Ho, H. & Huong, C. (2011). A multiple aspects quantitative indicator for ability of English vocabulary: vocabulary quotient..*Journal of Educational Technology Development and Exchange, 4*(1), 15-26.

A Multiple Aspects Quantitative Indicator for Ability of English Vocabulary: Vocabulary Quotient

Hong-Fa Ho
Chen Huong
National Taiwan Normal University

Abstract: *Students of EFL have a common problem with insufficient vocabulary. Some studies argue that English vocabulary is one of the most difficult parts. This paper applied the idea of KPI (Key Performance Indicator) from management science to EFL vocabulary learning and teaching. A vocabulary quotient (VQ in short) with four test models, including a listening comprehension test, was proposed as KPIs. Based on VQ, VQ testing software was developed. To test the validity and reliability of the assessment tool based on VQ, assessments with 54 junior high students (n=54) was conducted. The findings of this study were generalized as the following: (1) the relationships between the score of general English proficiency test and scores of VQ tests were significant positive correlations respectively and (2) the relationships among scores of VQ tests were significant positive correlations. Results of this study suggest that VQ could be considered to be a KPI of EFL vocabulary teaching and learning. The proposed method is suggested to estimate one's vocabulary sizes.*

Keywords: EFL, vocabulary size, English vocabulary, KPI, vocabulary quotient

1. Introduction

English has been increasingly important in recent years. Based on Johnson's (2003) statement, there were approximately one billion people learning English globally. Johnson also referred to most countries having adopted an English learning policy that they will use English as their first or official language. Besides, in most countries in the world, students have regarded English as their priority when they decide to select one foreign language to learn (Chang, Yeh, Joe, You, Chern, & Liao, 2007).

Reading abilities and vocabulary knowledge have received greater emphasis than they had in the past (Huckin, Haynes, & Coady, 1993; Walters & Bozkurt, 2009). The EFL (English as a Foreign Language) vocabulary was by far the most unmanageable part of language instruction (Tsai & Chang, 2009). Also, a poor vocabulary was acknowledged by most students at all levels (Lin, 2002; Segler, Pain, & Sorace, 2002). Therefore, vocabulary learning plays an important role in English-language acquisition as Laufer and Girsai (2008) suggest "when reading a text, or engaging in a group discussion, learners may come across unfamiliar words and look them up in a dictionary. The activity constitutes Focus on Form since the words attended to be necessary tools for task completion." Beglar and Hunt (2005, p.7) reminds us that "vocabulary acquisition is a crucial, and in some senses, the central component in successful foreign language acquisition." Chujo and Oghigian (2009, p.122) mentioned that vocabulary is the heart of a language. Learners depend on vocabulary as their

first resource (Huckin & Bloch, 1993) and a rich vocabulary makes the skills of listening, speaking, reading, and writing easier to perform (Nation, 1994).

Many researchers of vocabulary assessments focused on spelling, cognition, and usage. Some Yes/No tests as measures of receptive vocabulary knowledge were proposed and examined (Beeckmans, Eyckmans, Janssens, Dufranne, & Velde, 2001; Huibregtse, Admiraal, & Meara, 2002; Mochida & Harrington, 2006). However, the guess rate is 50%. Cameron (2002) proposed a method to measure the vocabulary size in English as an additional language. Eyckmans (2004) discussed how to measure receptive vocabulary size and Meara and Buxton (1987) proposed multiple-choice vocabulary tests. Meara and Jones (1988) proposed to use vocabulary test as a placement indicator. Meara (2011) provided various vocabulary test software tools on the Internet, such as P_Lex, X_Lex and Y_Lex, with different features. Studies discussed earlier did not consider vocabulary listening comprehension. Ho and Lin (2010) proposed Chinese character quotient and test models. A Chinese character listening comprehension test was included. However, no research has focused on the relationships between EFL vocabulary tests and the general English proficiency test (abbreviated as GEPT).

