



The Relationship Between English Vocabulary Size and Reading Comprehension Among EFL University Students

Karin Ariska¹✉, Devi Vionitta Wibowo²

¹UIN Raden Intan Lampung, Indonesia

²Sekolah Tinggi Pendidikan Holistik Berbasis Karakter, Indonesia

✉karinariska563@radenintan.ac.id

Article Information

Submitted March 10, 2026

Revised March 20, 2026

Accepted May 10, 2026

Keywords:

Vocabulary size;
Reading comprehension;
EFL learners;
Correlation; Language
proficiency.

Abstract

Background: Vocabulary knowledge is a fundamental component of language proficiency and plays a critical role in reading comprehension. A larger vocabulary enables learners to understand texts more effectively and supports overall language development.

Aim: This study aimed to examine the relationship between English vocabulary size and reading comprehension among EFL university students.

Method: A quantitative correlational design was employed involving 80 undergraduate EFL students. Data were collected using a Vocabulary Size Test (VST) and a reading comprehension test. Descriptive statistics and Pearson Product-Moment Correlation were used to analyze the data.

Results: The findings revealed a significant positive correlation between vocabulary size and reading comprehension ($r = .68, p < .001$). Students with larger vocabularies tended to achieve higher reading comprehension scores. The coefficient of determination indicated that vocabulary size accounted for approximately 46% of the variance in reading comprehension performance.

Conclusion: English vocabulary size is significantly associated with reading comprehension among EFL learners. The results suggest that enhancing learners' vocabulary knowledge may contribute to improved reading performance and overall language proficiency.

Introduction

Vocabulary knowledge is widely recognized as a fundamental component of language proficiency and plays a crucial role in second-language acquisition. In English as a Foreign Language (EFL) contexts, vocabulary serves as the foundation for effective communication and language development. Learners with limited vocabulary knowledge often experience difficulties in understanding and producing language, which may negatively affect their academic performance and overall language achievement (Schmitt, 2014; Vermeer, 2001).

Reading comprehension is one of the language skills most strongly influenced by vocabulary knowledge. Successful reading requires learners to recognize words, access their meanings, and integrate information across a text. Readers need sufficient lexical coverage to achieve adequate comprehension, and inadequate vocabulary knowledge may hinder effective understanding of written texts (Nation, 2006; Laufer & Ravenhorst-Kalovski, 2010).

Numerous studies have reported a significant relationship between vocabulary knowledge and reading comprehension. Learners with larger vocabularies tend to achieve better reading performance because they can process textual information more efficiently and accurately. Vocabulary knowledge has been identified as a strong predictor of academic reading success among second-language learners (Qian, 2002; Schmitt et al., 2011).

Evidence from previous research further supports the importance of vocabulary in reading performance. Vocabulary knowledge contributes substantially to reading comprehension and is consistently identified as one of the strongest predictors of successful second-language reading. In addition, vocabulary development plays a significant role in

reading achievement across different educational levels (Jeon & Yamashita, 2014; Verhoeven & Van Leeuwe, 2008).

Despite extensive research on vocabulary and reading comprehension, further investigation is needed in different educational contexts. University EFL students are frequently required to engage with academic texts containing sophisticated vocabulary and complex concepts. Therefore, examining the relationship between vocabulary size and reading comprehension among university learners remains important for improving language instruction and academic literacy development (Perfetti & Stafura, 2014; Kieffer & Lesaux, 2012).

Based on these considerations, the present study aimed to investigate the relationship between English vocabulary size and reading comprehension among EFL university students. Specifically, the study sought to determine whether a significant correlation exists between learners' vocabulary knowledge and their reading comprehension performance.

Literature Review

Vocabulary Size

Vocabulary size refers to the number of words a learner knows and can recognize in a language. It is considered one of the most important indicators of language proficiency because vocabulary knowledge supports learners' ability to communicate and comprehend language effectively. A larger vocabulary enables learners to understand a wider range of texts and participate more successfully in academic activities (Schmitt, 2014; Vermeer, 2001).

