Mohd Ayub A. F. et al. (Eds.) (2011). Workshop Proceedings of the 19th International Conference on Computers in Education. ChiangMai, Thailand: Asia-Pacific Society for Computers in Education.

The Developments of EFL Vocabulary Sizes of High School Students in Taiwan

Hong-Fa Ho
National Taiwan Normal University, Taipei, Taiwan, R.O.C.
jackho@ntnu.edu.tw

Abstract: EFL students tend to have a common problem of insufficient vocabulary. Some studies argued that English vocabulary was one of the most difficult areas in terms of learning English. The participants (*N*=1938) were asked to take the English vocabulary tests in spelling, reading, and listening. After the tests, vocabulary Quotient (*VQ*) was used to estimate the actual size of English vocabulary of the participants. The findings of this study were generalized as the following: (1) the growing rate of English vocabulary in spelling was about 1197 words per year, the growing rate of English vocabulary in reading was about 1443 words per year, and the growing rate of English vocabulary in listening was about 1402 words per year, (2) the growth of vocabulary size in spelling was the least, and (3) the growth of vocabulary size in reading was the greatest. Results of this study suggested that English learners did not tend to equally develop their vocabulary sizes in the areas of spelling, reading, and listening.

Keywords: EFL, vocabulary size, vocabulary quotient

1. Introduction

The importance of having proficient English ability has been emphasized nowadays. Based on Johnson's statement, there were about one billion people learning English in the world, and English has become a first or an official language in most countries [3]. In addition, English is considered as students' priority to learn a foreign language in the world [13].

Reading ability has an impact on learning ability. There was a reciprocal causation between students' individual cognition and their reading ability. Students' reading abilities was associated with their learning abilities [23]. This is the well-known Matthew's effects in reading.

If one wanted to learn English well, it would be necessary to acquire sufficient English vocabulary. The more sufficient English vocabulary one had, the more proficient one's English ability would be [1]. Most college students were required to read different kinds of English documents, such as online information and textbooks. Therefore, acquiring sufficient vocabulary was an important element of building English ability.

Several assessments were used to measure English learners' vocabulary size, such as the Eurocentres Vocabulary Size Test [16, 17], and the Vocabulary Levels Test [19, 20]. Meara proposed the concept of V-Size, and the computer software was developed to test English learners' vocabulary ability [15]. When the software e-rater was measuring English learners' writing ability, their vocabulary size was measured as well [1, 5]. It was noticed that the English vocabulary tests were mostly focused on spelling, recognition, and usage in previous studies. For instance, Yes/No question was used to test whether the participants

could recognize the word [2, 6, 11, 18]. However, since the guess rate was 50%, Yes/No question was not considered as a perfect measurement. Meara and Buxton presented another way of assessment: the method of multi-choice [16]; however, no study had yet included the test of listening to measure the vocabulary acquisition of English learners. Therefore, when Ho and Lin proposed the concept of Chinese character quotient [9], they included the test of listening as one of the evaluation methods.

In terms of reading, it was indicated that reading was one of the important methods of knowledge acquisition. A total of one million words was found in Brown corpus [21], and 200 million words in Collins Cobuild corpus were analyzed. If the number of learned vocabulary was about 1000 words, the text coverage would be 72%. Therefore, the context was commanded about 72%. If the number of learned vocabulary was 15851 words, the text coverage would be commanded more than 97.8% [8, 14]. Therefore, the vocabulary size really had a great impact on one's reading ability.

Listening comprehension was playing an important role in a conversational context [22]. The unique characteristic of a dialogue had an effect on listening comprehension, that is, the features of pronunciation-reduced [7, 12]. The word would not be recognized when different pronunciations were presented at the same time [4]. Therefore, listeners would need to have strong listening abilities in order to recognize the spoken word.

Ho and Huong proposed the vocabulary quotient (*VQ*) as a KPI (Key Performance Indicator) of EFL teaching and learning [10]. *VQ* was a quantitative indicator for the acquisition of English vocabulary in multiple aspects, including spelling, visual recognition, and audio recognition. Based on *VQ*, one's vocabulary sizes could be estimated.

