The Efficacy of ChatGPT "Voice" Mode in AI-Assisted Language Learning: EFL Students' Viewpoints

Mai Hoàng Việt[1], Nguyễn Trường Giang[1*], Lê Thị Xuân Quỳnh[1], Lê Phương Anh Thư[1], Lê Nguyễn Quyền Trang[1]

Saigon University

***Corresponding Author**

DOI: https://doi.org/10.47772/IJRISS.2026.1026EDU0338

Received: 17 May 2026; Accepted: 22 May 2026; Published: 16 June 2026

ABSTRACT:

The increasing development of artificial intelligence (AI) in education has created new opportunities for enhancing English speaking practice among EFL learners. This study investigates students' perceptions of the benefits and limitations of using ChatGPT "Voice" mode for English speaking practice among English-major sophomores at Saigon University. A qualitative research design was adopted with 15 second-year English-major students participating voluntarily in the study. Data were collected through semi-structured interviews to explore learners' experiences, attitudes, and viewpoints regarding the use of ChatGPT "Voice" mode as a virtual speaking partner outside the classroom. The findings revealed that ChatGPT "Voice" mode provided students with a convenient, low-anxiety, and flexible environment for English speaking practice, although several limitations related to feedback quality, technical issues, and over-reliance on AI were also identified. The study suggests that ChatGPT "Voice" mode should be integrated into speaking instruction with appropriate teacher guidance, prompt-training support, and continuous progress monitoring to maximize its educational effectiveness.

Keywords: ChatGPT "Voice" Mode, Speaking skills, AI-assisted Language Learning, Semi-structured interviews, Learner Perceptions

INTRODUCTION

In this day and age, effective English speaking skills are globally recognized as essential for personal and professional advancement. Concurrently, the rapid evolution of Artificial Intelligence (AI) has significantly transformed human activities, with a particularly profound impact on the educational sector. Generative AI (GenAI) models, most notably ChatGPT developed by OpenAI, have emerged as robust tools capable of reshaping language learning by providing instant feedback and highly personalized academic resources (Jo, 2024).

Despite the acknowledged importance of oral communication, university students in Vietnam frequently encounter psychological barriers, such as shyness and a lack of confidence, which hinder their willingness to speak in class (Le & Tran, 2020). Specifically, sophomore students enrolled in the "Listening-Speaking Skills 4" course at Saigon University require extensive practice to develop the fluency, coherence, and idea generation needed for complex tasks, such as IELTS Speaking Parts 2 and 3. However, the limited instructional time in the physical classroom and the lack of an authentic 1-on-1 English-speaking environment outside of school severely restrict their opportunities to practice, leading to persistent challenges in maintaining extended conversations.

To fulfill this increasing need for independent practice, integrating AI-powered virtual speaking partners presents a promising pedagogical solution. Unlike traditional text-based tools, the "Voice" mode of ChatGPT offers a dynamic, conversational interface that closely simulates real-life communication with a native speaker (Kohnke et al., 2023). This feature empowers learners to practice anytime and anywhere, providing a low-anxiety

environment where they can converse freely and continuously without the fear of making mistakes in front of their peers.

LITERATURE REVIEW

ChatGPT – an AI-assisted agent

Artificial intelligence (AI) refers to computer-based systems designed to mimic aspects of human intelligence, allowing machines to learn from experience, understand information, solve problems, make decisions, generate creative outputs, and operate with a degree of autonomous behavior (Kaplan & Haenlein, 2019). In educational contexts, Zawacki-Richter et al. (2019) demonstrate that AI has increasingly been applied to support personalized learning, automate instructional processes, and provide adaptive learning experiences for students. Within language education, previous studies have highlighted the potential of AI in enhancing learner interaction, promoting autonomy, and improving speaking practice through technology-mediated communication (L. Chen et al., 2020; Nugroho et al., 2023). Among recent AI developments, ChatGPT, developed by OpenAI, has emerged as a prominent generative AI tool capable of producing contextualized and human-like responses in real time (*OpenAI*, 2025). Research by Jo (2024) indicates that ChatGPT can support self-directed learning by offering personalized feedback and flexible learning opportunities. In ESL/EFL settings, it has been applied to vocabulary expansion, writing assistance, cultural understanding, and classroom interaction enhancement (Fadillah et al., 2025; Lo et al., 2024). Furthermore, several studies report that the integration of ChatGPT contributes positively to learner confidence, engagement, and motivation in language learning (Candrasari et al., 2024; Malik et al., 2025). Nevertheless, concerns regarding information accuracy, excessive dependence on AI, and data privacy remain important considerations in educational implementation (Munaye et al., 2025).

