Journal of Applied Linguistics and Language Research
Volume 2, Issue 6, 2015, pp. 154-164
Available online at www.jallr.ir
ISSN: 2376-760X

The Contribution of Different Aspects of Lexical Knowledge to Students' L2 Reading Performance

Mostafa Zamanian
Department of Foreign Languages, Human Faculty of Science, Shiraz Branch, Islamic Azad University, Shiraz, Iran

Rezvan Ghalebi *
Department of Foreign Languages, Human Faculty of Science, Khormuj Branch, Islamic Azad University, Khormuj, Iran

Abstract
The purpose of this study was to investigate the contribution of different aspects of lexical knowledge to students' L2 reading comprehension. To answer this question, a test of vocabulary, adapted from Zareva (2005), was administered to 157 Iranian B.A. university students majoring in English literature. The study also used the students' GPA in their reading courses. It was shown through regression that among different aspects of vocabulary knowledge studied in this study only vocabulary size and self-perceived vocabulary knowledge contributed to students' reading comprehension. It was also revealed that breadth, and receptive productive dimensions of vocabulary knowledge, accounted for high amount of the variability of students' reading scores. Moreover, considering that part of data collected from freshman students, it was demonstrated that Depth of vocabulary, had the highest contribution to reading after Breadth.
Keywords: lexical knowledge, reading comprehension, vocabulary size, depth of vocabulary, breadth of vocabulary

INTRODUCTION

Vocabulary knowledge is of great significance in language competence. In first language (L1) research, it has long been recognized that vocabulary knowledge makes an important contribution to reading comprehension (Anderson & Freebody, 1981, 1983; Mezynski, 1983; Stratton & Nacke, 1974; Tuinman & Brady, 1974). However a few numbers of researches have investigated the role of different components of vocabulary knowledge on reading performance in second language acquisition. According to Qian (2002) most of the studies performed in this area focus on breadth of vocabulary knowledge and little recognition is accorded to the roles other aspects of vocabulary knowledge play. The present study aims to explore what aspects of lexical knowledge

* **Correspondence:** Rezvan Ghalebi, Email: ms.ghalebie@yahoo.com

contribute more to reading performance of the students, employing the three-dimensional framework proposed by Henriksen (1999).

LITERATURE REVIEW

In this part studies that explore the role of vocabulary knowledge in reading comprehension are reviewed. Hawkins (1995) in a study examined the relation of grade, word recognition, listening comprehension, working memory and type of text (narrative and expository) on reading comprehension. Regression analyses revealed that a relation between word recognition and reading comprehension and a relation between listening comprehension and reading comprehension did exist for both text types. Results indicated that, for expository text, a relation did exist between listening comprehension and word recognition and reading comprehension moderated by grade.

Ramirez (2001) in a study examined the contributions of (a) English decoding skills, (b) English reading fluency, (c) English vocabulary knowledge, and (d) general Spanish reading proficiency on the English reading comprehension of fifth-grade Spanish-speaking English language learners.57 SELL students were examined in order to understand what factors contribute to Spanish-speaking English language learners' reading comprehension success. Four competing theoretical models of reading comprehension were explored through standard and stepwise multiple regression analysis techniques to develop a theory of reading comprehension specific to this population. The gained results indicated that English vocabulary was recognized as a significant predictor for reading comprehension performance of fifth-grade SELL students, but not the strongest predictor.

Robinson (2005) performed a non-experimental, multivariate correlation study that examined extant data from 51 third grade students to determine whether a relationship exists between vocabulary knowledge, oral reading fluency, and reading comprehension. Based on the results he concluded that there is a moderate relationship between vocabulary knowledge, oral reading fluency, and reading comprehension.

Lam (2005) in a research examined the role of grammatical skills, as well as vocabulary knowledge, and phonological processing skills in reading comprehension. Results regarding reading comprehension indicated that grammatical sensitivity was a relatively strong predictor for EL1 readers but not for ESL students. Phonological awareness was the most significant predictor for reading comprehension in ESL students, whereas vocabulary knowledge was most significant in predicting reading comprehension of EL1 students. For ESL students, grammatical skills and vocabulary knowledge may be subsumed under a broad construct of language proficiency. Hence, impact from grammatical skills upon reading comprehension in ESL readers, may be indirect.

