INTERACTION: Jurnal Pendidikan Bahasa Vol. 13, No.1; Maret 2026
P-ISSN: 2406-9558; E-ISSN: 2406-9566

The effectiveness of AI chatbot-based English Learning Methods in Improving Students' Speaking Skills

Fatmawati
fatmawatim934@gmail.com

Syamsu T
Syamsutang64@gmail.com

Nurul Faradillah
nurulfaradillah1401@gmail.com

Andi Sadapotto
sadapotto.andi@yahoo.co.id

Sam Hermansyah
sam.hermansyah82@gmail.com

Universitas Muhammadiyah Sidenreng Rappang

ABSTRACT

This mixed-methods study examined the effectiveness of AI chatbot-based English learning methods, specifically using TalkPal AI, in improving eleventh-grade students' speaking skills at SMAN 11 Sidrap, Indonesia. A quasi-experimental nonequivalent control group design was employed for the quantitative phase, while semi-structured interviews provided qualitative insights into students' perceptions. The experimental group (n = 18 for post-test) engaged in interactive speaking practice with TalkPal AI, while the control group (n = 14) received conventional teacher-centered instruction. Data were collected through pre- and post-speaking tests assessed using Brown's (2004) rubric across five components fluency, accuracy, pronunciation, vocabulary, and coherence and interviews with ten purposively selected students from the experimental group. Quantitative analysis revealed that the experimental group achieved a significantly higher post-test mean score (M = 23.17) compared to the control group (M = 12.57). The Mann-Whitney U test confirmed a statistically significant difference between groups ($p < 0.001$), indicating the effectiveness of the AI chatbot intervention. Qualitative findings showed overwhelmingly positive perceptions: students reported increased confidence, motivation, fluency, pronunciation, and reduced speaking anxiety, attributing these gains to the chatbot's interactive, non-judgmental, and flexible learning environment. Minor challenges included unstable internet connections and initial familiarity with the application. The results demonstrate that AI chatbot-based methods, such as TalkPal AI, constitute an effective and innovative approach for enhancing EFL learners' speaking proficiency in secondary

education contexts. Pedagogical implications for technology-integrated language teaching and recommendations for future research are discussed.
Keywords: AI Chatbot; TalkPal AI; Speaking Skills; EFL learning

INTRODUCTION

In the era of globalization and digital transformation, English speaking proficiency has become a critical competency for academic success, professional development, and international communication. As a foreign language (EFL) in Indonesia, English speaking remains one of the most challenging skills for students due to limited opportunities for authentic practice, high levels of speaking anxiety, insufficient vocabulary, and traditional teacher-centered instructional methods that emphasize grammar and reading over oral production (Sahara & Satria, 2025; Wang et al., 2024). These challenges are particularly evident at the secondary school level, where students often remain passive during speaking activities and prefer using their native language in classroom interactions.

Recent advancements in Artificial Intelligence (AI) have opened promising avenues for addressing these limitations. AI-powered chatbots, in particular, offer interactive, personalized, and judgment-free speaking practice that simulates real-life conversations. Unlike conventional methods, AI chatbots can provide immediate feedback on pronunciation, fluency, vocabulary, and grammar while operating 24/7 without time or location constraints (Alenezi & Alenezi, 2025; Li & Liang, 2019; Wiboolyasarin et al., 2024). Applications such as TalkPal AI enable learners to engage in voice-based dialogues with virtual partners, thereby reducing psychological barriers such as fear of making mistakes or being judged by peers and teachers.

Although previous studies have demonstrated the potential of AI chatbots in language learning, most research has been conducted at the university level or in developed countries, with limited empirical evidence from Indonesian secondary schools. Existing studies tend to focus on general language proficiency, motivation, or vocabulary acquisition, while fewer have employed rigorous experimental designs to measure specific improvements in core speaking components fluency, accuracy, pronunciation, vocabulary, and coherence particularly through a mixed-methods approach (Abimanto, 2023; Al-Jaf et al., 2024; Eisenring et al., 2024). Furthermore, the integration of AI chatbots in Indonesian high schools remains underexplored, despite the increasing accessibility of mobile technology among students.

This study addresses these gaps by investigating the effectiveness of AI chatbot-based English learning methods, specifically using TalkPal AI, in improving eleventh-grade students' speaking skills at SMAN 11 Sidrap, South Sulawesi, Indonesia. The research employed a mixed-methods design with a quasi-experimental nonequivalent control group for the quantitative phase and semi-structured interviews for the qualitative phase.

