



Examining the Relationship Between AI Chatbot Use Frequency and English Speaking Skills among the 2023 Cohort of English Education Students at UIN Datokarama Palu

Linda^{1*}, NurAsmawati², Ma'rifah Nurmala³, Zul Aini Rengur⁴

^{1,2,3,4} Universitas Islam Negeri Datokarama Palu, Indonesia

Corresponding e-mail : Lindabzie23@gmail.com

ARTICLE INFO	ABSTRACT
Keywords: Artificial Intelligence Chatbots; English Language Education; English Speaking Skills; Fluency; Pronunciation; Quantitative Correlational Study.	The rapid adoption of artificial intelligence (AI) chatbots in English language learning has created new opportunities for autonomous speaking practice; however, empirical evidence regarding whether the frequency of AI chatbot use is associated with learners' speaking skills remains limited. This study aimed to examine the relationship between AI chatbot use frequency and the English speaking skills of the 2023 cohort students in the English Language Education Program at UIN Datokarama Palu. A quantitative correlational design was employed involving 15 students selected through purposive sampling based on predetermined inclusion criteria. Data were collected using a validated questionnaire to measure AI chatbot use frequency and a speaking test assessing two indicators, namely fluency and pronunciation. The data were analyzed using descriptive statistics, the Shapiro-Wilk normality test, the linearity test, and the Pearson Product Moment correlation test with the assistance of IBM SPSS. The findings indicated that the participants demonstrated relatively high AI chatbot use frequency and fairly good English speaking skills. However, the Pearson Product Moment correlation analysis revealed no statistically significant relationship between AI chatbot use frequency and English speaking skills ($r = 0.010$, $p = 0.972$). These findings suggest that frequent AI chatbot use alone does not necessarily improve students' fluency and pronunciation. Therefore, AI chatbots should be integrated with structured pedagogical strategies and meaningful communicative activities to optimize their contribution to English speaking development in higher education.
Article History Received: March 30, 2026 Revised : May 24, 2026 Accepted : July 29, 2026	

This is an open access article under the [CC BY-SA](#) license



To cite this article : Examining the relationship between AI chatbot use frequency and English speaking skills among the 2023 cohort of English Education students at UIN Datokarama Palu. *Indonesian Technology and Education Journal*, 4(2), 120-133. Doi. 10.61255/itej.v4i2.1774

INTRODUCTION

Artificial Intelligence (AI) has evolved into one of the most transformative digital innovations, exerting a profound influence across various sectors, including education. Its development is no longer confined to computational functions or task automation but has expanded into intelligent technologies capable of supporting more adaptive, interactive, and learner-centered educational practices. Among the most rapidly advancing AI applications are AI chatbots, such as ChatGPT, Google Gemini, Microsoft Copilot, and Claude, which generate natural language responses and enable conversational interactions that closely resemble human communication (Bahari et al., 2025; Prayogo et al., 2025). Within educational settings, AI chatbots have moved beyond their traditional role as information providers to become learning companions that deliver immediate feedback, facilitate self-directed learning, and personalize learning experiences according to individual needs (Andrea et al., 2026). In line with this development, UNESCO emphasizes that AI has the potential to improve educational quality when it is implemented responsibly, ethically, and in ways that preserve teachers' central role in the learning process (Miao et al., 2021). Consequently, AI chatbots should no longer be viewed merely

as technological innovations but rather as components of a broader pedagogical transformation capable of reshaping how students acquire, practice, and develop language competencies.

This transformation has become particularly evident in English as a Foreign Language (EFL) education, where AI chatbots are increasingly adopted as tools for independent language practice. Students utilize AI chatbots to clarify course materials, expand their vocabulary, organize ideas in English, and engage in simulated conversations that can be accessed anytime and anywhere without temporal or geographical constraints (Laun & Wolff, 2025). Such flexibility is particularly valuable because opportunities for speaking practice in conventional classroom settings are often limited, whereas speaking proficiency can only be developed through sustained practice and continuous interaction. Speaking is recognized as a productive language skill that reflects learners' ability to communicate ideas effectively, making it one of the most essential competencies in English language learning (Brown, 2004; Astuti, 2019). Nevertheless, many learners continue to experience substantial challenges in developing speaking proficiency due to limited opportunities for authentic practice, inadequate vocabulary, low self-confidence, and difficulties in achieving accurate pronunciation and fluent speech production (Loewen & Sato, 2018). These challenges suggest that although AI chatbots offer promising opportunities to extend speaking practice beyond classroom environments, their actual contribution to improving students' speaking proficiency still requires rigorous empirical investigation.

Speaking proficiency is widely recognized as one of the primary indicators of successful English language learning because it reflects learners' ability to express ideas effectively through oral communication. In language education, speaking competence extends beyond grammatical knowledge to encompass two essential dimensions, namely fluency and pronunciation, both of which determine the effectiveness and intelligibility of spoken communication. Fluency refers to learners' ability to produce speech smoothly, coherently, and with minimal hesitation, whereas pronunciation concerns the accurate articulation of sounds to ensure that messages are clearly understood by listeners (Harmer, 2007). From a theoretical perspective, speaking competence develops through meaningful interaction that enables learners to receive comprehensible input, obtain corrective feedback, and negotiate meaning during communication, as explained by the Interaction Hypothesis (Shehadeh, 2023). Complementing this view, the Output Hypothesis argues that language acquisition is facilitated not only through exposure to the target language but also through repeated opportunities to actively produce language in authentic communicative contexts (Pannell et al., 2017). Given their ability to sustain conversations, provide immediate responses, and facilitate unlimited opportunities for language practice, AI chatbots possess considerable theoretical potential to support the development of both fluency and pronunciation in students' speaking performance.

