International Journal of Assessment and Evaluation in Education
Vol 6/Dec 2016 (17-26) ISSN 2232-1926
Submission Date: 20th January 2016
Acceptance Date: 13th February 2016

Vocabulary Levels Test and Word Associates Test: Can they Measure Language Proficiency?

Mostafa Janebi Enayat
Department of English Language and Literature
Hakim Sabzevari University, Sabzevar, Iran
mostafa.enayat@gmail.com

Seyed Mohammad Reza Amirian
Department of English Language and Literature
Hakim Sabzevari University, Sabzevar, Iran
smr.amirian@gmail.com

Abstract

The present study investigates whether it may be possible to use Vocabulary Levels Test (VLT) and Word Associates Test (WAT) for measuring language proficiency and placement purposes. In so doing, 85 upper-intermediate and 50 lower-intermediate EFL learners took Vocabulary Levels Test (VLT), Word Associates Test (WAT), and Oxford Placement Test (OPT). Results of correlation and regression analyses indicated that (a) Word Associates Test could not predict performance on Oxford Placement Test neither for the lower-intermediate nor the upper-intermediate group, (b) Vocabulary Levels Test could significantly predict performance on Oxford Placement Test for both proficiency groups, and (c) while the 2,000 word level correlated with and predicted performance on Oxford Placement Test for the lower-intermediate group, the 5,000 word level predicted performance on proficiency test for the upper-intermediate group. The findings point to the possibility of using VLT for measuring language proficiency and placement purposes. The results have implications for test developers, teachers, and researchers to use VLT as a placement test since it's easy to administer and score.

Keywords *Vocabulary Levels Test (VLT), Word Associates Test (WAT), language proficiency, placement test*

INTRODUCTION

Research on second language and foreign language vocabulary development has been thriving for the last decades or so and many studies have been conducted in applied linguistics journals to target this issue (Read & Chapelle, 2001). This wave of research has brought into focus many issues ranging from studying vocabulary knowledge and its aspects as key elements in reading comprehension (e.g., Kaivanpanah & Zandi, 2009; Nassaji, 2006; Qian, 1998, 2002; Shiotsu & Weir, 2007), in writing (Baba, 2009; Engber, 1995; Laufer & Nation, 1995; Lee, 2003), speaking (Batty, 2007), and listening (Mehrpour & Rahimi, 2010).

One of the main readers' variables that determine the learners' performance in different types of tests is vocabulary knowledge. According to Shen (2008), vocabulary is an important element in language and vocabulary acquisition has gained much attention in the research field of ELT. As Schmitt (2008) pointed out, vocabulary knowledge is of high importance and "one thing that students, teachers, materials writers, and researchers can all agree upon is that learning vocabulary is an essential part of mastering a second language" (p. 329). Two newly identified dimensions of vocabulary are breadth and depth of vocabulary knowledge (Henrikson, 1999) the effect of which has been mostly studied in reading comprehension (e.g., Kaivanpanah & Zandi, 2009; Nassaji, 2004, 2006; Qian, 1998, 2002; Shen, 2008). Depth is associated with syntagmatic and paradigmatic relations (Schoonen & Verhallen, 2008) which involve superordinates, synonyms, and collocations. Breadth of vocabulary knowledge on the other hand, refers to the size or quantity of the words a learner has at his/her disposal to use both receptively and productively (Read, 2000).

We cannot consider breadth and depth of vocabulary knowledge as polar opposites (Akbarian, 2010). This was earlier confirmed by Qian (1999) when he found that size is as valuable as depth of word knowledge since they are interconnected. In addition, Schoonen and Verhallen (1998) came to the conclusion that breadth tests correlated strongly with the depth tests in the reading comprehension performance of the students (cf. Read, 2004, in Bogaards & Laufer, 2004).

Vocabulary breadth or size is one of the crucial aspects of vocabulary knowledge and "tests of vocabulary size have been shown to predict success in reading, writing, and general language proficiency as well as academic achievement" (Laufer & Goldstein, 2004, p. 401). For measuring this aspect of vocabulary knowledge, different tests have been used and suggested by scholars of the field (see e.g. Laufer et al. 2004; Laufer & Goldstein, 2004; Laufer & Nation, 1999; Nation, 1993; Read, 2007; Schmitt, 2008; Schmitt et al. 2001; Xing & Fulcher, 2007). One of the best types of tests for measuring vocabulary size is Vocabulary Levels Test.

