A PRELIMINARY RESEARCH TO DEVELOP
A CUSTOMIZED SET OF VOCABULARY SIZE TEST

Laurentia Sumarni

This is a preliminary research to investigate useful words to function effectively in academic contexts, vocabulary size, the first-year students' word list, and the steps to design the customized set of entry-level vocabulary size test. This research was a library research. To find out the ELESP students' current vocabulary size, a vocabulary size test designed by Paul Nation was administered. The results show that useful words were (1) high-frequency words containing 2,000 word families, (2) academic words specified in the Academic Word List (Coxhead:2000), (3) technical words, (4) low-frequency words. The result of the test shows that the students' scores range between 33 and 96 words, with the mean score of 66. It means that the students' vocabulary size ranges between 3,300 and 9,600 word families. The average 6,600 word families imply that most ELESP students are ready to read authentic texts containing 88.7% word coverage. To increase vocabulary size by 10,000, they need to learn technical words and low-frequency words of a specialized subject area. Using Nation's specifications for making the test, some procedures of test design are: *Sampling the words for the items, Making the Stem, Writing the choices, The Order of the items in the Test, Piloting, Administering the Test.* The decisions on curriculum, materials and teaching strategies should be based on the results of vocabulary size to gain optimum learning outcome.
Keywords: *entry level vocabulary size test, high-frequency words, low-frequency level words*

INTRODUCTION

In English language learning and teaching, vocabulary is the currency without which communication and meaning-making are impossible. Ironically, in English language teaching, vocabulary is not considered important as seen from the very little attention and consideration of vocabulary learning in the curriculum. In university level, students' academic success is always measured by means of written and spoken assessments, in which they show their knowledge of a particular topic by the use of specialized vocabulary.

PBI (ELESP-*English Language Education Study Program*) does not have an entry-level vocabulary test to decide which level students are at the beginning of their programs. In addition, decisions on curriculum, syllabus and material development are still based on common sense, instead of research-based. In response to this, an entry-level vocabulary test to measure PBI students' vocabulary size is crucial and urgent to be conducted because the test can give a lot of important information regarding the current vocabulary size. The information obtained from the tests can also be used to decide the necessary action plans to address the problems.

Therefore, this research aims to investigate these questions:
1) Which words are useful to enable PBI students to function effectively in academic contexts?
2) Using Nation's Vocabulary Size Test, what is the size of the first-year students' vocabulary?
3) How many words should the first-year students learn?
4) What are the steps to design the set of customized entry-level vocabulary size tests?

It is hoped that this research can open opportunities for further research on vocabulary testing to explore its contribution in English language teaching and learning.

THEORETICAL REVIEW

Lehr, Osborn, & Hiebert (2004) cited in Read (2000) define vocabulary as words we use to communicate in spoken and written language. Receptive vocabulary refers to the words we understand through reading and listening, while productive vocabulary refers to words we use to communicate through writing and speaking (as quoted in Hanson and Padua, 2011). Ellis (1997) has shown that "vocabulary knowledge is indispensable to acquire grammar." Vocabulary is very crucial to understand spoken and written discourses and to express ideas orally as well as in writing. Coombe adds that "experts in the field of vocabulary development are in agreement that vocabulary is central to the language learning process, and as such, it is generally accepted that a focus on strengthening vocabulary is necessary at every stage of a learner's language development." Folse (2003) as cited in Coombe, states that "without syntax, meaning is hindered; but without vocabulary meaning is impossible".

In the English Language Education Study Program, the students are required to listen to spoken discourses, read written texts, speak up their ideas and write academic essays. In short, they have to be able to perform various communicative tasks in academic settings. The assessments are based on their spoken and written performance. Around 2,000 words are needed to reach 80% coverage of a written text. However, students must strive to get at least 98% coverage. Nation (2006) suggests the vocabulary sizes needed to get 98% coverage (including proper nouns) of various kinds of texts.

Nation says that 'the goal of around 8,000 word families is an important one for learners who wish to deal with a range of unsimplified spoken and written texts." In writing, Nation (2008) suggests that "a small number of words (around 2,000 to 3,000) can be used effectively to express an enormous number of ideas." Nation (1990, 1993, 2001) further adds that "the critical importance of developing an adequate high-frequency vocabulary since learner's skill in using the language is heavily dependent on the number of words they know, particularly in the early stages of learning a foreign language, with around 3,000 word families being a crucial threshold." Hirsch and Nation (1992) argue that "in order to reach text comprehension, readers need to be familiar with 95% of the words in a text." This text coverage percentage is logical because "to cope well in English, a second language learner would need around 5,000 words and preferably 10,000 words" (Nation, 2004).