Definition of ChatGPT's "Voice" Mode

ChatGPT's "Voice" mode is a feature that enables users to engage in natural spoken interaction with artificial intelligence across multiple languages through real-time communication (*OpenAI*, 2025). According to *OpenAI* (2025), the feature includes two interaction types: Standard Voice and Advanced Voice, which differ in processing systems and conversational capabilities. While Standard Voice converts speech into text before generating responses, Advanced Voice allows more direct audio interaction, creating a faster and more natural conversational experience. The Advanced Voice mode is also capable of recognizing users' speaking pace and generating emotionally responsive interactions, thereby enhancing the realism of communication (*OpenAI*, 2025). In addition, users can personalize their learning experience through various voice options, including Arbor, Breeze, Cove, Ember, Juniper, Maple, Sol, Spruce, and Vale (*OpenAI*, 2025). Compared to earlier AI communication systems, ChatGPT's "Voice" mode demonstrates stronger responsiveness, contextual understanding, and conversational naturalness (Kohnke et al., 2023). These features provide valuable opportunities for speaking practice through real-time interaction and conversational simulation (*OpenAI*, 2025). Previous studies suggest that repeated interaction with AI voice tools can support vocabulary development, communication reflexes, and language flexibility (Fitria, 2021), while immediate feedback on pronunciation, grammar, and expression may facilitate speaking improvement (Pratiwi et al., 2024). Moreover, researchers have reported positive effects on fluency, confidence, motivation, and anxiety reduction among language learners using AI-assisted speaking practice (Ding & Yusof, 2025; T. M. N. Nguyen & Pham, 2024). Nevertheless, research gaps still remain regarding long-term speaking development and the implementation of ChatGPT's "Voice" mode in authentic classroom contexts, particularly in Vietnam (Pangestu & Suwartono, 2024; Pham, 2024).

Benefits and Limitations of ChatGPT's "Voice" Mode in Developing Speaking Skills

AI-based speaking tools, particularly ChatGPT's "Voice" mode, have been increasingly recognized for their potential to support the development of English speaking skills through interactive and flexible communication practice. Regular interaction with AI systems may enhance learners' communication reflexes, vocabulary range, and sentence construction by providing continuous opportunities for spoken practice (Kohnke et al., 2023). In addition, ChatGPT's "Voice" mode supports various speaking activities, including role-play, thematic

discussions, and simulated conversations, which encourage more flexible language use in authentic communicative contexts (Pratiwi et al., 2024). Another significant advantage of AI voice tools lies in their ability to provide immediate feedback. According to Pratiwi et al. (2024), ChatGPT can instantly identify and correct errors related to pronunciation, grammar, and expression, thereby helping learners recognize weaknesses and improve more effectively. Similar findings were reported by Balla and Ramírez (2025), who argue that AI-generated feedback helps address several limitations commonly found in traditional language learning approaches. Previous studies have also highlighted the positive psychological effects of AI-assisted speaking practice. Ding and Yusof (2025) found that AI tools contribute to improvements in fluency, pronunciation, confidence, motivation, and anxiety reduction, while Bhar (2026) suggests that AI creates a non-judgmental speaking environment that lowers affective barriers for language learners. Earlier research by Fryer and Carpenter (2006) similarly indicates that learners often feel more comfortable communicating with chatbots than with peers or teachers. Furthermore, ChatGPT's "Voice" mode is considered accessible, flexible, and cost-efficient because learners can practice speaking anytime without depending heavily on instructors or paid speaking partners (Choi, 2025; Fryer & Carpenter, 2006). Such flexibility may also promote learner autonomy and sustained engagement in speaking practice. However, several limitations of AI voice tools have also been identified in the literature. Q. Chen (2025) notes that AI-generated feedback may sometimes be repetitive, generic, and insufficiently personalized, particularly regarding pronunciation and communicative effectiveness. Technical limitations, including the lack of audio storage and progress tracking, may further restrict learners' ability to monitor long-term improvement. In addition, concerns have been raised regarding over-reliance on AI. Kasneci et al. (2023) warns that excessive dependence on AI tools may weaken independent thinking and lead to cognitive offloading, while Tamayo et al. (2025) argue that uncontrolled AI use may negatively affect learner autonomy and active language production. Overall, the literature suggests that ChatGPT's "Voice" mode has considerable potential for supporting speaking development; nevertheless, its effectiveness depends on balanced and guided implementation within language learning contexts.

