IDEAS
Journal of Language Teaching and Learning, Linguistics and Literature

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Volume 13, Number 1, June 2025
pp. 2154 - 2164

Issued by English study program of IAIN Palopo

AI Chatbot to Improve Students' Vocabulary Mastery

Irmawati[1], Dzakiah[2], Fatima[3]
[1,2,3] Tadris Bahasa Inggris, Universitas Islam Negeri Datokarama Palu
Corresponding E-Mail: iwati4210@gmail.com

Received: 2025-05-31 Accepted: 2025-06-29
DOI: 10.24256/ideas.v13i1.6855

Abstract
Artificial Intelligence (AI) Chatbots are computer programs designed to simulate human conversation and offer interactive, personalized learning experiences. This quantitative study investigated the effectiveness of an AI Chatbot (implemented using the website platform https://www.aichatting.net) in improving students' vocabulary mastery at SMP Negeri 01 Dapurang Pasangkayu, Indonesia. The study employed a quasi-experimental design, comparing an experimental class (8B, n = 25) that utilized the AI Chatbot with a control class (8B, n = 25) that employed a conventional teaching method. The intervention was conducted over six sessions across three weeks, from April 26 to May 19, 2025. Students completed pre-tests and post-tests to assess vocabulary mastery. The experimental class mean score increased from 31.60 to 76.00 (gain = 44.40; Cohen's d = 4.71), while the control class improved from 32.00 to 46.00 (gain = 14.00; Cohen's d = 1.43). An independent samples t-test on post-test scores revealed a statistically significant difference between the classes, t (48) = 12.42, p < .0001, with a 95% confidence interval of \ [25.27, 34.73]. These results confirm that the AI chatbot had a substantial impact on vocabulary mastery in practice. While the study supports the potential of AI chatbots as effective vocabulary-learning tools in junior high school settings, it also acknowledges the limitations of the quasi-experimental design and recommends further research using randomized controlled trials.

Keywords: *AI Chatbot; Artificial Intelligence; Vocabulary Mastery*

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Introduction

Technology has advanced rapidly and significantly reshaped education, enabling more interactive, adaptive, and personalized learning experiences. In the 21st century, digital tools are pivotal in transforming conventional classrooms into dynamic learning environments. This shift urges educators to explore how such technologies, particularly Artificial Intelligence (AI), can enhance English learning (Fu, 2019).

As a global lingua franca, English is a major means of communication, trade, and scholarship. In an interconnected world, it is imperative for students to be fluent in the English language, and lack of proficiency can have serious consequences for their academic and professional success. Vocabulary acquisition is one of the most critical facets of second language acquisition, as it is the building block upon which successful communication and understanding lies (Batia Laufer, 2012). Learners with a rich vocabulary can share their ideas appropriately and comprehend challenging texts, improving their general language proficiency.

Vocabulary acquisition is one of the most fundamental steps in learning a new language; nonetheless, many students struggle to learn new words. However, traditional approaches to vocabulary teaching do not adequately capture learners' interests or develop word retention. Hence, innovative solutions would be needed to encourage vocabulary development in learners. A promising solution lies in leveraging Artificial Intelligence (AI) in academia, especially in the form of AI Chatbot. Such intelligent systems are expected to be able to provide adaptive learning experience for all the learners, immediate feedback, personalization that suit each individual's needs, creating more engaging and effective learning opportunities (Kukulska-Hulme & Lee, 2020).

AI Chatbot, defined as computer-based programs designed to simulate human conversation, have emerged as promising educational tools. Based on principles of constructivism and social learning theory, chatbots facilitate interactive learner-centered environments where students can engage in personalized, scaffolded language practice. Constructivism emphasizes that knowledge is actively constructed by learners through meaningful experiences, while social learning theory highlights the importance of interaction and feedback in the learning process.

Multiple studies have assessed the efficacy of AI Chatbots in improving vocabulary and language skills. A research about the effectiveness of AI Chatbots to improve students' vocabulary has been carried out by Hutauruk et al. which showed that by offering constant practice and immediate feedback, AI Chatbots can improve vocabulary learning significantly, leading students to internalize new words and phrases more readily (Hutauruk et al., 2017). A similar study investigating students' perceptions of how AI Chatbots could impact the vocabulary and grammar of EFL students' writing, and found that the students respected the potential for AI Chatbots to significantly increase the retention of vocabulary and

to improve writing in this way (Waziana et al., 2024).

