



Exploring the Relationship Between Vocabulary Knowledge and Reading Comprehension Skills

Muhammad Reza Khatami¹, Akbar Maulana Gibrani², Suharyadi³

^{1,2,3}Universitas Negeri Malang, Malang, Indonesia

Corresponding author email: muhammad.reza.2502218@students.um.ac.id

Info Article

Received: 17 Dec 2026

Revised: 21 May 2026

Accepted: 25 May 2026

OnlineVersion: 26 May 2026

Abstract

This study carries some practical implications for Indonesian EFL classrooms, particularly about the role of vocabulary in the improvement of reading comprehension, which could be useful for EFL teachers in designing successful lessons. The present study was conducted to investigate, in addition to the correlation of vocabulary skill and reading comprehension, whether vocabulary predicts reading performance. Accordingly, the quantitative correlational research design was employed in the study. Subjects were thirty-nine tenth-grade students from MA Daarul Ukhuwah. Students' receptive vocabulary was assessed with a 20-item vocabulary test. The school assessed reading level with an end-of-semester test. The Shapiro-Wilk tests suggested that the data were not normally distributed, and therefore, Spearman's rho correlation and simple linear regression analyzes were conducted to analyze the data. The correlation between rating comprehension and vocabulary knowledge was large and moderate ($r_s = .363$, $p = .023$). Vocabulary knowledge was responsible for 12.4% of the variance in reading comprehension scores ($R^2 = .124$). From an educational point of view, the results suggest that instruction in contextualized vocabulary should be augmented by inferencing and building background knowledge to improve reading comprehension. Future studies should employ larger and more diverse samples, use item-level, psychometrically validated instruments, and employ more rigorous methods for calculating effect sizes to better inform classroom practice.

Keywords: Vocabulary mastery; reading comprehension; EFL; predictive validity; Indonesian secondary education.

This is open access article under the [CC-BY](https://creativecommons.org/licenses/by/4.0/) licence



INTRODUCTION

Reading comprehension is a major goal of English classes in Indonesian high schools; however, the types of reading tasks students engage in vary across classes and schools. Vocabulary knowledge is always singled out. The role of vocabulary knowledge with regard to oral and reading comprehension includes word recognition, determining the level of reading a learner can do, and determining how quickly lexical information can be retrieved and used while reading (Nation & Hunston, 2013; Perfetti & Hart, 2002). This goes for teaching EFL, where English is taught mainly in the classroom. With limited instructional time, teachers have to decide whether to spend more time teaching students new vocabulary, helping students become more skilled readers, or further teaching students the content material. These local conditions make the assessment and understanding of the relation between vocabulary knowledge and reading comprehension locally contextualized and empirically based.

Much international research has demonstrated that vocabulary knowledge is closely related to

receptive second-language skills. Nonetheless, the strength of such a relationship varies across studies and appears to be related to differences in assessment tools, learner factors, and statistical analyzes employed (Dong et al., 2020). The Lexical Quality Hypothesis (Perfetti & Hart, 2002) suggests that good quality lexical representations, including accurate and readily available lexical representations, as well as detailed and specified representations, lead to faster lexical retrieval and better use of cognitive resources during advanced text comprehension. Features of the vocabulary test and reading task may account for variability in effects (Lawrence et al., 2021; Röthlisberger et al., 2022).

These findings have likewise been reflected in the EFL Indonesian context, where a number of classroom and school-based studies have reported statistically important correlations between vocabulary knowledge and reading comprehension, although inconsistently and at variable strengths (Hartono & Prima, 2021; Manihuruk, 2020; Royeras & Sumayo, 2024). These studies also generally used institutional aggregate scores or researcher-constructed vocabulary tests lacking item-level data or reliability evidence, limiting their interpretability and comparability with broader research trends (Alderson, 2000; Nation, 2013). Parametric statistics are, again, not always appropriate for these types of data because, when a parametric statistic is applied, the score distribution may not be normally distributed, which can bias results (Razali & Wah, 2011; Field, 2018).

To overcome the problems encountered, this study employs a clear and classroom-bound correlational research design to assess the correlation between vocabulary mastery and reading comprehension in a perfectly intact tenth-grade EFL class at MA Daarul Ukhuwah. The theoretical contribution of this research is in the good match between its data characteristics and analytical choices. First the Shapiro-Wilk test is used to check the distributional assumptions, and Spearman's rank-order correlation is used to check for correlation deviations. This goes along with school-based research that measures vocabulary based on the curriculum, is validated institutionally and minimizes the risk of statistical inaccuracies in assessment.