The importance of investigating this issue becomes even more evident when considering the current state of English proficiency in Indonesia. According to the EF English Proficiency Index (EF EPI) 2025, Indonesia continues to be classified within the Low Proficiency category, indicating that oral communication remains a significant challenge across different levels of English language education (Education First, 2025). This situation highlights the urgent need to provide learners with more intensive, flexible, and sustainable opportunities for speaking practice to improve their communicative competence. At the same time, the rapid development of AI chatbots offers an alternative learning environment in which students can engage in independent speaking practice beyond formal classroom instruction without relying entirely on face-to-face interactions with lecturers or peers. Nevertheless, increased exposure to educational technology does not necessarily translate into improved learning outcomes, as its effectiveness largely depends on how learners utilize the technology during the learning process. Therefore, empirical research is required not only to explore the pedagogical potential of AI chatbots but also to examine whether the frequency of AI chatbot use is genuinely associated with students' speaking proficiency within the context of English as a Foreign Language (EFL) learning.

Recent years have witnessed a substantial increase in research examining the integration of AI chatbots into English language learning. Most existing studies consistently report that AI chatbots create more interactive learning environments, enhance student engagement, and provide flexible opportunities for language practice beyond conventional classroom settings (Lo, 2023; Bahari et al., 2025; Prayogo et al., 2025). A systematic review conducted by Du and Kei (2024) concluded that AI-powered chatbots contribute positively to speaking practice by offering immediate feedback, facilitating conversational simulations, and increasing learners' exposure to meaningful target-language interactions. Similarly, Karataş et al. (2024) demonstrated that integrating ChatGPT into foreign language instruction significantly improved learners' motivation while expanding opportunities to use English in authentic communicative situations. Beyond technological capabilities, Han and Lee (2024) and Wiboolyasarin et al. (2024) argued that the effectiveness of AI chatbots is also shaped by instructional design and learners' interaction experiences rather than by the technology itself. Furthermore, several empirical studies have reported that AI chatbots contribute to the development of speaking skills by increasing learners' confidence, encouraging more frequent language practice, and fostering greater engagement in learning activities (Mojolou et al., 2025; Mushaddiq et al., 2024; Yella Pebriani et al., 2024). Research focusing specifically on pronunciation development has likewise produced encouraging findings, indicating that AI-supported learning environments, including chatbot-assisted instruction, can provide objective feedback that helps learners improve pronunciation accuracy and spoken language performance (Liu et al., 2023; Caiza et al., 2025; Stoughton et al., 2024).

Despite these promising findings, several limitations remain evident in the existing body of literature. Most previous studies have adopted experimental designs, instructional interventions, or perception-based surveys to evaluate the effectiveness of AI chatbots, thereby focusing primarily on treatment effects, learner satisfaction, motivation, technology acceptance, or instructional feasibility rather than examining the empirical relationship between students' actual chatbot usage behavior and their speaking proficiency (Du & Kei, 2024; Karataş et al., 2024; Mushaddiq et al., 2024; Safitri et al., 2025). Consequently, relatively little empirical evidence is available regarding whether the frequency of AI chatbot use is associated with speaking performance under authentic learning conditions, particularly within Indonesian higher education contexts. Moreover, previous studies have generally assessed speaking ability as a broad construct without distinguishing between fluency and pronunciation, despite the possibility that these dimensions may be influenced differently by AI-assisted language practice (Maulana Ocard Nasution, 2024). Research involving students enrolled in the English Language Education Program (Tadris Bahasa Inggris/TBI) at UIN Datokarama Palu also remains scarce, leaving limited empirical evidence regarding patterns of AI chatbot use within Islamic higher education institutions in Indonesia. These limitations reveal a clear research gap, namely the absence of studies that specifically investigate the relationship between the frequency of AI chatbot use and students' speaking proficiency, measured through fluency and pronunciation, under conditions of natural usage without experimental intervention. Addressing this gap is essential for extending current understanding of AI chatbot integration in English language education, not only from the perspective of technological effectiveness but also from the perspective of learners' authentic technology-use behavior in real educational settings.

Based on the foregoing discussion, this study aims to examine the relationship between the frequency of AI chatbot use and the speaking skills of 2023 cohort students enrolled in the English Language Education Program (Tadris Bahasa Inggris/TBI) at UIN Datokarama Palu. Specifically, the study evaluates speaking performance through two fundamental indicators, namely fluency and pronunciation, as these dimensions represent essential components of effective oral communication in English as a Foreign Language (EFL) learning. Unlike the majority of previous studies that have concentrated on instructional interventions, learner perceptions, or the

effectiveness of AI chatbot implementation, the present study investigates the association between students' natural AI chatbot usage behavior and their speaking proficiency using a quantitative correlational approach without experimental treatment. Theoretically, this study is expected to enrich the growing body of knowledge on AI-assisted language learning by providing empirical evidence regarding whether the intensity of AI chatbot use is associated with students' speaking performance. Practically, the findings are expected to inform lecturers, curriculum developers, and higher education institutions in designing pedagogical strategies that encourage not only more frequent use of AI chatbots but also more purposeful and pedagogically guided speaking practice. Ultimately, this study seeks to contribute to the development of more effective, contextually relevant, and evidence-based AI-supported English language learning practices while expanding current understanding of how AI chatbot usage relates to speaking proficiency among university students (Safitri et al., 2025).