Furthermore, Hawanti & Zubaydulloevna focused their research on AI Chatbots in relieving the anxiety of students in the English writing classroom. the research revealed the use of AI Chatbots enabled higher levels of participating in language tasks, with data indicating that students on tools all had higher mastery in vocabulary as well as lower anxiety levels (Hawanti & Zubaydulloevna, 2023). These findings are consistent with those have noticed that AI Chatbots were found to have a positive influence on students' self-efficacy, writing skills, and motivation, resulting in students feeling more confident using new vocabulary and grammar in their writing (Apriani et al., 2024).

Previous studies have demonstrated the effectiveness of chatbots in enhancing language skills, particularly in higher education or skill-specific contexts such as writing. However, research on their application for vocabulary mastery among junior high school students remains limited, especially in real classroom environments in Indonesia. Moreover, few studies have offered quantitative evidence based on robust experimental design.

To address this gap, the present study investigates the use of an AI Chatbot in improving students' vocabulary mastery in an Indonesian junior high school setting. Specifically, the Chatbot was implemented using the website platform (https://www.aichatting.net). This study defines "vocabulary mastery" operationally as a student's ability to understand, recall, and appropriately use vocabulary items introduced during instruction.

The research question of this study *is the use of AI Chatbot effective in improving students' vocabulary mastery?* In accordance to the research question, the objective of this study is *to find out the effectiveness of AI Chatbot in improving students' vocabulary mastery.*

In order to achieve this objective, the following hypotheses are formulated to be tested:

H0 (null): There is no significant improvement in vocabulary mastery among students using the AI Chatbot compared to those receiving conventional teaching methods.

Ha (alternative): There is a significant improvement in vocabulary mastery among students using the AI Chatbot compared to those receiving conventional teaching methods.

This research contributes to the growing body of knowledge on AI-supported language learning, offering practical implications for language educators, curriculum designers, and policymakers seeking to enhance English instruction through digital innovation.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Method

This study adopted a quantitative approach, which enables a standardized measurement and analysis of the effect of AI Chatbots on students' vocabulary skills. This approach enables an accurate and trajectory-based evaluation of the intervention's impact (Cohen et al., 2007). This study employed a quasi-experimental design, comparing two groups: the control class (Class 8A) and the experimental class (Class 8B). Since participants are not randomly assigned in this study, the design allows a controlled assessment of the effect of the AI Chatbot on vocabulary acquisition. The students in the experimental class used the AI Chatbot to learn, whereas the control class used the conventional teaching method (Telling/Lecture Method). This study took place over three weeks, from April 26 to May 19, 2025, and included six learning sessions.

The study population consists of 129 students from SMP Negeri 01 Dapurang Pasangkayu, specifically 20 students in class 7A, 20 students in class 7B, 25 students in class 8A, 25 students in class 8B, 19 students in class 9A, and 20 students in class 9B. Using purposive sampling, class 8B (n = 25) was designated as the experimental class and class 8A (n = 25) as the control class. These classes were selected by the researcher based on school schedule compatibility and teacher availability.

The instruments applied in this research were a pre-test and a post-test to measure the students' vocabulary mastery. The students were assessed on their vocabulary knowledge before the treatment (pre-test) and after the use of the AI Chatbot (post-test). To measure students' vocabulary mastery, the researcher developed a vocabulary test composed of 20 items: 10 multiple-choice and 10 matching-type questions. The test was content-validated by a panel of English language educators. Reliability analysis yielded a Cronbach's alpha coefficient of 0.81, indicating high internal consistency.

Both classes took the same pre-test to measure baseline vocabulary knowledge. The experimental class was taught using the AI chatbot hosted on (https://www.aichatting.net). This Chatbot offered interactive, adaptive learning sessions that included vocabulary drills, word usage examples, and immediate feedback. Sessions were conducted in the school's computer lab with supervision. Meanwhile, the control class received traditional lecture-based vocabulary instruction using printed materials.

