

A Bilingual Vocabulary Size Test of English for Vietnamese Learners

RELC Journal
42(1) 86–99
© The Author(s) 2011
Reprints and permission: sagepub.
co.uk/journalsPermissions.nav
DOI: 10.1177/0033688210390264
rel.sagepub.com


Le Thi Cam Nguyen

LALS, Victoria University of Wellington, New Zealand

Paul Nation

LALS, Victoria University of Wellington, New Zealand

Abstract

This article describes the development and validation of a Vietnamese bilingual version of the Vocabulary Size Test – a test which measures written receptive vocabulary size. The test can be used to measure the English vocabulary size of Vietnamese learners of English. A learner's total vocabulary size is calculated by multiplying their test result by 100. The research adds to our knowledge of vocabulary size testing in the following ways. First, it shows that a bilingual version of a monolingual test performs in much the same way as the monolingual test, distinguishing learners of different proficiency levels and returning lower scores at later levels of the test. Second, it shows that every level of the test should be sat, otherwise there will be a considerable underestimation of learners' vocabulary sizes. This suggests limitations to the previously accepted assumption that learners' vocabulary growth can be largely related to word frequency. Third, it shows that bilingual tests are feasible alternatives to more challenging and time-consuming monolingual tests.

Keywords

Vocabulary, testing, vocabulary size, frequency levels

Measuring learners' vocabulary size has a long history because vocabulary size is seen to be an important aspect of language knowledge that is essential for effective use of the language.

Vocabulary size tests are important for several reasons. First, they are important when designing a programme of study for a group of learners. If we do not know how many words learners know, then it is difficult to know what new words learners should focus on in their language learning programme. It is also very difficult to decide what level of graded reading they should do, and whether they are ready to read unsimplified texts.

Corresponding author:

Paul Nation, LALS, Victoria University of Wellington, PO Box 600, Wellington, New Zealand.
[email: Paul.Nation@vuw.ac.nz]

Table 1. Vocabulary Sizes Needed to Get 98% Coverage (Including Proper Nouns) of Various Kinds of Texts

Texts	98% coverage	Proper nouns
Novels	9,000 word families	1–2%
Newspapers	8,000 word families	5–6%
Children's movies	6,000 word families	1.5%
Spoken English	7,000 word families	1.3%

There is value in knowing a learner's total vocabulary size because this can then be related to the vocabulary demands of the material that the learner needs to work with. The data in Table 1 is taken from a study by Paul Nation (2006). The study assumes that 98% coverage of the running words in a text is needed to gain reasonable unassisted comprehension of the text (Hu and Nation, 2000).

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. Unless we know how close learners are to this goal, our planning of the vocabulary component of a language programme will be rather hit-and-miss.

In Vietnam students start learning English at a very young age from Year 3 in the 12-year system but English is not compulsory until students begin secondary school starting at Year 6. At this level there is very limited or no flexibility for teachers of English to develop their own language programmes and to use their own self-designed teaching materials because they have to follow textbooks designed by the Ministry of Education and Training. At the tertiary level however, there is more freedom for designing an English programme and developing teaching and learning materials that cater for the needs of students. Nevertheless, even in teaching situations where there is a set textbook and a central curriculum to follow, there is usually enough flexibility within a programme for teachers to help learners develop their vocabulary knowledge in an individualized way. First, learners can be encouraged to do deliberate learning using word cards. Second, they can do independent reading either of graded readers or of unsimplified texts. Third, they can work with supplementary materials in either hard copy or web-based form to increase their vocabulary size. The website www.lex-tutor.ca has a variety of useful vocabulary development resources, and there are other web-based programmes or free downloads that have been designed for vocabulary learning.

Second, vocabulary size tests are particularly important when learners are studying in an English medium system where English is their second language. That is, a Vietnamese bilingual test of English is also useful for Vietnamese learners studying outside of Vietnam. This is a very under-researched area, but it seems that learners who enter an English medium system around the age of five years old acquire vocabulary at the same rate as native-speakers, that is around 1,000 word families per year. A vocabulary size test thus can be a good tool to see if second language learners are indeed progressing at the same rate as native-speakers in their vocabulary growth.