The purpose of this study was to investigate the developments of English vocabulary sizes of junior and senior high school students in Taiwan. This study used *VQ* and its computer software tool to assess vocabulary sizes from the aspects of spelling, reading, and listening.

Three research questions are directing this study:
R1: How do EFL high school students' vocabulary sizes grow in multiple aspects?
R2: Which vocabulary size of spelling, reading, and listening do students grow the least?
R3: Which vocabulary size of spelling, reading, and listening do students grow the most?

The hypotheses of this study are:
H_1: The growths of EFL high school students' vocabulary sizes in multiple aspects are greater than 4800 words.
H_2: The growth of vocabulary size of spelling is the least.
H_3: The growth of vocabulary size of reading is the most.

2. Method

2.1 Participants

The participants in this study were selected from the contestants of the National Spelling Competition in 2010 (*N*=1938). 1156 elite students came from 169 junior high schools, and 782 elite students came from 107 senior high schools. Table 1 showed the composition of the participants. From each school, ten or less students with good English ability were recommended by the English teachers to participant in the spelling competition. The prerequisite for selection was: junior high school students had to complete at least three

years of mandatory English classes. Senior high school students had to complete at least five years of mandatory English classes. All the contestants had no known hearing problems, and had normal or corrected-to-normal vision.

Table 1: Participants of two groups (N=1938)

Grade	Grade 7th~9th	Grade 10th~12th
n	1156	782

2.2 Material

The Ministry of Education in Taiwan had recommended a vocabulary list of 2200 words for junior high school students. Basically, all the junior high school English textbooks were edited according to this vocabulary list. Thus, this vocabulary list was used as the lexicon to test the junior high school students. According to the College Entrance Examination Center in Taiwan, 7000 English words were recommended for senior high school students to acquire. Between the vocabulary list of 2200 words and the vocabulary list of 7000, there were 4800 words in difference. These 4800 new words were used as the lexicon to test the senior high school students in the competition.

2.3 Tools

The software of the English vocabulary test based on *VQ* was used to test the participants. Table 2 presents three test models of English Vocabulary Test software [10]. Symbols were defined as the following: E_t was the target English vocabulary. E_x, E_y, and E_z were non-target English vocabulary, and $x \neq y \neq z \neq t$. $C_p(E_t)$ was the corresponding Chinese meaning of E_t. $V(E_t)$ was the English pronunciation of E_t. $Len(E_t)$ is the length of E_t, that is, the number of letters used to spell a target English word.

The aim of these test models was to explore the vocabulary acquisition of EFL learners from multiple aspects. The purpose of the spelling test was to find out if the examinee could correctly spell the English vocabulary when the pronunciation was given. The reading test was to test whether the examinee could understand the meaning of the target English vocabulary or not. The listening test was to examine the examinee's listening ability and comprehension of the target English vocabulary.

Table 2: The test models of English vocabulary test

Test Model	Description
Spelling test	Given $C_p(E_t)$, $Len(E_t)$, and $V(E_t)$. Examinee is asked to key-in E_t.
Reading test	Given E_t and the choices of $C_p(E_t)$, C_q, C_r and C_s in random sequence. $C_p(E_t)$ was the target choice, and C_q, C_r and C_s were the wrong choices. Examinee was asked to choose the target choice.
Listening test	Given $V(E_t)$ and the choices of $C_p(E_t)$, C_q, C_r and C_s in random sequence. $C_p(E_t)$ was the best choice which matched the given $V(E_t)$. C_q, C_r and C_s were the wrong choices. Examinee was asked to choose the target choice.

The Statistical Package for the Social Science (SPSS) v.18 for Microsoft Windows was used to provide descriptive statistics and the distributions of scores. For research question 1, 2, and 3, descriptive statistics were utilized to provide mean scores and standard deviations of three tests for both samples.

