

An Investigation of the Relationship Between Vocabulary Mastery and Speaking Competence of Students at SMP Negeri 1 Pancarijang

Putri Chaerunnisa, Andi Sadapotto, Isumarni, Ibrahim Manda, Sam Hermansyah

¹²³⁴English Education, Universitas Muhammadiyah Sidenreng Rappang |

Corresponding author: putrichaerunnisa20@gmail.com | sadapotto.andi@yahoo.co.id | isumarni89@gmail.com | ibrahimmanda48@gmail.com | sam.hermansyah82@gmail.com

ARTICLE INFO

Keywords: vocabulary mastery, speaking competence, correlation, EFL, junior high school

©2026 Putri Chaerunnisa

: This is an open-access article distributed under the terms of the Creative Commons Attribution

[4.0 International License.](https://creativecommons.org/licenses/by/4.0/)



ABSTRACT

This study investigated the relationship between vocabulary mastery and speaking competence among students at SMP Negeri 1 Pancarijang. A quantitative correlational design was employed with 19 students of class VIII Java Script as the sample. Data were collected through a 30-item multiple-choice vocabulary test and an oral speaking test assessed in terms of fluency, pronunciation, vocabulary, comprehensibility, and accuracy. Descriptive statistics, the Shapiro-Wilk test, Pearson's product-moment correlation, Spearman's rank correlation, and a linear model summary were calculated using IBM SPSS Statistics. The students obtained a mean vocabulary score of 87.19 (SD = 10.50) and a mean speaking score of 70.95 (SD = 4.39). Pearson's analysis indicated a strong positive and statistically significant relationship ($r = .736$, $p < .001$). Because speaking scores did not meet the normality criterion, Spearman's rho was also computed and confirmed a positive significant association ($\rho = .597$, $p = .007$). The coefficient of determination was $R\text{-squared} = .542$, indicating substantial shared variation within the sample without establishing causality. The findings show that students with stronger vocabulary mastery tended to demonstrate better speaking competence. Teachers are therefore encouraged to integrate contextual vocabulary development with meaningful speaking activities while continuing to address pronunciation, grammar, fluency, confidence, and practice.

INTRODUCTION

English has a central role in education and international communication, and the ability to speak it allows learners to exchange ideas, negotiate meaning, and participate in authentic interaction. In Indonesian English as a Foreign Language (EFL) classrooms, instruction is therefore expected to develop not only knowledge of language forms but also the ability to use those forms appropriately in communication (Richards, 2006a). Speaking is particularly important because it makes learners' linguistic knowledge observable through oral performance. It requires them to select words, organize ideas, apply grammar, pronounce utterances intelligibly, and maintain fluency at the same time. Vazquez (2021) identifies fluency, pronunciation, vocabulary, grammar or accuracy, and comprehensibility as central aspects of speaking performance. These components work together: a learner may know a grammatical pattern, but communication can still break down when the words needed to express the intended meaning are unavailable.

Vocabulary mastery is consequently one of the principal linguistic resources supporting spoken communication. It includes knowledge of word meaning, recognition of words in spoken and written input, and the ability to use words appropriately in context (Aisyah, 2017). Vocabulary knowledge is commonly distinguished as receptive knowledge, involving words understood in listening or reading, and productive knowledge, involving words that learners can retrieve and use in speaking or writing (Maskor & Baharudin, 2016). Speaking depends especially on productive access because learners must retrieve suitable words under real-time pressure. Limited vocabulary can produce hesitation, repetition, avoidance, and incomplete expression, while a broader and more readily accessible vocabulary may help learners speak with greater fluency, precision, and confidence. Thus, vocabulary is not an isolated list of lexical items; it is a functional resource through which learners communicate meanings and intentions.

Previous research provides empirical support for this relationship. Alharthi (2019) reported an association between vocabulary knowledge and foreign-language speaking performance. Uchihara and Saito (2015) likewise found that productive vocabulary knowledge was related to second-language oral ability, particularly fluency. Koizumi and In'nami (2013) showed that vocabulary knowledge contributed to speaking proficiency from novice to intermediate levels, while Majeed and Hameed (2022) reported significant relationships between vocabulary breadth and depth and EFL speaking competence. Hartini and Suri Ardini (2024) also found a positive relationship between Indonesian EFL learners' vocabulary knowledge and English competencies. These studies suggest a consistent link, but relationships may differ across proficiency levels, learning settings, instruments, and learner groups. Evidence from Indonesian junior high schools remains valuable because learners at this level are still building foundational lexical resources and may have limited opportunities to use English outside the classroom.