The specific objectives of this study were to examine the effectiveness of AI chatbot-based learning in improving students' English speaking skills compared to conventional methods, and to explore students' perceptions toward the use of TalkPal AI in developing their speaking proficiency. It is expected that the findings of this study will contribute to the growing body of knowledge on technology-enhanced language learning and offer practical recommendations for English teachers, curriculum developers, and educational policymakers who are seeking innovative approaches to enhance EFL speaking instruction in similar contexts.

LITERATURE REVIEW

Artificial Intelligence (AI) refers to computer systems capable of performing tasks that typically require human intelligence, such as reasoning, learning, problem-solving, and natural language understanding (Besold et al., 2017). In educational contexts, AI has evolved from simple automation tools into sophisticated adaptive systems that personalize learning experiences, provide real-time feedback, and support individualized instruction (Holmes, 2023; Zawacki-Richter et al., 2019). The integration of AI in language education has gained significant momentum, particularly through conversational agents that simulate human-like interaction.

Speaking is a fundamental productive skill in second language acquisition that involves the ability to produce oral language fluently, accurately, and coherently in real-time communication (Harmer, 2001; Brown, 2004). Effective speaking requires mastery of several interrelated components: fluency (smooth delivery with minimal hesitation), accuracy (grammatical correctness), pronunciation (intelligible sound production), vocabulary (appropriate word choice and range), and coherence (logical organization of ideas) (Richards & Schmidt, 2002). For EFL learners in Indonesia, developing these aspects remains challenging due to limited practice opportunities and high anxiety levels (Sahara & Satria, 2025; Wang et al., 2024).

AI-powered chatbots are computer programs designed to engage users in natural language conversations through text or voice interfaces, utilizing Natural Language Processing (NLP) technology (Colace et al., 2018; Parina et al., 2022). In language learning, chatbots function as virtual tutors, conversation partners, instant feedback providers, and motivators (Winkler & Söllner, 2018; Li & Liang, 2019). Unlike traditional classroom settings, chatbots offer a non-judgmental, flexible, and always-available environment that allows learners to practice speaking without fear of embarrassment or time constraints.

Empirical evidence supports the effectiveness of AI chatbots in enhancing speaking skills. Li and Liang (2019) found that AI-based platforms significantly improved learners' oral proficiency through personalized practice and immediate feedback. Similarly, Chen et al. (2021) reported that virtual tutors and chatbots created immersive conversational environments that boosted learners' confidence and speaking fluency. Recent studies in EFL contexts further confirm that chatbot-

assisted learning leads to measurable gains in pronunciation, vocabulary, and overall communicative competence (Alenezi & Alenezi, 2025; Asrofi et al., 2025).

The advantages of chatbot integration in language learning are multifaceted. Chatbots provide accessibility anytime and anywhere, personalized learning paths, instant corrective feedback, and reduced speaking anxiety (Wiboolyasarin et al., 2024; Tai, 2024). Students often perceive chatbots as patient, non-judgmental partners that encourage more frequent practice and greater risk-taking in language production (Al-Jaf et al., 2024). These features align well with Communicative Language Teaching (CLT) principles, which emphasize meaningful interaction and authentic language use.

Nevertheless, the implementation of AI chatbots also presents certain challenges. Issues such as occasional inaccurate responses (hallucination), dependency on stable internet connections, initial technical difficulties, and limited contextual understanding have been reported (Zhai, 2023; Hockly, 2018). Additionally, the effectiveness of chatbots may vary depending on learners' digital literacy, proficiency levels, and the specific design of the application. In the Indonesian secondary school context, research on AI chatbot utilization for speaking development remains limited. Most existing studies focus on higher education or general language skills rather than specific speaking components at the high school level (Abimanto, 2023; Slamet, 2024). This creates a clear research gap, particularly regarding the use of free or accessible tools such as TalkPal AI in under-resourced educational settings.

This study is grounded in the theoretical framework that combines AI affordances with Communicative Language Teaching and Krashen's Affective Filter Hypothesis. It posits that AI chatbots can lower affective barriers while simultaneously providing abundant comprehensible input and output opportunities, thereby facilitating the development of EFL speaking skills. The conceptual framework illustrates the relationship between AI chatbot intervention (independent variable) and improvements in speaking performance and student perceptions (dependent variables).

METHOD

Design and Sample

This study adopted a mixed-methods approach to comprehensively investigate the effectiveness of AI chatbot-based English learning. The quantitative component employed a quasi-experimental nonequivalent control group design, which is appropriate in educational contexts where random assignment is not feasible (Hastjarjo, 2019; Sugiyono, 2019). In this design, the experimental group received an AI chatbot intervention using TalkPal AI, while the control group followed conventional teacher-centered instruction. Both groups were administered pre-test and post-test speaking assessments. The qualitative component involved semi-

structured interviews to explore students' perceptions and experiences, allowing for deeper insights and data triangulation.

