THE CORRELATION BETWEEN VOCABULARY SIZE AND THE READING COMPREHENSION OF THE ENGLISH EDUCATION DEPARTMENT STUDENTS

Cinthia Astan [4]
Siti Mina Tamah [5]

Abstract

The study explored the correlation between vocabulary size and the three levels of reading comprehension namely literal, inferential, and critical reading comprehension. The main intention was to discover what the correlation was between vocabulary size and literal reading comprehension, inferential reading comprehension, and critical reading comprehension in terms of direction and magnitude. The subjects were 30 students of the English Department at a university in Surabaya. A Vocabulary Size Test and a Reading Comprehension Test were administered to measure the subjects' vocabulary size and reading comprehension performance. The findings revealed that vocabulary size was positively, strongly, and significantly correlated to literal reading comprehension, inferential reading comprehension and critical reading comprehension.

Key Terms: *correlation, vocabulary size, literal reading comprehension, inferential reading comprehension, critical reading comprehension.*

Introduction

Vocabulary is plausibly related to reading. Numerous researchers have also acknowledged this relationship. Hancock (1998) as cited in Chou (2011) believes that in reading, comprehension involves understanding the vocabulary, seeing relationships among words and concepts, organizing ideas, recognizing the author's purpose, evaluating the context, and making judgments. Nation (2001) believes that students' reading comprehension will improve when their vocabulary size increases.

Moreover, Gray (1960) as cited in Alderson (2000) also states that to achieve comprehension the readers must concurrently process three levels, namely reading "the lines", reading 'between the lines", and reading "beyond the lines". The first refers to literal comprehension, the second refers to inferential comprehension, and the third refers to critical comprehension.

---

[4] Cinthia Astan *adalah* Mahasiswa S2 di Universitas Katolik Widya Mandala Surabaya.

[5] Siti Mina Tamah *adalah* Dosen Program Studi Pendidikan Bahasa Inggris di FKIP Universitas Katolik Widya Mandala Surabaya.

Generated by the aforementioned acknowledgment of the relationship between vocabulary knowledge and reading comprehension, the researchers are interested to conduct a more confined research in this area to answer the following questions:

1. What is the correlation between vocabulary size and literal reading comprehension?
2. What is the correlation between vocabulary size and inferential reading comprehension?
3. What is the correlation between vocabulary size and critical reading comprehension?

Vocabulary Size

Vocabulary size is termed by Qian (2002) as cited in Mehrpour, Razmjoo and Kian (2011) as vocabulary breadth. It refers to the number of words that a learner knows, at least the surface meaning. Meara (1996) as stated in Schmitt (2008) indicates the importance of vocabulary size. He endorses that the basic dimension of lexical competence is size, and states that students with big vocabularies are more proficient in a wide range of language skills than students with smaller vocabularies.

Vocabulary Threshold for Reading Comprehension

According to threshold hypothesis mentioned by Laufer (1997) as cited in Keshavarz and Mohammadi (2009), certain amount of vocabulary are necessarily acquired in order to be able to use higher level processing strategies to comprehend a text. If the threshold is crossed, adequate comprehension is possible. On the contrary, if the threshold is not crossed, the comprehension is consequently inadequate.

Hirsch and Nation (1992) as cited in Eyckmans (2004) assume that in order to reach text comprehension, readers need to be familiar with 95% of the words in a text. In the latter studies, Hu and Nation (2000), Schmitt, Jiang, and Grabe (2011) as cited in Nation and Anthony (2013) state that students need to understand around 98% of the running words in a text for unassisted comprehension, which equates to around 8,000 word families.

Reading Comprehension

Reading comprehension is generally defined as understanding what is read, where words have context and text have meaning. According to Lunzer and Dolan (1980) as cited in Hussein (2012), reading comprehension is a measure of ability and willingness to reflect on whatever it is being read. It means that the reader does two things in the reading process. One of them is certainly reading, and the other one is not only understanding but also in some sense thinking. Gray (1960) as cited in Alderson (2000) states that to achieve comprehension the readers must concurrently process three levels, which are reading 'the lines'

(literal comprehension), reading 'between the lines' (inferential comprehension), and reading 'beyond the lines' (critical comprehension).