Third, vocabulary size tests are useful diagnostic measures when learners are having problems with skills like reading or writing. Sometimes poor vocabulary knowledge is the cause of poor reading or writing, but often other causes may be more important. Vocabulary size tests can help a teacher decide if vocabulary size is indeed an issue in

poor performance of a particular language skill. For example, research with six year-old non-native speakers of English living in an English-speaking country found that poor reading performance was typically not the result of an inadequate vocabulary size but was caused by other factors (Ruffell, 2008).

Fourth, vocabulary size tests are useful research tools when doing studies based on factor analysis and when trying to divide learners into levels of proficiency. They are particularly useful in experiments where a measure is needed which is not strongly biased towards one of the four skills of listening, speaking, reading, and writing.

The purpose of this study is to show that it is possible to create effective bilingual versions of the Vocabulary Size Test and that the Vietnamese bilingual version works well and may be used with Vietnamese learners both inside and outside Vietnam to inform, monitor, and diagnose their vocabulary growth. It also examines if a shorter version of the test would be effective.

Measuring Vocabulary Size

The monolingual version of the Vocabulary Size Test, and its derived bilingual versions, are based on word frequency lists developed from the British National Corpus (Nation and Webb, 2011). The development of these word lists involved making word families for each of the words in the lists based on a set of word-building criteria described in Bauer and Nation (1993). In these lists, a word family consists of a headword plus its inflected forms and its closely related derived forms. The headword of a family must be a free form. That is, it can stand as a word in its own right, and the derived forms can only consist of affixes added to free forms. Here is an example:

Access

- Accessed
- Accesses
- Accessibility
- Inaccessibility
- Accessible
- Accessing
- Inaccessible

The headword is *access*, and the whole family contains eight members consisting of the headword, its inflected forms, and its derived forms which meet the rules described for level 6 in Bauer and Nation (1993).

The idea behind word families is that for receptive purposes very little extra knowledge is required to work out the meaning of an inflected or derived form, if the headword is already known and if the learners have a working knowledge of English inflections and frequent and productive derivational affixes. Tests based on word families are thus suited for receptive purposes, but are not suitable measures of productive knowledge. When measuring productive knowledge of vocabulary, the word family is not a suitable measure and single word forms or perhaps a headword and its inflected forms are a more appropriate unit of counting. The vocabulary size test is a measure of receptive vocabulary knowledge not productive use.

The Vocabulary Size Test

The Vocabulary Size Test (Nation and Beglar, 2007) was designed as a proficiency measure of total vocabulary size for learners of English as a second or foreign language.

This test consists of 140 items with 10 items from each of fourteen 1,000 word levels based on a frequency count of word families in the British National Corpus. The test has a four-item multiple-choice format. Here is an example from the third 1,000 level in the test:

3. jug: He was holding a **jug**.
 - a. a container for pouring liquids
 - b. an informal discussion
 - c. a soft cap
 - d. a weapon that explodes

The items were written so that the vocabulary used in the four choices were higher frequency words than the tested word. Each test word is put in a simple non-defining context. Because each item in the test represents 100 word families (there are 10 items from each 1,000 word frequency level), then the learner's score on the test is multiplied by 100 to get their total vocabulary size. So if a learner scores 68 out of 140 on the test, their total vocabulary size is 6,800 words.

Validating the Test

Beglar (2010) conducted a Rasch-based validation of the monolingual test finding that it has these features:

1. It can be used with learners with a very wide range of proficiency levels.
2. It measures what it is supposed to measure and does not measure other things. Beglar found that the test was very clearly measuring a single factor (presumably written receptive vocabulary knowledge) and other factors played a minor role in performance on the test.
3. It behaves in ways that we would expect it to behave, distinguishing between learners of different proficiency levels, having a range of item difficulties related to the frequency level of the tested words, and clearly distinguishing several different levels of vocabulary knowledge so that learners' vocabulary growth over time could be measured.
4. It performs consistently and reliably, even though circumstances change. In Beglar's trialling of the test, these changes included comparing the performance of male subjects with female subjects, comparing 70 item versions of the test with the 140 item version, and comparing learners of various proficiency levels. Rasch reliability measures were around .96.
5. It is easy to score and interpret the scores.
6. The items in the test are clear and unambiguous.
7. It can be administered in efficient ways. When learners sit the test, they need not sit 140 items but could sit a 70 item version of the test.