Francis and Kucera (1982) suggest that "the 2,000 most frequent word families of English make up 79.7% of the individual words in any English text, the 3,000 most frequent word families represent 84%, the 4,000 most frequent word families make up about 86.7%, and the 5,000 most frequent word families cover 88.6%. A much better reading comprehension power is ensured if a reader knows the meanings of at least 90% of the words in a text." Considering the importance of vocabulary size and knowledge to enable a foreign language learner to function effectively in academic settings, vocabulary testing is thus crucial to be implemented. Nation (2008:144) argues that "vocabulary testing can be used to work out what needs to be taught, to monitor and encourage learning, to place learners in the right class, to measure learners' achievement, to measure learners' vocabulary size and performance." He further adds that for research purposes, vocabulary testing can be used to "evaluate teaching and learning activities." Anderson and Freebody found that "vocabulary size is a good predictor of reading comprehension." Coady, Magott, Hubbard, Graney and Mokhtari (1993) also found that vocabulary is "an important factor for obtaining fluency in speech."

METHODOLOGY

The detailed library research will focus on such activities as follows:

1) reading the Nation's Vocabulary Levels Test and Productive Vocabulary Levels Test, Meara's Yes/No Test, Martinez's Phrase-Test, Read's Word Associates Test, Coxhead's Academic Word List, the West's General Service List, and Thorndike's 30,000-word list. This also involves researching the internet resources on vocabulary tests.
2) choosing the type of tests to be tested to the first-semester students
3) Administering Nation's Vocabulary Size Test to 156 students of PBI
4) Analyzing the Results of the Test
5) Drawing Conclusions and Recommendations on the design of a customized entry-level vocabulary size test.

RESEARCH FINDINGS AND DISCUSSION

To answer the first question, i.e. which words are useful to enable PBI students to function effectively in academic contexts?, the words that are useful to enable PBI students to function effectively in academic contexts are the high-frequency words consisting around 2,000 word families made up of 169 function words and 1831 content words; academic words (Coxhead (as cited in Nation, 2008:9)) consisting of 570 word families; technical words dealing with linguistics, literature and education which range up to 5,000 words; to acquire the low-frequency words as well in order to function effectively in academic settings. The low-frequency words make up the biggest and most diverse group of words, namely (1) words that are not quite frequent or wide range enough to be high frequency words, (2) technical words from other areas, and (3) words that just occur rarely. According to Nation (2008:11), English probably has a low frequency vocabulary of around 20,000 word families.

To answer the second question, a vocabulary size test designed by Nation and Beglar was administered to the ELESP 2013 students at the beginning of the program to find out the size of students' vocabulary. The test was given to 156 out of 165 first-semester students of the ELESP. The mean score 66 suggests that the students' average vocabulary size was 6,600 words. Out of 156 students, the score ranges between 33 and 96, which suggests that the students' lowest vocabulary size was 3,300 words and the highest vocabulary size was 9,600 words. The median score was 65 and the mode was 58 which imply that half of the students scored between 65 and 96. This means that more than 50% of the total students (78 people) already had a vocabulary size above 5,000 words. As stated in the literature review, most research indicates that knowledge of the most frequent 5,000 word families should provide enough vocabulary to enable learners to read authentic texts. From the results, it can be seen that actually most PBI students are lexically ready to read texts containing 88.7% text coverage. It means that with around 5,000 words, students are not familiar with 11.3%, or around 11 words per 100-word text. The following table dispays the result in detail.

From the table, it can be concluded that, in average, the 2013 ELESP students are already familiar with around 5,000 most frequent word families which cover 88.6% text coverage. The following table shows the relationship between the frequency level the vocabulary size score and what students need to do to increase their vocabulary size (Nation:2012).

Table 1.: The relationship between frequency level, vocabulary size and learning procedures.

Level	1000 word family lists	Learning procedures
High frequency	1000-2000	Reading graded readers Deliberate teaching and learning
Mid-frequency	3000-9000	Reading mid-frequency readers Deliberate learning
Low frequency	10,000 on	Wide reading

		Specialised study of a subject area

The result can be used to answer the third question, i.e. *How many words should the first-year students learn?* With the 6,600 word families in their hands, students are encouraged to acquire more technical words related to linguistics, literature and education; and low-frequency words. In order to increase their vocabulary size by 10,000, they need to learn words deliberately and widely on technical words and low-frequency words of a specialized subject area.

To answer the fourth question, i.e. What are the steps to design a customized set of entry-level vocabulary size tests?, here are the procedures of the vocabulary size test design adapted from Nation's Specification for making the test (Nation: 2012).