Recommendations for the Use of AI Chatbots in Language Learning

Previous studies suggest that AI chatbots should be integrated into language learning in a balanced and pedagogically guided manner rather than being used independently. L. A. D. Nguyen et al. (2025) report that although students generally hold positive attitudes toward AI chatbots because of their convenience, low-pressure learning environment, and personalized feedback, they also express concerns regarding limited corrective feedback, predictable responses, restricted discussion topics, and insufficient conversational depth. These findings indicate that the educational effectiveness of AI chatbots may depend not only on technological capabilities but also on how learners are guided to use them appropriately. In this regard, the literature emphasizes the importance of teacher mediation and AI literacy instruction in classroom implementation. According to Traga Philippakos and Rocconi (2025), students require explicit guidance on how to formulate and adjust prompts in order to obtain more relevant and accurate AI-generated feedback. Another major recommendation involves combining AI-supported learning with continuous teacher monitoring. Learners suggest that instructors should support speaking development through activities such as vocabulary logs, reflection tasks, and re-evaluation of AI-generated feedback, which may encourage more critical and effective use of AI tools in language learning. Furthermore, Revolusi and Febriandy (2025)argue that AI should function as a supplementary educational tool rather than a replacement for teachers, emphasizing that successful AI integration requires a balance between technological support and human interaction. Overall, the literature suggests that the effectiveness of ChatGPT's "Voice" mode depends not only on the technology itself, but also on teacher involvement, learner awareness, and appropriate instructional management.

Research gap

While there has been a considerable amount of research into the general applications of ChatGPT in education, empirical studies investigating its specific impact on speaking proficiency remain scant (Lo et al., 2024). Furthermore, there is a notable lack of qualitative inquiry delving into learners' in-depth personal experiences and attitudes toward adopting AI as a virtual speaking partner. Therefore, this study was conducted to profoundly explore students' perceptions, attitudes, and practical experiences regarding their use of ChatGPT's "Voice"

feature for independent English speaking practice outside the classroom, primarily drawing on qualitative data gathered through semi-structured interviews.

Research question

What benefits and limitations do students perceive when using ChatGPT Voice mode for English speaking practice?

METHODOLOGY

Pedagogical Setting and Participants

The primary setting of this research is the Faculty of Foreign Languages at Saigon University, a well-established public higher education institution located in Ho Chi Minh City, Vietnam. The target population includes second-year English-major students taking the "Listening–Speaking Skills 4" course. Among more than 200 students across six classes, one class consisting of 61 students was selected through convenience sampling due to its accessibility, availability, and efficiency in terms of time and cost (Memon et al., 2025; Obilor, 2023). From this class, 15 students voluntarily participated in the study. These participants were selected because they had already developed a fundamental English background during their first year and were continuing to improve their higher-level speaking proficiency.

Design of the study

This study adopted a qualitative research design to investigate the research question: "What benefits and limitations do students perceive when using ChatGPT Voice mode for English speaking practice?" The qualitative approach was considered appropriate because it enabled the researchers to gain an in-depth understanding of students' perceptions, attitudes, and personal experiences when interacting with ChatGPT "Voice mode" in English speaking activities. Rather than focusing on numerical measurements, this design emphasized the exploration of learners' individual viewpoints and interpretations regarding how the AI tool supported or hindered their speaking practice process. In line with the characteristics of qualitative inquiry, the study emphasized the collection of rich descriptive data from multiple sources in order to obtain a more comprehensive understanding of the research phenomenon (Creswell, 2009, p.176).