Some studies explored needed vocabulary sizes for different famous tests such as TOEIC and TOEFL based on the text coverage. Chujo and Oghigian (2009) claimed that in order to gain 95% coverage on TOEIC, a reader would need a minimum vocabulary size of 4,000 words. TOEFL requires a 4,500-word vocabulary, and a 5,500-word vocabulary is needed for EIKEN Pre-1st Grade. These needed vocabulary sizes could be defined as expected objective vocabulary KPIs. However, Chujo and Oghigian did not discuss about how to estimate the students' vocabulary sizes.

From the studies discussed earlier, we know that there are some vocabulary assessments focused on various aspects of vocabulary. We also know that teachers and students need a low-cost feasible method to estimate students' vocabulary sizes for EFL teaching and learning. Thus, one may ask what elements are included in the method and is the method good enough to estimate students' vocabulary sizes. If students' scores of GEPT are given, what is the relationship between the score of GEPT and the score of each element of the low-cost feasible method? Building on studies discussed earlier, the aim of this study is to propose some varied fundamental test models that can be used to estimate students' vocabulary abilities for EFL and to define some of the parameters used in EFL vocabulary teaching and learning, specifically regarding how variables such as the ability of word spelling, the ability of word recognition, and the ability of word listening comprehension relate to the variable of the score of GEPT, as well as to investigate the relationships among those variables. The independent variable (GEPT vs. English vocabulary spelling vs. English vocabulary reading comprehension vs. English vocabulary listening comprehension vs. English vocabulary spelling selection test) of this study was manipulated as a repeated measure. The dependent variable contains the scores of all tests.

More specifically, the research questions of this study are:

1. Is there a relationship between the score of English vocabulary spelling, reading comprehension, listening comprehension, or spelling selection test and the score of GEPT respectively for students of Chinese junior high school?
2. Are there relationships among scores of English vocabulary spelling, reading comprehension, listening comprehension and spelling selection test for students of Chinese junior high school?

The hypotheses of this study are stated as following:

H_1: There is a significant positive correlation between the score of English vocabulary spelling, reading comprehension, listening comprehension, or spelling selection test and the score of GEPT for students of Chinese junior high school respectively.

H_2: There are significant positive correlations among scores of English vocabulary spelling, reading comprehension, listening comprehension and spelling selection test for students of Chinese junior high school.

To test these hypotheses, we developed a vocabulary quotient with four test models for EFL Chinese students and the vocabulary test software that tests English vocabulary spelling, reading comprehension, listening comprehension and spelling selection. Findings from the study demonstrate that both hypotheses are supported. Thus, the vocabulary quotient of English for Chinese is suggested to be an objective quantification KPI for English teaching and learning. *VQ* can further be used for various *L*1*L*2 with appropriate modifications. In terms of costs for performing the assessment, the cost is low due to using developed computer software. Hence, the assessment is feasible in a computer classroom. A student's vocabulary size can be calculated from the student's *VQ*.

2. Method

2.1. Participants

Fifty-four students from the first grade in a junior high school, Chung Dau International School in Taiwan, were randomly selected to participate in the study. The basic prerequisite for selection was at least three years of mandatory English classes. Additionally, participants had no or minimal experience of using vocabulary testing software before participating in the assessments. The fifty-four participants included twenty-four male students and thirty female students, ranging in age from 13 to 14 years old. They had no known hearing problems, and had normal or corrected-to-normal vision.

2.2. Material

The testing sheets, which contain English words randomly selected from top 4000 high frequency words of the British National Corpus (BNC) (Burnard, 2000; Kennedy, 2003; Leech, Rayson, & Wilson, 2001), were designed by the researcher. The BNC is one of the largest electronically-accessible corpora consisting of over 100 million words in British English. It consists of an approximately 90 million-word written component of informative and imaginative text, and a 10 million word spoken component. All Chinese meanings of English vocabulary came from an English-Chinese dictionary. Pronunciations of vocabulary came from American announcers.