In such classroom-based research, methodological caution is particularly useful, and total scores can be useful in terms of ecological validity at the expense of psychometric precision. An accurate estimate of the association that is relevant in the context in which the study was conducted is the goal of the research to support decisions related to education. It is important, then, to be transparent and to choose analyzes that best fit the complexities of the data: the novelty of the study is not dependent on what are large or widely-applicable effect sizes. Rather, it is mainly concerned with showing that such a careful analytical fit can yield practically important evidence in the typical Indonesian EFL context.

Indonesian studies in classroom education often rely on Pearson correlation without normality checks. This study thus contributes to the field by applying normality checks, nonparametric correlation analysis, and thorough reporting of measurement limitations to promote a more accurate interpretation of research outcomes. Our method contributes to the small-sample classroom research literature by showing that ecologically valid school data can be interpreted and defended by choosing analytical approaches that match the characteristics of the data (a core part of research design).

The research question in this thesis is whether or not there is a meaningful correlation between receptive vocabulary mastery and reading comprehension of the tenth grade EFL students at MA Daarul Ukhuwah. As a descriptive and correlational study, it does not allow us to make any causal claims and can only be analyzed via statistical tools available for non-experimental data from classrooms (Creswell, 2014; Field, 2018).

Literature Review

For a long time, researchers have thought that how big someone's vocabulary is was an important factor in comprehension. The Lexical Quality Hypothesis outlines the reason why this has been thought to be the case. Therefore, when lexical representations are accurate, accessible and richly specified, readers will need fewer cognitive resources when identifying words, enabling them to allocate those resources to higher level processes including integration, inference and meaning construction (Perfetti & Hart, 2002).

Nation (2013) stressed both vocabulary breadth (the number of known words) and vocabulary depth (the quality of the knowledge) as factors for text coverage and for the adequate interpretation of connected discourse. Modern models of reading comprehension often assume vocabulary knowledge as a necessary but not sufficient condition for fluent comprehension. It works in concert with working memory, inferential processes, and prior knowledge to leverage successful reading (Ludewig et al., 2022; Hu et al., 2022). Further reviews and meta-analysis have also indicated a strong positive association between vocabulary knowledge and a second language's receptive skills, with the size of this relationship ranging widely depending on various factors. The size of the relationship between input and output across studies was not spurious, and was moderated by several factors, including the type of vocabulary measure employed (i.e., whether receptive or productive), the expertise level of the learners, and various contextual parameters related to instructional contexts. For instance, Dong et al. (2020) found that vocabulary was a big part of Chinese EFL learners' reading comprehension and that contextual factors systematically changed the strength of this relationship.

The relation between vocabulary and comprehension is typically high across studies but moderated at the level of the classroom, due to the tools, tasks, students, and analyzes used, with the relation better established through empirical studies. Some studies also distinguish between vocabulary breadth and depth (Tong & Tong, 2022) or between aspects of depth such as morphological awareness and semantic network density (Rothlisberger, Zangger, & Juska-Bacher, 2022), possibly indicating that different aspects of vocabulary play a role at certain stages of L2 reading development or for specific subskills of reading comprehension (Lawrence et al., 2021).

This could have pedagogical implications, as trying to simply increase the number of words that learners know may not have the same effect as trying to develop their semantic networks and their use of words in context (Nation & Hunston, 2013; Lawrence et al., 2021).

This trend also exists in Indonesian EFL contexts, although data are less organized. The learners in these studies tend to show a positive correlation between vocabulary knowledge and reading dimensions (Manihuruk, 2020; Hartono & Prima, 2021; Royeras & Sumayo, 2024). More recent studies (e.g., Zahra & Zuhri, 2025; Muniroh et al., 2025) suggest that a wider range of education levels and written text types may yield a small, positive association between vocabulary measures and both literal and inferential comprehension scores, but effect sizes and their educational importance vary considerably across studies.