Vocabulary knowledge plays a crucial role in language learning because words serve as the building blocks of communication. Learners with extensive vocabulary knowledge are more capable of processing information, understanding meanings, and interpreting written texts. Consequently, vocabulary size has been widely used as a predictor of language achievement in second-language research (Qian, 2002; Schmitt, 2014).

Reading Comprehension

Reading comprehension is the process of constructing meaning from written texts through the interaction of linguistic knowledge and cognitive processes. Effective reading comprehension requires readers to recognize words, understand sentence structures, and integrate information across a text. Therefore, reading comprehension involves not only decoding written symbols but also interpreting and evaluating textual information (Perfetti & Stafura, 2014; Cromley & Azevedo, 2007).

For EFL learners, reading comprehension is an essential academic skill because much of their learning depends on understanding written materials. Students with strong reading comprehension skills are better able to access academic information, develop critical thinking abilities, and achieve academic success. Conversely, limited comprehension may hinder learning and academic performance (Jeon & Yamashita, 2014; Droop & Verhoeven, 2003).

Relationship Between Vocabulary Size and Reading Comprehension

Previous studies have consistently demonstrated a positive relationship between vocabulary knowledge and reading comprehension. Learners with larger vocabularies tend to achieve better comprehension because they can recognize words more efficiently and process textual information more accurately. Vocabulary knowledge has been identified as one of the strongest predictors of reading achievement among second-language learners (Qian, 2002).

Research has also shown that adequate vocabulary coverage is necessary for successful reading comprehension. Learners who know a higher percentage of words in a text are more likely to understand its content, infer meanings, and identify key ideas. In contrast, insufficient vocabulary knowledge may limit comprehension and increase processing difficulty (Laufer & Ravenhorst-Kalovski, 2010; Schmitt et al., 2011).

Furthermore, a meta-analysis of second-language reading studies confirmed that vocabulary knowledge is among the most significant factors influencing reading comprehension performance. These findings suggest that vocabulary development should be prioritized in language instruction to support learners' reading achievement and overall language proficiency (Jeon & Yamashita, 2014).

Method

Research Design

This study employed a quantitative correlational research design to examine the relationship between English vocabulary size and reading comprehension among EFL university students. A correlational design is appropriate for investigating the strength and direction of the relationship between two variables without manipulating them (Qian, 2002; Verhoeven & Van Leeuwe, 2008). In this study, vocabulary size served as the independent variable, while reading comprehension functioned as the dependent variable. Pearson Product-Moment Correlation was employed to determine whether a significant relationship existed between the two variables (Kieffer & Lesaux, 2012; Jeon & Yamashita, 2014).

Research Site and Participants

The study was conducted at the English Education Department of a university in Indonesia. The participants consisted of 80 undergraduate EFL students selected through convenience sampling. Convenience sampling was employed because it enabled the researcher to recruit participants who were readily available and relevant to the objectives of the study (Alptekin & Erçetin, 2011). The participants ranged in age from 18 to 22 years and had studied English in formal educational settings for several years. They were selected because they possessed sufficient experience in English language learning and were actively enrolled in reading-related courses.

Instrument

Two instruments were employed to collect the data. The first instrument was the Vocabulary Size Test (VST), which was used to measure participants' English vocabulary knowledge. The VST is a widely recognized instrument for assessing learners' receptive vocabulary size and has demonstrated satisfactory validity and reliability in second-language research (Beglar, 2010; Schmitt, 2014). The test consisted of multiple-choice items designed to estimate the number of English word families known by the participants.

The second instrument was a Reading Comprehension Test adapted from academic English reading materials. The test included several reading passages followed by multiple-choice questions assessing learners' ability to identify main ideas, infer meanings, understand details, and interpret textual information. Reading comprehension tests are commonly used to evaluate learners' understanding of written texts and have been widely applied in EFL research (Jeon & Yamashita, 2014; Perfetti & Stafura, 2014).