Each session for the experimental class lasted approximately 40 minutes, with teachers facilitating technical access and monitoring student interaction with the Chatbot. The control class completed equivalent-duration sessions using conventional teaching methods, including direct explanation and paper-based vocabulary exercises.

Various techniques were used to analyze the data collected. Descriptive statistics (mean, standard deviation) were used to summarize pre-test and post-test scores. A paired samples t-test was used to assess within-group improvement.

An independent samples t-test was conducted to compare post-test scores between the two groups. The effect size was calculated using Cohen's d to determine practical significance. Assumptions of normality were verified using the Shapiro-Wilk test. Statistical analysis was performed using SPSS (Statistical Package for the Social Sciences), a widely used quantitative data analysis software (Pallant, 2011).

The study obtained school approval and informed consent from all student participants. No personally identifiable information was collected, and participants were assured that their performance data would be used solely for research purposes. Efforts were made to minimize confounding variables, including ensuring both classes had similar average English proficiency based on school records. The same teacher supervised both groups to eliminate instructor-related bias. The school's infrastructure ensured equal access to internet-connected devices. A post hoc power analysis indicated that the sample size (n=25 per group) was adequate to detect a large effect size (Cohen's $d > 0.8$) with power > 0.95 at $\alpha = 0.05$.

Results

This study employed three phases consisting of a pre-test, a treatment, and a post-test. For the first phase, the pre-test was distributed to all students in both control (8A) and experimental (8B) classes, which received an identical vocabulary pre-test. The test comprised 20 numbers with 10 multiple-choice and 10 matching items. It was carefully designed to assess students' foundational vocabulary knowledge before the intervention began.

The treatment was administered to the experimental class (8B), which participated in a specially designed learning program that utilized an AI Chatbot. This program offered an interactive and personalized learning experience tailored to individual needs. Meanwhile, the control class (8A) continued with the standard teaching methods typically used in the school. The implementation of both learning approaches proceeded smoothly and as planned.

After successfully completing the learning program, both groups were given the same vocabulary test. The post-test served to measure the improvement in vocabulary acquisition after the respective learning experiences. The administration of the post-test went smoothly and produced high-quality data suitable for robust analysis.

This section presents the findings based on descriptive statistics, paired samples t-test, independent samples t-test, and effect size (Cohen's d) to evaluate the effectiveness of the Al Chatbot in improving students' vocabulary mastery.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Descriptive Statistics of Pre-Test and Post-Test

Class	Test	N	Minimum	Maximum	Mean	Std. Deviation
Experimental	**Pre-test**	**25**	**20**	**50**	**31.60**	**9.43**
Experimental	**Post-test**	**25**	**55**	**85**	**76.00**	**6.77**
Control	**Pre-test**	**25**	**20**	**50**	**32.00**	**9.79**
Control	**Post-test**	**25**	**30**	**65**	**46.00**	**10.00**

The descriptive statistics indicate that both classes began with nearly identical pre-test mean scores (31.60 vs 32.00). However, the post-test scores show a substantial difference: the experimental class achieved a mean of 76.00, while the control class only reached 46.00.

Paired Samples t-Test (Experimental Class)

Mean Difference	T	Df	p-value
44.40	**-21.68**	**24**	**< .0001**

A paired samples t-test for the experimental class confirmed that the increase in vocabulary mastery was statistically significant (p<.0001). This suggests that the Chatbot intervention was highly effective.

Independent Samples t-Test (Post-test Comparison)

Mean Difference	t(48)	p-value	95% CI Lower	95% CI Upper
30.00	**12.42**	**< .0001**	**25.27**	**34.73**

An independent samples t-test comparing post-test scores between classes revealed a statistically significant difference in favor of the experimental class, t(48) 12.42, p<.0001.

Effect Size (Cohen's d)

Class	Cohen's d	Interpretation
Control	**1.43**	**Large**
Experimental	**4.71**	**Very Large**

Cohen's d for the experimental class was 4.71, indicating a huge effect size, while the control class also showed a significant but much smaller effect size (1.43). These results underscore the strong practical significance of the Al chatbot

intervention. Taken together, these findings support the conclusion that the use of an Al chatbot significantly and substantially enhances vocabulary mastery among junior high school students.