The test works very well because it covers a very wide range of frequency levels, it includes a large number of items (even half of this number would work well), the items have been very carefully designed, made, and trialled, and the test is designed to measure just one kind of vocabulary knowledge.

Users of the test need to be clear what the test is measuring and not measuring. It is measuring written receptive vocabulary knowledge, that is the vocabulary knowledge required for reading. It is not measuring listening vocabulary size, or the vocabulary knowledge needed for speaking and writing. It is also not a measure of reading skill, because although vocabulary size is a critical factor in reading, it is only a part of the reading skill.

The Value of Bilingual Versions of the Test

Let us look at another item from the monolingual Vocabulary Size Test.

3. marsupial: It is a **marsupial**.
 - a. an animal with hard feet
 - b. a plant that grows for several years
 - c. a plant with flowers that turn to face the sun
 - d. an animal with a pocket for babies

Notice that the grammar of the choices in this example involves relative clauses (*that grows for several years, that turn to face the sun*) and nouns followed by preposition groups. Notice that the choices are quite long. This means that instead of being a simple test of vocabulary knowledge, the monolingual Vocabulary Size Test also requires the learners to have some knowledge of complex grammar, and to have reasonable reading skills. Ideally a vocabulary size test should not depend on this kind of knowledge and should as far as possible focus only on vocabulary knowledge.

One way to avoid these problems is to write the choices in the first language of the learners sitting the test.

3. marsupial: It is a **marsupial**.
 - a. loài động vật có chân rất khỏe
 - b. cây lâu niên
 - c. hoa hướng dương
 - d. loài thú có túi

Such a bilingual test still remains a demanding test of receptive word knowledge, but it avoids the problems of the difficult grammar of definitions and reduces the demands on skill in reading English. From this perspective, a bilingual test is a more valid test of vocabulary knowledge, because it reduces the non-vocabulary factors influencing the result of the test.

Learners' scores are therefore likely to be slightly higher on a bilingual test (especially for low proficiency learners), and this will be a truer reflection of their vocabulary knowledge than their monolingual test results.

This test and the monolingual test are available from the following website <http://www.victoria.ac.nz/lals/staff/paul-nation.aspx>

The Development of the Bilingual Test

The development of a bilingual version of the Vocabulary Size Test involves translating the four English choices for each item into the learners' first language. This translation should not involve a word-for-word translation of each choice, but wherever possible should involve using the Vietnamese word for the correct choice and for each of the distractors. For example, in the following item from the thirteenth 1,000 word level, each of the distractors is the definition of an English word.

7. skylark: We watched a **skylark**.
- show with aeroplanes flying in patterns
 - man-made object going round the earth
 - person who does funny tricks
 - small bird that flies high as it sings

Choice *a* describes the word *airshow*, choice *b* describes the word *satellite*, choice *c* refers to someone who skylarks (behaves in a vigorous joking way), and choice *d* is the correct answer. If there are corresponding single words for each of these three distractors, then the translations used in the test should be translations of these words rather than their definitions. Often this is not possible, because there is no corresponding single word in the learners' first language which corresponds to the choice.

The test was translated by the first author and proofread by another native speaker of Vietnamese. They both were experienced teachers of English. The translators then went over each item together to ensure the accuracy and the intelligibility of the translations. They tried to find the single words in Vietnamese that best describe the items rather than translating the definitions of the items. They also checked the naturalness of the translation.

The purpose of the test and instructions on the interpretation of scores were also translated and attached to the test and served as explanations for learners before they did the test.