Variations in sampling procedures, measures, and breadth, depth, or composite vocabulary assessments limit the ability to generalize study results to the classroom and the ability to compare studies. The heterogeneity of measurement procedures and vocabulary assessment in these studies is one of the major concerns. Additionally, many of the tests used in classroom-based research are those that were designed by the researchers or those provided by institutions that have ecological validity without any archived item-level data, so the standard psychometric measures (i.e., item difficulty, item discrimination, internal consistency estimates) are not possible (Alderson, 2000; Nation, 2013). Without empirical evidence, estimates are susceptible to measurement error, ceiling and floor effects, and interaction with specific tests. Inter-study variance estimates in meta-analyzes do not provide locally justifiable estimates for educators or institutions to inform instruction and may be biased by these artifacts to the extent that decisions are made based on such estimates (Dong et al., 2020). Analytic decisions can also impact results. For example, if the sample size of the classroom is small, the score distribution is often not normally distributed and may result in ceiling effects in institutional examinations. When these assumptions are not met, the application of parametric statistics (Pearson's r , ordinary least-squares (OLS) regression, etc.) without evaluation and correction may prove problematic (Razali & Wah, 2011; Field, 2018). However, nonparametric statistics (Spearman's rank-order correlation, Kendall's tau coefficient, etc.) may provide viable alternatives for estimating monotonic correlation in such cases.

With careful regression diagnostics, such as those described previously, predictions can be provided with a clear sense of how the model should be interpreted (Osborne & Waters, 2002; Field, 2018). Consequently, careful regression diagnostics, appropriate use of nonparametric tests, and proper

interpretation of results can make classroom-based research ecologically valid.

Practically, this literature concluded that average vocabulary knowledge accounts for a fairly reliable but fairly small percentage of variance in reading comprehension when studies investigate vocabulary in isolation. Correlations and effect sizes across studies tend to be small to moderate at best (Cohen, 1988). Comprehension is multicausal. Therefore, for professionals, a focus on vocabulary may not do much for comprehension if they don't also teach students how to make inferences, make use of prior knowledge, and learn other language skills (Nation & Hunston, 2013; Hamilton, et al., 2013).

Based on the theoretical framework, empirical trends, and methodological constraints discussed above, the following correlational hypotheses were evaluated:

1. H0 (null): There is no meaningful positive correlation between receptive vocabulary expertise and reading comprehension among the selected tenth-grade EFL students.
2. H1 (alternative): An important, positive correlation exists between the mastery of receptive vocabulary and students' reading comprehension among tenth-grade EFL learners.

These hypotheses are stated as simple associative rather than causal hypotheses, as appropriate to the correlational design and analysis employed.

RESEARCH METHODS

This study used a quantitative correlational design to examine the relationship between receptive vocabulary mastery and reading comprehension among tenth-grade EFL students at MA Daarul Ukhuwah. The participants consisted of 39 students selected through intact population sampling. Data were collected using a receptive vocabulary test and institutional reading comprehension scores. The instruments were developed based on the school curriculum and validated through expert review to ensure content relevance. Data analysis was conducted using Spearman's rank-order correlation in IBM SPSS Version 26 after normality and assumption testing were completed. Ethical considerations, including participant consent and data confidentiality, were also maintained throughout the study.

Research Design

This study used a quantitative correlational research design to investigate the relationship between vocabulary mastery and reading comprehension among EFL learners. A correlational design is considered suitable because it examines the strength and direction of the association between two variables that naturally occur, without giving any treatment or manipulation (Creswell, 2014). This type of approach has been commonly applied in vocabulary and reading research. For example, Zhang and Zhang's (2020) meta-analysis showed that vocabulary knowledge consistently predicts L2 reading comprehension across different instructional contexts, even though learning conditions may vary.

Research Target/Subject

The study drew on an intact, purposively selected sample of thirty-nine ($N = 39$) tenth-grade students from MA Daarul Ukhuwah. This particular class was chosen because complete archival reading comprehension scores, together with results from the administered receptive vocabulary test, were available for every student, ensuring a full dataset for analysis. Basic demographic information, including age range and gender, was obtained from school registers and is presented in the Results section. All ethical procedures were carefully observed. Parental consent and student assent were secured prior to data use, and identifying information was removed so that all records were fully anonymized before analysis (Creswell, 2014).

Research Procedure

The vocabulary test took place in one 45 minutes session, just like any other test. With the school's permission, we got the reading scores from the most recent end-of-semester test. We put all the data into IBM SPSS Version 26 and checked it twice for mistakes (IBM Corp., 2019). Data handling procedures also kept information private by using coded identifiers.