According to Potts (1976) as cited in Mohamad (1999), literal comprehension is understanding the literal meaning of the text. It involves understanding surface meanings. At this level, students find information and ideas that are explicitly stated in the text   Inferential comprehension is defined by Durkin (1978) as cited in Hussein (2012) as understanding the meaning in the text that is not directly communicated. It may be a conclusion, and inference, a prediction, identification of a cause. Critical comprehension, according to Potts (1976) as cited in Mohamad (1999) includes ability to evaluate ideas and information, synthesize, and analyze.

Vocabulary Size and Reading Comprehension

Vocabulary is a key component of reading for meaning. If students know the meaning of a word, they will be able to read and digest it within a sentence. According to Anderson and Freebody (1981), who propose instrumentalist hypothesis, the presence or absence of vocabulary knowledge causes or hampers reading comprehension. Perfetti (1985) as cited in Chen (2011) declares about verbal efficiency theory that becoming efficient in processing lower level reading skills such as vocabulary knowledge and word recognition will facilitate readers in the processing of higher level reading skills in order to help them attain reading comprehension.

Hypotheses

The correlation between vocabulary and reading comprehension was depicted by instrumentalist hypothesis and verbal efficiency theory. Therefore, the formulated hypotheses were: (1) the correlation between vocabulary size and literal reading comprehension is positive and strong, (2) the correlation between vocabulary size and inferential reading comprehension is positive and strong, and (3) the correlation between vocabulary size and critical reading comprehension is positive and strong.

Research Method

This research is a quantitative correlational research, and it aims to investigate the correlation between vocabulary size and the three levels of reading comprehension. The data were gathered by using two instruments namely Vocabulary Size Test (Nation & Beglar, 2007) and Reading Comprehension Test. The research is planned (1) to measure the vocabulary size through Vocabulary Size Test, (2) to measure reading comprehension performance through Reading Comprehension Test, and (3) to draw conclusion based on the research questions formulated.

The subjects of this research were the second semester English Department students of a university in Surabaya, from two classes of

Reading 1 since this research was about reading comprehension. Each class consisted of 15 students. Therefore, the total sample was 30 students, in order to fulfill the accountability of the data gathered (McMillan, 2008).

In order to find the correlation between vocabulary size and reading comprehension in terms of literal, inferential, and critical comprehension, the researcher necessarily utilized a standardized test of vocabulary size and a self-developed reading comprehension test.

The Vocabulary Size Test used in this research was the 14,000 version containing 140 multiple-choice items, with 10 items from each 1,000 word family level (Nation & Beglar, 2007). This test was considerably suitable as the instrument of this research according to the goal and construction, sources of words, and distractors' quality.

In scoring, one point was given to each correct answer. The number of correct answers was then multiplied by 100 to find the vocabulary size. For example, if the number of correct answer is 35, the vocabulary size will be 35 x 100 or 3,500 word families.

A reading comprehension test was developed, tested for its realibility, and administered as instrument of this research. The passage was entitled *Solving Crimes with Modern Technology*, taken from Chapter 9 of I*ssues for Today* 3rd Edition (Smith & Mare, 2003). There were thirty-nine questions in the test. Nine questions represented literal comprehension, fifteen questions represented inferential comprehension, and 14 questions represented critical comprehension.

The scoring was provided for four parts individually: score of overall reading comprehension, score of literal comprehension, score of inferential comprehension, and score of critical comprehension. Each of correct answer was given 1 point. The maximum score for each part was 100 points.

Data Collection and Data Analysis Procedure

The Vocabulary Size Test and the Reading Comprehension Test (Research Instrument) were administered on April 16, 2014 to the subjects, taking Reading 1. Fifteen students in the first class started the test at 7.30 am, while another fifteen in the second class started at 11.40 am. Both classes did the tests in two sessions without intermission. In the first session, Reading Comprehension Test was administered. The time allocated was 32 minutes. The students read one passage and were asked to answer 38 questions about the passage. In the next session, the students' vocabulary size was tested using Vocabulary Size Test. The time allocated was 40 minutes. Both instuments use multiple choice type.