Cromley (2005) in an examination validates and refines a new model of reading comprehension in order to identify the components that have the largest effect on comprehension. The results indicated that all predictors made a significant contribution

to comprehension, with vocabulary, background knowledge, and strategies having significant indirect effects. Vocabulary and background knowledge made the greatest total contribution to comprehension.

Cromley and Azevedo (2007) in an experiment examined a model of the variables that make the largest contributions to comprehension. The results showed that vocabulary and background knowledge made the largest contributions to comprehension, followed by inference, word reading, and strategies.

Shiotsu and Weir (2007) in a study examines the relative contribution of knowledge of syntax and knowledge of vocabulary to L2 reading in two pilot studies in different contexts– a heterogeneous population studying at the tertiary level in the UK and a homogenous undergraduate group in Japan – followed by a larger main study, again involving a homogeneous Japanese undergraduate population. In contrast with previous findings in the literature, all three studies offer support for the relative superiority of syntactic knowledge over vocabulary knowledge in predicting performance on a text reading comprehension test.

As it was claimed by Qian (2002) most of the studies performed investigating the relationship between word knowledge and reading comprehension, consider word knowledge as vocabulary size and do not take multidimensional nature of vocabulary knowledge into consideration. Among these studies a few number of them pay attention to this gap in literature and focused on different aspects of vocabulary knowledge and its relationship to reading comprehension.

Qian (2002) conducted a study in the context of Test of English as a Foreign Language (TOEFL) 2000 research to conceptually validate the roles of breadth and depth of vocabulary knowledge in reading comprehension in academic settings and to empirically evaluate a test measuring three elements of the depth dimension of vocabulary knowledge, namely, synonymy, polysemy, and collocation. The study found that the dimension of vocabulary depth is as important as that of vocabulary size in predicting performance on academic reading and that scores on the three vocabulary measures tested are similarly useful in predicting performance on the reading comprehension measure used as the criterion. The study confirms the importance of the vocabulary factor in reading assessment.

Huang (2006) in an experiment explored the relationship between vocabulary size (i.e., breadth of knowledge), depth of vocabulary knowledge, and reading comprehension of Chinese-speaking ESL (English as a second language) university students in Canada. The results demonstrate that (1) test scores on vocabulary size, depth of vocabulary knowledge, and reading comprehension are positively correlated, (2) vocabulary size is a stronger predictor of reading comprehension than depth of vocabulary knowledge, and (3) breadth and depth of vocabulary knowledge are closely interrelated and mutually facilitative. The findings suggest the importance of vocabulary size in reading comprehension for the population tested.

Ouellette (2006) in a study distinguished between vocabulary breadth and depth of vocabulary knowledge to better explain the role of oral vocabulary in various reading skills. Concurrent analyses of the data revealed that each distinct reading skill was related to the vocabulary measures in a unique manner. Receptive vocabulary breadth was the only oral vocabulary variable that predicted decoding performance after controlling for age and nonverbal intelligence. In contrast, expressive vocabulary breadth predicted visual word recognition, whereas depth of vocabulary knowledge predicted reading comprehension.

Reviewing the studies performed on the role of vocabulary knowledge in reading comprehension it was discovered that there have been few articles exploring contribution of different aspects of lexical knowledge to reading performance of the students. Therefore the present study is going to explore the contribution of three dimensions of lexical knowledge, i.e. breadth, depth, and receptive-productive dimensions, to reading performance of Shiraz University students.

Theoretical Framework

In this study, knowledge of vocabulary is considered based on the three dimensional framework proposed by Henriksen (1999), according to her lexical knowledge has these dimensions: (a) partial to precise knowledge, (b) depth of knowledge, and (c) receptive to productive use ability. Among different popular vocabulary tests a modified version of Vocabulary Knowledge Scale was adapted from Zareva (2005).