METHOD

This study employed a quantitative approach with a correlational research design to examine the relationship between the frequency of AI chatbot use and students' English speaking skills. A correlational design was considered appropriate because the primary objective of the study was to determine whether an association existed between the two variables without introducing any experimental treatment or instructional intervention (Sugiyono, 2020). The research was conducted at the English Language Education Program (Tadris Bahasa Inggris/TBI), UIN Datokarama Palu, involving students from the 2023 cohort during the 2025 academic year (Rizal Bagus Syaifulloh, 2025). The target population consisted of 31 students, all of whom were invited to complete an online questionnaire measuring the frequency of AI chatbot use. A total of 22 students completed the questionnaire, after which purposive sampling was employed to identify participants who met the predetermined inclusion criteria (Mazhar et al., 2021). Participants were required to (1) have used AI chatbots such as ChatGPT, Google Gemini, Microsoft Copilot, or Claude for English conversation practice; (2) use AI chatbots for more than one hour per day and at least three times per week; (3) utilize AI chatbots specifically for speaking practice rather than merely completing academic assignments; and (4) be willing to participate in an English speaking assessment. Information regarding the duration, frequency, and purpose of AI chatbot use was verified through a brief follow-up interview after questionnaire completion. Based on these criteria, 15 students satisfied all eligibility requirements and were selected as the final research participants (Caiza et al., 2025). The participant selection process is summarized in Figure 1.

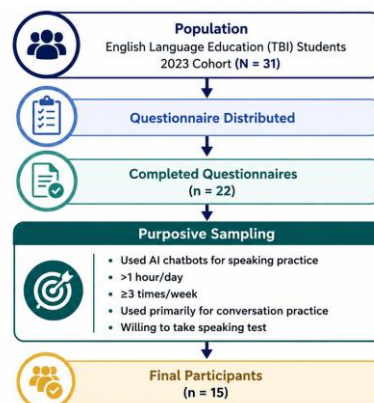


Figure 1. Research Participant Selection Procedure

Figure 1 illustrates the participant selection process employed in this study. The purposive sampling procedure ensured that only students with substantial experience using AI chatbots for

English speaking practice were included in the analysis, thereby increasing the relevance of the collected data to the research objectives.

The study examined two research variables, namely the frequency of AI chatbot use as the independent variable (X) and English speaking skills as the dependent variable (Y). The independent variable reflected the frequency, duration, and consistency with which students used AI chatbots to practice English conversations, whereas the dependent variable was assessed through two key speaking components, namely fluency and pronunciation (Stoughton et al., 2024). Data were collected using two research instruments. The first instrument was an eight-item questionnaire employing a five-point Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree) to measure the frequency of AI chatbot use. The second instrument was a recorded monologue-based speaking test assessed using an evaluation rubric adapted from Harmer (2007), focusing specifically on fluency and pronunciation, with scoring procedures supported by contemporary recommendations for speech assessment (Liu et al., 2023). The research variables and instruments are presented in Table 1.

Table 1. Research Variables and Instruments

Variable	Indicator(s)	Instrument	Measurement Scale
Frequency of AI Chatbot Use (X)	Frequency, duration, and consistency of AI chatbot use for English speaking practice	Eight-item questionnaire	Five-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree)
English Speaking Skills (Y)	Fluency and pronunciation	Speaking test rubric	Performance score (1–5)

Prior to data collection, the questionnaire underwent content validation through expert judgment conducted by a lecturer specializing in English language education. Following expert feedback, several items were revised before empirical testing was performed using the Pearson Product Moment Correlation in IBM SPSS. Questionnaire items were considered valid when the obtained correlation coefficient exceeded the critical value ($r = 0.361$) at the 0.05 significance level. As presented in Table 2, all eight questionnaire items satisfied the validity criteria, with correlation coefficients ranging from 0.639 to 0.917, indicating that every item was appropriate for measuring the intended construct. Instrument reliability was subsequently evaluated using Cronbach's Alpha, yielding a coefficient of 0.791, which exceeded the recommended threshold of 0.70 and demonstrated satisfactory internal consistency. These findings indicate that the questionnaire was both valid and reliable for measuring students' frequency of AI chatbot use.

Table 2. Validity Test Results of the AI Chatbot Use Frequency Questionnaire

Item	r-value	r-table	Sig.	Decision
X1	0.899	0.361	0.000	Valid
X2	0.727	0.361	0.000	Valid
X3	0.639	0.361	0.000	Valid
X4	0.842	0.361	0.000	Valid
X5	0.832	0.361	0.000	Valid
X6	0.917	0.361	0.000	Valid
X7	0.898	0.361	0.000	Valid
X8	0.760	0.361	0.000	Valid

Following the validity analysis, the internal consistency of the questionnaire was evaluated using Cronbach's Alpha to determine the reliability of the instrument. An instrument is considered reliable when the Cronbach's Alpha coefficient is equal to or greater than 0.70. The

results of the reliability analysis are presented in Table 3.