Trialling the Bilingual Test

The bilingual test was administered to 62 Vietnamese Year-3 English major learners in two classes at a university in Vietnam. Prior to the test, the instructions were given to the learners, and time was allowed for the learners to ask questions before they started the test. The learners were told to spend as much time on the test as they wished until they finished it. The first author invigilated the test and answered any questions learners had about the test. The results of the test were analysed to check the validity of the translated version. The data was used to see whether the test distinguished learners of different proficiency levels, and whether learners' scores dropped from one frequency level to the next?

Table 2. Three Groups of Learners Classified According to a Range of Proficiency Scores

Category of learners	Score range	Number of learners
Top	≥ 8	21
Middle	7	21
Lower	5-6	20

1. Does the test distinguish learners of different proficiency levels?

Because vocabulary knowledge is an important factor in language proficiency, high proficiency learners should gain a higher score on the Vocabulary Size Test than lower proficiency learners.

The learners were categorized into three equally sized proficiency groups, top, middle and lower learners according to the average score they achieved for the four subjects they took in the second semester of their second year of study. These four subjects were reading, writing, listening translation, and translation theory. At the time that the vocabulary size data was collected, the learners were studying in the third year of their Bachelor Degree in the four-year English course. The 21 learners who had the highest proficiency scores (8 or above out of 10) were placed in the top group. The middle group of learners consisted of the next 21 learners whose scores reached 7 but were less than 8. The remaining 20 learners were categorized as the lower learners who had scores ranging between 5 and 6. In cases where there were similar scores, the reading score was used to make a decision on whether or not to place learners in a higher or lower level of proficiency.

The mean scores on the test (test score multiplied by 100) achieved by lower, middle, and top learners were 6060.00, 6509.52, and 7385.71 respectively (Table 3). The top learners also achieved higher scores than the middle and lower learners in both the first and second seven levels of the vocabulary size test. The test thus distinguishes learners of different levels of proficiency. However, the standard deviation of the top learners group is the largest among the three groups of learners, indicating that the score differences among the top learners are the greatest. The opposite is true for the lower group of learners.

To determine if the differences among the three groups of learners were statistically significant, a one-way ANOVA was used. The results of the ANOVA indicate that the three categories of learners differed significantly from each other in their total score on the entire vocabulary test and in the scores for the first seven levels of the test. The F ratios for the total scores of the test and the first seven levels were $F(2, 61) = 3.081$, $p < .05$, and $F(2, 61) = 3.220$, $p < .05$ (Table 4). The differences among the three groups in the second seven levels of the test were not significant $F(2, 61) = 2.213$, *n.s.*, presumably because these later levels were largely beyond the vocabulary size of all the learners.

To identify differences among the three groups of top, middle and lower proficiency learners, post-hoc tests were performed. As shown in Table 5, the differences in the mean

Table 3. Descriptive Statistics for the Vocabulary Size Test for the Three Proficiency Levels of Learners

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Overall score								
Lower	20	6060.00	1481.607	331.297	5366.59	6753.41	3700	9900
Middle	21	6509.52	1612.422	351.859	5775.56	7243.49	4300	10600
Top	21	7385.71	2069.610	451.626	6443.64	8327.79	3300	11200
Total	62	6661.29	1802.126	228.870	6203.64	7118.94	3300	11200
Total 1st 7 levels								
Lower	20	4025.00	836.581	187.065	3633.47	4416.53	2900	5900
Middle	21	4347.62	884.658	193.048	3944.93	4750.31	3100	6400
Top	21	4757.14	1041.908	227.363	4282.87	5231.41	2900	6600
Total	62	4382.26	959.768	121.891	4138.52	4625.99	2900	6600
Total 2nd 7 levels								
Lower	20	2035.00	841.849	188.243	1641.00	2429.00	300	4000
Middle	21	2161.90	859.928	187.652	1770.47	2553.34	400	4200
Top	21	2628.57	1132.759	247.188	2112.95	3144.20	400	4600
Total	62	2279.03	974.787	123.798	2031.48	2526.58	300	4600