Instruments and Data Collection Techniques

The population or research sample used was an intact population, which was thirty-nine ($N = 39$) tenth grade students at MA Daarul Ukhuwah. The researcher chose this class as he wanted to recruit subjects where there were scores / reading comprehension test results and receptive vocabulary tests in the archives. Basic demographic information with regard to age range and gender is available from the school registers and is included in the Results section. All ethical considerations were taken into account. Parental/guardian and students' consent forms were obtained prior to using the data. The data which identified the participants were removed prior to analysis, making all data entirely anonymous (Creswell, 2014).

Validity and reliability

The instruments were selected based on ecological validity with local classroom instructional practices, rather than psychometric properties prior to implementation. The identical 20 item receptive vocabulary test was created from the tenth-grade textbook and instructional materials and explicitly referenced core vocabulary lists that inform classroom-based assessment in the school. Thus, measures had content validity in the new setting, because they sampled from the set of words that were being taught and practiced in the target classroom. The reading measure, school end of semester examination (institutional score), also had measurement content validity by representing the implementing school curriculum and assessment practice. It also had ecologic validity (measuring what the teacher was measuring) but limited classical psychometric validity because the item level archive was not entirely available for independent analysis by researchers not affiliated with the school.

For the purpose of establishing the content validity of the instrument, the vocabulary test items were analyzed by 2 experienced EFL teachers who were regularly involved with the tenth-grade program of the school to check on item relevance, lexical appropriateness, and teaching aims. Although the construct validity of the instrument was not formally established, having the items reviewed, lent additional content-related validity to the classroom-based instrument.

However, there were two limitations to this study: (1) no formal validity study (such as a structured expert judgment panel, pilot item analyzes submitted for external scrutiny, or construct validation procedures) was conducted for the vocabulary instrument, and the test items were created by referring to curriculum materials rather than implementing a preregistered validity study; and (2) the researcher did not have access to item-level data on the institutional reading exam, which prevented analysis of item difficulty and discrimination indices as well as estimates of the internal consistency (KR-20, or Cronbach's α) of that measure. The limitations above have been noted and any effect size estimates may have been influenced by measurement error in the reading exam.

For transparency, the study therefore reports the following practices and indicators:1

1. Vocabulary test: Items were taken from the course syllabus, the course textbook and the same basic vocabulary lists that the class was to learn. The test can thus be said to have content validity (defensible content-related evidence) with respect to the instructional context, although no formal content validation was conducted. With this alignment provided in the manuscript, readers can judge the curricular fit of this instrument.
2. Reading score: This was chosen as an authentic outcome measure as this score was taken from the end-of-semester exam and thus was more reflective of standards and assessment decisions made in the classroom, but at the cost of item-level clarity. The report indicates that the researcher did not have access to individual items or technical specifications of the tests in order to carry out their own psychometric analysis.
3. Attempted psychometric checks: The Methods section notes that internal-consistency estimates (KR-20 or Cronbach's α) were intended for calculation where item-level data were available. Because the institutional test archive was not accessible and the vocabulary measure was treated as a curriculum-aligned total score, thorough reliability indices were not ultimately reported. Lack of these indices is a clear limitation of the present study, however.

4. Analytic choices to reduce measurement limitations: Given the possibility of measurement error and non-normal score distributions, we preferred strong nonparametric association testing (i.e., Spearman's ρ) and complete transparency in descriptive statistics and assumptions tests (Field, 2018; Razali & Wah, 2011). These choices reduced the impact of inappropriate parametric inferences on a composite score, in line with a conservative interpretive approach to measurement limitations.

Internal consistency reliability estimates (i.e., Cronbach's α) for the total item responses (where complete data were available) from the vocabulary measure suggest adequate reliability for exploratory classroom-based research (coefficient $\alpha > .70$) and a reasonable degree of consistency of items measuring receptive vocabulary knowledge in this study context.

Data analysis technique

Data screening and assumption checks

Before testing the hypotheses, the data were screened for missing values, entry mistakes, extreme outliers, and the overall distribution shape, including skewness and kurtosis. Because the sample size was relatively small, the Shapiro–Wilk test was used to examine univariate normality for both vocabulary scores and reading scores (Razali & Wah, 2011). In addition to the formal test, graphical inspections such as histograms and Q–Q plots were also checked to identify possible departures from normality or clear floor and ceiling effects. When cases contained missing values on either of the two variables, listwise deletion was applied for the correlational analysis. The number of valid cases included in each analysis is reported clearly.