Vocabulary Levels Test

The first attempt to design a vocabulary levels test goes back to the research conducted by Nation (1983) in which he developed a vocabulary size test at Victoria University of Wellington which assesses the number of word knowledge at 2000 word level, 3000 word level, 5000 word level, 10000 word level, and university word level. He pointed out that the main idea behind the Vocabulary Levels Test is that it is useful to view the vocabulary of English as consisting of a series of levels based on frequency of occurrence. After the seminal work of Nation (1983), many other studies tried to either validate or find the reliability of the test or even revise it (e.g. Ishii & Schmitt, 2009; Laufer et al. 2004; Laufer & Goldstein, 2004; Schmitt et al. 2001; Xing & Fulcher, 2007).

Two revised and expanded versions of Vocabulary Levels Test were later developed by Schmitt et al. (2001). They used Rasch analysis, item analysis, and factor analysis to validate those versions. The results of their study indicated that the newly developed versions of VLT were highly predictive of students' breadth of vocabulary knowledge. In later attempts, Laufer and Goldstein (2004) and Laufer et al. (2004) tried to develop computerized versions of vocabulary tests which could test breadth and depth of vocabulary knowledge simultaneously. A more recent study by Xing and Fulcher (2007) also analyzed the reliability of two versions of Vocabulary Levels Test i.e. A & B which found more reliability for version B at 5000 word level. The following figure shows an example of the items in VLT (taken from Akbarian, 2010, p. 395).

Participants must choose the right word that goes with each meaning. They must write the number of that word next to its meaning. Here is an example.

1 business
2 clock ----- part of a house
3 horse ----- animal with four legs
4 pencil ----- something used for writing
5 shoe
6 wall

Participants answer it in the following way.
1 business
2 clock ---6--- part of a house
3 horse ---3--- animal with four legs
4 pencil ---4--- something used for writing
5 shoe
6 wall

Figure 1 A Sample of Vocabulary Levels Test

Vocabulary knowledge can be conceptualized not only as the number of lexical items known (vocabulary size), but also as how well these items are mastered (vocabulary depth or quality) which is highly important for productive use of language (Schmitt et al. 2011). This has put a great attention on the assessment of this aspect of vocabulary knowledge which in Read's (2007) opinion, less progress has been obtained in measuring quality (or depth) of word knowledge in comparison with measuring vocabulary size or breadth.

Word Associates Test

Primarily based on the main relations of syntagmatic, paradigmatic, and analytic relationships between two lexical units (Batty, 2007), this test was developed by Read (1993) at Victoria University of Wellington to measure the receptive aspect of depth of vocabulary knowledge. It was later called Word Associates Test (WAT) and Depth of Vocabulary Knowledge Test (DVKT). The building block of this test was the concept of word association and lexical network. It was developed to measure three vocabulary elements: synonymy, polysemy, and collocation (Qian, 2002). The test has been revised and validated since first developed (e.g. Ishii & Schmitt, 2009; Laufer et al. 2004; Qian, 1999; Schmitt et al. 2011).

The test is composed of 40 items each of which consists of one stimulus word, which is an adjective, and two boxes each containing four words (Figure 2). The left box shows the synonymous words and the right one shows the collocations of the stimulus word. Among the four words in the left box, one to three words can be synonymous to one aspect of, or the whole meaning of the stimulus word. Among the four words in the right box, there can be one to three words that collocate with the stimulus word. In other words, three situations are possible (Qian, 2002): (a) the left and right boxes both contain two correct answers; (b) the left box contains one correct answer, and the right box contains three correct answers; and (c) the left box contains three correct answers, and the right box contains only one correct answer. As Schmitt et al. (2011) explained, the reason for varying the distribution of responses was to prevent guessing on the part of the test takers.

A **sudden**

beautiful quick surprising thirsty	change doctor noise school

Figure 2 Sample of Word Associate Test (WAT) Item

VLT and WAT as measures of language proficiency

Vocabulary levels tests have been regarded as proper means of placement and admission purposes in ELT curriculum (Laufer & Nation, 1999). They have "proved to be useful in helping teachers to determine the kind of attention they should be giving to vocabulary for particular groups of learners" (p. 34).