Table 4. The Results of an ANOVA For the Three Proficiency Level Groups

	Sum of Squares	df	Mean Square	F	Sig.
Total score					
Between Groups	1.874E7	2	9367643.625	3.081	.053
Within Groups	1.794E8	59	3040200.161		
Total	1.981E8	61			
1st 7 levels					
Between Groups	5529174.347	2	2764587.174	3.220	.047
Within Groups	5.066E7	59	858666.263		
Total	5.619E7	61			
2nd 7 levels					
Between Groups	4044860.983	2	2022430.492	2.213	.118
Within Groups	5.392E7	59	913862.389		
Total	5.796E7	61			

score for the entire test and the first seven levels of the test between the top and lower groups were significant ($p < .05$). The scores of the top learners also differed significantly ($p = .052$) from those of the lower learners in the second seven levels of the vocabulary test. The scores of the top learners did not differ significantly from the scores of the middle proficiency learners.

Because vocabulary knowledge is such a central factor in proficiency in a foreign language, it is reassuring to see that the Vocabulary Size Test can distinguish different proficiency levels. Because vocabulary knowledge is not the only factor in proficiency, this distinguishing effect is by no means perfect.

2. Do scores drop from one frequency level to the next?

There is a relationship between frequency of occurrence and the likelihood of a word being known (Read, 1988). That is, learners are more likely to know high-frequency words than low frequency words, because they meet high frequency words more often. If the bilingual Vocabulary Size Test is working well, it should show that learners perform better on the higher frequency words than the lower frequency words.

Table 6 shows a drop in scores from level 1 to level 14, although the drop in scores is not perfectly consistent. The learners achieved higher scores for the first seven levels and lower scores for the second seven levels. These mean scores in Table 6 then were sorted to show the means from smallest to largest (Table 7).

It can be seen from Table 7 that the means for the second (8.69) and the first (8.63) levels were the highest, followed by means for the fourth (6.63), third (5.92), fifth (5.06), sixth (5.03), and eighth (4.56) levels. The means dropped from levels seven, ten, eleven, 14, nine, 12 and 13 with 3.85, 3.76, 3.48, 3.13, 2.76, 2.68, and 2.42 respectively. The scores generally dropped from the first seven levels to the next second seven levels. However, the score did not necessarily drop from one frequency level to the next within

Table 5. Multiple Comparisons Between the Vocabulary Size Test Scores for Learners at the Top, Middle and Lower Proficiency Levels

Dependent Variable	(I) Academic rating	(J) Academic rating	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Overall score	Lower	Middle	-449.524	544.777	.413	-1539.62	640.57
		Top	-1325.714*	544.777	.018	-2415.81	-235.62
	Middle	Lower	449.524	544.777	.413	-640.57	1539.62
		Top	-876.190	538.092	.109	-1952.91	200.53
	Top	Lower	1325.714*	544.777	.018	235.62	2415.81
		Middle	876.190	538.092	.109	-200.53	1952.91
Total 1 st 7 levels	Lower	Middle	-322.619	289.521	.270	-901.95	256.71
		Top	-732.143*	289.521	.014	-1311.47	-152.81
	Middle	Lower	322.619	289.521	.270	-256.71	901.95
		Top	-409.524	285.968	.157	-981.74	162.70
	Top	Lower	732.143*	289.521	.014	152.81	1311.47
		Middle	409.524	285.968	.157	-162.70	981.74
Total 2 nd 7 levels	Lower	Middle	-126.905	298.681	.672	-724.56	470.75
		Top	-593.571*	298.681	.052	-1191.23	4.09
	Middle	Lower	126.905	298.681	.672	-470.75	724.56
		Top	-466.667	295.016	.119	-1056.99	123.66
	Top	Lower	593.571*	298.681	.052	-4.09	1191.23
		Middle	466.667	295.016	.119	-123.66	1056.99

*The mean difference is significant at the 0.05 level.