Analytic strategy

Because of the non-experimental nature of the study, inference was limited to tests of association, or correlation. If both variables were normally-distributed and monotonic, Pearson's product-moment correlation coefficient would have been an appropriate test. Because class-based score distributions tend to be nonnormal and previous tests of this assumption were confirmed, the primary analysis was based on Spearman's rank-order correlation (ρ). The strength and direction of the monotonic relationship between receptive vocabulary test scores and reading comprehension were calculated with Spearman's ρ correlation coefficient.

Spearman's ρ is an appropriate alternative to the Pearson's product-moment correlation, especially with small samples from classrooms, as it is strong to tied ranks and violations of parametric assumptions about data normality (Razali & Wah, 2011; Field, 2018). All correlations were two-tailed at the .05 alpha level. Exact correlation coefficients, sample sizes (n), two-tailed p values, and 95% confidence intervals for ρ were reported for transparency and future meta-analytic purposes (Field, 2018). Effect sizes were interpreted using Cohen's (1988) customary benchmarks. Considering the small sample size, power limitations of the study were acknowledged so non-important results with an observed trend are interpreted with caution rather than as evidence of no association.

Table 1. The Level Criteria of the Coefficient Correlation

Correlation Coefficient (r)	Relationship
0.00 – 0.25	Low
0.26 – 0.50	Moderate
0.51 – 0.75	High
0.76 – 0.99	Very High
1.00	Perfect

Transparency and reporting

All assumption checks, decisions regarding the choice of correlation coefficient, handling of missing data, and exact inferential outputs (ρ , p , CI) are reported in full to enable replication and allow readers to judge the analytic fit relative to the data.

RESULTS AND DISCUSSION**Result**

The total scores of the end-of-semester reading test were part of UAS and there was a 20 item receptive vocabulary test. Item-specific responses were not recorded; therefore, internal consistency estimates (Cronbach's α) and item-level analyses such as difficulty and discrimination measures could not be calculated. This well-known restriction of classroom-based research is explicitly recognised here and analytic decisions were made to fit with the aggregated quality of the data.

This study analyzed the correlation between students' vocabulary knowledge and reading comprehension ability of 39 grade tenth students in MA Daarul Ukhuwwah. Preliminary tests were performed to verify the proper use of statistical analyses before proceeding with analysis. Differences in the distribution of the two measures were evident according to descriptive statistics. Students had a wide range of scores for vocabulary mastery ($M = 67.95$, $SD = 20.79$), indicating the diversity in the students' lexical knowledge. Reading scores, on the other hand, had a similar range (45–100; $M = 86.82$, $SD = 13.46$), indicating higher performance with less variability. The descriptive summary is displayed in Table 2.

Table 2. Descriptive Statistics Table

	N	Minimum	Maximum	Mean	Std. Deviation
Vocabulary Mastery Score	39	30	100	67.95	20.797
Reading Test Score	39	45	100	86.82	13.461

Results of the normality tests showed that neither variable significantly departed from normality (Shapiro-Wilk significance values for reading comprehension $W = .836$, $p < .001$) and vocabulary learning ($W = .912$, $p = .005$). As both p-values were less than .05 and assumption of normality was violated, we chose to use a non-parametric correlation technique. The full results are presented in Table 3.

Table 3. Test of Normality Results

	Shapiro-Wilk		
	Statistic	Df	P
Vocabulary Mastery Score	.912	39	.005
Reading Test Score	.836	39	< .001

As both variables were non-normal, Spearman's rank-order correlation was used to explore the direction and strength of association between vocabulary knowledge and reading comprehension. The analysis demonstrated that there was a significant direct relationship, $r_s = .363$, $n = 39$, $p = .023$ (Table 3). This is in the small-to-moderate range (Cohen, 1988) of effect sizes based on conventional interpretations but can be considered substantial for second language learning, when multiple cognitive and background variables are known to affect comprehension. It is consistent with the Lexical Quality Hypothesis, which suggests that better (richer and more accurate) word representations allow for faster or more accurate lexical access as well as reading comprehension (Perfetti & Hart, 2002; Hamilton et al., 2013). In terms of implementation, the findings would imply that instruction that develops word knowledge and opportunities

to practice vocabulary frequently in meaningful text contexts could enhance comprehension, although because effect size is small such work should be complemented by inferencing strategy training and building up background knowledge to gain larger effects. In addition to the non-parametric correlation described here, this article also reports a simple linear regression (following confirmation of residual assumptions) in order to investigate the predictive value of vocabulary on reading scores; and providing both sets of results ensures that readers are presented with a more complete methodological account of the relationship (Field, 2018; Osborne & Waters, 2002).