Devising a word association test (WAT) as a means of assessing proficiency in a foreign language has always had something of an inherent appeal to it. If one could in fact be developed, it would (1) be relatively quick and easy both to administer and to score, (2) be a nice complement to other methods of assessing learner performance by giving researchers a means of assessing proficiency that is not based specifically on 'correct' language performance per se, and (3) tend to suggest that there may be something of a connection between psycholinguistic knowledge and more general proficiency in a foreign language (Wolter, 2002). In respect to this last point, the underlying argument is that we would expect learners of higher proficiency to have more highly developed semantic networks in the L2 mental lexicon. And if we can view responses on a WAT as evidence of connections between words in the mental lexicon, then it may be possible to make claims about the sophistication of a learner's semantic network and the implications this has for the learner's general proficiency in the foreign language.

While it is still certainly too early to suggest that a word association test can function as an indication of proficiency, there does not really seem to be enough evidence to the contrary. It is for this reason that this study was undertaken, in the hope that a more carefully constructed test may lead to better results and that with subsequent consideration and refinement, a word association test may one day be developed as a means of assessing proficiency. The present study aimed at examining the possibility of using VLT and WAT for assessing language proficiency. In light of the significance of these tests as measures of vocabulary size and depth and the importance of having a proficiency test which is easy to administer and score, this study investigates the following research questions:

1. Can Vocabulary Levels Test measure language proficiency and be used for placement purposes?
2. Can Word Associates Test measure language proficiency and be used for placement purposes?

METHOD

Participants

135 Iranian undergraduate students majoring in English Translation Studies and English Literature participated in this study. They were classified into 50 lower-intermediate and 85 upper-intermediate level based on the results obtained from the Oxford Quick Placement Test (2004). The selected participants were males and females ranging from 18 to 24 years of age.

Instruments

Oxford Quick Placement Test (2004)

This test was administered to determine the language proficiency level of the participants and classify them into lower-intermediate and upper-intermediate levels. Also, it was used as a criterion to compare the scores on VLT and WAT to find out if they can be used as measures of language proficiency and/or placement purposes. This test is consisted of 60 items which was developed by Oxford University Press and University of Cambridge Local Examinations Syndicate. The test has been validated in 20 countries by more than 6,000 students and its reliability has reached 0.90 (Geranpayeh, 2003).

Word Associates Test (WAT)

WAT is based on the main relations of syntagmatic, paradigmatic, and analytic relationships between two lexical units which was developed by Read (1993) to measure the receptive aspect of depth of vocabulary knowledge. The reliability obtained for this measure was 0.91 with a sample of 74 Korean and Chinese speakers (Qian, 1998). The split half reliability of the test was found to be 0.89 in a study done by Qian (2002) and Nassaji (2006) also reported its split half reliability as 0.89.

Vocabulary Levels Test (VLT)

VLT has been designed based on the frequency levels of words or frequency word lists. It has been used to make inferences about the test takers' vocabulary size (Read, 2000) by measuring single meanings of content words at four frequency levels (2,000, 3,000, 5,000, and 10,000 words). Nation (1983) designed the first vocabulary level test and many other studies tried to either validate, find its reliability, or even revise it as mentioned before. Two revised and expanded versions of Vocabulary Levels Test were later developed by Schmitt et al. (2001). They found Cronbach alpha reliability of 0.92 for all the word levels of this measure. The present study used the second version of Vocabulary Levels Test developed and revised by Schmitt et al. (2001).

Data Collection

To collect the data necessary for this study, several steps were taken. First, the Oxford Quick Placement Test was administered to determine the students' level of proficiency and homogenize the groups. Second, Word Associates Test (WAT) and Vocabulary Levels Test (VLT) were taken. As for VLT, the participants had to choose the right word that went with each meaning. The time allotted was 30 minutes for each test. As for WAT, the participating students were instructed to read each of the target words and then circle the four words closely related to the target word even if they were not sure whether their answer was correct or not (Read, 1993). The time allotted to WAT was 30 minutes too. Finally, the results of VLT and WAT were compared with Oxford Placement Test (OPT).

RESULTS

Lower-intermediate level

The lower-intermediate group of this study consisted of 50 students. To find out if Word Associates Test and Vocabulary Levels Test could correlate with the scores on Oxford Quick Placement Test, Pearson correlation coefficient and multiple regression analyses were run. Table 1 shows the descriptive statistics of students' scores on each test. The results of Correlation analysis is then provided in Table 2.