Data collection and analysis

Interviews

The data collection instrument employed in this study was the semi-structured interview. The primary rationale for utilizing this qualitative method is its capacity to strike a balance between structural consistency and conversational freedom (Ruslin et al., 2022; Sahoo, 2022). This methodological adaptability enabled the research team to profoundly explore the learners' perspectives, attitudes, and personal experiences regarding their use of ChatGPT's "Voice" feature for English speaking practice. By navigating through a predefined set of questions while organically adapting to the participants' spontaneous responses, the researchers could gather rich, authentic, and in-depth qualitative data without deviating from the core research focus.

To operationalize this approach, the interview protocol was systematically designed around core guiding questions, each accompanied by a series of open-ended probing questions. While the main questions ensured that the primary research objectives were addressed, the follow-up questions provided the vital flexibility to delve deeper into the students' underlying insights, thought processes, and emotional responses. This targeted yet adaptable structure allowed the researchers to clarify nuances, expand upon initial answers, and extract a highly comprehensive understanding of the learners' practical experiences with the AI tool.

Data analysis

The qualitative data derived from the semi-structured Google Meet interviews and the students' Google Form journals underwent a systematic thematic analysis following the six-phase framework established by Braun and Clarke (2006). To ensure the credibility, trustworthiness, and rigor of the findings, a strict investigator triangulation procedure was implemented during the coding and theme generation phases.

First, the primary researchers conducted manual open coding on the verbatim transcripts to identify recurrent patterns and conceptual categories. Independently, the identical textual dataset was processed through NotebookLM, which served as an automated, objective secondary coder to generate a parallel set of initial codes and provisional themes.

Following this independent phase, the codebooks from both the human researchers and the AI tool were cross-examined and contrasted to establish inter-coder consensus. Any ambiguities or divergent thematic categorizations were subsequently reviewed and resolved through a validation session with the research supervisor, who provided expert oversight to finalize the definitive thematic framework. This collaborative verification ensured that the extracted themes accurately and comprehensively reflected the participants' authentic experiences with ChatGPT "Voice" mode.

FINDINGS AND DISCUSSION

Findings

The semi-structured interviews were conducted, collected data, and analyzed to address the second research question: What benefits and limitations exist when students use the "Voice" feature to practice speaking in the context of the Listening - Speaking 4 course? These preliminary codes were grouped into key themes to better support the ideas of the tool's benefits and limitations, and the interviewees' recommendations.

Benefits

A convenient speaking partner

The "Voice" mode functions as a convenient speaking partner, providing a natural, comfortable, and accessible environment for regular practice. Participant 6 showed personal opinion: "Practicing speaking with this tool brings a more natural and comfortable feeling compared to communicating with friends." Learners highlighted the practicality of the application; for instance, Participant 3 said: "This tool is more convenient and faster than the recording function, because it feels like communicating with a real person." In addition, Participant 13 considered the "Voice" mode as an ideal environment for practicing English speaking, noting: "It's useful for those who lack an English-speaking environment." Furthermore, the technology effectively compensates for inadequate learning conditions. As Participant 9 remarked: "For those who face difficulties in learning English or do not have adequate study conditions, it is a suitable choice". Supporting this idea, Participant 1 also shared: "The tool is like having someone to speak English with."

Creating a comfortable and low-anxiety speaking environment

ChatGPT's "Voice" mode helps create a low-anxiety speaking environment that fosters learner confidence. Participant 5 preferred the tool for this reason, noting: "Some friends might be shy to communicate face-to-face with another person, so we can use ChatGPT to practice beforehand to build up our confidence." Participant 11 agreed: "We don't need to communicate with a real person, so it will reduce anxiety and we will talk more comfortably." Echoing this sentiment, Participant 6 stated: "Practicing speaking with the 'Voice' mode brings a more natural and comfortable feeling compared to communicating with friends." Participant 15 further emphasized: "Practicing with ChatGPT makes me less afraid of speaking English."

Constructive feedback

The "Voice" mode of ChatGPT offers constructive feedback that significantly aids in linguistic improvement and idea generation. Participant 2 appreciated the assistance in expanding vocabulary, expressing: "Especially in speaking, with such assistance, I have partly been able to fix common mistakes, which helps me avoid word repetition." Regarding content development, Participant 7 stated: "It can suggest ideas for me or different ways to develop my speech." Furthermore, learners valued the AI's ability to identify errors. Participant 9 added: "It also offers feedback like more vocabulary, or it suggests corrections for mistakes in my speech." Similarly, Participant 10 highlighted: "ChatGPT also contributes to correcting the speaker's mistakes and helps expressions become more natural."

