in recent literature dealing with vocabulary acquisition, there can be seen increasing advocacy for explicitly teaching words out of context at an early stage of language acquisition, with more context-based vocabulary learning taking place at later stages of language development” (Nielson, n.d.). Nation supports this by writing that “there has long been substantial evidence that deliberately learning vocabulary can result in large amounts of well retained useable knowledge” (2001, p. 296). Further, vocabulary must be given explicit attention as the natural occurrence of words will not provide learners with even a minimum vocabulary. In an overview of studies showing vocabulary increase from reading, Waring and Nation (2004) state that research data indicates a vocabulary growth of only about 3-6 words per hour of reading, or between 150 to 300 words per year (where students spend an average of 50 hours a year in reading). This is not to suggest that reading is not a good source of vocabulary growth. However “there is a fragility to this kind of learning” (Nation, 1995-6, p. 7) and therefore it must be supported by other explicit language-focused instruction, which in Nation’s view, is an essential part of a language course.

Clearly students need to accelerate their learning in a foundations programme in order to be ready to cope with the demands of tertiary level texts. As many teachers will agree, “one of the major barriers to reading in a second language is vocabulary size” (Nation, 1995-6, p.7).
Decontextualised or Contextualized Instruction?

In summarizing the debate about the use of word lists, Nielsen (n.d.) refers to recent research which supports the notion that “teachers of beginner-level learners need to include greater amounts of decontextualized vocabulary instruction (e.g. word lists), gradually increasing toward more context-based vocabulary learning (e.g. extensive reading) as the language ability of their learners develop” (p. 6). He adds that explicit, decontextualised study is an effective way of rapidly increasing vocabulary size, and that the learning can last for a very long time.

Selection and Presentation of Vocabulary

If students are expected to acquire vocabulary knowledge through meaning-focused input from reading and listening they need to have an adequate vocabulary to do so. Therefore, as Coady (1997) suggests, it is logical to help beginners explicitly learn the basic 3,000 word families, thought to represent the fundamental lexical competence by which “learners can read independently and acquire language in a natural manner” (p. 235). This is supported by Nation (1997) who wrote that “the learner needs to know 3,000 or so high frequency words” (p. 11). He is referring here to the 3,000 head words, and thus also to the word families.

Dubai Women’s College Higher Foundations vocabulary programme assumed a knowledge of the first thousand words and focused on the second thousand word list. Research indicates that this level of knowledge will only provide 86-90% coverage, that is, the percentage of known words in an unsimplified, written text. To put this in a context, it has been found that even 98% coverage does not give learners adequate comprehension (Nation, 2006, p. 61). In view of this, even a secure knowledge of 2,000 words is inadequate preparation for entry to a bachelor programme.

As has been discussed, learning decontextualised vocabulary can be beneficial for learners at an early stage of acquisition, but this is not to suggest that care should not be taken when presenting vocabulary items. Research indicates that “learning related words at the same time makes learning them more difficult” (Nation, 2000, p. 6). Thus words presented for the first time should not be grouped as lexical sets, synonyms, antonyms or items with similar written or spoken form. There is only value in deliberately grouping similar words when knowledge is established and learners can see the difference between them.

Vocabulary Acquisition as a Cumulative Process

The acquisition of a new vocabulary item into a learner’s lexicon is a gradual process. Gains in all the aspects of knowledge of a word are incremental. It is not a simple case of “I know that word / I don’t know it” but more a case of partial knowledge being gradually strengthened. In the words of Nation (2005):

We need to see learning any particular word as being a cumulative process where knowledge is built up over a series of varied meetings with the word. At best, teaching can provide only one or two of these meetings. The others involve deliberate study, meeting through meaning-focused input and output, and fluency development activities.