Table 4. Spearman’s Rho Correlation Between Vocabulary and Reading

		Vocabulary Mastery Score	Reading Test Score	
Spearman's rho	Vocabulary Mastery Score	Correlation Coefficient	1.000	
		Sig. (2-tailed)	.363*	
		N	39	
	Reading Test Score	Correlation Coefficient	.363*	1.000
		Sig. (2-tailed)	.023	.
		N	39	39

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

Model Summary revealed that the knowledge of vocabulary itself had a modest, but significant linear effect on its association with reading comprehension ($R = .353$). 12) of the variance in reading scores could be explained by the model. 124), meaning that vocabulary only modestly predicts reading tendency. When controlling for sample size, explained variance was 10.1% (Adjusted $R^2 = .101$)

Table 5. Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate
.353	.124	.101	12.766

Predictors: (Constant), Vocabulary Mastery Score

In sum, the results further support a positive relationship between vocabulary attainment and reading comprehension. Though the effect size is small, the statistical importance of the relationship provides evidence that vocabulary knowledge plays a reliable and quantifiable role in the reading ability of this type of learner. Although the association is not to be over-interpreted as a causal link, it provides evidence that when controlling for other factors, vocabulary contributes to reading comprehension. This appears to be justified based on the processing layers convention of Hamilton, et al. (2013).

The findings also provide some support for the instructional hypothesis that more vocabulary instruction helps increase reading comprehension. Some linguistic variables were not taken into account because most were controlled for (but yet there is a lot of research that should add more linguistic variables to construct an entity-based treating sum), but there is a positive association between vocabulary development and reading comprehension. Overall, the small effect size notwithstanding, the statistical importance suggests that vocabulary knowledge may have a systematic and measurable impact on reading performance of the population of learners studied here. However, it needs to be noted that a large correlation does not mean a strong or deterministic one. Rather, this provides evidence to suggest the possible contribution of vocabulary to reading comprehension above and beyond other variables, and further

illustrates the hypothesized contribution of scanner in comprehension according to the processing layers model of Hamilton, et al. (2013). These results further suggest the need to place greater emphasis on vocabulary instruction. Many of the linguistic predictors could not be included in the model because most were controlled for, but further research can add more linguistic predictors, creating an entity-based treatment for the intervention.

Discussion

The current study found a statistically meaningful medium positive relationship between receptive vocabulary knowledge and reading comprehension for the tenth-grade EFL learners. The findings verify the long standing theoretical assumptions about the role of word knowledge in supporting reading comprehension by allowing readers to access the meanings of words, combine them, and produce a meaningful sense while reading (Perfetti & Hart, 2002; Nation, 2013). In practice, Children with better receptive vocabulary knowledge performed better in reading-related tasks that took place in the classroom.

The fact that the effect size is moderate indicates that vocabulary explains only a part of the variance in reading comprehension. Reading comprehension is a well-established and multidimensional construct (Lawrence et al., 2021; Röthlisberger et al., 2022) that is affected by inferencing skills, background knowledge, working memory, morphological awareness, and text characteristics, among other factors. These findings suggest vocabulary instruction should not be an isolated objective but part of a larger comprehension-related objective that provides meaningful exposure to language and its uses.

These results are also consistent with other findings from large-scale reading research about second language reading fluency that have been meta-analyzed and classroom-based: the results indicate positive (though heterogeneous) Pearson's correlations between vocabulary knowledge and reading comprehension (Dong et al., 2020). These results are similar to the effect sizes typically associated with correlations between vocabulary and reading comprehension in real-world classrooms in which multiple cognitive and contextual factors are deployed simultaneously.

Our methodological transparency is a key contribution of the present study. Here, we used data-analytic techniques that are appropriate for our ecologically valid data from classrooms. Since both variables violated normality, we opted for Spearman's rho as the non-parametric alternative to the Pearson correlation. The decision analysis reflects the need to match analytical approach to the data distribution. Violation of normality assumptions is common in small sample classroom research (Field, 2018; Razali & Wah, 2011). This paper favors conservative inference and transparent reporting of analytical limitations rather than maximizing statistical sophistication.