Discussion

The results of this study provide compelling evidence for the effectiveness of the AI chatbot in improving students' vocabulary mastery. The experimental group not only showed a statistically significant increase in post-test scores ($t = -21.68$, $p < .0001$), but the effect size was also very large (Cohen's $d = 4.71$), indicating substantial practical significance. The independent samples t-test comparing post-test scores between groups further confirmed the superiority of the AI chatbot intervention, with a mean difference of 30.00 and 95% CI \[25.27, 34.73].

This improvement can be attributed to several key features of the AI chatbot platform (https://www.aichatting.net), including: 1) Immediate Feedback: Students received real-time correction and clarification, which is aligned with the principle of reinforcement in social learning theory; 2) Personalized Learning: Adaptive responses allowed learners to proceed at their own pace, reducing frustration and enhancing engagement; and 3) Gamification Elements: The Chatbot incorporated interactive tasks and mini-challenges, increasing student motivation and retention.

This finding aligns with previous research highlighting the positive impact of AI chatbots on language learning (Hutauruk et al., 2022; Waziana et al., 2021; Hawanti & Zubaydulloevna, 2021; Apriana et al., 2022). However, this study provides specific quantitative evidence within an Indonesian junior high school context, a gap previously identified in the literature.

The findings of this study are consistent with recent Indonesian research demonstrating the effectiveness of AI-powered Chatbots in enhancing students' vocabulary mastery and overall English language skills. For example, a systematic review by Syaifulloh et al. (2025) concluded that AI Chatbots serve as effective cognitive tools in EFL learning by improving vocabulary retention, grammar accuracy, and writing fluency. This study also highlighted positive psychological effects such as increased learner confidence and reduced language anxiety, which align with the self-efficacy improvements observed in our research.

Moreover, the MicroLingo Chatbot by Wijaya et al. (2025) employed a microlearning approach via WhatsApp to help young Indonesian entrepreneurs improve English vocabulary and communication skills. The study emphasized the flexibility and ease of access of chatbot-based microlearning, which contributed to sustained vocabulary practice and learner motivation.

Contrary to concerns about cognitive overload from excessive Chatbot interaction, studies such as Olugbade et al. (2024) suggest that careful Chatbot design and pre-implementation training mitigate such risks, ensuring that Chatbots complement rather than hinder language learning processes. This aligns

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

with our findings where no adverse effects related to cognitive fatigue were detected.

Additional studies by Kingdom & Qammar, (2024) confirm that AI Chatbots positively influence vocabulary mastery by providing immediate feedback and personalized learning paths, thus fostering self-regulated learning behaviors critical for language acquisition.

Despite these promising results, limitations remain. Lee at al. (2020) caution that excessive Chatbot use without adequate pedagogical guidance may reduce learner autonomy and lead to superficial vocabulary learning, highlighting the necessity of integrating Chatbots with traditional instructional methods to ensure depth and sustainability of learning. From a theoretical perspective, this study aligns with constructivist learning theory, where learners build knowledge actively through interaction, and social learning theory, where feedback and observation enhance learning. The Chatbot functioned as a digital scaffolding agent, helping students move from recognition to productive use of vocabulary.

However, alternative explanations such as the novelty effect (students are more motivated by new tools) or the Hawthorne effect (improvement due to increased attention) cannot be ruled out entirely. Nevertheless, the magnitude of improvement and consistency with previous research suggest that the Chatbot's pedagogical design played a more crucial role than temporary novelty.

The findings of this study suggest that Chatbot can serve as effective supplementary tools in educational settings, particularly in large classrooms where providing individual feedback is challenging. Chatbot offer immediate, personalized responses that can enhance student engagement and support autonomous learning. Furthermore, schools located in rural or underserved areas may particularly benefit from Chatbot-assisted learning due to the technology's low cost and accessibility. Integrating Chatbot into blended learning models allows educators to complement traditional teaching method with adaptive, technology-driven support, thereby enriching the overall learning experience.