The study was conducted at SMAN 11 Sidrap, South Sulawesi, Indonesia, during the 2025/2026 academic year. The population consisted of all eleventh-grade students. Two intact classes were purposively selected as samples: Class XI.A served as the experimental group (n=19 in the pre-test and n=18 in the post-test), while Class XI.B functioned as the control group (n=14 for both tests). The slight variation in participant numbers was due to attendance differences during data collection. These classes were selected based on their relatively similar English proficiency levels and the teacher's willingness to participate.

Instruments and Procedures

Speaking skills were assessed using a rubric adapted from Brown (2004), which evaluated five aspects: fluency, accuracy, pronunciation, vocabulary, and coherence. Each aspect was scored on a scale of 1 to 5, with a total maximum score of 25. The same speaking task format with equivalent difficulty levels was used in both the pre-test and post-test to ensure consistency. In addition, qualitative data were collected through semi-structured interviews with 10 purposively selected students from the experimental group, representing high, medium, and low achievers. The interviews consisted of seven open-ended questions focusing on students' experiences, perceived improvements, advantages, challenges, and overall perceptions of using TalkPal AI.

The data collection process was carried out in three stages. First, both groups completed the pre-test to establish baseline speaking ability. Second, the experimental group participated in four sessions of interactive voice-based practice using TalkPal AI over a two-week period, while the control group continued with conventional speaking instruction. Third, post-tests were administered to both groups, followed by interviews with selected participants from the experimental group. All speaking performances were audio-recorded and evaluated independently by two raters to ensure scoring reliability.

Data Analysis

Quantitative data were analyzed using both descriptive and inferential statistics. Descriptive statistics, including mean scores and standard deviations, were calculated to summarize students' speaking performance. Since the data were not normally distributed, as indicated by the Kolmogorov-Smirnov test, the non-parametric Mann-Whitney U test was used to compare post-test scores between the experimental and control groups at a significance level of 0.05. Qualitative data from the interviews were analyzed thematically following the framework proposed by Sugiyono (2019) and Miles and Huberman. The analysis involved data reduction, coding, theme development, and interpretation to identify key patterns and insights related to students' experiences with AI chatbot-based learning.

RESULT AND DISCUSSION

The pre-test results indicated that both groups had low initial speaking proficiency. The control group (n=14) obtained a mean score of 5.57, while the experimental group (n=19) achieved a slightly higher mean of 6.89. These low scores reflect limited baseline abilities in fluency, accuracy, pronunciation, vocabulary, and coherence among eleventh-grade students.

Table 1. Pre-test Results

Group	N	Mean	Total Score
Control	14	5.57	78
Experimental	19	6.89	131

After the treatment period, post-test results showed marked differences between groups. The control group (n=14) recorded a mean score of 12.57, indicating modest improvement from conventional instruction. In contrast, the experimental group (n=18) achieved a substantially higher mean of 23.17, demonstrating strong gains following TalkPal AI intervention.

Table 2. Post-test Results

Group	N	Mean	Total Score
Control	14	12.57	176
Experimental	18	23.17	417

Descriptive statistics further highlighted the superior performance of the experimental group. The mean gain in the experimental group was approximately 16.28 points, compared to only 7.00 points in the control group. Standard deviation in the experimental post-test was low (SD = 0.84), indicating consistent improvement across participants. Normality testing using the Kolmogorov-Smirnov test revealed that the data were not normally distributed ($p < 0.05$). Therefore, a non-parametric Mann-Whitney U test was conducted to compare post-test scores between groups. The test yielded a significance value of Asymp. Sig. (2-tailed) < 0.001, confirming a statistically significant difference between the experimental and control groups.

Qualitative findings from interviews with 10 students from the experimental group revealed overwhelmingly positive perceptions. All participants reported increased confidence and reduced anxiety when speaking English. Students frequently mentioned feeling “more comfortable” and “not afraid of making mistakes” due to the non-judgmental nature of the AI chatbot. Students also perceived clear improvements in specific speaking components. The most frequently mentioned gains were in fluency (reduced pausing), pronunciation (clearer articulation), and vocabulary range. Several students noted better grammar control and faster responses during conversations with TalkPal AI. Minor challenges were acknowledged, including unstable internet connections and initial difficulty navigating the application. However, these technical issues diminished after the

first session, and students adapted quickly. Overall, participants strongly recommended wider integration of AI chatbots in English classes.