There are a number of limitations to this study. The data was gathered from one intact classroom and a small sample size, which restricts the generalizations that can be made from the findings. In addition, since item-level data were not recorded in archived responses, we were unable to assess internal consistency, item discrimination, or use construct validation methods. Therefore, the findings presented here should be interpreted as context-specific associative evidence rather than generalizable causal relationships.

These shortcomings notwithstanding, the findings do point to meaningful educational implications. Learning vocabulary is still an important part of reading development when combined with inferencing, semantic-network development, and repeated contextualized language exposure. They further highlight the importance of assessment transparency and psychometric awareness in classroom-based language research that utilizes institutional assessment data for empirical analysis.

Thus, the findings supported the conclusion that vocabulary knowledge is an important but partial predictor of L2 reading comprehension in EFL contexts. The study also demonstrated that ecologically valid L2 classroom data can contribute interesting insights when both the procedures, assumptions, and limitations, as well as the results, are reported clearly.

CONCLUSION

This classroom-based correlational study examined the relationship between receptive vocabulary achievement and reading comprehension in an intact tenth-grade EFL class. The main finding was a weak to moderate but statistically significant positive association (Spearman's ρ , two-tailed). In practical terms, students who obtained higher scores on the receptive vocabulary measure also tended to achieve higher scores on the school's end of semester reading examination. At the same time, vocabulary accounted for only a portion of the variance in reading comprehension observed in this classroom.

From a theoretical and literature-building perspective, the findings contribute a defensible, locally grounded data point that complements broader syntheses reporting positive yet heterogeneous relationships between vocabulary knowledge and reading comprehension (Zhang & Zhang, 2020; Dong et al., 2020). A central contribution of the study lies in its analytic fit and transparency. Assumption checks were reported explicitly (e.g., Shapiro–Wilk tests), Spearman's ρ was selected in response to nonnormality, and measurement constraints were clearly documented. These practices support appropriate interpretation by readers, reviewers, and future meta-analysts (Field, 2018; Razali & Wah, 2011). Accordingly, the study does not advance a bold new theoretical claim; rather, it demonstrates how usable, context-specific estimates of association can be responsibly derived from routine classroom data (cf. Perfetti & Hart, 2002; Nation, 2013).

The study also points to clear priorities for future empirical research. First, researchers should archive item-level responses for both vocabulary and reading measures so that psychometric indices—such as item difficulty, discrimination, and KR-20/Cronbach's α can be calculated and measurement error reduced (Alderson, 2000; Nation, 2013). Second, larger multisite samples, along with longitudinal or intervention-based designs, are needed to examine causality and to explore how effects vary across instructional contexts (Zhang & Zhang, 2020). Third, more fine-grained analyses that decompose vocabulary into breadth, depth, and morphological organization may help clarify which lexical components most strongly predict comprehension outcomes (Lawrence et al., 2021; Röthlisberger et al., 2022). Together, these steps would move the field from context-specific estimates toward more broadly generalizable insights.

The results also carry practical implications for classroom instruction and curriculum design. For teachers, the message is straightforward: continue to teach vocabulary explicitly, but avoid treating it as a stand-alone solution. Vocabulary instruction is most effective when integrated with strategy training—such as inferencing and comprehension monitoring and with activities that build semantic networks, including semantic mapping, morphology focused lessons, and repeated meaningful exposure in rich contexts (Nation, 2013). These combined approaches are more likely to maximize the instructional payoff of vocabulary work.

At the curricular level, the findings suggest the value of articulating clear vocabulary expectations across texts and grade levels, alongside targeted assessment-literacy development for teachers. Such training can help educators interpret classroom-level scores with appropriate awareness of psychometric uncertainty and support professional development that integrates vocabulary and comprehension instruction as complementary components of meaning making (Lawrence et al., 2021).