In order to obtain the vocabulary size of the research subjects, the vocabulary size from 1,000 word families level up to 14,000 word families level was counted. The calculation was conducted by accumulating the subjects' correct answers multiplied by 100 at each

level. Furthermore, the mean of the subjects' vocabulary size at the levels was calculated to disclose the average of the vocabulary.

As for the reading test, the score was based on the number of items correctly answered in the reading comprehension test. The analysis was conducted based on overall reading comprehension score, literal comprehension score, inferential comprehension score, and critical comprehension score. Furthermore, the mean of the subjects' scores at the levels was calculated to show the average of the reading score.

Three steps were taken for the correlation analysis to answer the research questions and test the hypotheses: (1) Pearson correlation coefficient was calculated to discover the correlation between vocabulary size and the three levels of reading comprehension in terms of its direction and magnitude, (2) the result of the Pearson correlation coefficient was cross-examined with the standard critical value of correlation coefficient at significant level of .05 (the standard for educational research), and (3) the result of the correlation coefficient was squared to calculate the coefficient of determination.

Findings and Discussion

From 30 research subjects, the overall vocabulary size ranged from 5,900 word families to 8,400 word families. The average was 7,203 word families. According to the result, the average vocabulary size of the research subjects were apparently approaching the vocabulary threshold for reading comprehension postulated by Nation, which was around 8,000 word families (Nation, 2006 as cited in Nation and Anthony, 2013).

The overall reading comprehension scores of the research subjects ranged from 37 to 79. The average was 55. This result showed that reading comprehension of the subjects was moderately low. The average score of literal, inferential, and critical reading comprehension was respectively 70, 53, and 47.

Correlation between Vocabulary Size and the Three Levels of Reading Comprehension

Below table displays the correlation between vocabulary size and the three levels of reading comprehension namely literal, inferential, and critical reading comprehension.

Table 1
Correlation between Vocabulary Size and the Three Levels of
Reading Comprehension (N=30)

Variable	VS	Interpretation
L Pearson Correlation (r)	.761**	Positive and strong
Sig 2-tailed (p)	.001	
Critical Value (r_{table})	.361	Significant ($r > r_{table}$)
Coefficient of	.579	57.9% of variations in L is
Determination (r^2)	30	predicted by variation in VS

Number of subjects			
I Pearson Correlation (r)	.822**	Positive and strong	
Sig 2-tailed (p)	.001		
Critical Value (r_{table})	.361	Significant ($r > r_{table}$)	
Coefficient of Determination (r^2)	.676	67.6% of variation in I is predicted by variation in VS	
Number of subjects	30		
C Pearson Correlation (r)	.839**	Positive and strong	
Sig 2-tailed (p)	.001		
Critical Value (r_{table})	.361	Significant ($r > r_{table}$)	
Coefficient of Determination (r^2)	.704	70.4% of variation in C is predicted by variation in VS	
Number of subjects	30		

Note: **Correlation is significant at 0.001 level (2-tailed)

VS: Vocabulary Size at 14,000 word families level; **L**: Literal Reading Comprehension
I: Inferential Reading Comprehension; **C**: Critical Reading Comprehension

Correlation between vocabulary size and literal reading comprehension

The Pearson correlation coefficient between vocabulary size and literal reading comprehension was .761, p< .001. The positive number indicated that the direction of the correlation was positive. The magnitude of the correlation was categorized as strong since it was in the range of .70 and 1.00 (Creswell, 2008). Therefore, the null hypothesis for the first research question was rejected, and the alternative hypothesis was accepted.

Moreover, the correlation of vocabulary size and literal reading comprehension (r = .761, p< .001) was higher than the standard critical coefficient value of (r_{table} = .361). It suggested that the correlation was statistically significant. The coefficient of determination, showed the value of .579, which means 57.9% of the variation in literal reading comprehension was accounted for by variation in vocabulary size. In other words, vocabulary predicted literal reading comprehension by 57.9% accuracy.

Correlation between vocabulary size and inferential reading comprehension

The Pearson correlation coefficient between vocabulary size and inferential reading comprehension was .822, p< .001. The correlation was categorized as positive and strong since the number is positive and in the range of .70 and 1.00 (Creswell, 2008). Therefore, the null hypothesis for the second research question was rejected and the alternative hypothesis was accepted.