As referred to earlier, the first step towards acquisition is noticing the word as an unfamiliar item, and this may be through the deliberate study of word lists. Having noticed a word and being aware of its written form (spelling), the next step is to assign meaning to it. It is this association
of form and meaning that is the essence of word knowledge. But assigning meaning to form is only an initial aspect of word knowledge. Such partial knowledge must be strengthened by repeated encounters in a variety of contexts to develop a further deeper level of knowledge, including collocations, inflections, connotations, formality, frequency and limitations on use. In one of many such studies Waring and Takaki (2003) show how incidental learning from reading occurs at several levels by enriching word knowledge through repeated exposures to words. The question of how many exposures are necessary has been the subject of various studies. Waring and Nation (2004) argue that knowledge decay over time should be accounted for and suggest that the number of encounters with the word may need to be as high as twenty. Repetition of items therefore is another essential condition for learning. Not only should there be repetitions, but they should be sufficiently frequent to ensure the word is encountered again before it is forgotten.

The length of time that the memory of meeting a word is retained has been the subject of research. Nation (2001) makes the very important point that:

If a learner has a vocabulary of around 1,000 words and is thus expanding her vocabulary at the 1,001-2,000 word level, on average each word at this level will appear once in every 10,000-15,000 running words… If, for example, the memory of meeting a word lasts for one week, then the learner will need to read at least 10,000 words per week (40 pages of 250 words per page) to ensure that there is another meeting with the word before the memory of it is lost. This is the equivalent of one graded reader every week at the 2,000 word level. (p. 169)

Vocabulary Testing

Laufer (2005) recommends frequent testing as a sound principle in teaching vocabulary. She claims that, after meeting words in class, they are “remembered much better after an additional stage of intentional memorization, and testing is one way to encourage students to do this” (p. 4). If testing does nothing else, it exposes the learner to the word after a space of time.

Research Methodology

In view of this body of literature, the evaluation of our vocabulary programme raised the two following questions:

1. In what way did our programme reflect this research?
2. How effective was our programme for vocabulary growth?

Research Question 1

An analysis of the programme showed it met many of the conditions for vocabulary learning as described in the literature. The words were introduced in a decontextualised list of approximately 60 words which had been selected from the second thousand word list. Although selected randomly, care was taken to avoid words which were likely to cause interference for the learner. For example, the words ‘consistent’ and ‘constant’ were introduced separately because of their similarity in form and meaning. Each list of 60 words formed the basis of two weeks learning.

The initial knowledge was strengthened through the other strands of the vocabulary programme.
For example, meaning-focused input was achieved through reading graded readers, of which students were required to read a minimum of one per week. Although initial testing was carried out to ascertain an appropriate level of reader, this could not always be pursued throughout the year, so some students might not have benefitted as much as they could have. Little was done to ensure fluency development, so the introduction of a speed reading course allowing students to make use of their knowledge of the first thousand words was advisable.

Language-focused learning also occurred with the carefully designed worksheets that directed attention to such aspects as word form, collocations, connotations, inflections and usage. The worksheets also provided opportunity for meaning-focused output, including speaking and writing activities. Examples included productive cloze exercises, paired discussions involving the target words, and ranking or grouping activities.

Research Question 2

A pre- and post-test of vocabulary size was carried out using the Vocabulary Levels Test, initially developed by Paul Nation (1990) and subsequently validated by Schmitt and Clapham (2001). This tool is designed to measure vocabulary size and is divided into frequency levels. For the purposes of this study, we were only interested in a comparison of the students' knowledge of the second thousand words, which is the focus of our programme. Each level of the test contained 36 words, and a score of 83% (30 words correctly recognized out of 36) was considered adequate for the student to move to the next level, given that she would meet these words in her usual study.

The online version of the test was first administered to 295 students early in the academic year (pre-test), after they had become familiar with using their laptops. To measure growth, a different version of the same test was administered at the end of the semester in January (post-test). The 295 students in this study were all those who completed the full semester.

Findings

Test Scores

A one-tailed t-test revealed a statistically significant difference between the two scores (p < 0.001).

Table 1. Change in vocabulary size.

<table>
<thead>
<tr>
<th>Test score mean</th>
<th>Pre-test (out of 36 words)</th>
<th>Post-test (out of 36 words)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test score mean</td>
<td>25.69 (71%)</td>
<td>28.55 (79%)</td>
</tr>
<tr>
<td>Range of test scores</td>
<td>10-36</td>
<td>13-36</td>
</tr>
<tr>
<td>Estimate of average size of 2nd thousand word vocabulary based on % correct</td>
<td>710 words</td>
<td>790 words</td>
</tr>
</tbody>
</table>

Table 1 shows that the mean change in vocabulary size was 2.86. In other words, an average of approximately 80 words had been acquired in 16 weeks of instruction.
Figures 1 and 2 above show that while there was a relatively normal distribution of scores in the pre-test, more students scored closer to the score of 36 in the post-test, which in educational terms is a desirable outcome.
Range of Vocabulary Change

Figure 3. Vocabulary growth by number of students.