2.4 Design

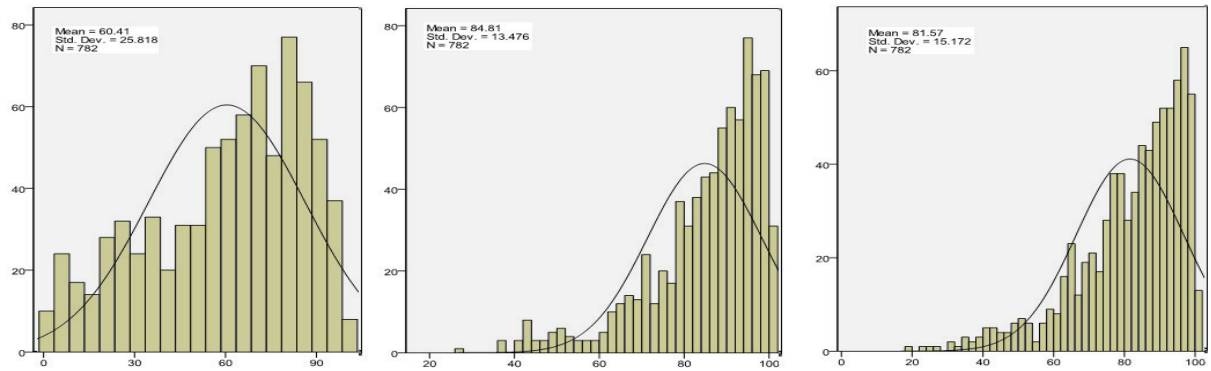
This study was a normative survey. The researcher was interested to investigate the current phenomena of English vocabulary acquisition among elite Taiwanese junior and senior high school students. *VQ* would be used to estimate the size of English vocabulary of the participants.

2.5 Procedure

The experiment was embedded in the National Spelling Competition in 2010. Three different types of English vocabulary tests were used to evaluate the participants. Every test had 100 questions and was worth 100 points. Before the tests, participants had already known how to use the software. Spelling test had to be finished in 20 minutes, while the reading and the listening test had to be finished in 10 minutes, respectively. After the test, the scores of English vocabulary tests would be collected by computers.

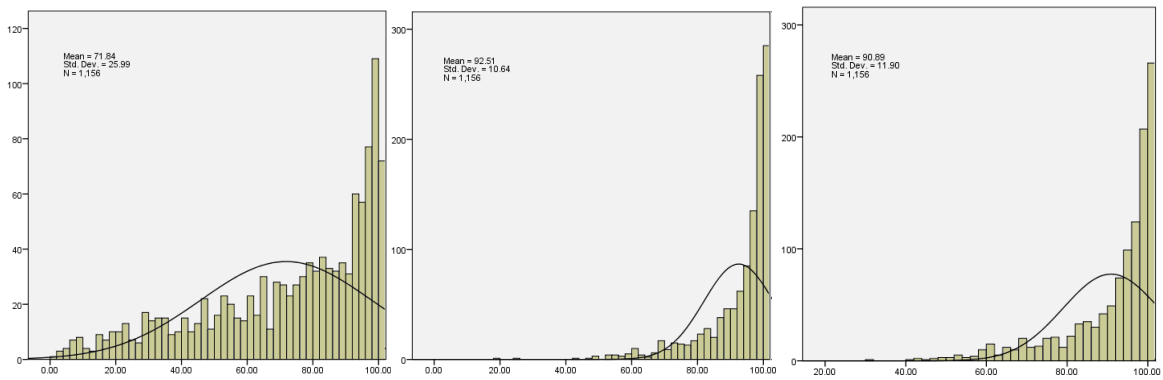
3. Results

Fig. 1 presented the histograms of scores of three English vocabulary tests for all senior high participants. Fig. 2 illustrated the test results of all junior high participants. The horizontal axis was the scores of the tests and the vertical axis was the number of participants. Distribution curves were also illustrated in Fig. 1 and 2. Results showed that the distribution of scores of spelling, reading and listening tests were all negative skewed.