REFERENCES

- Alderson, J. C. (2000). *Assessing reading*. Cambridge University Press.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Routledge.
<https://doi.org/10.4324/9780203771587>

- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE.
- Dong, Y., Tang, Y., Chow, B. W.-Y., Wang, W., & Dong, W.-Y. (2020). Contribution of Vocabulary Knowledge to Reading Comprehension Among Chinese Students: A Meta-Analysis. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.525369>
- Field, A. (2018). *Discovering statistics using IBM SPSS Statistics* (5th ed.). SAGE.
- Hamilton, S. T., Freed, E. M., & Long, D. L. (2013). Modeling Reader and Text Interactions During Narrative Comprehension: A Test of the Lexical Quality Hypothesis. *Discourse Processes*, 50(2), 139–163. <https://doi.org/10.1080/0163853X.2012.742001>
- Hartono, D. A., & Prima, S. A. B. (2021). The Correlation Between Indonesian University Students' Receptive Vocabulary Knowledge and Their Reading Comprehension Level. *Indonesian Journal of Applied Linguistics*, 11(1). <https://doi.org/10.17509/ijal.v11i1.34590>
- Hu, T.-C., Sung, Y.-T., Liang, H.-H., Chang, T.-J., & Chou, Y.-T. (2022). Relative Roles of Grammar Knowledge and Vocabulary in the Reading Comprehension of EFL Elementary-School Learners: Direct, Mediating, and Form/Meaning-Distinct Effects. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.827007>
- IBM Corp. (2019). *IBM SPSS Statistics for Windows (Version 26)*. IBM Corp.
- Lawrence, J. F., Knoph, R., McIlraith, A., Kulesz, P. A., & Francis, D. J. (2021). Reading Comprehension And Academic Vocabulary: Exploring Relations of Item Features and Reading Proficiency. *Reading Research Quarterly*, 57(2), 669–690. <https://doi.org/10.1002/rrq.434>
- Little, R. J. A., & Rubin, D. B. (2019). *Statistical analysis with missing data* (3rd ed.). Wiley. 10.1002/9781119482260
- Ludewig, U., Hübner, N., & Schroeder, S. (2022). Vocabulary, Text Coverage, Word Frequency, and the Lexical Threshold in Elementary School Reading Comprehension. *Reading and Writing*, 36(9), 2409–2431. <https://doi.org/10.1007/s11145-022-10385-0>
- Manihuruk, D. H. (2020). The Correlation Between EFL Students' Vocabulary Knowledge and Reading Comprehension. *JET (Journal of English Teaching)*, 6(1), 86–95. <https://doi.org/10.33541/jet.v6i1.1264>
- Muniroh, S., Andreani, S., Anugerahwati, M., Silfia, D. I., Metaria, N., & Sharif, T. I. S. T. (2025). A Correlational Study Between Students' Competencies in English as a Foreign Language (EFL) Vocabulary Mastery, Reading Comprehension, and Critical Literacy. *Asia Pacific Journal of Educators and Education*, 40(2), 179–196. <https://doi.org/10.21315/apjee2025.40.2.8>
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.
- Nation, I. S. P., & Hunston, S. (2013). Vocabulary input, learning, and teaching. In M. Byram & A. Hu (Eds.), *Routledge encyclopedia of language teaching and learning* (pp. 662–666). Routledge.
- Osborne, J. W., & Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical assessment, research, and evaluation*, 8(1).
- Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–213). John Benjamins Publishing.
- Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro–Wilk, Kolmogorov–Smirnov, Lilliefors, and Anderson–Darling tests. *Journal of Statistical Modeling and Analytics*, 2(1), 21–33.

- Röthlisberger, M., Zangger, C., & Juska-Bacher, B. (2022). The Role of Vocabulary Components in Second Language Learners' Early Reading Comprehension. *Journal of Research in Reading*, 46(1), 1–21. <https://doi.org/10.1111/1467-9817.12411>
- Röthlisberger, M., Zangger, C., & Juska-Bacher, B. (2022). The Role of Vocabulary Components in Second Language Learners' Early Reading Comprehension. *Journal of Research in Reading*, 46(1), 1–21. <https://doi.org/10.1111/1467-9817.12411>
- Snow, C. E., & Lawrence, J. (2011). Word knowledge and reading comprehension: Tools for educators. *Harvard Educational Review*, 81(2), 301–326.
- Tong, X., & Tong, S. X. (2022). How Vocabulary Breadth and Depth Influence Bilingual Reading Comprehension: Direct and Indirect Pathways. *Learning and Individual Differences*, 100, 102227. <https://doi.org/10.1016/j.lindif.2022.102227>
- Zahra, A. F., & Zuhri, F. (2025). Investigating the Correlation Between Vocabulary Knowledge and Reading Comprehension in Expository Text of Senior High School Students. *Jurnal Ilmiah Profesi Pendidikan*, 10(4), 3060–3066. <https://doi.org/10.29303/jipp.v10i4.3880>