Time and cost efficiency

This tool is considered time and cost-efficient, as it eliminates the need for paid instructors and allows flexible, immediate practice. Highlighting its convenience, Participant 3 reported: "I find it very convenient; it is faster than using the recording function because it feels like communicating with another person." Emphasizing the financial benefits, Participant 10 added: "Firstly, we don't have to spend time or money on hiring someone like a real instructor." Participant 15 also stated: "The tool is free, useful, and allows daily practice without taking too much time." In terms of flexible access, Participant 11 mentioned: "There is no limitation on practice time; learners can practice anytime without waiting for teachers."

Limitations

Incomplete feedback

The feedback provided by the tool is sometimes incomplete and lacks accuracy or depth. Sharing concerns about repetitive patterns, Participant 1 expressed concerns about repetitive and non-responsive feedback, stating: "There is feedback that is exactly the same even though I have made changes, like a pre-written script." Participant 3 similarly noted: "It frequently repeats the same questions from previous practice sessions." In terms of specific language components, accuracy remains a concern. Participant 4 pointed out: "For pronunciation and intonation feedback, its 'Voice' feature is not very accurate yet." Furthermore, learners seeking highly specialized evaluations found the tool insufficient. Participant 14 highlighted the lack of detailed evaluation, particularly in an IELTS context: "Sometimes not optimal at correcting vocabulary. For example, in IELTS speaking, it does not guarantee detailed corrections for every single criterion." This lack of reliability was echoed by Participant 6, who explained: "Sometimes I feel that my speaking scores are not really stable."

Technical issues and lack of history storage

Beyond pedagogical constraints, system stability and data retention were identified as notable drawbacks. Disruptions during practice sessions were reported by Participant 13, who pointed out: "When the Internet connection is unstable, the system responds slowly or is interrupted, which affects the practice experience." Additionally, the absence of an archiving feature hinders learners' ability to track their long-term progress. Participant 8 mentioned limitations in tracking spoken performance, explaining: "It doesn't record the audio, so I don't know if I spoke incorrectly or not... it cannot capture the pauses." The same participant further added: "Without audio recordings for comparison, it is difficult for me to identify my pronunciation mistakes accurately." The inconvenience of losing previous practice histories was also emphasized by Participant 12: "When I practice speaking again, it doesn't save everything... it resets, which is a bit inconvenient because I have to start over." Summarizing this challenge, Participant 1 stated: "It is difficult to review my improvement process across multiple practice sessions."

Over-reliance leading to a loss of independent thinking

Finally, the integration of artificial intelligence raises concerns regarding cognitive autonomy. Participant 2 warned about the potential negative impact of excessive dependence on the tool, stating: "They over-rely on AI

too much, and as a result, they are gradually losing their conscious ability to think and form their own words." The same participant further noted: "Dependence on AI may gradually make learners lose the ability to think and express ideas independently." To mitigate this risk, Participant 14 suggested a more balanced approach: "We should only teach how to use AI for the right purposes to help ourselves develop and progress, rather than over-exploiting it." In addition, Participant 13 pointed out another technological limitation: "When the Internet connection is unstable, the system responds slowly or is interrupted, which affects the practice experience."

Recommendations

Training students on proper prompt usage

To maximize the efficacy of the "Voice" mode, learners highlighted the necessity of explicit instruction on prompt formulation. Participant 14 recommended that teachers should guide students in this area, saying: "Lecturers can teach students how to correctly read, use, and regulate the prompts." This reflects a clear need to adapt and optimize prompts to align with specific lesson objectives. Participant 15 agreed and added "Teachers will teach us how to talk to ChatGPT so that it can give appropriate feedback." Ultimately, both participants recognized the tool's immense potential to enhance oral proficiency, suggesting that mastering prompt adjustments would lead to more satisfying and productive practice sessions.