This study has several limitations. The quasi-experimental design didn't involve random participant, which restricts the ability to draw causal inferences from the findings. Additionally, John Creswell et al., (2017) discusses potential biases inherent in quasi-experimental designs, such as that employed in this study, suggesting that future research should utilize randomized controlled trials to strengthen validity.

Similar to limitations discussed by Santoso (2014) regarding quasi-experimental designs, these studies also note potential biases and call for more rigorous randomized controlled trials to strengthen evidence on Chatbot efficacy in language learning. Moreover, the absence of a delayed post-test limits the study's capacity to evaluate long-term vocabulary retention. The intervention was also conducted in a single school, which may affect the generalizability of the results to broader educational contexts. These limitations highlight the need for

caution in interpreting the outcomes and suggest directions for more rigorous future research.

To enhance the robustness of findings, future studies should consider employing randomized controlled trials (RCTs) to strengthen internal validity. Longitudinal research is also encouraged to investigate long-term vocabulary retention, including the use of delayed post-tests. Gathering qualitative feedback from both students and teachers could offer deeper insights into the usability and effectiveness of Chatbot-based instruction. Additionally, future studies should explore the cost-effectiveness and scalability of Chatbot implementation across diverse educational settings and examine its cross-cultural applicability in varying learning environments.

In summary, the current body of evidence indicates that AI Chatbots are effective tools for improving students' vocabulary mastery by providing interactive, personalized, and scaffolded learning experiences that enhance both receptive and productive vocabulary knowledge while maintaining manageable cognitive load and fostering learner engagement.

Conclusion

This study provides strong empirical and practical evidence that AI chatbots, particularly those implemented via web-based platforms like (https://www.aichatting.net), can significantly enhance junior high school students' vocabulary mastery. The experimental group showed a substantial improvement (mean gain = 44.40, Cohen's d = 4.71) compared to the control group (mean gain = 14.00, Cohen's d = 1.43), and an independent samples t-test confirmed a statistically significant difference in post-test scores between the groups ($t(48) = 12.42$, $p < .0001$).

These results highlight the chatbot's potential to offer an interactive, personalized, and feedback-rich learning experience, addressing the shortcomings of traditional vocabulary teaching methods. The intervention not only yielded statistically significant outcomes but also demonstrated large practical effects, aligning with constructivist and social learning theories.Despite the promising results, the study is not without limitations. The use of a quasi-experimental design without random assignment restricts the generalizability of the findings.

Additionally, the absence of a delayed post-test means the long-term effects of the chatbot intervention remain unknown. Future research should adopt randomized controlled trials, expand to diverse educational settings, include retention assessments, and consider the integration of qualitative data to capture student experiences more holistically. Nevertheless, this study contributes valuable insights into the integration of AI Chatbot in junior high school English instruction.

With careful design and implementation, AI-powered tools can supplement traditional instruction, support student-centered learning, and promote vocabulary development effectively and equitably across different learning contexts.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Acknowledgement

First and foremost, I, Irmawati, would like to express my sincere gratitude to myself for the dedication, perseverance, and hard work invested in this research. My mother (Musniati) and my sisters (Nurmina and Sriwati) deserve special recognition for their unwavering support and encouragement throughout this journey. Deep appreciation extends to my family for their patience and understanding. I am also deeply thankful to my supervising lecturers, Dzakiah, M.Pd. and Fatima, M.Pd., for their invaluable guidance and mentorship. My heartfelt thanks also go to Ummy, Sida, Liana, Sappe, Raya, Zain, Yasin, Dini, Lina and Karin for their prayers, support, and assistance. Finally, I thanks to Nanda for her support and companionship throughout this research. Their contributions were invaluable to the successful completion of this study.

References

Apriani, E., Cardoso, L., & Obaid, A. J. (2024). *Impact of AI-Powered ChatBots on EFL Students ' Writing Skills , Self-Efficacy , and Self-Regulation : A Mixed-Methods Study*. *1*(2), 57–72.