2.3. Vocabulary Quotient

For better managing the ability of English vocabulary of learners, objective quantification data helped. Introducing the concept of KPI in the management science to EFL, objective quantification data not only indicated the current status but also helped plan better teaching and learning plans. Many researchers pointed out that English vocabulary ability was an important factor in English ability. An English vocabulary quotient with four basic test models was stated to provide an alternative to get an objective quantification data. By using the item response theory (Baker, 1992; Hambleton & Swaminathan, 1985; Hulin, Drasgow, & Parsons, 1983), item characteristic functions of four basic test models had been derived later.

An English vocabulary quotient, or in the abbreviation form of *VQ*, is a set of quotients derived from several different standardized test models designed to assess the ability of English vocabulary.

Definition: The English vocabulary quotient is a function, $VQ\ (p, m, L, t, z) = \{VQ_i \mid 1 \leqq i \leqq d\}$, where

- p is the examinee,
- m is the native language of p, such as Chinese in this study,
- L is the set of target English vocabulary,
- t is the date of the test,
- z is the number of test items for each test model, and z is a positive integer,
- VQ_i is the quotient of the number of correct answers divided by z of some specific test model i about English vocabulary in L, and
- d is the number of different test models.

Note that, the target English vocabulary is selected randomly from L by using the uniform distribution random function.

Definition can be applied to any language. Both *L*1*L*1 and *L*1*L*2 are suitable. Different languages have different number of test models. *VQ* is time-variable for individual. If fhe parameter z is larger, the accuracy of *VQ* is higher. Examinees will spend more time to carry out tests, however.

Considering vocabulary abilities, there are various aspects that relate to this, for example, spelling, reading, listening, speaking, usages, collations, and others. Different languages have different aspects of vocabulary abilities. Each aspect needs at least a test model in the English vocabulary quotient. d stands for the total number of test models in *VQ*. If d is larger, the number of needed test items in the database is larger. The cost of establishing test items is higher. In addition, the time cost of the assessment is proportional to d and z. Because of the budget limitation and unacceptable voice recognition technology, the basic abilities of English vocabulary considered in this study only include spelling, reading comprehension, listening comprehension, and spelling selection.

For EFL, four basic test models ($d = 4$) are proposed. They are described in Table 1. It is easy to change Chinese into other language for these four basic test models.

Table 1: Four Basic test Models of English Vocabulary

Test Model	Description
Test Model 1	Given $C_p(E_t)$ with $V(E_t)$. Examinee is asked to key-in or write the target English vocabulary E_t.
Test Model 2	Given E_t and 4 choices $C_p(E_t)$, C_q, C_r and C_s in random sequence. $C_p(E_t)$ is the best choice, and C_q, C_r and C_s are wrong choices. Examinee is asked to choose the best choice.
Test Model 3	Given $V(E_t)$ and 4 choices $C_p(E_t)$, C_q, C_r and C_s in random sequence. $C_p(E_t)$ is the best choice which matches the given $V(E_t)$, and C_q, C_r and C_s are wrong choices. Examinee is asked to choose the best choice.
Test Model 4	Given $V(E_t)$ and 4 choices E_t, E_i, E_j and E_k in random sequence. E_t is the best choice which matches the given $V(E_t)$, and E_i, E_j and E_k are wrong choices. Examinee is asked to choose the best choice.

For reading easily and accurately, some symbols are defined in advance as follows:

- E_t stands for the target English vocabulary,
- $C_p(E_t)$ stands for Chinese meaning of E_t,
- $V(E_t)$ stands for the pronunciation of E_t,
- C_q, C_r and C_s stand for other Chinese meanings, where $p \neq q \neq r \neq s$, and
- E_i, E_j and E_k stand for incorrect English words which are similar to E_t, where $t \neq i \neq j \neq k$.

