

# The Use and Abuse of Vocabulary Tests

Arthur McNeill

Hong Kong University of Science and Technology

## Introduction

Vocabulary tests appear to be enjoying enormous popularity across a range of stakeholders in education, including researchers, administrators, classroom teachers, and, perhaps most surprisingly, language learners. For many years, researchers have attached importance to the development and application of vocabulary tests, particularly for investigating the size and growth of learners' vocabulary (e.g. Laufer & Nation 1995, Laufer 1998, Meara & Fitzpatrick 2000). Some of the tests and test formats developed by these scholars have become widely used in applied linguistics research and now provide 'industry standard' assessment instruments to the language education community.

In particular, three vocabulary testing systems are now widely used in studies of language learning and teaching: the Vocabulary Levels Test (VLT) (Nation 1990), X-Lex: Swansea Vocabulary Levels Test (X-Lex) (Meara & Milton 2003) and the Lexical Frequency Profile (LFP) (Laufer & Nation 1995). VLT and X-Lex are tests of vocabulary size based on words drawn from different levels of frequency of occurrence in English. LFP measures lexical richness and requires users to submit a written text, which is analyzed by separating the words of the text into categories according to frequency levels. The merits of these tests and their approaches to assessing vocabulary have been examined in detail in works by Read & Chappelle (2001) and Daller, Milton & Treffers-Daller (2007). While the tests were originally developed for use in research, they have come to assume a more central position within language education. This increase in the use of the tests may be attributed, at least in part, to the acknowledged reliability of the measures. However, their popularity for a whole range of uses no doubt owes much to their availability, simplicity and

convenience. This paper considers the suitability of vocabulary tests within the mainstream language curriculum and as extra-curricular learning support.

### **Making Tests Attractive to Teachers and Learners**

When tests are readily available in simple, easy-to-use formats, they are bound to hold a special appeal for teachers and learners, especially when they are accessible on-line and provide immediate feedback to test-takers. Feedback has long been recognized as one of the essential conditions of successful second language acquisition. The three tests mentioned above all meet these criteria and can be accessed free-of-charge via the authors' university websites and elsewhere. Unlike other components of language proficiency, vocabulary allows itself to be handled in discrete units which can be counted and classified into categories, such as frequency of occurrence, and domains such as general, academic and technical vocabulary. However, as all teachers and curriculum designers are aware, there is much more to acquiring the lexicon of a second language than mastery of word forms. Not surprisingly, a number of scholars have voiced concern about a possible over-reliance on tests which adopt a discrete, uncontextualized approach to vocabulary. In fact, the nature of vocabulary as a construct and component of proficiency appears to be somewhat ill-defined (Read & Chapelle 2001:1).

### **Vocabulary as a Construct**

Is there a distinct vocabulary construct? This assumption appears to be made by many researchers in vocabulary studies who are interested in assessing knowledge of individual word forms. Singleton (1999) has challenged the assumption and argues in favour of expanding the scope of both vocabulary teaching and assessment to reflect what Read and Chapelle refer to as the "pervasiveness of lexical phenomena in language" (2001:2). They point out that scholars who treat vocabulary as a separate construct tend to use tests that fit comfortably within the psychometric-structuralist tradition in language testing, assessing knowledge of content words with such relatively decontextualized item types as multiple choice, word-definition matching, word completion and

the checklist. Pearson, Hiebert & Kamil (2007) have expressed similar concerns about the dominance of context-independent approaches to vocabulary assessment, particularly in relation to vocabulary and reading:

“... vocabulary assessment is grossly undernourished, both in its theoretical and practical aspects – (that) it has been driven by tradition, convenience, psychometric standards, and a quest for economy of effort rather than a clear conceptualisation of its nature in relation to other aspects of reading expertise.” (p281)

Read (2000) proposed three dimensions, which can help to highlight the characteristics of a vocabulary test (Figure 1).