Table 6. Descriptive Statistics for Scores on Each of the Fourteen Levels of the Test for All Learners

	N	Minimum	Maximum	Mean	Std. Deviation
Level 1	62	7	9	8.63	.550
Level 2	62	5	10	8.69	1.125
Level 3	62	2	10	5.92	2.335
Level 4	62	3	10	6.63	1.952
Level 5	62	1	10	5.06	2.353
Level 6	62	1	9	5.03	2.239
Level 7	62	1	8	3.85	1.880
Level 8	62	1	10	4.56	2.420
Level 9	62	0	9	2.76	2.186
Level 10	62	0	8	3.76	2.163
Level 11	62	1	10	3.48	2.022
Level 12	62	0	7	2.68	1.871
Level 13	62	0	7	2.42	1.834
Level 14	62	0	6	3.13	1.769
Overall score	62	3300	11200	6661.29	1802.126
Total 1 st 7 levels	62	2900	6600	4382.26	959.768
Total 2 nd 7 levels	62	300	4600	2279.03	974.787
Valid N (listwise)	62				

Table 7. Means for the Fourteen Levels of the Test from Smallest to Largest

	N	Minimum	Maximum	Mean	Std. Deviation
Level 13	62	0	7	2.42	1.834
Level 12	62	0	7	2.68	1.871
Level 9	62	0	9	2.76	2.186
Level 14	62	0	6	3.13	1.769
Level 11	62	1	10	3.48	2.022
Level 10	62	0	8	3.76	2.163
Level 7	62	1	8	3.85	1.88
Level 8	62	1	10	4.56	2.42
Level 6	62	1	9	5.03	2.239
Level 5	62	1	10	5.06	2.353
Level 3	62	2	10	5.92	2.335
Level 4	62	3	10	6.63	1.952
Level 1	62	7	9	8.63	0.55
Level 2	62	5	10	8.69	1.125
Total 2 nd 7 levels	62	300	4600	2279.03	974.787
Total 1 st 7 levels	62	2900	6600	4382.26	959.768
Overall score	62	3300	11200	6661.29	1802.126
Valid N (listwise)	62				

Table 8. Average Scores for the First and Second Seven Levels for Four Groups Divided According to Total Score on the Test

	Score on 1 st 7 levels	Score on 2 nd 7 levels and (range)	Total score and (range)
Group 1 (15 learners)	3440	1120 (300-2200)	4560 (3300-5300)
Group 2 (17 learners)	3982	1958 (1200-2400)	5940 (5600-6600)
Group 3 (16 learners)	4475	2537 (1900-3000)	7012 (6700-7400)
Group 4 (14 learners)	5771	3043 (2500-4600)	8814 (7500-11200)

the first and the second seven levels of the test. There is a rough relationship between frequency level and the number of correct items, showing that this test is largely performing as it should. There is further explanation of the small inconsistencies in the drop in scores in the following section of this article.

Using the Test

Is it necessary to sit all fourteen levels of the test?

If the learners have a low proficiency in English, do they need to sit all 14 levels of the test if their vocabulary size is likely to be only a few thousand words? This question was answered by looking at the average score of the first seven of the 14 frequency based levels in the test and the average score of second seven levels. If most learners know a reasonable number of items in the later levels of the test, it is important that they sit the whole test in order to get the most valid measure of their vocabulary size. If lower scoring learners get few if any correct in the later levels, then sitting the later levels is a waste of their time. The total score on the vocabulary test was used to put learners into four different groups according to their total scores. Each group is roughly equal in size. The highest possible score on the test is 14,000.

As Table 8 shows, all learners answered some items correctly in the second half of the test with items correctly answered ranging from 3 out of 70 to 46 out of 70. Clearly, it is worth sitting all the 14 levels of the test. If 140 items are too many, then a 70 item version made by taking five items from each of the 14 levels is much better than only giving the first seven levels of the test. Beglar (2010), when looking at the monolingual Vocabulary Size Test, found that a 70 item version would perform as well as a 140 item version.

There are several reasons why learners get items correct at the later levels.

1. Learners could be guessing. It is important when using the Vocabulary Size Test that scores are not corrected for guessing, because the test is meant to be a measure of total vocabulary size. Adjusting for guessing distorts the vocabulary size measure. In future research, it would be useful to go through some learners' tests individually with them getting them to explain how they managed to answer some of the lower frequency items correctly. It is likely, as in the movie *Slumdog Millionaire*, that there are good reasons why they were able to answer correctly.