Prior to data collection, both instruments were reviewed to ensure clarity and suitability for the participants' proficiency level. The scores obtained from the Vocabulary Size Test and Reading Comprehension Test were used for subsequent statistical analyses.

Data Collection

Data collection was conducted during the academic semester at the selected university. Participants completed the Vocabulary Size Test (VST) and the Reading Comprehension Test in a classroom setting under the supervision of the researcher. Before the administration of the tests, participants received instructions regarding the purpose and procedures of the study.

The Vocabulary Size Test was administered first and required approximately 20–30 minutes to complete. Following a short break, participants completed the Reading Comprehension Test, which required approximately 30–40 minutes. All participants completed

the tests individually and were instructed not to use dictionaries, mobile phones, or other external resources during the testing session.

Upon completion of the tests, participants' responses were collected and scored according to the established scoring procedures. The resulting scores represented the participants' vocabulary size and reading comprehension performance and were subsequently used for statistical analysis. Data collection procedures were conducted in accordance with ethical principles to ensure participant confidentiality and voluntary participation (Alptekin & Erçetin, 2011).

Data Analysis

The collected data were analyzed using IBM SPSS Statistics. Descriptive statistics, including means, standard deviations, minimum scores, and maximum scores, were calculated to provide an overview of participants' vocabulary size and reading comprehension performance. These statistical measures were used to summarize participants' performance and identify general patterns in the data (Verhoeven & Van Leeuwe, 2008; Kieffer & Lesaux, 2012).

To examine the relationship between English vocabulary size and reading comprehension, Pearson Product-Moment Correlation analysis was performed. Correlational analysis has been widely used in previous studies investigating the association between vocabulary knowledge and reading performance (Qian, 2002; Stæhr, 2008). The correlation coefficient (r) was interpreted to determine the strength and direction of the relationship between the two variables.

Prior to conducting the correlation analysis, the assumptions of normality and linearity were examined to ensure the appropriateness of Pearson correlation. Statistical significance was established at the .05 level ($p < .05$). In addition, the coefficient of determination (r^2) was calculated to determine the proportion of variance in reading comprehension that could be explained by vocabulary size (Kieffer & Lesaux, 2012; Verhoeven & Van Leeuwe, 2008).

Result and Discussion

Result

Descriptive Statistics

Descriptive statistics were calculated to summarize participants' performance on the Vocabulary Size Test and Reading Comprehension Test. The results are presented in Table 1.

Table 1. Descriptive Statistics of Vocabulary Size and Reading Comprehension

Variable	Mean	SD	Minimum	Maximum
Vocabulary Size	67.45	8.72	48	85
Reading Comprehension	73.18	9.15	52	91

As shown in Table 1, the participants obtained a mean score of 67.45 ($SD = 8.72$) on the Vocabulary Size Test and 73.18 ($SD = 9.15$) on the Reading Comprehension Test. The results indicate that the participants demonstrated moderate to high levels of vocabulary knowledge and reading comprehension performance.

Correlation Analysis

Pearson Product-Moment Correlation analysis was conducted to examine the relationship between English vocabulary size and reading comprehension. The results are presented in Table 2.

Table 2. Pearson Correlation Between Vocabulary Size and Reading Comprehension

Variables	r	p
Vocabulary Size and Reading Comprehension	.68	< .001

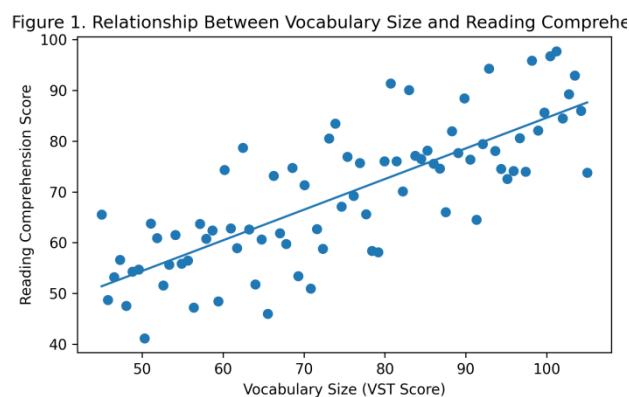


Figure 1. Scatter Plot of the Relationship Between Vocabulary Size and Reading Comprehension