METHOD

Participants

The participants of the study included a total of 106 Iranian B.A. university students (79 female and 27 male) majoring in English literature and all of them in their 20s. They were native speakers of Persian and studied English as a foreign language. They were selected from all educational years. They did not have any special courses on vocabulary but, they passed at least two reading comprehension course in the university. All of the participants agreed willingly to complete the test. They were motivated as some reward for the first five students who gained the best result were considered. At first, 157 students took part in the study, but after the test was administered and the data were collected, it was found that 35 questionnaires were incomplete and they were deleted from the study. Also, 16 students were left out from the study as there were not enough data on their reading comprehension and proficiency courses.

Instruments

The necessary data were collected through a vocabulary test that was adopted from Zareva (2005). It contains 73 target words that are taken from Oxford Student's Dictionary of Current English (Hornby, 1978) by means of a spaced sampling procedure, i.e., selecting words at a specific interval from a randomly determined starting point in the dictionary.

Based on what Zareva mentioned in his or her study the words are a reliable representative of the 23,996-word content of the dictionary in terms of word frequency, lexical categories, and morphological word types. It was also mentioned that the sample is not biased towards any frequency band, reflecting natural language in terms of lexical class distribution, and is a good representation of several frequency bands.

It contains 13 target words (18%) from a frequency band with SFI between 20.0 and 29.9, 23 items (31%) from the frequency band with SFI between 30.0 and 39.9, 19 items (27%) with SFI from 40.0 to 49.9, and 18 items (24%) with SFI from 50.0 to 70.0+. All lexical categories were represented in the sample, i.e., nouns (n = 41), verbs (n = 16), adjectives (n = 13), and adverbs (n = 3). (See appendix A)

Each TW is accompanied by a modified version of Dale's (1965) and Paribakht and Wesche's (1993) word familiarity scale, adopted from Zareva (2005), in which the four steps identifying the four degrees of familiarity were preserved and the fifth step was modified and intended to collect WA data. The steps are as follows:

(1) ''I have never seen this word before''

(2) ''I have seen this word before, but I don't remember what it means''

(3) ''I think this word means _____ (synonym, translation, or brief explanation)''

(4) ''I know that this word means _____ (synonym, translation, or brief explanation)''

(5) ''I associate this word with _____, _____, _____.''

Internal consistency reliability of the test as a measuring instrument was calculated by Zareva (2005) through using Kuder–Richardson 21 (K–R 21) formula. He asserted that the results were comparable with the reliability values of other instruments for assessment of lexical knowledge (e.g., Forms A and B from the University Word Level Test [Xue and Nation, 1984]). The reliability of the test was checked with the data of the present study as well and it was founded As follows:

Table 1. The reliability of the Test as a Measuring Instrument

	N	Mean	Std. Deviation	Reliability
V.V.K	106	165.8962	19.81627	1.5589
V.F	106	38.0660	13.56836	0.9135
S.P.V	106	184.8113	25.93145	1.4407
V.A	106	33.5377	19.63872	0.9662
VSIZE	106	8127.8775	2102.84546	1.2195
Valid N (list wise)	106			

According to Zareva (2005) the test has lowest values of total bias across the different proficiency groups and also has a strong potential to account for the variation in the vocabulary knowledge of language users who are at different levels of proficiency.

Procedures

The researcher herself attended all the classes in which the vocabulary test was administered. The students were provided with some instruction on completing the scale and the scoring procedures at the beginning of each session. The students were requested to answer the questionnaire without any time limit. The instructors introduced the researcher to the class, and then left the room for the duration of the study. The researcher who was not previously acquainted with any of the participants, informed students of the confidential nature of the study and that their participation was voluntary. The students were encouraged to ask any questions with regard to the language (word or structure) of the questionnaire to make sure that they do not have any problem with the language. They were also asked to write their sex, name, and educational year on both the test and the questionnaire. The students were also informed that their participation in the study would not affect their course grades. In summary, the following conditions were common to all cases: a) the students were given as much time as was needed to complete the test, b) the students answered the questions without conferring with classmates, and c) students were motivated to try their best in order to achieve the assumed rewards.