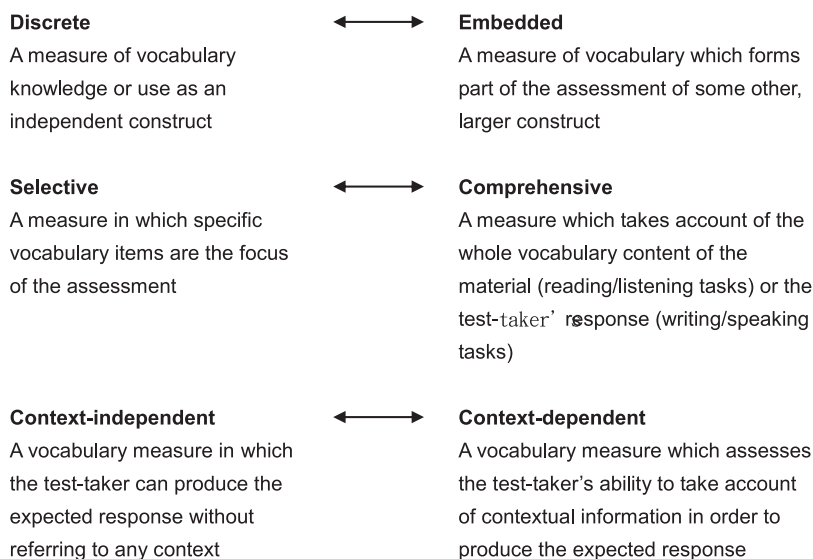


Figure 1: Three dimensions of vocabulary assessment (Read 2000:9)

According to Read's three dimensions, two of the popular vocabulary tests mentioned in the Introduction, VLT and X-Lex, can be characterized as discrete, selective and context-independent, i.e. placed at the left-hand extreme of each of the three continua. The third popular test mentioned in the Introduction, LFP, also falls at the left-hand extreme in terms of discrete/embedded and selective/comprehensive, but is less context independent than the two other tests,

since learners submit a piece of continuous text for analysis. So, does it matter that the most popular vocabulary tests appear to isolate vocabulary from other components of language and may, therefore, be limited in terms of their uses? If we accept that vocabulary cannot always be regarded a separate component of language knowledge and that there are lexical dimensions of most language tasks (Skehan 1996, Foster & Skehan 1996), then multiple measures are required for assessment of learning, especially where high-stakes decisions are based on their results.

## Different Tests for Different Purposes

Read & Chapelle (2001) argue that the design of a vocabulary test should reflect the test's intended purpose and propose a framework for vocabulary testing (Figure 2).

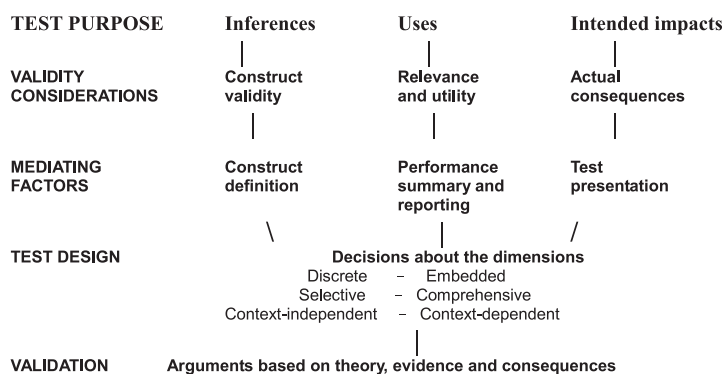


Figure 2: A framework for vocabulary testing (Read & Chapelle 2001: 10)

According to the framework, vocabulary assessment needs to take account of test purpose in the design and validation of tests. It follows from the framework that vocabulary tests should require learners to perform tasks under contextual constraints relevant to the inferences to be made about lexical ability. It would appear that some of the popular vocabulary tests are used for purposes for which they were not originally intended. For example, there is obviously considerable scope for the development of new tests that provide reliable and

valid assessment of the lexical elements of integrated language tasks. However, not all tests have a single clear purpose. Indeed, vocabulary tests are often used by teachers and students to promote learning rather than to inform decisions about students' progress. In the light of the above concerns about the uses of vocabulary tests for purposes for which they were not designed, it is interesting to consider to what extent the widely used popular vocabulary tests can help to promote learning.

## Tests and Vocabulary Learning

The literature on L2 vocabulary learning provides overwhelming support for the need to provide learners with multiple exposures to target vocabulary items. It is estimated that a learner needs to encounter a word between six and twenty times before the item is acquired (Rott 1999, Zahar, Cobb & Spada 2001). It is difficult for a curriculum and individual teachers to provide students with sufficient repeated encounters with the words they are expected to learn. Vocabulary tests can help to provide these. For this purpose, popular tests such as VLT and X-Lex with their discrete, selective, uncontextualized formats can provide efficient and valuable exposure to essential vocabulary items. Since the items are grouped according to their frequency of occurrence in English, teachers can recommend the most appropriate levels to their students.