The correlation analysis revealed a significant positive relationship between vocabulary size and reading comprehension ($r = .68$, $p < .001$). The correlation coefficient indicates a strong positive association between the two variables. Therefore, students with larger vocabulary sizes tended to achieve higher reading comprehension scores.

Furthermore, the coefficient of determination ($r^2 = .46$) indicated that approximately 46% of the variance in reading comprehension performance could be explained by vocabulary size. This finding suggests that vocabulary knowledge is an important factor contributing to reading comprehension among EFL university students.

Based on the results, the null hypothesis (H_0) was rejected, while the alternative hypothesis (H_1) was accepted. The findings confirm that a significant positive relationship exists between English vocabulary size and reading comprehension among EFL university students.

Discussion

The present study aimed to investigate the relationship between English vocabulary size and reading comprehension among EFL university students. The findings revealed a significant positive correlation between the two variables ($r = .68$, $p < .001$), indicating that students with larger vocabularies tended to achieve higher reading comprehension scores. This result suggests that vocabulary knowledge plays a substantial role in facilitating learners' understanding of written texts.

The findings are consistent with previous research demonstrating that vocabulary knowledge is a strong predictor of reading comprehension performance. Learners with broader vocabulary knowledge are better able to recognize words, access meanings, and interpret textual information effectively. As a result, they can comprehend texts more accurately and efficiently than learners with limited vocabulary knowledge (Qian, 2002; Schmitt et al., 2011).

The results also support previous studies suggesting that adequate vocabulary knowledge is necessary for successful reading comprehension. Students who possess sufficient vocabulary knowledge are more likely to understand the content of a text, identify main ideas, and infer meanings from context. In contrast, insufficient vocabulary knowledge may create comprehension difficulties and hinder academic reading performance (Laufer & Ravenhorst-Kalovski, 2010; Hsueh-Chao & Nation, 2000).

Furthermore, the coefficient of determination ($r^2 = .46$) indicated that vocabulary size explained approximately 46% of the variance in reading comprehension scores. This finding suggests that vocabulary knowledge is a significant contributor to reading achievement, although other factors such as background knowledge, reading strategies, motivation, and

language proficiency may also influence comprehension performance (Jeon & Yamashita, 2014; Perfetti & Stafura, 2014).

From a pedagogical perspective, the findings highlight the importance of vocabulary instruction in EFL classrooms. Teachers should provide learners with opportunities to expand their vocabulary through extensive reading, contextualized vocabulary activities, and explicit vocabulary-learning strategies. Strengthening students' vocabulary knowledge may enhance their reading comprehension and overall language proficiency (Schmitt, 2014; Carlo et al., 2004).

Overall, the findings reinforce the view that vocabulary knowledge is a fundamental component of successful reading comprehension. Therefore, vocabulary development should remain a central focus of language teaching and learning, particularly in academic EFL contexts (Kieffer & Lesaux, 2012; Verhoeven & Van Leeuwe, 2008).

Implications

The findings of this study have important implications for English language teaching and learning. Since vocabulary size was found to be significantly associated with reading comprehension, language teachers should place greater emphasis on vocabulary development in EFL classrooms. Explicit vocabulary instruction, extensive reading activities, and contextualized vocabulary-learning strategies may help students expand their lexical knowledge and improve their reading performance (Schmitt, 2014; Carlo et al., 2004).