Tracking practice progress via Google Forms or in-class checks

Beyond initial instruction, continuous monitoring of students' independent practice emerged as a crucial recommendation to sustain motivation. Participant 6 suggested establishing a systematic tracking method, declaring: "Lecturers can create a Google Form or something, for example, for some new words that I find while practicing speaking so I can write them down." Meanwhile, Participant 15 highlighted the importance of teacher oversight in evaluating the AI's effectiveness: "Re-check the students' skills to see if they have improved or not. Also, check the feedback that ChatGPT gives, to see if the student has improved upon that feedback." These insights indicate that rather than being replaced by AI, teachers play an essential complementary role by validating progress and providing ongoing pedagogical support.

DISCUSSION

Drawing upon the qualitative data, the findings of this research indicate that learners generally held positive perceptions toward integrating ChatGPT "Voice" mode into their English speaking practice, particularly valuing its convenience, comfort, and accessibility. The application functioned as a natural and supportive conversational partner that simulated authentic communication while mitigating affective barriers such as speaking anxiety. This is consistent with (Bhar, 2026), who highlights that AI technologies can expand speaking opportunities and lower emotional barriers in EFL contexts. Furthermore, the non-judgmental nature of AI interaction fostered increased comfort and confidence, which aligns with Fryer and Carpenter (2006) and Zou et al. (2023). The technology also promoted learner autonomy by enabling flexible and regular practice, an especially crucial benefit for those lacking an English-speaking environment. Moreover, the immediate feedback on vocabulary use and idea development was highly appreciated, which supports findings from Choi (2025). Finally, this virtual partner proved to be both time- and cost-efficient, offering accessible speaking practice without the financial burden of paid instruction.

Conversely, several limitations of the ChatGPT "Voice" mode were identified. Participants reported that the feedback was occasionally repetitive, incomplete, and lacking in accuracy, particularly in areas such as pronunciation, intonation, and IELTS-specific assessment criteria. This aligns with Q. Chen (2025), who notes that AI-generated feedback can be overly generic and insufficiently personalized. Technical constraints were equally apparent; the lack of audio recording and progress-tracking features restricted learners' ability to monitor their improvement over time. Additionally, there were significant concerns regarding over-reliance on AI, suggesting that excessive dependence might diminish independent thinking and language production ability. This concern is supported by Tamayo et al. (2025), who emphasize the risks of cognitive offloading and reduced learner autonomy in AI-supported learning environments. Additionally, although privacy issues were not directly

reported in the present study, ethical issues related to data security and personal information should still be considered when integrating ChatGPT "Voice" mode into language learning. Similarly, Alzamil et al. (2025) found that a majority of users expressed concern about privacy and the transparency of ChatGPT's data collection practices. Therefore, future implementation of AI-assisted learning tools should address privacy protection and ethical considerations more carefully to maintain learners' trust and confidence. Furthermore, though reduced human interaction was not explicitly identified in the current study, it remains important ethical concerns in AI-assisted language learning. The study Exploring the Impact of AI Chatbots on EFL Learners' Conversational Proficiency by Aliakbari et al. (2025) also suggested that AI chatbots cannot fully replace the sociocultural understanding, pragmatic competence, and contextual depth provided by human communication. The study further emphasizes that real-life connection remains irreplaceable for developing holistic communicative capacity and meaningful conversational skills.

In terms of recommendations, the data underscored a strong need for active teacher mediation and structured learning support when using ChatGPT "Voice" mode. First, it was suggested that educators should explicitly instruct students on how to design and adjust prompts to obtain more relevant and accurate AI feedback, highlighting the importance of AI literacy in language learning contexts (Traga Philippakos & Rocconi, 2025). Second, teachers are expected to play a key role in monitoring learners' progress by integrating independent AI use with classroom-based tracking activities such as vocabulary logs and feedback review. This reflects the view of Revolusi and Febriandy (2025), who argue that AI should function as a pedagogical support tool rather than a replacement for teacher guidance. Overall, these recommendations suggest that the ultimate efficacy of AI voice tools depends heavily on structured teacher involvement to ensure balanced, guided, and meaningful learning outcomes.