The aim of these aforementioned test models is to find out different aspects of the basic ability of English vocabulary recognition for EFL learners. The purpose of Test Model 1 is to find out if the examinee can spell the English vocabulary correctly with or without pronunciation. Test Model 2 is to find out whether the examinee understands the meaning of the target English vocabulary or not. Test Model 3 is to examine the ability of listening and comprehension of the target English vocabulary. Test Model 4 is to test if the examinee can distinguish between right and wrong spelling. In Test Model 4, vowels of E_t are replaced by different and similar vowel characters. These four basic test models could be easily implemented by software. By using the software, it is a low-cost way to find out different aspects of the basic ability of English vocabulary recognition for EFL learners.

2.3.1. Analysis based on Item Response Theory

Because all test items were selected randomly by computers, every participant had different test items. Traditional reliability and validity cannot be derived in this case. Four test models are analyzed by Item Response Theory as following. The analysis provides item characteristic functions which describe the feature of test models.

For Test Model 1, two-parameter logistic model is considered because there is no chance to guess. The formula for the item characteristic function with two-parameter is

$$P_i(\theta)=\frac{e^{a_i(\theta - b_i)}}{1 + e^{a_i(\theta - b_i)}}, \quad i = 1, 2, ..., z, \qquad (1)$$

where $P_i(\theta)$ stands for the probability of any examinee with ability θ correctly answer test i; a_i stands for item discrimination parameter; b_i stands for difficulty parameter; z is the number of test items.

For Test Model 1, a_i is a constant. Formula (1) becomes one-parameter logistic model.

Considering b_i of (1), there are two cases in the Test Model 1. The pronunciation of the target word is given or not. In practice, examinee responded that it becomes harder if the pronunciation is not given. Both Ho (2006) and Chen and Chung (2008) took the length of word as a factor of b_i. Chen and Chung also took phonetic and weight parameters of a word as two more factor of b_i. In addition, if a word is brand new for the examinee, it is very difficult to spell; if a word has been learned, it is easier. Based on the above, the formula of b_i of Test Model 1 is

$$b_i = (L_i \times 0.7 + h_i \times 0.3) \times G_i \times B_i \times K_i, \quad (2)$$

where b_i is the difficulty parameter of the ith test, L_i is the length parameter of the ith test, h_i is the phonetic parameter of the ith test, G_i is the weight parameter of the corresponding GEPT grading level of the ith vocabulary, B_i is the never-learned parameter of the ith vocabulary, and K_i is the pronunciation parameter of the ith vocabulary. If the vocabulary is brand new for the examinee, then B_i is 1; else B_i is 0.1. If the pronunciation of the vocabulary is given, then K_i is 0.1; else K_i is 1. Both B_i and K_i dramatically determine the difficulty parameter b_i.

For Test Models 2, 3 and 4, three-parameter logistic model is considered because the examinee is able to guess. The formula for the item characteristic function with three-parameter is

$$P_i(\theta) = c_i + (1 - c_i)\frac{e^{a_i(\theta - b_i)}}{1 + e^{a_i(\theta - b_i)}},$$

$$i = 1, 2, ..., z, \quad (3)$$

where c_i stands for guessing parameter, c_i =0.25.

Considering b_i of (3), there is nothing to do with both the length and the phonetic of the word because the test is not about spelling. The formula of b_i of (3) is:

$$b_i = G_i \text{ x } B_i \text{ x } K_i \quad (4)$$

where G_i, B_i and K_i are the same definitions in formula (2). If a_i is a constant and c_i=0.25, the formula (3) becomes a one-parameter logistic model.

Different test models have different item characteristic functions. These functions show features of test models.