Table 3. Reliability Test Results of the AI Chatbot Use Frequency Questionnaire

Instrument	Cronbach's Alpha	Number of Items	Criterion	Interpretation
AI Chatbot Use Frequency Questionnaire	0.791	8	≥ 0.70	Reliable

Data collection was conducted in two sequential stages. During the first stage, participants completed the questionnaire measuring the frequency of AI chatbot use for English speaking practice. In the second stage, each participant completed an individual English monologue task, which was audio-recorded and evaluated using the predetermined speaking rubric focusing on fluency and pronunciation. All statistical analyses were performed using IBM SPSS. Descriptive statistics, including the minimum, maximum, mean, and standard deviation, were first calculated to summarize participant characteristics. Prior to hypothesis testing, data normality was examined using the Shapiro–Wilk test, while the assumption of linearity was assessed using the Test for Linearity. After these assumptions were satisfied, the relationship between AI chatbot use frequency and English speaking skills was analyzed using the Pearson Product Moment Correlation with a significance level of 0.05. A significance value below 0.05 indicated a statistically significant relationship between the two variables, whereas a value greater than 0.05 indicated that no significant relationship existed.

RESULTS AND DISCUSSION

Results

This section presents the empirical findings of the study in a sequential manner, beginning with descriptive statistics, followed by assumption testing, including normality and linearity analyses, and concluding with the Pearson Product Moment correlation analysis. These analyses were conducted to provide a comprehensive overview of the research data and to examine the relationship between AI chatbot use frequency and students' English speaking skills.

Descriptive Statistics

Descriptive statistical analysis was conducted to summarize the characteristics of the research variables prior to hypothesis testing. The analysis included the number of participants, minimum and maximum scores, mean values, and standard deviations for both the frequency of AI chatbot use and English speaking skills. The results are presented in Table 4.

Table 4. Descriptive Statistics of the Research Variables

Variable	N	Minimum	Maximum	Mean	SD
AI Chatbot Use Frequency	15	20.00	35.00	28.67	4.08
English Speaking Skills	15	62.00	100.00	78.53	11.77
Valid N (listwise)	15				

As shown in Table 4, the mean score for AI chatbot use frequency was 28.67 (SD = 4.08), with scores ranging from 20.00 to 35.00. Meanwhile, the English speaking skills variable obtained a mean score of 78.53 (SD = 11.77), with scores ranging from 62.00 to 100.00. The larger standard deviation observed for speaking skills indicates greater variability in students' speaking performance than in their frequency of AI chatbot use. Overall, the descriptive statistics demonstrate sufficient variation in both variables to support subsequent assumption testing and correlation analysis.

Normality Test

Before conducting the correlation analysis, the normality assumption was examined using the Kolmogorov–Smirnov and Shapiro–Wilk tests in IBM SPSS. Because the sample size consisted of fewer than 50 participants ($n = 15$), the interpretation primarily referred to the Shapiro–Wilk results. Data were considered normally distributed when the significance value exceeded 0.05.

Table 5. Results of the Normality Test

Variable	Kolmogorov–Smirnov (Sig.)	Shapiro–Wilk (Sig.)	Distribution
AI Chatbot Use Frequency	0.200	0.832	Normal
English Speaking Skills	0.177	0.087	Normal

The results presented in Table 5 indicate that the Shapiro–Wilk significance values were 0.832 for AI chatbot use frequency and 0.087 for English speaking skills. Since both significance values exceeded the threshold of 0.05, the two variables satisfied the assumption of normality and were considered appropriate for subsequent parametric statistical analyses.

Linearity Test

A linearity test was subsequently performed to determine whether the relationship between AI chatbot use frequency and English speaking skills met the assumption of linearity required for Pearson Product Moment correlation analysis. The relationship was considered linear when the significance value of the Deviation from Linearity exceeded 0.05.

Table 6. Results of the Linearity Test

Source	Sum of Squares	df	Mean Square	F	Sig.
Between Groups (Combined)	1034.438	11	94.040	107.474	.000
Linearity	1002.307	1	1002.307	1145.494	.000
Deviation from Linearity	32.131	10	3.213	3.672	.111
Within Groups	3.500	4	0.875		
Total	1037.938	15			

As presented in Table 6, the Deviation from Linearity significance value was 0.111, which exceeded the criterion of 0.05. This result indicates that no significant deviation from linearity was detected, suggesting that the relationship between AI chatbot use frequency and English speaking skills satisfies the linearity assumption. Therefore, the data met the statistical requirements for Pearson Product Moment correlation analysis.

Pearson Product Moment Correlation

After confirming that the assumptions of normality and linearity had been satisfied, Pearson Product Moment correlation analysis was conducted to examine the relationship between AI chatbot use frequency and English speaking skills. The analysis was performed using a significance level of 0.05, and the results are summarized in Table 7.