(a) Distribution of Spelling. (b) Distribution of Reading. (c) Distribution of Listening.

Figure 1. Distribution of three aspects for senior high participants



(a) Distribution of Spelling. (b) Distribution of Reading. (c) Distribution of Listening.

Figure 2. Distribution of three aspects for junior high participants

Table 3 presented the means of the two groups, their standard deviations, and the scores of spelling, reading, and listening. It was found that the standard deviations of spelling were the greatest among the three tests, and that the means of spelling were the smallest among the three tests. From the viewpoint of the mean scores, the sequence of hard-level was spelling, listening, and reading. These outcomes appeared both in the group of junior high school students and senior high school students.

Table 3: Descriptive statistics for the all participants (*N*=1938)

Group	*n*	Spelling		Reading		Listening	
		M	*SD*	*M*	*SD*	*M*	*SD*
Grade7th~9th	1156	71.84	25.99	92.51	10.64	90.89	11.90
Grade10th~12th	782	60.41	25.82	84.81	13.48	81.57	15.17

Based on data in Table 3 and the method of estimating vocabulary sizes proposed by Ho and Huong [10], the estimated vocabulary sizes were calculated in Table 4 for both groups. Note that estimated vocabulary sizes of spelling, reading, and listening were all different. Senior high school students and junior high school students also differed in the areas of spelling, reading, and listening.

The difference of the means showed the growth of average vocabulary size between senior high students and junior high students. These results were for research questions 1, 2 and 3. The growth of vocabulary size of spelling between junior and senior high school students was 3591 words. The growth of vocabulary size of reading between two groups of the participants was 4328 words. The growth of vocabulary size of listening between the two groups was 4206 words. Thus, H_1 was not supported. H_2 and H_3 were supported.

Table 4: Estimated vocabulary sizes for the all participants (*N*=1938)

	n	Spelling			Reading			Listening		
		min	*max*	*M*	*min*	*max*	*M*	*min*	*max*	*M*
Grade7th~9th	1156	22	2200	1509	418	2200	1943	682	2200	1909
Grade10th~12th	782	2248	7000	5100	3496	7000	6271	3112	7000	6115
Difference of *M*				3591			4328			4206

Figure 3 presented the minimum, the maximum, and the mean of vocabulary sizes of three tests for each group of the participants. It was found that some participants performed better in reading and listening, but not in spelling. The range between the maximum and the minimum was large in every aspect of each group of the participants.