Table 1 Descriptive statistics for lower-intermediate group

Test/Level	MPS	Mean	Std. deviation	*N*
Oxford	60	30.72	3.67	50
WAT	100	60.82	10.14	50
VLT	120	69.80	12.78	50
2,000	30	27.82	2.76	50
3,000	30	22.78	4.39	50
5,000	30	14.20	5.77	50
10,000	30	4.96	5.36	50

MPS = Maximum possible score.

Table 2 Pearson correlation for WAT and VLT levels on Oxford Quick Placement Test

Independent	Info.	Placement test
WAT	Pearson correlation	-.040
	Sig (2-tailed)	.784
VLT Total	Pearson correlation	.398**
	Sig (2-tailed)	.004
VLT (2,000)	Pearson correlation	.607*
	Sig (2-tailed)	.000
VLT (3,000)	Pearson correlation	.323*
	Sig (2-tailed)	.022
VLT (5,000)	Pearson correlation	.206
	Sig (2-tailed)	.152
VLT (10,000)	Pearson correlation	.145
	Sig (2-tailed)	.314

**: Correlation is significant at the 0.01 level (2-tailed)
*: Correlation is significant at the 0.05 level (2-tailed)

As the results of Pearson correlation indicates, the scores of the participants on Word Associates Test (WAT) did not have any significant correlation with the students' performance on Oxford Placement Test, $r = -.04$, $p > .05$. However, the participants' performance on VLT was found to have a significant correlation with their score on Oxford Placement Test, $r = .39$, $p < .01$. The scores of the participants at 2,000 word level had a significant correlation with the students' performance on OPT ($r = .60$, $p < .001$). Also, the scores of the students at 3,000 word level correlated significantly with OPT ($r = .32$, $p < .05$). In contrast, the 5,000 and 10,000 word levels as the more complex parts of the test, didn't have a significant correlation with OPT, $r = .20$, $p > .05$, $r = .14$, $p > .05$, respectively. However, to have a better analysis of the relationship between students' scores on WAT and different levels of VLT and their performance on OPT, multiple regression was run to find how the students' scores on WAT and each level of VLT contribute to the prediction of their performance on OPT.

Table 3 Multiple Regression analysis for WAT and VLT levels on OPT

Dependent	Model		Method	F	t	R^2	Sig	ß
OPT	1	(Constant)	Enter	.076	9.810	.002[a]	.000	
		WAT			- .275		.784	- .040
	1	(Constant)	Enter	9.008	8.417	.158[b]	.000	
		VLT Total			3.001		.004**	.398
	1	(Constant)	Stepwise	27.966	1.949	.368[c]	.057	
		Level 2,000			5.288		.000**	.607
	1	(Constant)	Enter	6.829	1.895	.378[d]	.065	
		Level 2,000			4.341		.000**	.617
		Level 3,000			- .409		.684	- .072
		Level 5,000			.584		.562	.091
		Level 10,000			.450		.655	.055

**Significant at *p* = .01.
a. Predictors: (Constant), WAT
b. Predictors: (Constant), VLT
c. Predictors: (Constant), level 2,000
d. Predictors: (Constant), level 2,000, level 3,000, level 5,000, level 10,000

The results of multiple regression indicated that the test takers' performance on WAT could not predict the scores on OPT (R^2 = .002, F = .076, p > .05). In contrast, the scores on VLT could significantly explain the variance on OPT performance. VLT could explain about 16% of variance in OPT (R^2 = .158, F = 9.008, p < .01). As for the levels of VLT, when regression was run using stepwise method, just the 2,000 level was entered in the model and other levels were excluded. The 2,000 level could explain nearly 37% of the variance in OPT test (R^2 = .368, F = 27.966, p < .001). When other levels were included in the model using enter method, only 1% was added to the predictive power of VLT which was not significant (R^2 = .378, F = 6.829, p > .05). However, in this model, the 2,000 word level significantly predicted performance on OPT (ß = .617, p < .001). The 3,000, 5,000 and 10,000 word levels could not predict the performance on OPT (p > .05).

Upper-intermediate level

The upper-intermediate group of this study consisted of 85 EFL students. The same statistical procedures were conducted to answer the research questions of this study. The general profile of students' achievements is provided in the table of descriptive statistics.

Table 4 Descriptive statistics for upper-intermediate group

Test/Level	MPS	Mean	Std. deviation	*N*
Oxford	60	42.21	3.32	85
WAT	100	69.65	9.52	85
VLT	120	88.00	12.36	85
2,000	30	29.52	.78	85
3,000	30	27.67	2.53	85
5,000	30	22.42	5.16	85
10,000	30	8.36	6.07	85

MPS = Maximum possible score.