Table 7. Pearson Product Moment Correlation between AI Chatbot Use Frequency and English Speaking Skills

Variables	Pearson's r	Sig. (2-tailed)	N	Interpretation
AI Chatbot Use Frequency and English Speaking Skills	0.010	0.972	15	Very Weak, Not Significant

The Pearson Product Moment correlation analysis produced a correlation coefficient of $r = 0.010$ with a significance value of 0.972. These findings indicate a very weak positive relationship between AI chatbot use frequency and English speaking skills. Because the significance value exceeded 0.05, the null hypothesis (H_0) was accepted, indicating that no statistically significant relationship existed between the frequency of AI chatbot use and the English speaking skills of the 2023 cohort students in the English Language Education Program at UIN Datokarama Palu. The findings also indicate that variations in students' AI chatbot use frequency were not accompanied by corresponding variations in their English speaking performance within the sample examined.

The results presented in this section indicate that all statistical assumptions for Pearson Product Moment correlation analysis were fulfilled. The descriptive analysis summarized the characteristics of the research variables, whereas the assumption tests confirmed that the data were normally distributed and exhibited a linear relationship. The correlation analysis further revealed that the relationship between AI chatbot use frequency and English speaking skills was positive but statistically insignificant. These findings are discussed in greater detail in the following section.

Discussion

The present study found that the frequency of AI chatbot use was not significantly associated with students' English speaking skills, as indicated by the Pearson Product Moment correlation coefficient of $r = 0.010$ with a significance value of $p = 0.972$. This finding suggests that merely increasing the frequency of AI chatbot use does not necessarily translate into improvements in speaking performance, particularly in terms of fluency and pronunciation. Rather than functioning as an independent determinant of speaking proficiency, AI chatbot use appears to represent only one component within a broader language learning process influenced by multiple cognitive, linguistic, and pedagogical factors. This result extends the current understanding of AI-assisted language learning by indicating that the effectiveness of AI chatbots cannot be adequately explained through usage intensity alone. Consequently, the educational value of AI chatbots should be interpreted not in terms of how often learners interact with the technology, but in terms of how those interactions facilitate meaningful language production and communicative practice.

From a theoretical perspective, these findings provide an interesting interpretation of both the Interaction Hypothesis and the Output Hypothesis. The Interaction Hypothesis proposes that language acquisition develops through meaningful interaction in which learners receive comprehensible input, negotiate meaning, and obtain corrective feedback (Shehadeh, 2023). Similarly, the Output Hypothesis argues that language development is promoted when learners actively produce language and subsequently reflect upon or modify their linguistic output (Pannell et al., 2017). The absence of a significant relationship in the present study does not contradict these theories; rather, it suggests that interaction frequency alone is insufficient to activate the learning mechanisms proposed by both frameworks. Productive interaction requires meaningful negotiation, immediate feedback, and opportunities for linguistic refinement, whereas repetitive or superficial interactions may contribute little to language development. This

interpretation is also consistent with Brown's (2004) view that speaking proficiency represents a complex productive skill requiring the integration of linguistic knowledge, communicative competence, and continuous practice, as well as Astuti's (2019) argument that speaking performance is influenced by numerous internal and external factors beyond learning frequency alone. Therefore, the findings imply that the pedagogical quality of AI-supported interaction may be considerably more influential than the quantity of interaction itself.

The descriptive findings provide additional insight into why no significant correlation was identified between AI chatbot use frequency and English speaking skills. Although participants reported a relatively high level of AI chatbot use ($M = 28.67$), their speaking performance exhibited greater variability ($M = 78.53$; $SD = 11.77$), suggesting that differences in speaking proficiency were influenced by factors beyond the frequency of technology use. This pattern indicates that repeated exposure to AI chatbots does not necessarily ensure equivalent learning outcomes among students with diverse linguistic backgrounds and learning characteristics. In language learning, speaking proficiency develops through the interaction of multiple dimensions, including linguistic competence, communicative confidence, prior learning experiences, and learners' ability to regulate their own learning processes. Consequently, students who use AI chatbots with similar frequency may still demonstrate substantially different levels of speaking performance because the quality of their engagement with the technology varies considerably. These findings reinforce the view that technology functions primarily as a learning facilitator, whereas the effectiveness of language acquisition ultimately depends on how learners utilize available learning opportunities.

The findings of this study differ from those reported in several previous investigations that generally describe AI chatbots as effective tools for improving English speaking performance. Studies by Yella Pebriani et al. (2024) and Mushaddiq et al. (2024) found that students perceived AI chatbots as beneficial for increasing speaking practice, learning motivation, and classroom engagement. Likewise, systematic reviews conducted by Lo (2023), Du and Kei (2024), Bahari et al. (2025), and Safitri et al. (2025) concluded that AI chatbots provide meaningful opportunities for language practice by facilitating interactive communication, immediate feedback, and autonomous learning. However, these previous studies primarily evaluated learners' perceptions, instructional interventions, or the pedagogical effectiveness of chatbot integration rather than examining the empirical relationship between usage frequency and actual speaking performance. Consequently, positive perceptions of AI chatbots should not be interpreted as direct evidence that more frequent use automatically leads to higher speaking proficiency. The present study therefore contributes a complementary perspective to the existing literature by demonstrating that the intensity of AI chatbot use should be distinguished from its educational effectiveness, highlighting that the quality, purpose, and context of chatbot interaction may be more important determinants of speaking development than usage frequency alone.