At SMP Negeri 1 Pancarijang, students were observed to experience difficulty expressing opinions, participating in classroom discussion, and selecting appropriate English words. This context created a need to examine whether students who demonstrated stronger vocabulary mastery also tended to achieve better speaking performance. Accordingly, the present study investigated the relationship between vocabulary mastery as the independent variable and speaking competence as the dependent variable among eighth-grade students. It addressed the question: Is there a significant relationship between students' vocabulary mastery and their speaking competence at SMP Negeri 1 Pancarijang? The study was designed to provide context-

specific evidence and pedagogical insight for vocabulary-based speaking instruction. It did not seek to establish causality, but to determine the direction, strength, and statistical significance of the association between the two measured variables.

LITERATURE REVIEW

Vocabulary is a fundamental component of language proficiency because words carry the meanings learners need to understand input and produce output. Richards and Schmidt (2010) define vocabulary as the set of words known and used in a language, while Jones and Waller (2017) emphasize its influence on both comprehension and meaningful language production. Vocabulary mastery extends beyond recognizing a definition. It includes form, meaning, use, collocation, grammatical behavior, and the ability to select a word appropriate to a particular context. Receptive knowledge allows learners to recognize words, whereas productive knowledge enables active retrieval. Productive vocabulary is particularly relevant to speaking because oral interaction gives learners little time to search deliberately for lexical items. The more efficiently words can be accessed, the more attention speakers can allocate to organizing ideas, pronunciation, grammar, and interaction.

Speaking competence refers to the ability to convey ideas and meanings orally in a way that is appropriate and understandable. Speaking is a productive skill requiring simultaneous control of several processes (Brown, 2007). In this study, competence was operationalized through five observable criteria: fluency, pronunciation, vocabulary, comprehensibility, and accuracy. Fluency concerns continuity and the avoidance of excessive hesitation; pronunciation concerns clarity of sounds, stress, and intonation; vocabulary concerns appropriate lexical choice; comprehensibility concerns how readily a listener understands the message; and accuracy concerns the appropriate use of grammatical forms. These dimensions are interdependent. Vocabulary facilitates the construction of messages, but successful performance also depends on pronunciation, grammatical control, processing speed, confidence, and opportunities for practice. Speaking competence is therefore broader than vocabulary mastery even when vocabulary is a significant supporting resource.

The relationship between the variables can be explained through communicative competence. Linguistic resources support grammatical competence by enabling appropriate word selection and sentence construction. They contribute to discourse competence by helping speakers establish cohesion and coherence, and they support strategic competence when speakers paraphrase, negotiate meaning, or compensate for lexical limitations (Savignon, 1976). Insufficient vocabulary may cause communication breakdowns, reduce willingness to speak, and make it difficult for learners to maintain interaction (Meara, 2014). Conversely, learners with larger productive vocabularies are more likely to express meanings precisely and sustain speech. This theoretical position predicts a positive association: as vocabulary mastery rises, speaking performance also tends to rise. Nevertheless, the relationship is probabilistic rather than deterministic because learners with similar vocabulary scores may differ in pronunciation, grammar, anxiety, motivation, exposure to English, or

speaking practice.

Empirical studies have examined the association using different measures. Alharthi (2019) focused on productive vocabulary and foreign-language speaking performance, demonstrating that lexical knowledge is relevant to oral outcomes. Uchihara and Saito (2015) connected productive vocabulary with second-language oral ability and highlighted its relationship with fluency. Koizumi and In'nami (2013) examined learners from novice to intermediate levels and showed that the contribution of vocabulary knowledge may vary with proficiency. Majeed and Hameed (2022) considered both vocabulary breadth and depth, reinforcing the importance of knowing many words and knowing them well. The present study extends this line of inquiry to one Indonesian junior high school. Its conceptual framework positioned vocabulary mastery (X) as statistically associated with speaking competence (Y), while acknowledging other unmeasured influences. The alternative hypothesis proposed a significant relationship; the null hypothesis proposed no significant relationship.

METHOD

This study employed a quantitative approach with a correlational research design. A correlational design was appropriate because the objective was to measure the degree and direction of association between two naturally occurring variables without manipulating either one (Fraenkel et al., 2011). The research was conducted at SMP Negeri 1 Pancarijang. The population comprised all eighth-grade students distributed across five classes: Java Script, Python, Adobe Photoshop, Microsoft Excel, and Canva, with a total population of 135 students. The sample consisted of the 19 students in class VIII Java Script. The class was selected through purposive sampling because it met the study requirements and was accessible for the administration of both instruments. Since the sample came from one intact class, the findings are interpreted as specific to the participants and setting rather than as broad causal or population-wide claims.