Table 5 Pearson correlation for WAT and VLT levels on Oxford Quick Placement Test

Independent	Info.	Placement test
WAT	Pearson correlation	-.029
	Sig (2-tailed)	.793
VLT (Total)	Pearson correlation	.457**
	Sig (2-tailed)	.000
VLT (2,000)	Pearson correlation	.163
	Sig (2-tailed)	.137
VLT (3,000)	Pearson correlation	.308**
	Sig (2-tailed)	.004
VLT (5,000)	Pearson correlation	.457**
	Sig (2-tailed)	.000
VLT (10,000)	Pearson correlation	.390**
	Sig (2-tailed)	.000

**: Correlation is significant at the 0.01 level (2-tailed)
*: Correlation is significant at the 0.05 level (2-tailed)

As the results of Pearson correlation in Table 5 indicates, the scores of the participants in Word Associates Test (WAT) did not have any significant correlation with the students' performance on Oxford Placement Test, r = .79, p > .05. However, the participants' performance on VLT was found to have a significant correlation with their score on Oxford Placement Test (OPT), r = .45 p < .001. In contrast, the scores of the students at 2,000 word level could not correlate significantly with OPT, r = .16, p > .05. The students' scores at 3,000 level had a significant correlation with scores on OPT (r = .30, p < .01). In contrast with lower-intermediate level analysis, the correlation between 5,000 and 10,000 levels and OPT was significant, r = .45, p < .001, r = .39, p < .001 respectively. However, to have a better analysis, multiple regression was run to find how the students' scores at WAT and each level of VLT contribute to the prediction of their performance on OPT.

Table 6 Multiple Regression analysis for WAT and VLT levels on OPT

Dependent	Model		Method	F	t	R^2	Sig	ß
OPT	1	(Constant)	Enter	.069	15.941	.001[a]	.000	
		WAT			- .263		.793	- .029
	1	(Constant)	Enter	21.874	13.467	.209[b]	.000	
		VLT (Total)			4.677		.000**	.457
	1	(Constant)	Stepwise	21.887	24.664	.209[c]	.000	
		Level 5,000			4.678		.000**	.457
	1	(Constant)	Enter	5.819	2.565	.225[d]	.012	
		Level 2,000			.316		.753	.035
		Level 3,000			- .106		.915	- .015
		Level 5,000			2.411		.018*	.351
		Level 10,000			1.261		.211	.164

*Significant at p = .05
**Significant at p = .001.
a. Predictors: (Constant), WAT
b. Predictors: (Constant), VLT
c. Predictors: (Constant), level 5,000
d. Predictors: (Constant), level 2,000, level 3,000, level 5,000, level 10,000

The results of multiple regression indicated that the test takers' performance on WAT could not predict the scores on OPT (R^2 = .001, F = .069, p > .05). In contrast, the scores on VLT as a whole could explain about 21% of the variance in Oxford Placement Test (R^2 = .209, F = 21.874, p < .001). As for the predictive ability of each word level, using the stepwise method, just the 5,000 word level was entered in the model which could explain about 21% of the variance in OPT (R^2 = .209, F = 21.887, p < .001). When all levels were entered in the model, the model could explain about 22% of the variance (R^2 = .225, F = 5.819, p > .05) which means adding other levels could not add to the predictive power of VLT and level 5,000 had the most influential effect. The 2,000, 3,000 and 10,000 word levels could not predict the students' performance on OPT (p > .05).

DISCUSSION

The present study made an attempt to investigate the power of Vocabulary Levels Test (VLT) and Word Associates Test (WAT) in predicting the performance on Oxford Placement Test (OPT). The purpose was to find out if we can use VLT and WAT as placement tests or measures of language proficiency. It has been argued that these tests could measure language proficiency (Schmitt et al., 2001; Schmitt et al., 2011; Wolter, 2002).