The findings demonstrate that AI chatbot-based learning using TalkPal AI is significantly more effective than conventional methods in improving students' speaking skills. The large mean difference between experimental (23.17) and control (12.57) groups, supported by $p < 0.001$, provides strong empirical evidence of the intervention's impact (Alenezi & Alenezi, 2025). This effectiveness can be attributed to the interactive and repetitive nature of chatbot practice. Unlike traditional classrooms with limited speaking time, TalkPal AI allowed students to engage in unlimited voice-based conversations, promoting fluency and confidence (Li & Liang, 2019; Tai, 2024). Immediate feedback features further accelerated skill development in pronunciation and accuracy.

The results align with Communicative Language Teaching (CLT) principles by creating authentic, meaningful interaction in a low-anxiety environment. Krashen's Affective Filter Hypothesis is also supported, as reduced anxiety enabled students to produce more output and notice linguistic forms more effectively. Qualitative data reinforce quantitative results. Students' reports of increased motivation, confidence, and enjoyment confirm that AI chatbots address psychological barriers common among Indonesian EFL learners (Sahara & Satria, 2025; Wang et al., 2024). The non-judgmental AI partner encouraged risk-taking in language production.

However, the study also identified practical limitations. Dependence on stable internet and initial technical adaptation highlight infrastructure challenges in rural or semi-urban Indonesian schools. These findings echo concerns raised by previous researchers regarding digital divides in technology-enhanced learning (Hockly, 2018; Zhai, 2023). The mixed-methods design strengthens the credibility of the conclusions by triangulating statistical significance with rich student voices. This approach provides a more holistic understanding than purely quantitative studies common in the field.

Pedagogically, the results suggest that English teachers should integrate accessible AI tools like TalkPal AI as supplementary speaking practice. Schools need to improve internet infrastructure and provide teacher training to maximize benefits. Future research should address current limitations by using larger samples, longer intervention periods, and including objective pronunciation analysis through speech recognition software. Comparative studies across different regions of Indonesia would further validate the generalizability of these findings.

CONCLUSION

This study concludes that AI chatbot-based English learning methods, particularly using TalkPal AI, are highly effective in improving eleventh-grade students'

speaking skills at SMAN 11 Sidrap. The quasi-experimental results demonstrated a statistically significant difference between groups, with the experimental group achieving a post-test mean score of 23.17 compared to 12.57 in the control group ($p < 0.001$). These findings confirm that interactive voice-based practice with AI chatbots substantially enhances students' performance across key speaking components: fluency, accuracy, pronunciation, vocabulary, and coherence. Qualitative data from student interviews further support the quantitative results. Participants reported increased confidence, reduced speaking anxiety, greater motivation, and more positive attitudes toward English learning. The non-judgmental, flexible, and always-available nature of TalkPal AI created a supportive environment that encouraged active speaking practice beyond traditional classroom limitations. Students particularly valued the immediate feedback and repetitive practice opportunities provided by the chatbot. The integration of AI chatbots offers promising pedagogical implications for EFL instruction in Indonesian secondary schools. English teachers are encouraged to adopt accessible AI tools as supplementary resources to complement conventional methods. Future researchers should explore the long-term effects of AI chatbot use with larger samples, different proficiency levels, and additional language skills. Schools and policymakers should also prioritize digital infrastructure and teacher training to maximize the benefits of AI-enhanced language learning in the 21st century.

REFERENCES

Abimanto, R. (2023). Enhancing students' English speaking skills through AI chatbots. *Journal of Language and Technology, 5*(2), 45–53.

Aisa, S., Hermansyah, S., & Faradillah, N. (2025). The effectiveness of audio storytelling in improving students' listening skills in EFL classrooms. *IDEAS: Journal of Language Teaching and Learning, Linguistics and Literature* (Sinta 3).

Alenezi, A., & Alenezi, M. (2025). Emotional effects of AI chatbot interaction on student motivation and confidence. *Journal of Language Education Research, 14*(2), 88–102.

Alenezi, A., & Alenezi, M. (2025). The effectiveness of chatbot-assisted learning in improving speaking skills among EFL learners. *International Journal of Emerging Technologies in Learning, 20*(1), 34–49.

Al-Jaf, A., Al-Khalifa, H., & Rahman, M. (2024). AI chatbots as digital learning assistants in high schools. *Computers & Education, 15*(3), 99–110. https://doi.org/10.1016/j.compedu.2024.104567

Apriyani, R. D., & Hermansyah, S. (2026). Enhancing students' speaking fluency through roleplaying segmented scenes from Inside Out 2: A descriptive qualitative study. *INTERACTION: Jurnal Pendidikan Bahasa, 13*(1), 732–750.

