

Estimating vocabulary size

David Beglar

What's the best way to estimate an EFL learner's vocabulary size? Are there any effective tests to estimate how many words EFL learners know? Are there any problems with standard vocabulary-estimating tests we should be aware of? This article explores these questions.

One institution in Japan which has made a rigorous attempt to measure the vocabulary of their students is Temple University Japan's Corporate Education Program (CEP). This program has been using a version of Paul Nation's Vocabulary Levels Tests (Nation, 1990). These tests are relatively straightforward, and are made up of sets of six words and three definitions, as in the following example:

- | | |
|-----------------------------|-------------|
| 1. _____ first | a. royal |
| 2. _____ not public | b. slow |
| 3. _____ all added together | c. original |
| | d. sorry |
| | e. total |
| | f. private |

This test is designed to estimate examinees' basic knowledge of common word meanings, and, specifically, the extent to which they know the common meanings of words at the 2,000, 3,000, 5,000, 10,000 and university word levels. The test can be classified as a sensitive vocabulary test, which means that the format is sensitive to partial word knowledge. A less sensitive test (e.g., a multiple-choice cloze test focused on specific content words only) would result in lower scores even if the same words were tested.

Originally, the test had 90 items, but after being trailed with Japanese learners, the best performing 60 items chosen. It is fast and can be easily administered in twenty minutes. It is also reliable (Cronbach's alpha = .95 and Rasch reliability estimate = .97). In short, this test gives a general idea of the number of words an English speaker knows.

Paul Nation and Batia Laufer have both utilized versions of the Vocabulary Levels Tests to estimate vocabulary this way: if learner A scores 9 out of 12 (75%) on the 2,000 word level, s/he probably knows approximately 75% (1,500) of the first 2,000 words of English. If you continue to apply this logic to the results of the rest of the test (i.e., the 3,000, 5,000, University Word Level, and the 10,000 word level), you can arrive at an approximate estimate of vocabulary size.

Another way to measure vocabulary is to focus on words which are of greater importance and to test only those words. The advantage is that by focusing on a more narrow range of words, you can test more items and presumably arrive at more accurate estimates of what learners know. Beglar and Hunt (1999) did that with several versions of the 2,000 word level and University Word List tests. They trialed original pools of 72 items with native speakers of Japanese, selected the best performing 54 items for each test and made two 27-item parallel forms. You can find these forms in the appendix of Beglar and Hunt (1999), or e-mail to receive copies as attachments. In the same article, we also briefly discuss why the 2,000 and University Word List levels are important words for learners to know.

In addition to these tests, Paul Meara and several of his colleagues and students (e.g., Meara and Jones, 1987; Meara and Jones, 1990) have worked extensively with the Eurocentres Vocabulary Size Test. This is a checklist test in which the examinee checks the words he thinks he knows. I have seen published research which sometimes paints these tests in a very good and at other times in a very bad (e.g., low reliability) light. You can find more information on the tests at the Vocabulary Acquisition Research Group's homepage at <http://www.swansea.ac.uk/cals/calsres.html> [Expired link]. Several tests can be downloaded from the "freebies" section of this web page, such as the EVST (a basic vocabulary size test) and the LLEX 2.21 (a basic recognition vocabulary test). In addition, Paul Meara's students often have information about ongoing research posted on this webpage. In my experience, there is almost always an article posted concerning the Eurocentres Vocabulary Size tests. Finally, you can obtain a large number of these tests through the ERIC document reproduction service (see the Meara, 1992 reference below).

Finally, if you are interested in reading more about vocabulary testing, you might wish to take a look at the new text by John Read (2000) which has just been published by Cambridge. I can also recommend a book manuscript on vocabulary acquisition and teaching by Paul Nation (1999) which can be ordered by e-mail. In addition to learning everything about teaching vocabulary that you ever wanted to know, there is a good chapter on vocabulary testing which brings up a number of interesting issues such as vocabulary tests which are sensitive to differing levels of word knowledge.

References

Beglar, D., & Hunt, A. (1999). Revising and validating the 2000 word level and university word level vocabulary tests. *Language Testing, 16* (2), 131-162.

Shiken: JALT Testing & Evaluation SIG Newsletter, 4 (1) April 2000 (p. 2 - 4)

Meara, P., & Jones, G. (1987). Tests of vocabulary size in English as a foreign language. *Polyglot*, 8 (1), 1-40.

Meara, P., & Jones, G. (1990). *Eurocentres Vocabulary Size Test* (version E1.1/K10,MSDOS). Zurich: Eurocentres Learning Service.

Meara, P. (1992). *EFL vocabulary tests*. Wales University: Swansea Centre for Applied Language Studies. (ERIC Document Reproduction Service No. ED 362 046).

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

Nation, I. S. P. (1999). *Learning vocabulary in another language*. English Language Institute Occasional Publication No. 19. Wellington, NZ: Victoria University of Wellington.

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.

HTML: http://www.jalt.org/test/beg_1.htm **PDF:** <http://www.jalt.org/test/PDF/Beglar1.pdf>