The results of this study indicated that WAT could not correlate with and explain the variance in OPT. In other words, WAT cannot be used as a placement test as the probability has been mentioned before. This was true for both lower-intermediate and upper-intermediate students. In Contrast, the findings showed that VLT could significantly predict the performance on OPT. This was true for both levels except for some minor differences. The overall performance on VLT could significantly correlate with OPT and predict the students' performance on this test. This was a little more significant for the upper-intermediate level. In addition, while the 2,000 word level could highly correlate with and predict performance on OPT for lower-intermediate level; this was not true for the upper-intermediate level. As for the upper-intermediate level, the 5,000 level could predict performance on OPT. This difference indicated that the upper-intermediate students had a more breadth of vocabulary knowledge and VLT could differentiate between the lower-intermediate and upper-intermediate students. However, this was not true for the WAT.

The findings of this study confirmed the results of Wolter (2002) which indicated that word association in a foreign language is not linked to proficiency. He used C-test as a measure of language proficiency to be correlated with WAT. However, the researcher claimed that there might be correlations between WAT and C-test. The findings of this study provided evidence for the lack of any kind of correlation between WAT and OPT as a measure of language proficiency and placement test.

The difference between the two groups for the results of correlation between VLT and OPT indicated that VLT could be used as a predictor of language proficiency. In the lower-intermediate group, the 2,000 word level was entered as the predictor variable while for the upper-intermediate group, the 5,000 word level had the highest predictive power. The difference between the lower- and upper-intermediate groups is consistent with two experiments conducted by Coady et al. (1993). The authors showed that more proficient students had more vocabulary size in comparison with the low proficient ones. In this study, the 5,000 word level acted as the predictor variable for the upper-intermediate group while for the lower group, the 2,000 word level was entered into the model. This can be justified by the discriminatory power of VLT for determining the proficiency level of the students. In other words, in line with the argument of Laufer (1999), VLT is a valid measure of vocabulary growth and can be used for placement purposes. The Vocabulary Levels Test is designed to give an estimate of vocabulary size for second language (L2) learners of general or academic English. The rationale for the test stems from research which has shown that vocabulary size is directly related to the ability to use English in various ways (Schmitt et al., 2001).

This study demonstrated that Word Associates Test (WAT) could not be used as a measure of language proficiency or placement purposes. In contrast, Vocabulary Levels Test (VLT) could correlate well and predict the performance on Oxford Placement Test. The findings provided evidence for the possibility of using VLT as a measure of proficiency to be used especially for placement purposes. Vocabulary size test has been considered suitable for placement and admission purposes and is thus able to tap vocabulary knowledge of the learners (Beglar & Hunt, 1999; Laufer & Nation, 1999; Zhang & Anual, 2008). It has been found to be cost-effective in terms of the class time. Therefore, Vocabulary Levels Test can be used as placement tests in ELT curriculum to find the extent of the students' vocabulary size prior to the course so that more effective decisions can be made. In addition, it can be used as a placement test for research purposes as it's easy to administer and score. Akbarian (2010) used the 2,000 word level to find the proficiency level of the students and classify them into low and high proficient groups.

The Vocabulary Levels Test should be used at all time periods of the curriculum so that the teachers know at what stage of vocabulary development their learners are since this test assesses the amount of the students' knowledge at different word levels. In other words, different levels could be used to assess the proficiency as well as the vocabulary knowledge of the students at different levels of language proficiency.

REFERENCES

Akbarian, I. (2010). The relationship between vocabulary size and depth for ESP/EAP learners. *System, 38*(3), 391-401.

Baba, K. (2009). Aspects of lexical proficiency in writing summaries in a foreign language. *Journal of Second Language Writing, 18*(3), 191-208.

Batty, A.O. (2007). Vocabulary depth in written and oral assessment. In K. Brandford-Watts (Ed.), JALT *2006 Conference Proceedings* (pp. 1100-1108). Tokyo: JALT.

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

Bogaards, P., & Laufer, B. (2004). *Vocabulary in a Second Language*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Coady, J., Magoto, J., Hubbard, P., Graney, J., & Mokhtari, K. (1993). High frequency vocabulary and reading proficiency in ESL readers. In Huckin, T., Hynes, M., & Coady, J. (Eds.), *Second Language Reading and Vocabulary Acquisition* (pp. 217-226). Ablex, Norwood, NJ.

Engber, C.A. (1995). The relationship of lexical proficiency to the quality of ESL compositions. *Journal of Second Language Writing, 4*(2), 139-155.

Geranpayeh, A. (2003). A quick review of the English Quick Placement Test. *UCLES Research Notes, 12*, 8-10.

Henriksen, B. (1999). Three Dimensions of Vocabulary Development. *Studies in Second Language Acquisition, 21*(2), 303-317.