2.3.2. Estimating Vocabulary Sizes based on VQ

For a learner whose native language is not English, *VQ* is a four-tuple function based on above four basic test models. *VQ* function can be expressed as following:

$$VQ(p, m, L, t, z) = \langle VQ_1, VQ_2, VQ_3, VQ_4 \rangle. \quad (5)$$

Note that each test model estimates a vocabulary size of a specific aspect. Let $|L|$ denote the number of vocabulary in L. There are four different estimated vocabulary sizes which are calculated by:

Estimated vocabulary size of Test Model

$$i = |L| \text{x} VQ_i, \quad 1 \leqq i \leqq 4. \quad (6)$$

How can one estimate an individual's English vocabulary size based on the *VQ*? First of all, one must take the *VQ* assessment and get one's $\langle VQ_1, VQ_2, VQ_3, VQ_4 \rangle$. Secondly, one can get four estimated vocabulary sizes by formula (6). Based on Definition, vocabulary size is an array of numbers.

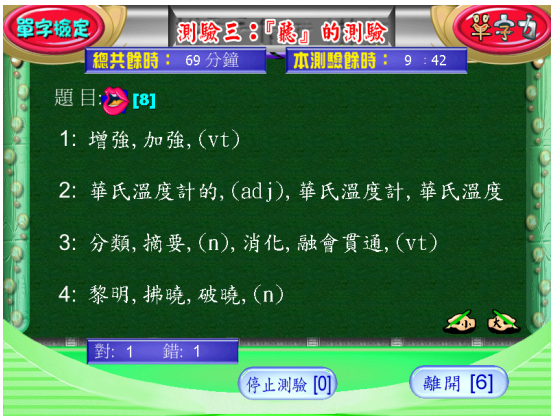
2.4. Tools

An English vocabulary testing tool based on *VQ* had been developed by researchers. All material mentioned above was implemented in the testing tool. Four test models' screen captures of the testing tool are shown in Figure 1.

In Figure 1(a), Chinese meanings, lexical categories, and the number of characters of the target English word are shown, the pronunciation of the target English word is pro-

(a) Test Model 1

(b) Test Model 2



(c) Test Model 3

(d) Test Model 4

Figure 1. Screen captures of English vocabulary testing tool

vided, and examinees are asked to key-in the spelling on a keyboard. In Figure 1(b), the target English word and four choices of Chinese meanings with lexical categories are shown, the pronunciation of the target English word is provided, and examinees are asked to click the answer by the mouse of computer. In Figure 1(c), the pronunciation of the target English word is pronounced, four choices of Chinese meanings with lexical categories are shown, and examinees are asked to click the answer with the mouse of computer. In Figure 1(d), the pronunciation of the target English word is pronounced, four choices of similar spellings are shown, and examinees are asked to click the answer with the mouse. Note that vowels of the target English word are replaced by different and similar vowel characters to form wrong choices.

In the assessment, the testing tool provided stimuli for participants through computers based on four test models stated in Table 1. Each test model had 100 tests. The range of the score of each test is 0~100. Scores of all tests were automatically collected with the testing tool.

Multimedia computers with headphones were used by participants to carry out the English vocabulary tests. The Statistical Package for the Social Science (SPSS) 18.0 for Microsoft Windows was used to analyze the collected data.

2.5. Design

A single factor design was used. The independent variable (GEPT vs. English vocabulary spelling vs. English vocabulary reading comprehension vs. English vocabulary listening comprehension vs. English vocabulary spelling selection test) was manipulated as a repeated measure. The dependent variable contains the scores of tests. The scores of five tests range from 0 to 100. Scores of assessments were analyzed by descriptive statistics and the Pearson product-moment correlation.

2.6. Procedure

At the very beginning, participants were trained to have the ability of typing on a computer for nine hours within three weeks. Before using the testing tool, participants were all familiar with the computer keyboard. Participants had a regular GEPT held by the school. Researchers assume that the regular GEPT had good validity and reliability. Scores of the regular GEPT of participants, ranging from 0 to 100, were recorded manually. Three days after the regular GEPT, a course comprised a half-hourly teaching, ten-minute rehearsal, and a computer simulation of the assessment for thirty minutes

was conducted. After that participants were invited to perform the formal assessment for thirty minutes to assess their vocabulary abilities. Moreover, all of them had sufficient time to take these four tests. All scores of these tests were collected automatically by computers.