Data Analysis

The students' tests of vocabulary were corrected and scored in appropriate way. The scoring procedures for quantifying each of the variables were as follows:

(1) Verified (actual) vocabulary knowledge

Options (1) through (4) on the test were assigned numerical values dependent on how a participant self-reported his/her familiarity with the TWs. Options (1) and (2) yielded a score of 1 and 2 points, respectively. Option (3) leaded to a maximum score of 3, if a participant have provided an acceptable synonym, brief definition, or translation of the TW (for the L2 learners), or 2 points if a participant have claimed some familiarity but the response have showed that he/she did not recognize the TW correctly. Option (4) yielded maximum 4 points, if knowledge of the TW has been demonstrated as required, or 2 points, if the response has revealed some sort of a meaning misinterpretation of the TW or has not reflected its lexical category.

(2) Self-perceived vocabulary knowledge

The answers were assigned a numerical value from 1 to 4 according to the option the participants have chosen.

(3) Vocabulary size

Vocabulary size was estimated by multiplying the number of known words obtained from the correct responses to option (3) or (4) by the number of words counted in the dictionary (23,996), divided by the number of words used in the test (73).

(4) Word frequency effects

The TWs in the test represent four word frequency bands, which were defined by reference to The Educator's Word Frequency Guide (Zeno et al., 1995). To examine the relationship between word frequency and vocabulary knowledge, a numerical value from 1 to 4 was assigned to each TW the participants knew, following the principle: the lower the frequency of occurrence of a TW (i.e., its SFI), the higher the numerical value assigned. A major consideration in the quantification of the responses was to give credit to participants who knew words from the lower frequency bands, in addition to knowing high frequency words.

(5) Number of associations

The responses obtained from the participants were lemmatized and tallied on a list by combining in one item the base form of a word and its regularly inflected forms (e.g., school and schools, think and thinks, low – lower – lowest, etc.). The multiword responses were scored as one item and their classification was based on the head of the phrase (e.g., baseball game, wedding ring, etc.). All derivations (e.g., disadvantage, famished, starving, amoral, immoral, etc.) and irregularly inflected forms (e.g., men, worse, began, etc.) were treated as separate items. Each response was assigned a numerical value of commonality based on its frequency of occurrence in the collected data.

A numerical value of 0–3, reflecting the absolute number of associations generated to a TW by a participant, was assigned to the TWs with which participants demonstrated familiarity in an acceptable way. The total number of word associations generated by a participant was considered to reflect the size of his/her associative domain.

Also, students' scores of their reading courses were standardized and changed into t score in order to make them comparable. Then average of the standardized scores was used as an index of participant's reading performance.

Regression analysis was used to examine the relation between the criterions, i.e., participants' reading performance and their different aspects of vocabulary knowledge measured by their verified responses to the 73 TWs, and the five predictors, i.e.: (1) Verified (actual) vocabulary knowledge, (2) self-perception of vocabulary knowledge, (3) knowledge of words from various frequency bands, (4) vocabulary size, (5) number of associations.

RESULTS AND DISCUSSION

A regression analysis between vocabulary knowledge and reading performance was administered to see if the overall model is significant or not**.** Referring to the p-value of the F-test it was observed that the model was statistically significant with a p-value of zero to three decimal places. The R-squared is 0.25, meaning that approximately 25% of the variability of students' reading scores is accounted for by the variables in the model. In this case, the adjusted R-squared indicates that about 22% of the variability of students' reading scores is accounted for by the model.

Table 2. Regression Analysis between Vocabulary Knowledge and Reading Performance

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.500(a)	.250	.220	7.62283

In order to examine the contribution of different aspects of lexical knowledge to reading performance of the students, regression analysis was employed. The entry method used was "enter". The attained data is presented in Table 3.