Several plausible explanations may account for the absence of a significant relationship between AI chatbot use frequency and English speaking skills. First, the frequency of technology use does not necessarily reflect the quality of learning engagement. Students may interact with AI chatbots frequently while using them primarily to complete assignments, translate texts, or seek explanations rather than engaging in sustained English conversations that promote speaking development. Such interaction patterns provide limited opportunities for learners to practice spontaneous speech production, negotiate meaning, or refine pronunciation, all of which are fundamental processes in second language acquisition. This explanation is consistent with recent studies emphasizing that the educational benefits of AI chatbots depend largely on the design of learning activities and the quality of learner-system interaction rather than on technology availability alone (Han & Lee, 2024; Wiboolyasarin et al., 2024). Similarly, comprehensive reviews by Andrea et al. (2026), Laun and Wolff (2025), and Prayogo et al. (2025) conclude that AI chatbots are most effective when integrated into structured pedagogical practices that

encourage active participation, meaningful interaction, and continuous feedback. Therefore, the present findings suggest that how students use AI chatbots may be considerably more important than how often they use them.

Another possible explanation relates to the inherent characteristics of English speaking as a complex productive skill. Speaking proficiency is not developed solely through repeated exposure to language but also requires continuous practice involving pronunciation accuracy, fluency, vocabulary retrieval, grammatical control, and communicative interaction within authentic contexts (Brown, 2004; Harmer, 2007). While AI chatbots are capable of facilitating written dialogue and simulated conversations, they may not consistently provide the level of corrective feedback, pronunciation modeling, or interactional authenticity necessary for improving oral performance. Previous studies have demonstrated that AI-assisted pronunciation training can be effective when supported by speech recognition technology, structured pronunciation tasks, or immersive learning environments specifically designed for oral language development (Liu et al., 2023; Caiza et al., 2025; Stoughton et al., 2024). Similarly, Karataş et al. (2024) found that the positive impact of ChatGPT emerged when its use was embedded within well-designed instructional activities rather than through independent technology use. Consequently, the findings of this study do not imply that AI chatbots are ineffective for language learning; instead, they indicate that frequency of use alone is an insufficient predictor of speaking proficiency unless accompanied by purposeful instructional design and high-quality communicative practice. This perspective expands the current literature by emphasizing that pedagogical integration, rather than technology exposure alone, remains the critical factor in realizing the educational potential of AI-assisted language learning.

The findings of this study also contribute to the ongoing discussion regarding the role of AI in language education by suggesting that technology should be viewed as a pedagogical enabler rather than an autonomous driver of learning outcomes. Although AI chatbots provide learners with flexible access to language practice, immediate responses, and opportunities for autonomous learning, these technological affordances do not automatically translate into measurable improvements in speaking proficiency. This perspective is consistent with recent literature emphasizing that successful AI integration depends on the alignment between technological capabilities, instructional objectives, and learner engagement (Andrea et al., 2026; Bahari et al., 2025). Furthermore, UNESCO has highlighted that AI should complement, rather than replace, the pedagogical role of educators by supporting meaningful learning experiences that foster critical thinking, communication, and active participation (Miao et al., 2021). The present findings therefore reinforce the notion that AI chatbots are most beneficial when embedded within well-structured instructional activities that deliberately encourage authentic communication and guided language practice. In this regard, the study extends current research by demonstrating that the educational value of AI chatbots lies not merely in increasing learners' access to technology but in enhancing the quality of learning experiences facilitated through appropriate pedagogical strategies.

Despite the absence of a statistically significant relationship, this study provides several important theoretical and practical contributions. From a theoretical perspective, it broadens the current understanding of AI-assisted language learning by demonstrating that the frequency of AI chatbot use should not be interpreted as a direct indicator of speaking development, thereby complementing previous studies that have predominantly emphasized learners' perceptions, motivation, or the effectiveness of instructional interventions. From a practical perspective, the findings suggest that educators should move beyond encouraging frequent AI chatbot use and instead design learning activities that promote purposeful interaction, reflective language production, and constructive feedback. Such an approach may involve integrating AI chatbots into structured speaking tasks, classroom discussions, peer collaboration, and instructor-guided reflection to maximize their pedagogical value. Nevertheless, several limitations should be

acknowledged when interpreting these findings. The relatively small sample size limits the generalizability of the results, while the assessment of speaking skills was restricted to fluency and pronunciation through a monologue task, without considering additional dimensions such as grammar, vocabulary, interactional competence, or spontaneous conversational performance. These limitations provide opportunities for future studies to employ larger samples, experimental or mixed-methods designs, and more comprehensive measures of speaking proficiency to further investigate the complex relationship between AI chatbot use and language learning outcomes across diverse educational contexts.