3. Results

Table 2 presents the means and standard deviations of the GEPT (*M*=59.4630, *SD*=19.62475), Test Model 1 (*M*=9.2037, *SD*=7.04334), Test Model 2 (*M*=38.7037, *SD*=9.74898), Test Model 3 (*M*=36.0741, *SD*=8.43365), and Test Model 4 (*M*=50.3704, *SD*=11.98788). Because the standard deviation of GEPT was greater than Test Models 1, 2, 3 and 4, the difference of GEPT scores of participants existed. In addition, because the mean and standard deviation of Test Model 1 were smallest among Test Models 1, 2, 3 and 4, the difference of Test Model 1 scores was less than Test Models 2, 3 and 4. From the view point of mean scores, the sequence of hard-level was Test Model 1, 3, 2, 4 and GEPT.

Table 2: *M* and *SD* of GEPT and the Four Tests (n = 54)

	M	*SD*
GEPT	59.4630	19.62475
Test Model 1	9.2037	7.04334
Test Model 2	38.7037	9.74898
Test Model 3	36.0741	8.43365
Test Model 4	50.3704	11.98788

Pearson correlation matrix is presented in Table 3. There were significant positive correlations (r = .791, .826, .866, .784) among scores of GEPT, Test Models 1, 2, 3 and 4. Thus, hypothesis H_1 was supported. There were significant positive correlations (r = .829, .695, .736, .659, .752, .722) among scores of Test Models 1, 2, 3 and 4. Thus, hypothesis H_2 is supported.

Table 3: Correlations among Score of GEPT and scores of the Four Tests

Pearson correlation analysis **(*n* = 54)**

Context	Pearson correlation / significance	GEPT	Test Model 1	Test Model 2	Test Model 3	Test Model 4
GEPT	*r*	1				
	p	—				
Test Model 1	*r*	.791**	1			
	p	.000	—			
Test Model 2	*r*	.826**	.829**	1		
	p	.000	.000	—		
Test Model 3	*r*	.866**	.695**	.736**	1	
	p	.000	.000	.000	—	
Test Model 4	*r*	.784**	.659**	.752**	.722**	1
	p	.000	.000	.000	.000	—

**$p < .01$

4. Discussion

According to the hypotheses, the discussions were presented as the following:

1. This study assumed the validity and reliability of the general English proficiency test were good. Therefore, the validity and reliability of Test Models 1, 2, 3 and 4 were

good because Pearson correlations were over .659. Although item characteristic functions were derived, it is hard to tell the features of *VQ* with four test models directly. Because hypothesis H_1 was supported, it gives *VQ* with four test models a good support.

2. There were significant positive correlations among scores of English vocabulary spelling, reading comprehension, listening comprehension, and spelling selection test for students of Chinese junior high school. The study found that the Pearson correlation of Test Models 1 and 3 was greater than the Pearson correlation of Test Models 1 and 4. In addition, Test Model 1 is a spelling test and Test Model 4 is a spelling selection test. Spelling is harder than spelling selection. Therefore, Test Model 4 can be omitted if the cost of time is a major concern.

Results also demonstrated that spelling ability was the worst. This implies that ability of writing has no good base. On the other hand, abilities of reading comprehension and listening comprehension have better bases. EFL teachers and students could use *VQ* to find out their weaknesses of vocabulary. *VQ* could also be a KPI for managing vocabulary ability of EFL teaching and learning.

From the view point of mean scores in Table 2, the sequence of hard-level was Test Model 1, 3, 2, 4 and GEPT. Spelling test (Test Model 1) is always a difficult problem for most students. Note that Test Model 2 assessed the reading comprehension ability and Test Model 3 assessed the listening comprehension ability, and mean scores of them were close. The reason for this phenomenon is that participants can never know the meanings of unknown vocabulary even thought participants see the word and listen to the voice of the word. For unknown vocabulary, the probability of clicking the right answer is .25. Thus, the discriminations of Test Models 2 and 3 are close. However, some participants' scores of Test Model 2 were obviously greater than their scores of Test Model 3; some were in the converse. By taking both Test Models 2 and 3, teachers could differentiate between abilities of reading comprehension and listening comprehension of vocabulary. Because Test Model 4 provided voice of the target word and four choices, participants chose the answer with the help of the voice and choices. This was why the mean score of Test Model 4 was significantly greater than that of Test Model 1.