Table 3. Regression Analysis between Vocabulary Knowledge and Reading Performance

Model		Unstandardized Coefficients		Standardized Coefficients	t	Significance
		B	Std. Error	Beta		
1	(Constant)	30.435	10.441		2.915	.004
	V.V.K	.250	.136	.574	1.836	.069
	S.P.V	-.201	.080	-.550	-.498	.014
	V.F	-.302	.257	-.475	-.174	.243
	VSIZE	.003	.002	.818	2.074	.041

Dependent Variable: READING

As it is shown in the table the results of the analysis were only significant for two of the independent variables. According to the results students' reading performance can be predicted by their self-perceived vocabulary knowledge and vocabulary size with Beta values of (-.550) and (.818). Consequently, it can be said that none of other independent variables, i.e.: (1) Verified (actual) vocabulary knowledge, (2) knowledge of words from various frequency bands, (3) number of associations can predict the achievement of students in their reading comprehension courses. Analysis of the gained results revealed that vocabulary size had the highest contribution to the reading performance of the students, as it is implied by the Beta value of (0.818).

In an attempt to explore the contribution of different dimensions of vocabulary knowledge to reading performance another regression analysis was performed in which the contribution of different models was investigated. The first model consisted of students' scores on vocabulary size and knowledge of words from different frequency bands as independent variables. Both characteristics have been found by previous research to tap well into learners' breadth of knowledge.

In another model, the receptive-productive dimension was investigated in for its contribution to the reading performance.

The results are presented in Table 4. As it is shown through the analysis the R – square that is gained for breadth, and receptive – productive vocabulary knowledge indicates that these models approximately account for 20%, and 14% of the variability of students' reading scores. Among these models one that consists of breadth of vocabulary knowledge has the greatest contribution to reading performance of students. Referring to the p-value of the F-test to see if the overall model is significant or not it was observed that the models were statistically significant with a p-value of zero to three decimal places.

Table 4. Regression Analysis between Breadth, and Receptive – Productive Vocabulary Knowledge and Reading Performance

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
Breadth	.451(a)	.203	.188	7.77890
Receptive –productive	.374(a)	.140	.132	8.04417

This result is in line with those gained through previous studies reviewed in chapter two, e.g., Qian, 2002; Huang ,2006, in which breadth of vocabulary knowledge has the greatest contribution to reading comprehension in comparison to other aspects of vocabulary knowledge. Moreover in the present study another aspect of vocabulary, i.e. receptive – productive dimension was investigated. And it was observed that it had a good amount of contribution to reading comprehension.

Since the collected data on students' depth of vocabulary knowledge, knowledge of word associations, was not statistically significant, the associated part was not included in the regression analysis.

In order to compensate for the defective data on students' knowledge of word associations and to include this part in regression analysis, that part of the data which was collected from freshman students (N = 42) was entered in the regression analysis. Data collected from these students were normally distributed and valid. The results are presented in Table 5.

Table 5. Regression Analysis between Breadth, Depth, and Receptive – Productive Vocabulary Knowledge and Reading Performance

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
Breadth (V Size, VF)	.458(a)	.209	.168	7.23210
Depth (VA)	.368(a)	.136	.113	7.46457
Receptive –productive (V.V.K)	.328(a)	.108	.085	7.58419

Referring to the p-value of the F-test to see if the overall models are significant or not, it was observed that the models were statistically significant with p-values of .012, .018, and .036 for Breadth, Depth, and Receptive – productive vocabulary knowledge. This result is also in line with previous studies that were mentioned. It was demonstrated that Depth of vocabulary, with R- Square of (.136), had the highest contribution to reading after Breadth, with R- Square of (.209).

CONCLUSION

Based on the results gained students' reading performance can be predicted by their self-perceived vocabulary knowledge and vocabulary size with Beta value of (-.550) and (.818). Consequently vocabulary size had the highest contribution to the reading performance of the students among other aspects of vocabulary knowledge. According to regression analysis performed on whole data, Breadth and Receptive – Productive dimensions of vocabulary knowledge account for 20%, and 14% of the variability of

students' reading scores, respectively. Based on regression analysis performed on data collected from freshman students (n = 42), Breadth, Depth and Receptive – Productive dimensions of vocabulary knowledge account for 20%, 13%, and 10% of the variability of students' reading scores, respectively. Therefore, Depth of vocabulary, with R- Square of (.136), had the highest contribution to reading after Breadth, with R- Square of (.209).