1. *Sampling the words for the items*
   The items in the test need to represent the various frequency levels of the language without a bias towards any particular frequency levels, which are based on word families which occurred in the British National Corpus according to Bauer and Nations (1993 as cited in Nation 2012).
2. *Making the Stem*
   The test uses a stem plus a 4 choice in multiple-choice format. The non-defining sentence has the roles of (1) indicating the part of speech of the word, (2) limiting the meaning of the word where words may have a homograph or very different senses, and (3) slightly cueing the meaning by presenting an example of use.
3. *Writing the choices*
   The distracters are the same part of speech as the correct answer, and in most cases the distracters are the meanings of words from around the same 1,000 word frequency level as the correct answer. The choices must be written in a much easier language than the tested word.
4. *The Order of the items in the Test*
   The test items in the test are usually arranged in frequency order.
5. *Piloting*
   Piloting the test can be done by getting applied linguists who are native speakers of English to individually read and critique the test, replacing the target word with the nonsense word and getting a test-wise native speaker to try to choose the correct answer, and running the tests through the Range program to check the frequency levels of words used in the contexts and choices.

CONCLUSIONS

First, useful words to enable the ELESP students to function effectively in academic contexts were (1) high-frequency words containing 2,000 word families made up of 169 function words and 1831 content words, (2) academic words which are used in the specialized area containing at least 570 word families called the Academic Word List (Coxhead:2000), (3) technical words which deal with linguistics, literature and education which range up to 5,000 word families, (4) low-frequency words in order to understand conversations, to read newspapers and understand academic texts.

Second, the result of the administration of Nation's Vocabulary Size Test shows that the students' scores range between 33 and 96 words, with the mean score of 66 and median of 65 and mode of 58. It means that students' vocabulary size ranges between 3,300 and 9,600

word families. The average students' vocabulary size was 6,600 word families, which implies that most ELESP students are lexically ready to read texts containing 88.7% word coverage.

Third, the average students of PBI are familiar with the 6,600 word families which represent high-frequency words in the English language, students are encouraged to acquire more technical words related to linguistics, literature and education. Currently, they are in their mid-frequency level, with the vocabulary size ranging from 3,000 to 9,000 word families. In order to increase their vocabulary size by 10,000, they need to learn words deliberately and widely on technical words and low-frequency words of a specialized subject area.

Fourth, using Nation's specifications for making the test, the researcher listed some procedures of test design (Nation: 2012). They are: 1) *Sampling the words for the items; 2) Making the Stem; 3) Writing the choices; 4)The Order of the items in the Test; 5) Piloting; 6) Administering the Test.*

REFERENCES

Anderson, R.C., & P. Freebody. (1983). Reading comprehension and the assessment and acquisition of word knowledge. In B. Houston (ed.), *Advances in Reading/Language Research*. Volume 2 (pp. 231-256). Greenwich, CT:JAI Press.

Coombe, Christine. *Assessing vocabulary in the language classroom*. In Malaysian English Language Teaching Association, pp. 111- 124.

Coxhead, A., (2010). Grabbed early by vocabulary: Nation's ongoing contributions to vocabulary and reading in a foreign language. In *Reading in a Foreign Language*, April 2010, Vol. 22, No. 1, pp. 1-14.

Francis, N., & H. Kucera. (1982). Frequency analysis of English Usage. Boston: Houghton Mifflin Company.

Hanson, S., & Padua, Jennifer F. M. (2011). *Teaching Vocabulary Explicitly*. Honolulu: Pacific Resources for Education and Learning (PREL).

Hirsch, D., & Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? In *Reading in a foreign language*, Vol. 8 (2), 1992.

Nation, I.S.P. & Beglar, D. (2007) A vocabulary size test. *The Language Teacher, 31*(7), 9-13.Nation, I. S. P. (1990). Teaching and Learning vocabulary. New York: Newbury House.

Nation, I. S. P. (1995-1996). Best Practice in vocabulary teaching and learning. EA Journal, 3 (2), 7-15.

Nation, P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy*. Cambridge: Cambridge University Press, pp. 6-19.

Nation, I., & Heatley, A. (2002). Range: A program for the analysis of vocabulary in texts [Computer Software]. Available from http://www.vuw.ac.nz/lals/staff/paul-nation/nation.aspx

Natin, I. S. P., & Beglar, D. (2007). A vocabulary size test. In *The Language Teacher*, 31(7), 9-13.

Nation, I. S. P. (2008). Teaching Vocabulary: Strategies and Techniques. Boston: Heinle Cengage Learning.