Figure 3 shows that of the 295 students, 12 students made impressive gains of 12 to 20 words. However, one student decreased her vocabulary size by two words, 113 students recorded a zero growth and 29 students increased their vocabulary size by one word only, as measured by this test.

Zero Growth

Of the 113 students who recorded zero growth, 57 (50.4%) had a score of 30 or more. This suggests a ceiling effect. One could assume that these students had learnt the most frequent of the second thousand word list and their growth slowed down as lower frequency words were encountered. Nation (2006) reports on two research studies which found that learners’ scores dropped on the Vocabulary Levels Test as they moved from higher to lower frequency levels.

Is There a Relationship Between Pre- and Post-Test Scores?

There was a very high correlation between performance on the pre-test and performance on the post-test. The correlation between the two scores was 0.826, which shows that students who performed well on the pre-test also performed well on the post-test, while those who performed poorly on the first test performed poorly on the second test. Perhaps this could be explained by the more successful learning strategies adopted by the better performing group, and conversely the inadequate strategies used by the poorer performing group.
What is the Size of the Effect?

As we wanted to find out how well the intervention of our vocabulary programme worked, we needed to quantify the effectiveness by measuring effect size. According to Coe (2002), “effect size is simply a way of quantifying the size of the difference between two groups.”

Using Cohen’s d measure, the result shows $d=0.24$, which can be considered a small effect. This suggests that an average student in the post-test would score higher than 58% of the pre-test group. Although it may be more meaningful to see this measure in relation to other interventions which have the same aim, Coe suggests the practical importance of an effect also has to be seen in terms of its relative costs and benefits. Even a small effect size could mean a significant improvement over a period of time.

Conclusions and Recommendations

There are limitations to this research which are related to the time frame, the type of knowledge tested and the test design.

This research was carried out over one semester only. This was because we wanted to measure the impact while still teaching in order for us to reflect on our practice before the end of the year. It would have been useful to see if growth continued throughout the whole year.

As the Vocabulary Levels Test was initially designed to assist teachers in selecting appropriate reading material for students, it only measures receptive knowledge. A second test is available to measure productive knowledge and in future it would be of interest to measure growth in productive knowledge and possibly to compare both receptive and productive growth.

The test used does not provide a sensitive measure. In other words it does not measure incremental increases in partial word knowledge. Neither was it triangulated in any way, as it was a post-hoc study of the data. Therefore claims that vocabulary growth is a direct result of our programme need to be cautious and limited.

In spite of these limitations, this research has revealed data of interest. It highlighted the number of students who recorded zero growth. These students fall into two groups: those who started with a high vocabulary and had reached a plateau, and those who had a poorer vocabulary and made no progress. It would be interesting in a further study to consider what factors the latter students had in common. There does not appear to be any obvious patterns relating to their section or class level.

It also identified a small group of 12 students who increased their vocabulary size noticeably more than others, with a growth of between 12 and 20 words (remembering that the mean growth was 2.86). In order to find an explanation for this, an informal interview was carried out with these successful learners. However it did not appear that these students were employing any particular strategies that others were not. Future research could carry out a more formal interview with such a group, or an ethnographic study including students’ test taking strategies, motivation, time management and study habits.

Overall, however, the results were disappointing, indicating that the reward for effort was not as great as could have been hoped for, and suggested the need to scrutinize the implementation of our vocabulary programme.
The question about an effective vocabulary programme remains. I believe making use of a list to introduce new words serves the vital function of noticing or raising awareness of words. But an equally essential component of the programme must provide opportunity for students to have repeated exposures these words by encountering them in a variety of contexts. These findings send a strong message that students need to be encouraged to support their vocabulary learning with systematic revision and increased reading at an appropriate level.

Acknowledgement

I would like to thank Michael Fitze for his assistance in the data analysis.

References