Asrofi, M., Sari, D. P., & Wibowo, T. (2025). Integrating speech recognition technology with AI chatbots for EFL learners. *Language and Technology Journal, 12*(2), 14–28.

Besold, T. R., Schorlemmer, M., & Smaill, A. (2017). *Artificial intelligence and reasoning systems: Foundations and applications*. Cognitive Science Press.

Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Pearson Education.

Chen, X., Li, J., & Wang, Y. (2021). Virtual tutors and chatbots in English language learning. *TESOL Quarterly, 55*(3), 605–627. https://doi.org/10.1002/tesq.678

Colace, F., De Santo, M., & Greco, L. (2018). Chatbot-based learning systems: A new educational perspective. *International Journal of Digital Learning, 6*(2), 111–120.

Eisenring, C., Keller, R., & Braun, F. (2024). AI chatbot feedback and learner confidence in English speaking practice. *Journal of Computer-Assisted Language Learning, 38*(1), 15–27.

Faradillah, N., & Usman, M. (2025). Artificial intelligence as a learning tool to improve vocabulary skills of students at SMP Negeri 2 Panca Rijang. *Macca: Journal of Linguistic Applied Research, 2*(3), 45–58.

Harmer, J. (2001). *The practice of English language teaching* (3rd ed.). Pearson Education.

Hastjarjo, T. (2019). Rancangan eksperimen kuasi dan penerapannya dalam penelitian pendidikan. *Jurnal Penelitian Sosial dan Pendidikan, 4*(2), 125–137.

Hermansyah, S., Kencana, N., & Wahyuddin, N. R. (2026). Exploring EFL lecturers' and students' perceptions of Spotify as a digital tool for enhancing listening skills. *LangLit: An International Peer-Reviewed Open Access Journal, 11*(2), 120–135.

Hermansyah, S., Syahrir, L., & Sadapotto, A. (2025). Analyzing of the use of ELSA Speak application with a video tutorial approach among Generation Z students. *Macca: Journal of Linguistic Applied Research, 2*(2), 78–92.

Hockly, N. (2018). Technology and English language teaching. *ELT Journal, 72*(1), 98–102. https://doi.org/10.1093/elt/ccx052

Holmes, W. (2023). Artificial intelligence as a cognitive partner in education. *Oxford Review of Education, 49*(2), 155–170. https://doi.org/10.1080/03054985.2022.2145678

Li, J., & Liang, Y. (2019). AI-based language learning and proficiency improvement. *Journal of Educational Technology, 8*(1), 32–47.

Nasruddin, M. A., Syamsu T., Faradillah, N., & Manda, I. (2026). An analysis of factors affecting speaking confidence of secondary school students at UPT SMA Negeri 1 Sidrap. *INTERACTION: Jurnal Pendidikan Bahasa, 13*(Maret), 698–715.

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

Sahara, D., & Satria, R. (2025). Personalized learning through AI and its impact on speaking skills. *Journal of Digital Education Studies, 11*(1), 55–67.

Slamet, T. (2024). The impact of AI chatbot applications on EFL learners' communication skills. *Indonesian Journal of Educational Technology, 6*(4), 100–113.

Sugiyono. (2019). *Metode penelitian pendidikan: Pendekatan kuantitatif, kualitatif, dan R&D*. Alfabeta.

Tai, W. (2024). Generative AI chatbots and speaking fluency among elementary EFL learners. *International Journal of TESOL Studies, 6*(1), 122–136.

Wahyuddin, N. R., Yanti, N. E., Arnas, R., & Hermansyah, S. (2025). Utilization of artificial intelligence in EFL learning from a digital literacy perspective. *IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature*.

Wang, L., Chen, Y., & Zhou, F. (2024). Speaking accuracy and fluency among EFL learners: The role of emotional intelligence and motivation. *Journal of Applied Linguistics, 15*(2), 75–88.

Wiboolyasarin, S., Lee, C., & Park, J. (2024). Adaptive virtual tutors using text and voice chatbots for language learning. *Journal of Interactive Learning Environments, 32*(2), 177–190. https://doi.org/10.1080/10494820.2023.2289456

Winkler, R., & Söllner, M. (2018). Unleashing the potential of chatbots in education: A research agenda. *Journal of Educational Technology & Society, 21*(4), 275–288.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education, 16*(1), 1–27. https://doi.org/10.1186/s41239-019-0171-0

Zhai, X. (2023). Hallucination in large language models and its impact on education. *AI Education Review, 5*(2), 66–79.