5. Conclusion

Students of EFL have a common problem of insufficient vocabulary. Some studies argue that English vocabulary is one of the most difficult parts in teaching and learning. Teachers and students of EFL need a way to manage vocabulary. For managing vocabulary teaching and learning, a low-cost feasible assessing method, an objective quantification, a vocabulary size estimating method, and an assessment tool are needed.

This paper applied the idea of KPI (Key Performance Indicator) from management science to EFL vocabulary learning and teaching. A vocabulary quotient (*VQ* in short) with four test models including a listening comprehension test was proposed as KPIs. The proposed *VQ* could be applied to other languages with modifications.

Based on *VQ*, a *VQ* testing software was developed. To test the validity and reliability of the assessment tool based on *VQ*, assessments with 54 junior high students (n=54) was conducted. The findings of this study were generalized as the following: (1) the relationships between the score of general English proficiency test and scores of *VQ* tests were significant positive correlations respectively, and (2) the relationships among scores of *VQ* tests were

significant positive correlations. Results of this study suggested that *VQ* could be considered to be a KPI of EFL vocabulary teaching and learning. The proposed method is suggested to estimate one's vocabulary size. Both the number of participants (*n*=54) included in this study and relying only on one well-known test, that is GEPT, are the limitations of this study.

VQ is suggested to play one of the KPIs of EFL teaching and learning. The most unmanageable part of the EFL language instruction could be indicated by *VQ*. It could be applied to other *L*1*L*2 with modifications. Test Model 3 assessing the ability of vocabulary listening comprehension could be useful to find out students' basic building block of general listening comprehension skills. The *VQ* tool is also a low-cost assessment tool because tests are handled by computers within thirty minutes.

VQ with four test models does not cover the speaking ability because the accuracy of English vocabulary voice recognition technology is far from the minimal request. In the future, if English vocabulary voice recognition technology matures, speaking ability of vocabulary should be added to *VQ*. Furthermore, four basic test models of *VQ* are far from various needs. One can later add necessary testing models to the *VQ* to fulfill one's needs easily.

References

Baker, F. B. (1992). *Item response theory: Parameter estimation techniques*. New York: Marcel Dekker.

Beeckmans, R., Eyckmans, J., Janssens, V., Dufranne, M., & Velde, H. (2001). Examining the Yes/No vocabulary test: some methodological issues in theory and practice. *Language Testing, 18,* 235-274.

Beglar, D., & Hunt, A. (2005). Six principles for teaching foreign language vocabulary: A commentary on Laufer, Meara, and Nation's "ten best ideas." *The Language Teacher, 29*(7), 7-10.

Burnard, L. (2000). *Reference guide for the British National Corpus (World Edition)*. Retrieved February 16, 2009 from the World Wide Web http://www.natcorp.ox.ac.uk/World/HTML/thebib.html

Cameron, L. (2002). Measuring vocabulary size in English as an additional language. *Language Teaching Research, 6,* 145-173.

Chang, W. C., Yeh, H. N., Joe, S. G., You, Y. L., Chern, C. L., & Liao, M. L. (2007). The investigation of the influence of English teaching from the policy of Taiwan English education. *Selected papers from the proceedings of 2007 international conference and workshop on TEFL & applied linguistics compiled by Department of Applied English Ming Chuan University* (pp.672-686). Taipei, Taiwan: Crane.

Chen, C. M., & Chung, C. J. (2008). Personalized mobile English vocabulary learning system based on item response theory and learning memory cycle. *Computers & Education*, *51*, 624-645.