Two instruments were used. Vocabulary mastery was measured through a 30-item multiple-choice test addressing word meaning, synonym and antonym recognition, and contextual usage. Each correct response received one point, and the total was converted to a percentage score. Speaking competence was assessed through an oral speaking task. Students' performances were scored with an analytic rubric covering fluency, pronunciation, vocabulary, comprehensibility, and accuracy. The speaking total was converted to a score on a 100-point scale. These measures generated a vocabulary score (X) and a speaking score (Y) for each participant. Data collection involved preparing the instruments, explaining the procedures, administering the vocabulary test, conducting the speaking assessment, scoring responses according to the answer key and rubric, and organizing the scores for analysis.

Data were analyzed with IBM SPSS Statistics using descriptive and inferential procedures. Minimum, maximum, mean, and standard deviation described the two score distributions. Because the sample contained fewer than 50 participants, the Shapiro-Wilk test was used to assess normality at the .05 level. Pearson's product-

moment correlation was the planned primary analysis for the linear relationship. The speaking scores did not satisfy the normality criterion, so Spearman's rank correlation was also calculated as a supporting nonparametric analysis. The hypotheses were tested at $\alpha = .05$: a two-tailed significance value below .05 led to rejection of the null hypothesis. The coefficient of determination, R-squared, was reported to indicate the proportion of score variation associated with the fitted linear relationship. Because the design was correlational, R-squared was not interpreted as evidence that vocabulary caused changes in speaking competence.

RESULT

Nineteen paired sets of vocabulary and speaking scores were obtained. Vocabulary scores ranged from 60.00 to 100.00, with a mean of 87.19 and a standard deviation of 10.50. These results indicate that performance on the vocabulary test was generally high, although students differed substantially across the observed range. Speaking competence scores ranged from 60.00 to 80.00, with a mean of 70.95 and a standard deviation of 4.39. The lower standard deviation and narrower range show that speaking scores were more closely clustered than vocabulary scores. Descriptively, the students therefore demonstrated stronger average performance on the vocabulary measure than on the speaking measure, while the two scales captured different amounts of variability.

The Shapiro-Wilk test produced a significance value of .054 for vocabulary mastery and .027 for speaking competence. Vocabulary scores did not significantly deviate from normality at the .05 threshold, whereas speaking scores fell below that threshold. This mixed result required care in selecting and interpreting the correlation procedure. Pearson's correlation was retained as the primary analysis because it reflected the original analytical plan and measured the linear association between the score variables. Spearman's rho was then used as a robustness check because it does not require normally distributed scores in the same way. Reporting both analyses allowed the consistency of the conclusion to be evaluated rather than relying on a single coefficient.

Pearson's product-moment analysis showed a correlation coefficient of $r = .736$ with $p < .001$ ($N = 19$). The positive sign indicates that students with higher vocabulary scores tended to receive higher speaking competence scores. The magnitude represents a strong positive linear relationship within the sample, and the probability value is below the predetermined .05 level. The null hypothesis was therefore rejected and the alternative hypothesis was accepted. In substantive terms, vocabulary mastery and speaking competence were not independent in the observed class; better lexical performance was associated with better oral performance.

The supporting Spearman analysis produced $\rho = .597$ with $p = .007$ ($N = 19$). Although the rank-order coefficient was smaller than the Pearson coefficient, it remained positive and statistically significant. Both procedures consequently led to the

same central conclusion. This convergence is important because one variable did not meet the normality criterion. It suggests that the detected relationship was not solely an artifact of using Pearson's parametric coefficient. At the same time, the difference between .736 and .597 shows that the estimated strength depends partly on the statistical representation of the scores and should not be overstated.

The fitted linear model yielded $R\text{-squared} = .542$, $F = 20.093$, $p < .001$. Approximately 54.2% of the observed variation in speaking scores was statistically associated with vocabulary mastery through this sample's linear relationship. The remaining 45.8% was not represented by that simple model and may relate to pronunciation, grammatical accuracy, confidence, speaking anxiety, practice, instructional exposure, motivation, or measurement error. The result does not mean that vocabulary causally determines 54.2% of speaking competence. Instead, it quantifies the shared variation observed in this correlational data set. Overall, the descriptive, Pearson, Spearman, and model-summary results consistently support a meaningful positive association.