2. Some of the items on the test are loan words. In the Vietnamese bilingual version examined here, the following words are loan words in Vietnamese. They are spread through the levels of the test.

Level 1: 9. standard

Level 2: 9. microphone; 10. pro

Level 5: 7. miniature

Level 7: 1. olives; 10. yoghurt

Level 8: 5. eclipse

Level 11: 3. yoga

Level 12: 7. caffeine

Level 13: 3. rouble

This is a rather small number but is likely to increase as the influence of English on Vietnamese increases. It is important that loan words are kept in the test because they are an important part of a learner's English vocabulary. Getting them correct on the test reflects their transparency or low learning burden when meeting them in reading. There is also the possibility that some words are cognates or loan words in other languages known to some learners, like French or Russian. In bilingual test versions, the cognate L1 word form should not be used in the choices. The choice should be an L1 definition.

3. The sequencing of the items in the test depends on the range and frequency of occurrence of the words in the British National Corpus. The British National Corpus is necessarily a rather poor representation of Vietnamese learners' English language needs and experience with English. So, low frequency in the British National Corpus does not necessarily mean that the words have not been met before by the learners. The Vocabulary Size Test is divided into frequency levels partly to ensure that each frequency level is properly represented. A major problem with the dictionary-based tests used many years ago was that their sampling method was faulty and they were strongly biased towards high-frequency words. Sampling based on frequency level avoids this problem. However, it is not essential in the Vocabulary Size Test that the items are ordered by frequency levels, and it may be better, in order to make sure that learners do not give up on the lower frequency words, that there is a mixture of high and low frequency words spread through the test.
4. There may be poorly designed items in the test which give away the correct answer. The monolingual version of this test has had a lot of trialling to avoid this, and the results of Beglar's (2010) study suggest that the items are well written. However, in future research on the bilingual test, it would be useful to do a thorough item analysis, rewriting (but not replacing) items where necessary. It would also be useful to check the reliability of the test.

Future research could compare monolingual and bilingual test scores (within subjects and between subjects) to see to what degree the bilingual test results in higher scores and whether this is dependent on the proficiency level of the learners sitting the test. It would

also be useful to check if a 70 item test (five from each of 14 levels) gives the same scores as the 140 item test.

The Vietnamese bilingual version of the Vocabulary Size Test has been shown to be working well, and it is hoped that teachers of Vietnamese learners both inside and outside Vietnam will use it as a part of a needs analysis to inform the design of their English programmes. Without having a clear idea of how much vocabulary learners know, it is difficult to advise and support learners on their best options for increasing their vocabulary size.

Links

You can find the monolingual Vocabulary Size Test at these sites. The Vietnamese bilingual version and other bilingual versions are available at the first website listed.

<http://www.victoria.ac.nz/lals/staff/paul-nation.aspx>

<http://jalt-publications.org/tlt/resources/2007/0707a.pdf>

<http://www.lex tutor.ca/>

Funding acknowledgement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

- Bauer L, Nation ISP (1993) Word families. *International Journal of Lexicography* 6(4): 253-79.
- Beglar D (2010) A Rasch-based validation of the Vocabulary Size Test. *Language Testing* 27(1): 101-18.
- Hu M, Nation ISP (2000) Vocabulary density and reading comprehension. *Reading in a Foreign Language* 13(1): 403-30.
- Nation ISP (2006) How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review* 63(1): 59-82.
- Nation ISP, Beglar D (2007) A vocabulary size test. *The Language Teacher* 31(7): 9-13.
- Nation ISP, Webb S (forthcoming 2011) *Researching and Analysing Vocabulary*. Boston, MA: Heinle Cengage Learning.
- Read J (1988) Measuring the vocabulary knowledge of second language learners. *RELC Journal* 19(2): 12-25.
- Ruffell SK (2008) Reading recovery and the vocabulary knowledge of young children. Unpublished Masters thesis in Applied Linguistics, Victoria University of Wellington, New Zealand. C