Moreover, the correlation of vocabulary size and inferential reading comprehension (r = .822, p< .001) was higher than the standard critical coefficient value of (r_{table} = .361). It suggested that the correlation was statistically significant. As predictor of inferential reading comprehension, the variation in vocabulary size contributed 67.6% to the variation in inferential reading comprehension. It was discovered by the coefficient of determination of .676.

Correlation between vocabulary size and critical reading comprehension

The correlation between vocabulary size and critical reading comprehension was .839, p< .001. It was seemingly higher than the correlation between vocabulary size and the two previous levels. The positive number indicated that the direction of the correlation was positive. The magnitude was categorized as strong, since it was in the ranged of $\pm$ .70 - 1.00 (Creswell, 2008). Therefore, the null hypothesis for the third research question was rejected, and the alternative hypothesis was accepted.

The correlation (r = .839, p< .001) was also considered statistically significant since it was higher than .361, which was the standard critical coefficient value of (r_{table} = .361). The coefficient of determination showed the value of .704. It suggested that 70.4% of variation in critical reading comprehension could be accounted for by variation in vocabulary size.

Conclusion

The findings reveal that vocabulary size has significantly positive and strong correlation with the three levels of reading. This correlation plausibly supports many researchers' acknowledgment about the correlation between vocabulary and reading. Koda (1989) and Qian (1999) as cited in Chen (2011) mention that vocabulary knowledge heavily relates to reading comprehension more than other factors such as grammar knowledge.

Vocabulary size is apparently also a good predictor to the three levels of reading comprehension since the variation in vocabulary size could predict the variation in literal, inferential, and critical reading comprehension with 57.9%, 67.6%, and 70.4% accuracy respectively. Hence, it is crucial for readers to continually increase their vocabulary size in order to facilitate their reading comprehension especially in the three levels which are literal, inferential, and critical reading comprehension.

It is worth-noticing that the correlation between vocabulary size and the three levels of reading comprehension were all positive and strong, and statistically significant, despite the subjects' low scores in the reading comprehension test. Seemingly, the correlation was not affected

by the low scores. It is presumably because correlational research only determines whether an increase and decrease in one variable corresponded to an increase or decrease in the other variable. When vocabulary size of the subjects increased, their literal, inferential, and critical reading comprehension also increased.

References

Alderson, C. (2000). *Assessing reading.* Cambridge: Cambridge University Press.

Anderson, R. C. & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews,* (pp. 77-81). Newark, DE: International Reading Association.

Chen, K.Y. (2011). The impact of EFL students' vocabulary breadth of knowledge on literal reading comprehension. *Asian English Foreign Language Journal, 51*, 30-40.

Chou, P. (2011). The effects of vocabulary knowledge and background knowledge on reading comprehension of Taiwanese EFL students. *Electronic Journal of Foreign Language Teaching, 8*(1), 108-115.

Eyckmans, J. (2004). *Measuring receptive vocabulary size.* Utretch, Netherlands: LOT Company. Retrieved from https://www.dare.ubn.kun.nl/bitstream/2066/19469/1/.pdf

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

Hussein, B.A. (2012). Analysis of the real situation of teaching reading comprehension to first year students at the department of English language and literature at Al Zaytoonah Private University of Jordan. *Asian Social Science, 8*(4), 237-251.

McMillan, J.H. (2008). *Educational Research.* New Jersey: Pearson Education, Inc.

Mehrpour, M., Razmjoo, S.A., & Kian, P. (2011). The relationship between depth and breadth of vocabulary knowledge and reading comprehension among Iranian EFL learners. *Journal of English Language Teaching and Learning, 53*(22), 97-127.

Mohamad, A. (1999). What do we test when we test reading comprehension? *Journal of Teaching English as Second Language, 5*(12). Retrieved from iteslj.org/Techniques/Mohamad-TestingReading.html

Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

Nation, I. S. P. & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, *31*(7), 9-13.

Nation, I. S. P. (2012). The vocabulary size test. Retrieved from http://www.victoria.ac.nz

Nation, I. S. P. & Anthony, L. (2013). Mid-frequency readers. *Journal of Extensive Reading,*

Schmitt, N. (2008). Review article instructed second language vocabulary learning. *Language Teaching Research*, *12*(3), 329-363.