Ishii, T., & Schmitt, N. (2009). Developing an integrated diagnostic test of vocabulary size and depth. *RELC Journal, 40*(1), 5-22.

Kaivanpanah, S., & Zandi, H. (2009). The role of depth of vocabulary knowledge in reading comprehension in EFL context. *Journal of Applied Sciences, 9*(4), 698-706.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: size, strength, and computer adaptiveness. *Language Learning, 54*(3), 399-436.

Laufer, B., & Nation, P. (1995). Vocabulary size and use: lexical richness in L2 writing production. *Applied Linguistics, 16*(3), 307-322.

Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing, 16*(1), 33-51.

Laufer, B., Elder, C., Hill, K., & Congdon, P. (2004). Size and strength: do we need both to measure vocabulary knowledge? *Language Testing, 21*(2), 202-226.

Lee, S.H. (2003). ESL learners' vocabulary use in writing and the effects of explicit vocabulary instruction. *System 31*(4), 537-561.

Mehrpour, S., & Rahimi, M. (2010). The impact of general and specific vocabulary knowledge on reading and listening comprehension: A case of Iranian EFL learners. *System, 38*(2), 292-300.

Nassaji, H. (2004). The Relationship between Depth of Vocabulary Knowledge and L2 Learners' Lexical Inferencing Strategy Use and Success. *The Canadian Modern Language Review, 61*(1), 107-134.

Nassaji, H. (2006). The Relationship between Depth of Vocabulary Knowledge and L2 Learners' Lexical Inferencing Strategy Use and Success. *The Modern Language Journal, 90* (3), 387-401.

Nation, P. (1983). Testing and teaching vocabulary. *Guidelines, 5*(1), 12-25.

Nation, P. (1993). Using dictionaries to estimate vocabulary size: essential, but rarely followed, procedures. *Language Testing, 10*(1), 27-40.

Qian, D.D. (1998). Depth of vocabulary knowledge: assessing its role in adult's reading comprehension in English as a second language. *Unpublished PhD Dissertation*.

Qian, D.D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *The Canadian Modern Language Review,* 56(2), 282-307.

Qian, D.D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: an assessment perspective. *Language Learning*, *52*(3), 513-536.
Read, J. & Chapelle, C.A. (2001). A framework for second language vocabulary assessment. *Language Testing, 18*(1), 1-32.
Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing*, *10*(3), 355-371.
Read, J. (2000). *Assessing Vocabulary*. Cambridge: Cambridge University Press.
Read, J. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined? In Bogaards, P., & Laufer, B. (Eds.), *Vocabulary in a Second Language* (pp. 209-227). Amsterdam/Philadelphia: John Benjamin Publishing Company.
Read, J. (2007). Second language vocabulary assessment: Current practices and new directions. *International Journal of English Studies, 7(2)*, 105-125.
Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, *12*(3), 329-363.
Schmitt, N., Ching Ng, J.W., & Garras, J. (2011). The Word Associates Format: Validation evidence. *Language Testing, 28*(1), 105-126.
Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing, 18*(1), 55-88.
Schoonen, R. & Verhallen, M. (1998). Aspects of vocabulary knowledge and reading performance. *Paper presented at the Annual Meeting of the American Educational Research Association, San Diego.*
Schoonen, R. & Verhallen, M. (2008). The assessment of deep word knowledge in young first and second language learners. *Language Testing, 25* (2), 211-236.
Shen, Z. (2008). The Roles of Depth and Breadth of Vocabulary Knowledge in EFL Reading Performance. *Asian Social Science, 4* (12), 135-137.
Shiotsu, T., & Weir, C.J. (2007). The relative significance of syntactic knowledge and vocabulary breadth in the prediction of reading comprehension test performance. *Language Testing, 24*(1), 99-128.
University of Cambridge: *ESOL Examinations (2004). Quick Placement Test* (Version 1). Oxford University Press.
Wolter, B. (2002). Assessing proficiency through word associations: is there still hope? *System, 30*(3), 315-329.
Xing, P., & Fulcher, G. (2007). Reliability assessment for two versions of Vocabulary Levels Tests. *System, 35*(2), 182-191.
Zhang, L.J., & Anual, S.B. (2008). The Role of Vocabulary in Reading Comprehension: The Case of Secondary School Students Learning English in Singapore. *RELC Journal, 39*(1), 51-76.