Batia Laufer, I. S. P. N. (2012). *The Routledge Handbook of Second Language Acquisition* (1st ed.). https://www.taylorfrancis.com/chapters/edit/10.4324/9780203808184-13/vocabulary-batia-laufer-nation

Bani, M., & Masruddin, M. (2021). Development of Android-based harmonic oscillation pocket book for senior high school students. JOTSE: Journal of Technology and Science Education, 11(1), 93-103.

Cohen, L., Manion, L., & Morrison, K. (2007). Book Reviews Research Methods in Education. *The Austr Alian Educational Researcher*, *2*, 147–156.

Fu, G.-J. H. &Qing-K. (2019). Trends in the research design and application of mobile learning: a review of 2007-2016 publications in selected SSCI journals. *Interactive Learning Environments*, *27*(4), 567–581. https://doi.org/10.1080/10494820.2018.1486861

Hawanti, S., & Zubaydulloevna, K. M. (2023). AI chatbot-based learning: alleviating students' anxiety in english writing classroom. *Bulletin of Social Informatics Theory and Application*, *7*(2), 182–192. https://doi.org/10.31763/businta.v7i2.659

Hutauruk, B. S., Purba, R., & Sihombing, S. (2017). *THE EFFECTIVENESS OF ARTIFICIAL INTELLIGENCE BY CHATBOT IN ENHANCING THE STUDENTS ' VOCABULARY*. *6*.

John Creswell et al. (2017). John W, Creswell, Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. In *Journal of Social and Administrative Sciences* (Vol. 4, Issue June, pp. 3–5).

Kingdom, U., & Qammar, M. Z. (2024). *Chatbots as Conversational Partners : Their*

Effectiveness in Facilitating Language Acquisition and Reducing Foreign Language Anxiety Journal of Applied Linguistics and TESOL (JALT) Chatbots as Conversational Partners : Their Effectiveness in Facilitatin. *October*.

Kukulska-Hulme, A., & Lee, H. (2020). Intelligent assistants in language learning: an analysis of features and limitations. *CALL for Widening Participation: Short Papers from EUROCALL 2020*, *2020*(2020), 172–176. https://doi.org/10.14705/rpnet.2020.48.1184

Lee, Jang Ho and Yang, Hyejin and Shin, Dongkwang and Kim, H. (2020). Chatbots. *ELT Journal*, *74*, 338–344. https://doi.org/10.1093/elt/ccaa035

Masruddin, Hartina, S., Arifin, M. A., & Langaji, A. (2024). Flipped learning: facilitating student engagement through repeated instruction and direct feedback. Cogent Education, 11(1), 2412500.

Olugbade, D., Edwards, B. I., & Ojo, O. A. (2024). Facilitating Cognitive Load Management and Improved Learning Outcomes and Attitudes in Middle School Technology and Vocational Education Through AI Chatbot. *Journal of Technical Education and Training*, *16*(3), 114–131. https://doi.org/10.30880/jtet.2024.16.03.009

Pallant, J. (2011). *This is what readers from around the world say about the SPSS Survival Manual: 'Whenever*. Routledge.

Rizal Bagus Syaifulloh, S. N. Y. (2025). AI CHATBOTS IN EFL: SYNTHESIZING COGNITIVE, PSYCHOLOGICAL, AND ETHICAL IMPLICATIONS. *Urnal Inovasi Pendidikan Berbantuan Teknologi*, *5 no.* https://jurnalp4i.com/index.php/edutech

Santoso, S. (2014). SPSS 22 From Essensial to Ekpert Skills, Jakarta: PT. *Elex Media Komputindo*.

Waziana, W., Andewi, W., & Hastomo, T. (2024). *Students ' Perceptions of the Impact of AI Chatbots on Vocabulary and Grammar in EFL Writing*. *17*(02), 352–382. https://doi.org/10.18326/rgt.v17i2.352-382

Wijaya, S. M., El, M., Az, Q., & Prawita, F. N. (2025). *MicroLingo : Chatbot WhatsApp Berbasis AI untuk Pelatihan Bahasa Inggris dengan Pendekatan Microlearning untuk Pelaku UMKM Generasi Z di Indonesia*. *11*(1), 72–76.