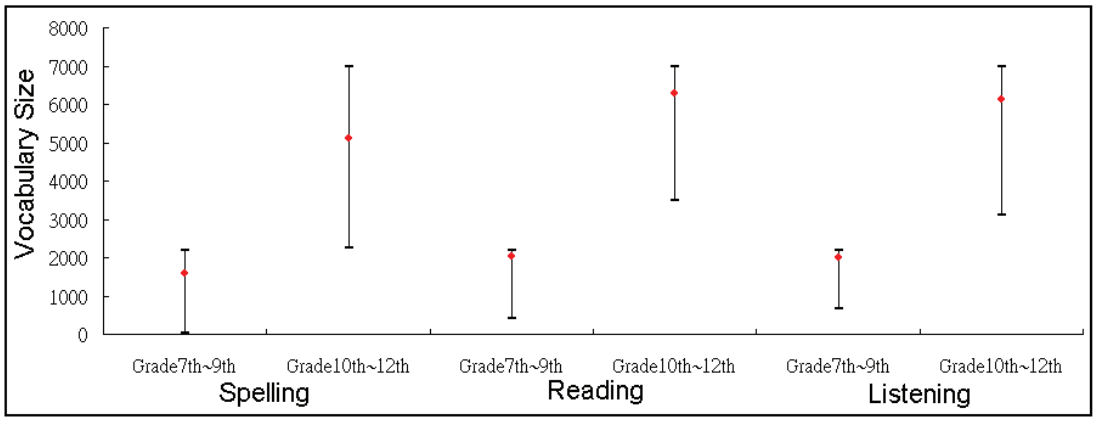


Figure 3. The min, max and mean of vocabulary sizes of three aspects

4. Discussion

Based on data in Table 4, the growth of vocabulary size of spelling was 3591 words. The growth of vocabulary size of reading was 4328 words. The growth of vocabulary size of listening was 4206 words. The growing rate of spelling was about 1197 words/year. The growing rate of reading was about 1443 words/year. The growing rate of listening was about 1402 words/year. They were far from the native speakers. The growths of EFL high school students' vocabulary sizes in multiple aspects were not found to be greater than 4800 words. Thus, H_1 was not supported. The growth of vocabulary size of spelling was found to be the least. H_2 was supported. The growth of vocabulary size of reading was the greatest. H_3 was supported.

Based on Fig. 1, 2, and 3, there were many junior and senior high students who had low vocabulary sizes. These students might have Matthew effects in reading [23]. Using *VQ* software tool, EFL teachers could find out if the students had low vocabulary sizes. After identifying the problems, necessary teaching and learning could be carried out to solve problems.

There were some limitations in this study. The number of participants was small, and only three linguistic aspects of English vocabulary were investigated. Since the lexicons in this study were set as 2200 words for junior high school students, and 7000 words for senior high school students, the estimated vocabulary sizes were within these boundaries. If any participants had the lexicon greater than 7000 words, this study would not be able to measure the actual the size of the lexicon.

This study could have both theoretical and practical contributions. Vocabulary quotient could become an effective indicator to estimate the actual vocabulary size of English learners. As the research results had showed that English learners might not grow equally in the areas of spelling, reading, and listening, English teachers might want to consider offering more instructions for students to strengthen their weakness in spelling.

Acknowledgements

This work was supported in part by the "Aim for the Top University Plan" from National Taiwan Normal University and the Ministry of Education, Taiwan, R.O.C.

References

[1] Attali, Y., & Burstein, J. (2006). Automated essay scoring with e-rater v. 2. *Journal of Technology, Learning, and Assessment, 4*(3), 1–30.

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

[3] 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.

[4] Chen, H. M., & Cheng, S. H. (2007). An investigation on the listening difficulties of technical college students in Taiwan. *Journal of China Institute of Technology, 36*, 335-361.

[5] Enright, M., & Quinlan, T. (2010). Complementing human judgment of essays written by English language learners with e-rater scoring. *Language Testing, 27*(3), 317–334.

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

[7] Field, J. (2003). Promoting perception: lexical segmentation in L2 listening. *ELT Jounral, 57*(4), 325-334.

[8] Hirsh, D., & Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language, 8*(2), 689-696.

[9] Ho, H. F., & Lin, P. Z. (2010). Chinese character Learning Review System with Item Response Theory Analysis, *Proc. of Research in Reading Chinese (RRC) Conference.*

[10] Ho, H. F., & Huong, C. (2011). A multiple aspects quantitative indicator for ability of English vocabulary:vocabulary quotient, *Journal of Educational Technology Development and Exchange, 4*(1), 15-22.

[11] 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.

[12] Ito, Y. (2001). Effect of reduced forms on ESL learners' input-intake process. *Second Language Studies, 20*(1), 99-124.

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

[14] Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316-323). Clevedon: Multilingual Matters.

[15] Meara, P. (2005). Designing vocabulary tests for English, Spanish and other languages. In C.S. Butler, M. G. -Gonzalez & S. D. -Suarez (Eds.), *The Dynamics of Language Use: functional and contrastive perspectives* (pp. 271-285). Amsterdam: Benjamins.

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

[17] Meara, P., & Jones, G. (1990). *The Eurocentres Vocabulary Size Test.* Zurich: Eurocentres.

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

[19] Nation, P. (1983). Testing and teaching vocabulary. *Guidelines, 5*, 12–25.

[20] Nation, P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.

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

[22] Rosa, M. (2002). Don't cha know? A survey of ESL teachers' perspectives on reduced forms instruction. *Second Language Studies, 21*(1), 49-78.

[23] Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 360-407.