Chujo, K., & Oghigian, K. (2009). How Many Words Do You Need to Know to Understand TOEIC, TOEFL & EIKEN? An Examination of Text Coverage and High Frequency Vocabulary. *The Journal of Asia TEFL 6*(2), 121-148.

Eyckmans, J. (2004). *Measuring receptive vocabulary size*. Utrecht, the Netherlands: LOT.

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications.* Kluwer Academic Publisher.

Ho, H. F. (2006). A Scientific Method of Memory, *Proc. of IADAT International Conference on Automation, Control and Instrumentation*, 79-83.

Ho, H. F., & Lin, P. Z. (2010). Chinese character Learning Review System with Item Re-

sponse Theory Analysis, *Proc. of Research in Reading Chinese (RRC) Conference*, 84-100.

Huckin, T., & Bloch, J. (1993). Strategies for inferring word-meanings in context: A cognitive model. In T. Huckin, et al. (Eds.) *Second language, reading and vocabulary acquisition* (pp. 153-180). Ablex, NJ: Norwood.

Huckin, T., Haynes, M., & Coady, J. (1993). *Second language reading and vocabulary learning*. Norwood, NJ: Ablex Publishing Corporation.

Huibregtse, I., Admiraal, W., & Meara, P. (2002). Scores on a yes/no vocabulary test: correction for guessing and response style. *Language Testing 19*, 227-245.

Hulin, C. L., Drasgow, F., & Parsons, C. K. (1983). *Item response theory: Applications to psychological measurement*. Homewood, IL: Irwin.

Johnson, K. (2003). *Designing language teaching tasks*. Basingstoke, HA: Palgrave Macmillan.

Kennedy, G. (2003). Amplifier collocations in the British National Corpus: Implications for English language teaching. *TESOL Quarterly, 37*(3), 467-487.

Laufer, B., & Girsai, N. (2008). Form-focused Instruction in Second Language Vocabulary Learning: A Case for Contrastive Analysis and Translation. *Applied Linguistics, 29*(4), 694-716.

Leech, G., Rayson, P., & Wilson, A. (2001). *Word frequencies in written and spoken English*. Harlow, CA: Pearson Education Limited.

Lin, Z. (2002). Discovering EFL learners' perception of prior knowledge and its role in reading comprehension. *Journal of Research in Reading, 25*(2), 172-190.

Meara, P. (2011, March). Free software from_ lognostics. Retrieved from: http://www.lognostics.co.uk/tools/index.htm

Meara, P., & Buxton, B. (1987). An alternative to multiple choice vocabulary tests. *Language Testing, 4*, 142-154.

Meara, P., & Jones, G. (1988). Vocabulary size as a placement indicator. In P Grunwell (Ed.), *Applied Linguistics in Society*. London: CILT.

Mochida, A., & Harrington, M. (2006). The Yes/No test as a measure of receptive vocabulary knowledge. *Language Testing, 23*, 73-98.

Nation, P. (1994). *New ways in teaching vocabulary*. Alexandria, VA: TESOL, Inc.

Segler, T. M., Pain, H., & Sorace, A. (2002). Second language vocabulary acquisition and learning strategies in ICALL environment. *Computer Assisted Language Learning, 15*(4), 409-422.

Tsai, C.C., & Chang, I. C. (2009). An examination of EFL vocabulary learning strategies of students at the university of technology of Taiwan. *International Forum of Teaching and Studies, 5*(2), 32-38.

Walter, J., & Bozkurt, N. (2009). The effect of keeping vocabulary notebooks on vocabulary acquisition. *Language Teaching Research, 13*(4), 403–423.

Contact the Authors

Hong-Fa Ho, Ph.D.
Department of Applied Electronics Technology
National Taiwan Normal University
E-mail: jackho@ntnu.edu.tw

Chen Huong, Ph.D.
Graduate Institute of Political Science
National Taiwan Normal University
E-mail: hc0815@gmail.com