The findings also offer a distinct contribution to the current literature on AI-assisted language learning by clarifying that frequent chatbot use should not be treated as a direct proxy for language development. While earlier studies have emphasized the potential of AI chatbots to increase access, motivation, learner autonomy, and engagement, the present study demonstrates that these benefits do not necessarily produce measurable gains in speaking performance when usage is examined under natural learning conditions (Maulana Oscar Nasution, 2024; Rizal Bagus Syaifulloh, 2025). This distinction is important because it shifts the focus from technological exposure to the pedagogical quality of interaction. In the Indonesian context, where English proficiency remains a continuing challenge, merely expanding access to AI tools may be insufficient unless students are guided to use them for sustained oral production, corrective feedback, and meaningful communicative practice (Education First, 2025). Conceptually, this study modifies the common assumption that greater technology use leads to better learning outcomes by showing that the relationship is conditional upon instructional design, learner readiness, and interactional quality. Practically, lecturers should therefore integrate AI chatbots into structured speaking activities that include clear communicative goals, guided prompts, reflection, and follow-up feedback rather than relying on unguided independent use. In this way, the study positions AI chatbots as pedagogical resources whose effectiveness depends on how they are embedded within the broader language learning environment.

The interpretation of these findings should nevertheless consider several methodological and contextual limitations. The final sample consisted of only 15 students from a single cohort and institution, which restricts the extent to which the findings can be generalized to broader populations of EFL learners. In addition, speaking performance was assessed only through fluency and pronunciation, while other important dimensions, including grammar, vocabulary, interactional competence, responsiveness, and discourse management, were not examined. The use of a monologue task may also have captured individual oral production more effectively than genuine conversational ability, since interactive speaking requires turn-taking, negotiation of meaning, and spontaneous responses to interlocutors. Furthermore, the study measured frequency of use without directly evaluating the quality, duration, linguistic complexity, or pedagogical purpose of each chatbot interaction. Future research should therefore involve larger and more diverse samples, incorporate interactive speaking assessments, and examine variables such as prompt quality, feedback type, learner motivation, self-regulation, and prior proficiency. Experimental and mixed-methods designs would also be valuable for identifying how specific forms of chatbot-assisted speaking practice influence different dimensions of oral competence. These directions may provide a more comprehensive explanation of when, how, and under what conditions AI chatbots contribute meaningfully to EFL speaking development.

CONCLUSIONS

This study examined the relationship between the frequency of AI chatbot use and the English speaking skills of the 2023 cohort students in the English Language Education Program at UIN Datokarama Palu. The findings demonstrated that the frequency of AI chatbot use was not significantly associated with students' speaking performance, indicating that frequent interaction

with AI chatbots alone does not necessarily improve fluency and pronunciation. These results suggest that speaking proficiency is shaped by multiple interrelated factors beyond technology use, including the quality of learning experiences, learners' individual characteristics, and the pedagogical strategies employed during language instruction. Consequently, the educational value of AI chatbots should be viewed not merely in terms of usage frequency but in terms of how effectively they are integrated into structured learning activities that promote meaningful communication, active language production, and constructive feedback. This study contributes empirical evidence to the growing body of research on AI-assisted language learning by providing a different perspective on the relationship between natural AI chatbot usage and speaking proficiency in higher education. The findings also offer practical insights for educators in designing more purposeful AI-supported speaking activities and provide a foundation for future studies to investigate additional variables and instructional approaches that may further enhance English speaking development.

ACKNOWLEDGEMENT

The author expresses sincere gratitude to the Almighty God for His blessings, guidance, and strength throughout the completion of this research. The author also extends sincere appreciation to the Coordinator of the English Language Education Program at UIN Datokarama Palu for granting permission and facilitating the conduct of this study. Special thanks are addressed to the author's supervisors for their invaluable guidance, constructive feedback, continuous encouragement, and insightful suggestions throughout every stage of the research process. Their expertise and support contributed significantly to the quality and completion of this article.

The author is also grateful to the students of the English Language Education Program (Tadris Bahasa Inggris/TBI), 2023 Cohort, at UIN Datokarama Palu for their voluntary participation and valuable contributions during the data collection process. Their cooperation was essential to the successful completion of this study. Finally, the author sincerely appreciates the support and encouragement provided by colleagues and friends throughout the research process, which contributed to the successful completion of this research.

REFERENCES

- Andrea, P., Palencia, E., Patricia, K., Arteaga, A., Alfonso, T., & Cama, P. (2026). Artificial intelligence chatbots for inclusive pedagogical practices: A PRISMA systematic review and bibliometric mapping (2021–2025). *Journal of Artificial Intelligence and Technology*, 6(2), 1–10. <https://doi.org/10.37965/jait.2026.1252>
- Astuti, E. S. (2019). Faktor-faktor yang mempengaruhi *speaking performance* mahasiswa jurusan pendidikan bahasa Inggris. *Paradigma: Jurnal Filsafat, Sains, Teknologi, dan Sosial Budaya*, 25(2), 27–33. <https://doi.org/10.33503/paradigma.v25i2.543>
- Bahari, A., Han, F., & Strzelecki, A. (2025). Integrating CALL and AIALL for an interactive pedagogical model of language learning. *Education and Information Technologies*, 30, 14305–14333. <https://doi.org/10.1007/s10639-025-13388-w>
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Pearson Education.
- Caiza, G., Villafuerte, C., & Guanuche, A. (2025). Interactive application with virtual reality and artificial intelligence for improving pronunciation in English learning. *Applied Sciences*, 15(17), 1–18. <https://doi.org/10.3390/app15179270>
- Du, J., & Kei, B. (2024). Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. *Computers and Education: Artificial Intelligence*, 6, Article 100230. <https://doi.org/10.1016/j.caeai.2024.100230>