CONCLUSION

In conclusion, the findings of this study suggest that ChatGPT's "Voice" mode has considerable potential to support the development of English speaking skills among EFL learners. Through flexible and real-time interaction, the tool provided students with a convenient, low-pressure, and accessible speaking environment that encouraged regular practice, increased confidence, and promoted learner autonomy. Participants also perceived the tool as beneficial for improving vocabulary use, idea development, and overall speaking performance through immediate feedback and conversational practice. However, several limitations were also identified, including repetitive or insufficiently detailed feedback, technical constraints related to progress tracking, and the risk of over-reliance on AI. These findings indicate that while ChatGPT's "Voice" mode can function as an effective supplementary learning tool, its educational value depends greatly on balanced implementation and continuous teacher guidance. Therefore, integrating AI voice tools into speaking instruction should involve not only technological support but also pedagogical management, learner training, and critical awareness to ensure more meaningful and sustainable language learning outcomes.

Conflict of Interest: Authors have no conflict of interest to declare.

REFERENCES

1. Aliakbari, M., Barzan, P., & Sayyadi, M. (2025). Exploring the Impact of AI Chatbots on EFL Learners' Conversational Proficiency. Interdisciplinary Research in English Language Communication, 1(2), 66–80. https://doi.org/10.30470/irelc.2025.2058800.1022
2. Alzamil, L. M., Alhasani, A. M., & Alshehri, S. (2025). Privacy Concerns in ChatGPT Data Collection and Its Impact on Individuals. Future Internet, 17(11), 511. https://doi.org/10.3390/fi17110511
3. Balla, F. R. C., & Ramírez, E. A. V. (2025). Redefining speaking practice through ChatGPT's voice conversation: Towards autonomous and personalized practice. Pacha Revista de Estudios Contemporáneos Del Sur Global, 6. https://doi.org/http://dx.doi.org/10.46652/pacha.v6i18.430
4. Bhar, S. K. (2026). Artificial Intelligence in EFL Speaking Instruction: A Systematic Review of Pedagogical Design, Affective Conditions and Instructional Input. Encyclopedia, 6(4), 74. https://doi.org/10.3390/encyclopedia6040074

5. Candrasari, R., Makulua, J., Noviasmy, Y., & Makulua, K. (2024). GPT Chat: Useful or Not in Supporting Learning in Higher Education. International Journal of Language and Ubiquitous Learning, 2(2), 114–124. https://doi.org/10.70177/ijlul.v2i2.963
6. Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in Education: A Review. IEEE Access, 8, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510
7. Chen, Q. (2025). Students' Perceptions of AI-Powered Feedback in English Writing: Benefits and Challenges in Higher Education. International Journal of Changes in Education, 1–11. https://doi.org/10.47852/bonviewIJCE52025580
8. Choi, W. (2025). The role of changes in pronunciation ability and anxiety through Gen-AI on EFL learners' self-directed speaking motivation and social interaction confidence—A CHAT perspective. System, 133, 103788. https://doi.org/10.1016/j.system.2025.103788
9. Creswell, J. W. (2009). Research design: Qualitative, quantitative, and mixed methods approaches (3rd ed.). SAGE Publications. https://www.researchgate.net/profile/Rulinawaty-Kasmad/publication/342328830_THIRD_EDITION/links/5eed839ca6fdcc73be8d869e/THIRD-EDITION.pdf
10. Ding, D., & Yusof, A. M. B. (2025). Investigating the role of AI-powered conversation bots in enhancing L2 speaking skills and reducing speaking anxiety: A mixed methods study. Humanities and Social Sciences Communications, 12, 1223. https://doi.org/https://doi.org/10.1057/s41599-025-05550-z
11. Fadillah, E. N., Saridah, S., Kamilasari, M., & Nur'aida, A. (2025). Enhancing Prior Knowledge Development in English Language Education through Chat GPT-Assisted Learning. Journal of Languages and Language Teaching, 13(1), 390–401. https://doi.org/10.33394/jollt.v13i1.13356
12. Fitria, T. N. (2021). ARTIFICIAL INTELLIGENCE (AI) IN EDUCATION: USING AI TOOLS FOR TEACHING AND LEARNING PROCESS. Prosiding Seminar Nasional & Call for Paper STIE AAS, 4(1), 134–147.
13. Fryer, L., & Carpenter, R. (2006). EMERGING TECHNOLOGIES Bots as Language Learning Tools. Language Learning & Technology, 10(3), 8–14.
14. Jo, H. (2024). From concerns to benefits: A comprehensive study of ChatGPT usage in education. International Journal of Educational Technology in Higher Education, 21(35). https://doi.org/https://doi.org/10.1186/s41239-024-00471-4
15. Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. Business Horizons, 62(1), 15–25. https://doi.org/10.1016/j.bushor.2018.08.004
16. Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274
17. Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for Language Teaching and Learning. RELC Journal, 54(2). https://doi.org/10.1177/00336882231162868
18. Le, Q. D., & Tran, L. H. (2020). Speaking Anxiety and Language Proficiency among EFL at A University in Vietnam. International Journal of Social Science and Human Research, 03(09), 124–133. https://doi.org/10.47191/ijsshr/v3-i9-01
19. Lo, C. K., Yu, P. L. H., Xu, S., Ng, D. T. K., & Jong, M. S. (2024). Exploring the application of ChatGPT in ESL/EFL education and related research issues: A systematic review of empirical studies. Smart Learning Environments, 11(50). https://doi.org/https://doi.org/10.1186/s40561-024-00342-5
20. Malik, N., Kousar, A., & Arshad, K. (2025). Chat-GPT in Education: Learning Outcomes and Facilitating Knowledge Acquisition. Inverge Journal of Social Sciences, 4(2), 97–104. https://doi.org/10.63544/ijss.v4i2.131
21. Memon, M., Ramayah, T., Ting, H., & Cheah, J.-H. (2025). CONVENIENCE SAMPLING: A REVIEW AND GUIDELINES FOR QUANTITATIVE RESEARCH. Journal of Applied Structural Equation Modeling, 9(2), 1–15. https://doi.org/10.47263/JASEM.9(2)01