DISCUSSION

The strong positive Pearson correlation supports the theoretical expectation that lexical knowledge facilitates oral communication. Students cannot formulate and deliver a message unless they can retrieve words that represent the intended ideas. A richer vocabulary provides alternatives for expressing meaning, reduces the need to abandon an utterance, and may allow learners to speak more continuously. This interpretation is consistent with Alharthi (2019), who found a relationship between vocabulary knowledge and foreign-language speaking performance, and with Uchihara and Saito (2015), who connected productive vocabulary knowledge with second-language oral ability. It also aligns with Koizumi and In'nami (2013), whose findings indicate that vocabulary knowledge is relevant to speaking proficiency across developing proficiency levels.

The mean vocabulary score of 87.19 was comparatively high, but the mean speaking score was 70.95 and the speaking distribution was narrower. This difference reinforces the distinction between knowing vocabulary and mobilizing it in real-time speech. A multiple-choice vocabulary test can measure recognition and contextual knowledge, whereas an oral test requires immediate retrieval along with pronunciation, fluency, grammatical control, and comprehensibility. Learners may recognize a word when options are presented yet hesitate to produce it spontaneously. For that reason, the positive relationship should not be interpreted as equivalence between the constructs. Vocabulary is an important foundation, but speaking performance represents the integrated operation of several competencies.

The significant Spearman result strengthens the interpretation because it confirms the positive association when scores are treated as ranks. Its moderate-to-strong coefficient also introduces an appropriate degree of caution. With only 19 participants, individual score patterns can influence the magnitude of a correlation. The sample was drawn from one class through purposive sampling, and the speaking scores occupied

a relatively restricted range. These characteristics limit generalization. The study nevertheless provides useful local evidence: within class VIII Java Script, learners who achieved higher lexical scores generally performed better in speaking. Similar conclusions reported by Majeed and Hameed (2022) and Hartini and Suri Ardini (2024) place this local pattern within a broader body of research linking vocabulary knowledge with English performance.

Pedagogically, the findings support instruction that connects vocabulary learning directly to meaningful speaking. Teachers can introduce lexical items in context and then require students to retrieve and use them through role-plays, short presentations, information-gap tasks, discussions, and guided conversations. Activities should address pronunciation, collocation, grammatical patterns, and appropriateness rather than memorized definitions alone. Repeated retrieval and recycling can help transfer words from receptive recognition to productive use. Assessment should likewise combine vocabulary checks with oral tasks so that teachers can identify whether learners understand words and whether they can deploy them effectively. Such an approach acknowledges vocabulary's importance without neglecting fluency, accuracy, comprehensibility, confidence, and interactional practice.

CONCLUSION

This study investigated the relationship between vocabulary mastery and speaking competence among 19 eighth-grade students of SMP Negeri 1 Pancarijang. Vocabulary scores averaged 87.19, while speaking competence scores averaged 70.95. Pearson's analysis revealed a strong, positive, and statistically significant relationship ($r = .736$, $p < .001$). Spearman's supporting analysis reached the same conclusion ($\rho = .597$, $p = .007$), and the fitted linear relationship produced $R\text{-squared} = .542$. The null hypothesis was rejected. Thus, students with stronger vocabulary mastery tended to demonstrate stronger speaking competence in the observed class.

The findings indicate that vocabulary mastery is an important component of oral English development, but they do not establish a causal effect. Speaking also depends on pronunciation, grammar, fluency, comprehensibility, confidence, anxiety, practice, and exposure. The small purposively selected sample, single-school context, restricted range of speaking scores, and use of one set of instruments limit generalization. These limitations should be considered when applying the findings to other classes or educational settings.

Students are encouraged to expand and actively use vocabulary through regular speaking practice. English teachers should integrate contextual vocabulary instruction with interactive oral activities and provide feedback on word choice as well as other dimensions of speaking. Future researchers should use larger and more diverse samples, examine productive vocabulary separately from receptive vocabulary, and consider additional predictors such as motivation, confidence, anxiety, and exposure to English. Longitudinal or intervention studies may also clarify how vocabulary development and speaking performance change over time.

References

- Aisyah, D. N. (2017). The vocabulary teaching and vocabulary learning: Perception, strategies, and influences on students' vocabulary mastery. Unpublished doctoral dissertation, 9(2), 293-318.