- Education First. (2025). *EF English Proficiency Index*. <https://www.ef.com/wwen/epi/>
- Han, J., & Lee, D. (2024). Research on the development of principles for designing elementary English speaking lessons using artificial intelligence chatbots. *Humanities and Social Sciences Communications*, 11(1), 1–15. <https://doi.org/10.1057/s41599-024-02646-w>
- Harmer, J. (2007). *How to teach English* (2nd ed.). Pearson Education Limited.
- Karataş, F., Abedi, F. Y., Ozek Gunyel, F., Karadeniz, D., & Kuzgun, Y. (2024). Incorporating AI in foreign language education: An investigation into ChatGPT's effect on foreign language learners. *Education and Information Technologies*, 29(15), 19343–19366. <https://doi.org/10.1007/s10639-024-12574-6>
- Laun, M., & Wolff, F. (2025). Chatbots in education: Hype or help? A meta-analysis. *Learning and Individual Differences*, 119, Article 102646. <https://doi.org/10.1016/j.lindif.2025.102646>
- Liu, J., Wumaier, A., Fan, C., & Guo, S. (2023). Automatic fluency assessment method for spontaneous speech without reference text. *Electronics*, 12(8). <https://doi.org/10.3390/electronics12081775>
- Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4). <https://doi.org/10.3390/educsci13040410>
- Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching*, 51(3), 285–329. <https://doi.org/10.1017/S0261444818000125>
- Maulana Oscar Nasution, A. (2024). Exploring the role of AI chatbot in English language teaching and learning. *Journal of English Language and Education*, 9(5), 34–43. <https://doi.org/10.31004/jele.v9i5.513>
- Mazhar, S. A., Anjum, R., Anwar, A. I., & Khan, A. A. (2021). Methods of data collection: A fundamental tool of research. *Journal of Integrated Community Health*, 10(1), 6–10. <https://doi.org/10.24321/2319.9113.202101>
- Miao, F., Holmes, W., Huang, R., & Zhang, H. (2021). *AI and education: Guidance for policy-makers* (F. Miao, Ed.). United Nations Educational, Scientific and Cultural Organization (UNESCO). <https://doi.org/10.54675/PCSP7350>
- Mojolou, D. N., Madli, F., Abdullah, A. S., Thomas, H., & Nayve, R. (2025). Transforming student engagement in higher education: The impact of artificial intelligence on learning involvement. *ASEAN Artificial Intelligence Journal*, 3(1), 58–65. <https://karyailham.com.my/index.php/aaaj/index>
- Mushaddiq, M., Noni, N., Abdullah, A., & Tahir, M. (2024). Exploring EFL students' creativity and engagement with AI chatbot in web-based English course. *GLENS: Global English Insights Journal*, 2(1), 19–27. <https://doi.org/10.61220/glens.v2i1.591>
- Pannell, J., Partsch, F., & Fuller, N. (2017). The output hypothesis: From theory to practice. *TESOL Working Paper Series*, 15, 126–159.
- Pebriani, Y., & Serasi, R. (2024). The correlation between AI (artificial intelligence) and students' speaking performance in the EFL classroom. *Pendas: Jurnal Ilmiah Pendidikan Dasar*, 9. <https://doi.org/10.23969/jp.v9i3.18365>
- Prayogo, L., Agustin, R. K., & Puspita, B. (2025). Integrating artificial intelligence in English language learning: A systematic literature review. *Journal of English Education*, 5(2), 75–88. <https://doi.org/10.30998/jedu.v5i2.14647>
- Rizal Bagus Syaifulloh, S. N. Y. (2025). AI chatbots in EFL: Synthesizing cognitive, psychological, and ethical implications. *EDUTECH: Jurnal Inovasi Pendidikan Berbantuan Teknologi*, 5(1), 113–119. <https://doi.org/10.51878/edutech.v5i1.4701>

-
- Safitri, E. I., Hidayati, S., & Ciptaningrum, D. S. (2025). The impact of AI chatbots on English language learners' speaking proficiency: A systematic review. *Journal of Research on English and Language Learning (J-REaLL)*, 6(2), 317–329. <https://doi.org/10.33474/j-reall.v6i2.23866>
- Shehadeh, A. (2023). Review of *Interaction, feedback and task research in second language learning: Methods and design* (A. Mackey, 2020). *Canadian Journal of Applied Linguistics*, 26(1). <https://doi.org/10.37213/cjal.2023.32147>
- Stoughton, A. M., & Kang, O. (2024). A systematic review of empirical mobile-assisted pronunciation studies through a perception–production lens. *Languages*, 9. <https://doi.org/10.3390/languages9070251>
- Sugiyono. (2020). *Metode penelitian kuantitatif, kualitatif, dan R&D*. Alfabeta.
- Wiboolyasarini, W., Wiboolyasarini, K., Tiranant, P., Boonyakitanont, P., & Jinowat, N. (2024). Designing chatbots in language classrooms: An empirical investigation from user learning experience. *Smart Learning Environments*, 11(1). <https://doi.org/10.1186/s40561-024-00319-4>