In addition, curriculum developers should consider integrating vocabulary-focused materials into reading courses. Strengthening learners' vocabulary knowledge may facilitate text comprehension and contribute to overall language proficiency. Therefore, vocabulary instruction should be systematically incorporated into language-learning programs to support students' academic success (Kieffer & Lesaux, 2012; Perfetti & Stafura, 2014).

Limitations and Future Research

This study has several limitations that should be acknowledged. First, the participants were limited to undergraduate EFL students from a single university. Therefore, the findings may not be generalizable to learners from different educational levels or institutions (Alptekin & Erçetin, 2011; Droop & Verhoeven, 2003).

Second, the study focused only on the relationship between vocabulary size and reading comprehension. Other factors that may influence reading performance, such as reading strategies, motivation, background knowledge, and language proficiency, were not examined. Future studies should investigate these variables to provide a more comprehensive understanding of reading comprehension among EFL learners (Jeon & Yamashita, 2014; Perfetti & Stafura, 2014).

Third, the study employed a correlational design, which does not allow conclusions about causal relationships between variables. Future research may utilize experimental or longitudinal designs to examine how vocabulary development influences reading comprehension over time (Verhoeven & Van Leeuwe, 2008; Lervåg & Aukrust, 2010).

Future studies are also encouraged to include larger and more diverse samples from different educational contexts. Such research may provide broader insights into the role of vocabulary knowledge in second-language reading and contribute to the development of more effective language-learning practices (Schmitt, 2014; Kieffer & Lesaux, 2012).

Conclusion

This study investigated the relationship between English vocabulary size and reading comprehension among EFL university students. The findings revealed a significant positive correlation between the two variables, indicating that students with larger vocabularies tended to achieve better reading comprehension performance. The results suggest that vocabulary

knowledge plays an important role in facilitating learners' ability to understand written texts and process textual information effectively.

Furthermore, vocabulary size was found to account for a substantial proportion of the variance in reading comprehension scores, highlighting its importance as a predictor of reading success. These findings are consistent with previous studies emphasizing the critical role of vocabulary knowledge in second-language reading development.

Overall, the study underscores the importance of vocabulary development in EFL learning contexts. Enhancing learners' vocabulary knowledge through appropriate instructional practices may contribute to improved reading comprehension and overall language proficiency. Therefore, vocabulary instruction should remain a central component of English language teaching and learning.

Acknowledgment

The author would like to thank all participants for their valuable contribution to this study. Appreciation is also extended to the lecturers, academic staff, and colleagues whose support, guidance, and feedback contributed to the completion of this research.

References

- Alptekin, C., & Erçetin, G. (2011). The effects of working memory capacity and content familiarity on literal and inferential comprehension in L2 reading. *TESOL Quarterly*, 45(2), 235–266. <https://doi.org/10.5054/tq.2011.247705>
- Beglar, D. (2010). A Rasch-based validation of the Vocabulary Size Test. *Language Testing*, 27(1), 101–118. <https://doi.org/10.1177/0265532209340194>
- Biemiller, A. (2003). Vocabulary: Needed if more children are to read well. *Reading Psychology*, 24(3–4), 323–335. <https://doi.org/10.1080/02702710390227297>
- Binder, K. S., Cote, N. G., Lee, C., Bessette, E., & Vu, H. (2017). Beyond breadth: The contributions of vocabulary depth to reading comprehension among skilled readers. *Journal of Research in Reading*, 40(3), 333–343. <https://doi.org/10.1111/1467-9817.12069>
- Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, 96(1), 31–42. <https://doi.org/10.1037/0022-0663.96.1.31>
- Carlo, M. S., August, D., McLaughlin, B., Snow, C., Dressler, C., Lippman, D., Lively, T., & White, C. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners. *Reading Research Quarterly*, 39(2), 188–215. <https://doi.org/10.1598/RRQ.39.2.3>
- 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. <https://doi.org/10.1037/0022-0663.99.2.311>
- Dickinson, D. K., Golinkoff, R. M., & Hirsh-Pasek, K. (2010). Speaking out for language: Why language is central to reading development. *Educational Researcher*, 39(4), 305–310. <https://doi.org/10.3102/0013189X10370204>
- Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly*, 38(1), 78–103. <https://doi.org/10.1598/RRQ.38.1.4>
- Hsueh-Chao, M. H., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403–430. <https://nflrc.hawaii.edu/rfl>

- Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning*, 64(1), 160–212. <https://doi.org/10.1111/lang.12034>
- Kieffer, M. J., & Lesaux, N. K. (2012). Direct and indirect roles of vocabulary knowledge in English reading comprehension. *Reading and Writing*, 25(5), 1165–1194. <https://doi.org/10.1007/s11145-011-9318-7>
- Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1), 15–30. <https://nflrc.hawaii.edu/rfl/item/223>
- Lervåg, A., & Aukrust, V. G. (2010). Vocabulary knowledge is a critical determinant of reading comprehension growth. *Journal of Child Psychology and Psychiatry*, 51(5), 612–620. <https://doi.org/10.1111/j.1469-7610.2009.02185.x>
- Nassaji, H. (2003). Higher-level and lower-level text processing skills in advanced ESL reading comprehension. *The Modern Language Journal*, 87(2), 261–276. <https://doi.org/10.1111/1540-4781.00189>
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, 63(1), 59–82. <https://doi.org/10.3138/cmlr.63.1.59>
- 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. <https://doi.org/10.1037/0022-0663.98.3.554>
- Perfetti, C. A., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, 18(1), 22–37. <https://doi.org/10.1080/10888438.2013.827687>
- Proctor, C. P., Silverman, R. D., Harring, J. R., & Montecillo, C. (2012). The role of vocabulary depth in predicting reading comprehension among English language learners. *Reading and Writing*, 25(7), 1635–1664. <https://doi.org/10.1007/s11145-011-9336-5>
- Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, 52(3), 513–536. <https://doi.org/10.1111/1467-9922.00193>
- Ricketts, J., Nation, K., & Bishop, D. V. M. (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading*, 11(3), 235–257. <https://doi.org/10.1080/10888430701344306>
- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 64(4), 913–951. <https://doi.org/10.1111/lang.12077>
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, 95(1), 26–43. <https://doi.org/10.1111/j.1540-4781.2011.01146.x>
- Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, 36(2), 139–152. <https://doi.org/10.1080/09571730802389975>
- Tannenbaum, K. R., Torgesen, J. K., & Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension. *Scientific Studies of Reading*, 10(4), 381–398. https://doi.org/10.1207/s1532799xssr1004_3
- Tilstra, J., McMaster, K., Van den Broek, P., Kendeou, P., & Rapp, D. (2009). Simple but complex: Components of the Simple View of Reading across grade levels. *Journal of Research in Reading*, 32(4), 383–401. <https://doi.org/10.1111/j.1467-9817.2009.01401.x>
- Verhoeven, L., & Van Leeuwe, J. (2008). Prediction of the development of reading comprehension: A longitudinal study. *Applied Cognitive Psychology*, 22(3), 407–423. <https://doi.org/10.1002/acp.1414>

- Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics*, 22(2), 217–234. <https://doi.org/10.1017/S0142716401002041>
- Wise, J. C., Sevcik, R. A., Morris, R. D., Lovett, M. W., & Wolf, M. (2007). The relationship among receptive and expressive vocabulary and reading comprehension. *Journal of Speech, Language, and Hearing Research*, 50(4), 1093–1109. [https://doi.org/10.1044/1092-4388\(2007/076\)](https://doi.org/10.1044/1092-4388(2007/076))
- Zhang, D., & Annual, S. B. A. (2008). The role of vocabulary in reading comprehension. *RELC Journal*, 39(1), 51–76. <https://doi.org/10.1177/0033688208091140>