Although it may take some time before the issue of a vocabulary construct is fully resolved, existing descriptions of L2 vocabulary knowledge can help to determine whether particular vocabulary tests contribute to language learning. For example, descriptions of L2 vocabulary knowledge (e.g. Nation 1990, McNeill 1994) identify the following dimensions of 'knowing a word': semantic, phonological, grammatical, orthographic, collocational and idiomatic. Learners typically know only some of the dimensions of the words they would claim to have learned and need to acquire the others. Vocabulary tests can be a valuable means of providing learners with the opportunity to build on their incomplete word knowledge, particularly when the tests specifically address the dimensions of word knowledge.

LFP, the third of the popular vocabulary tests mentioned in the Introduction, is probably the most useful for promoting vocabulary learning. When students submit their own texts to the software program, they receive an immediate profile of their writing, with the lexical content divided into different frequency bands. Students can see the extent to which their writing consists of high and low frequency words. If the LFP indicates that the lexical profile of a text is weak, the student can edit it by substituting some low frequency items for the high frequency words used. This practice of drafting a text, then revising it in the light of feedback is educationally desirable and promotes good study habits as well as learner autonomy. Since the wordlists used by LFP can be changed to suit local contexts, the test software can be adapted to include prescribed wordlists. For example, the Hong Kong University of Science and Technology has developed a version of LFP based on the new English wordlists compiled for different key stages of the Hong Kong school curriculum. This version of LFP can be used by teachers to assess students' productive command of the prescribed wordlists and by learners to monitor and reflect on their use of the vocabulary content of the curriculum.

## **Conclusion**

This chapter has examined the widespread use of three English vocabulary tests which have become regular features of language programs and research studies throughout the world. In spite of the popularity of the tests, it was noted that they focus on words as discrete, context-independent items and appear to support the notion of a distinct vocabulary construct. In order to assess vocabulary adequately in an English program, a wider range of assessment types has been called for, including embedded and context-dependent tests. Nevertheless, the three tests can be expected to provide learners with incentives to deepen their knowledge of vocabulary and should support some of the key principles of vocabulary teaching, such as providing multiple exposures to target words and focusing on dimensions of word knowledge. Above all, the accessibility, simplicity and inter-activity of the tests serve as powerful motivating factors to learners. Ultimately, for successful vocabulary acquisition, every encounter counts.

## References

- Daller, H., Milton, J. and Treffers-Daller, J. (eds.) (2007). *Modelling and assessing vocabulary knowledge*. Cambridge: Cambridge University Press.
- Foster, P. and Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition* 18, 299-323.
- Laufer, B. (1998). The development of active and passive vocabulary in a second language: same or different? *Applied Linguistics* 19, 255-271.
- Laufer, B. and Nation, P. (1995). Vocabulary size and use: lexical richness in L2 written production. *Applied Linguistics* 16, 33-51.
- McNeill, A. (1994). Vocabulary testing: defining word knowledge. In Falvey, P. and Boyle, J. (eds.) *Language Testing in Hong Kong*, Hong Kong: Chinese University Press.
- Meara, P. (1992). *EFL vocabulary tests*. Swansea: Centre for Applied Language Studies, University of Wales, Swansea.
- Meara, P. and Fitzpatrick, T. (2000). Lex30: an improved method of assessing productive vocabulary in an L2. *System* 28, 19-30.
- Meara, P. and Milton, J. (2003). *X-Lex: Swansea Vocabulary Levels Test*. Swansea: Lognostics.
- Nation, I.S.P (1990). *Teaching and learning vocabulary*. Boston, MA: Heinle & Heinle
- Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Pearson, P.D., Hiebert, E.H. and Kamil, M.L. (2007). Vocabulary assessment; What we know and what we need to know. *Reading Research Quarterly* 42/2, 282-296.
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Read, J. and Chapelle, C. A. (2001). A framework for second language vocabulary assessment. *Language Testing* 18 (1), 1-32.
- Rott, S. (1999). The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition and retention through reading. *Studies in Second Language Acquisition*, 21 (4), 589-619.
- Singleton, D. (1999). *Exploring the second language mental lexicon*. Cambridge: Cambridge University Press.
- Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics* 17, 38-62.

Zahar, R., Cobb, T. and Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *Canadian Modern Language Review*, 57 (4), 541-573.

Author's e-mail: [lcamcneill@ust.hk](mailto:lcamcneill@ust.hk)