The results accomplished through this study have some implications on material development issues that are followed: First, considering the significant role of vocabulary breadth in reading comprehension, as it was demonstrated in the previous and present researches, it is implied that material developers pay more attention to plan reading materials that are graded based on students' vocabulary size, in order to increase the efficiency of pedagogical texts. Second, as it was revealed in the present study, and based on the great contribution of depth of vocabulary to reading comprehension, it is implied that EFL reading comprehension texts include exercises and activities that are focused on deepening and elaborating knowledge of new words along with extensive reading materials.

REFERENCES

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J.T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77-117). Newark, DE: International Reading Association.

Anderson, R. C., & Freebody, P. (1983). Reading comprehension and the assessment and acquisition of word knowledge. In B. Hutson (Ed.), *Advances in reading/language research: A research annual* (pp. 231-256). Greenwich, CT: JAI

Barrow, J. Nakanishi, Y., & Ishino, H. (1999). Assessing Japanese college students' vocabulary knowledge with a self-checking familiarity survey. *System, 27*(2), 223-247.

Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology, 99*(2), 311-325.

Cromley, J. G. (2005). Reading comprehension component processes in early adolescence. Ph.D. dissertation, University of Maryland, College Park, MD. Retrieved December 26, 2007, from ProQuest Digital Dissertations database. (Publication No. AAT 3178579).

Dale, E., (1965). Vocabulary measurement: Techniques and major findings. *Elementary English*, *42*, 395–401.

Hawkins, C. W. (1995). *The differential effects of word recognition and listening comprehension of narrative and expository text in grades one and three*. Ph.D. dissertation, University of Houston, Texas, United States of America. Retrieved December 26, 2007, from ProQuest Digital Dissertations database. (Publication No. AAT 9542602).

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition, 21*, 303–317.

Huang, C. (1999). *The effects of vocabulary knowledge and prior knowledge on reading comprehension of EFL students in Taiwan*. Ph.D. dissertation, Ohio University, Ohio,

United States of America. Retrieved December 26, 2007, from ProQuest Digital Dissertations database. (Publication No. AAT 9929306).

Huang, H. (2006). *Breadth and depth of English vocabulary knowledge: Which really matters in the academic reading performance of Chinese university students?* M.A. dissertation, McGill University, Canada. Retrieved December 26, 2007, from ProQuest Digital Dissertations database. (Publication No. AAT MR24873).

Lam, M. (2005). *The acquisition of grammatical skills and its relation to reading comprehension in ESL students*. Ph.D. dissertation, University of Toronto, Canada. Retrieved December 30, 2007, from ProQuest Digital Dissertations database. (Publication No. AAT NR02622).

Mezynski, K. (1983). Issues concerning the acquisition of knowledge: Effects of vocabulary training on reading comprehension. *Review of Educational Research, 53*, 253-279.

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension? *Journal of Educational Psychology, 98* (*3*), 554-566.

Qian, D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, *52* (3), 513-536.

Qian, D., & Schedl, M. (2004). Evaluation of an in-depth vocabulary knowledge measure for assessing reading performance. *Language Testing*, *21* (1), 28–52.

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.

Stratton, R. P., & Nacke, P.L. (1974). The role of vocabulary knowledge in comprehension In P. L. Nacke (Ed.), *Interaction: Research and Practice for College-Adult Reading*, Twenty-Third Yearbook of the National Reading Conference. (pp. 176-184). Clemson, SC: National Reading Conference.

Tuinman, J. J., & Brady, M. E. (1974). How does vocabulary account for variance on reading comprehension tests? A preliminary instructional analysis. In P. L. Nacke (Ed.), *Interaction: Research and Practice for College-Adult Reading*, Twenty-Third Yearbook of the National Reading Conference. (pp. 176-184). Clemson, SC: National Reading Conference.

Zareva, A. (2005). Models of lexical knowledge assessment of second language learners of English at higher levels of language proficiency. *System*, *33*, 547–562.