- Alharthi, T. (2019). Investigating the relationship between vocabulary knowledge and FL speaking performance. *International Journal of English Linguistics*, 10(1), 37. <https://doi.org/10.5539/ijel.v10n1p37>
- Askhatova, A. (2020). Importance of vocabulary and spoken grammar in teaching spontaneous speaking skill. *Sciences of Europe*, 49, 3-6.
- Brown, H. D. (2007). *Teaching by principles: An interactive approach to language pedagogy*. Pearson Education.
- Burkxonovna Sharipova, S., & Ergashevich Kodirov, B. (2021). The role of vocabulary in developing communicative competence. *Scientific Journal Impact Factor*, 2(1), 2181-1385.
- Diamond, L., & Gutlohn, L. (2006). Teaching vocabulary. *Reading Rockets*. <http://www.readingrockets.org/article/9943>
- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2011). *How to design and evaluate research in education*. McGraw-Hill.
- Gardner, D. (2024). *Exploring vocabulary: Language in action* (Vol. 2). Routledge.
- Hartini, L. W., & Suri Ardini, A. (2024). The relationship between Indonesian EFL learners' vocabulary knowledge and English competencies. *Journal of English and Education*, 10(1), 55-66. <https://doi.org/10.20885/jee.v10i1.33651>
- Jones, C., & Waller, D. (2017). The effect of input enhancement on vocabulary learning: Is there an impact upon receptive and productive knowledge? *TESOL International Journal*, 12(1), 48-62.
- Koizumi, R., & In'nami, Y. (2013). Vocabulary knowledge and speaking proficiency among second language learners from novice to intermediate levels. *Journal of Language Teaching and Research*, 4(5), 900-913. <https://doi.org/10.4304/jltr.4.5.900-913>
- M U, Sari H, Hermansyah S, Maming K, Kahar A, Hasan, Elfahmi FK (2025), "Understanding Indonesian students' reading knowledge in digital literacy within socio-cultural of rural middle schools". *International Journal of Information and Learning Technology*, Vol. 42 No. 5 pp. 432-448, doi: <https://doi.org/10.1108/IJILT-12-2023-0239>
- Maskor, Z. M., & Baharudin, H. (2016). Receptive vocabulary knowledge or productive vocabulary knowledge in writing skill, which one is important? *International Journal of Academic Research in Business and Social Sciences*, 6(11), 261-271. <https://doi.org/10.6007/ijarbss/v6-i11/2395>
- Meara, P. M. (2014). Vocabulary research in *The Modern Language Journal: A bibliometric analysis*. *Vocabulary Learning and Instruction*, 3(1), 1-28. <https://doi.org/10.7820/vli.v03.1.meara>
- Majeed, N. T., & Hameed, Y. Q. (2022). The correlation between EFL learners' vocabulary mastery and speaking competence. *Journal of Tikrit University for Humanities*, 27, 119-141.
- Rao, P. S. (2019). The role of English as a global language. *Research Journal of English*,

4(1), 65-79.

Richards, J. C. (2006a). Communicative language teaching today. Cambridge University Press.

Richards, J. C., & Schmidt, R. (2010). Longman dictionary of language teaching and applied linguistics. Pearson Education.

Sam Hermansyah, Usman M., Syamsu T. (2026). Developing Digital Learning Media for Teaching Public Speaking to Improve English Skills of Junior High School Students. *International Journal of Special Education*, 41(16s), 502–509. Retrieved from <https://www.internationalsped.com/index.php/ijse/article/view/5027>

Savignon, S. J. (1976). Communicative competence: Theory and classroom practice. *System*, 13(2), 167-168. [https://doi.org/10.1016/0346-251X\(85\)90026-0](https://doi.org/10.1016/0346-251X(85)90026-0)

Stratton, S. J. (2021). Population research: Convenience sampling strategies. *Prehospital and Disaster Medicine*, 36(4), 373-374. <https://doi.org/10.1017/S1049023X21000649>

Uchihara, T., & Saito, K. (2015). Exploring the relationship between productive vocabulary knowledge and second language oral ability. *Language Learning*, 65(1), 123-151.

Vazquez, A. (2021). Testing for language teachers. *English Australia Journal*, 37(2), 89-91. <https://doi.org/10.60161/2501-001-002-005>

Wilkins, D. A. (1972). Linguistics in language teaching. Edward Arnold.

Brown, D., & Lord, R. G. (2004). Leadership processes and follower self-identity. Lawrence Erlbaum Associates.

Davies, A. (2014). An introduction to applied linguistics. Edinburgh University Press.

Dudeney, G., & Hockly, N. (1990). How to teach English with technology. Pearson Longman.

Hu, C.-P., & Chang, Y.-Y. (2017). Variables and measurement in social research. *Journal of Social and Administrative Sciences*, 4, 3-5.

Russo, R. (2003). Statistics for the behavioural sciences. Psychology Press.