22. Munaye, Y. Y., Admass, W., Belayneh, Y., Molla, A., & Asmare, M. (2025). ChatGPT in Education: A Systematic Review on Opportunities, Challenges, and Future Directions. MDPI Open Access Journals. https://doi.org/https://doi.org/10.3390/a18060352
23. Nguyen, L. A. D., Dang, V. M. T., & Le, M. T. (2025). An Investigation into Students' Perception of Using Call Annie as a Virtual Conversational Partner. Proceedings of the AsiaCALL International Conference, 6. https://asiacall.info/proceedings/index.php/articles/article/view/129/94
24. Nguyen, T. M. N., & Pham, T. T. N. (2024). Voice Mode of ChatGPT: a New Support Tool for Teaching English in University Programs. EPRA International Journal of Research and Development (IJRD), 9(7). https://doi.org/10.36713/epra2016
25. Nugroho, A., Putro, N. H. P. S., & Syamsi, K. (2023). The Potentials of ChatGPT for Language Learning: Unpacking its Benefits and Limitations. REGISTER JOURNAL, 16(2), 224–247. https://doi.org/10.18326/register.v16i2.224-247
26. Obilor, E. (2023). Convenience and Purposive Sampling Techniques: Are they the Same? International Journal of Innovative Social & Science Education Research, 11(1), 1–7.
27. OpenAI. (2025, September 13). https://help.openai.com/en/articles/8400625-voice-mode-faq
28. Pangestu, H. P., & Suwartono, T. (2024). Exploring the Role of Artificial Intelligence (AI) Support in Assisting Students' English-Speaking Skills. Educalitra: English Education, Linguistics, and Literature Journal, 3. https://doi.org/https://doi.org/10.5281/zenodo.13134979
29. Pham, T. T. N. (2024). Investigating the effectiveness of practicing English conversation with ChatGPT in improving non-English majored students' English speaking skills at Nguyen Tat Thanh University. Journal of Science and Technology - Nguyen Tat Thanh University, 7(4). https://doi.org/https://doi.org/10.55401/b1rnww21
30. Pratiwi, N., Efendy, A. G., Rini, H. C., & Ahmed, N. A. (2024). Speaking Practice using ChatGPT's Voice Conversation: A Review on Potentials and Concerns. Journal of Language Intelligence and Culture, 6(1), 59–72. https://doi.org/http://dx.doi.org/10.35719/jlic.v6i1.149
31. Revolusi, P., & Febriandy, R. (2025). HUMAN-AI EDUCATIONAL COLLABORATION: FACING LEARNING CHALLENGES IN THE DIGITAL AGE. International Journal of Management Entrepreneurship Social Sciences and Humanities, 8(2), 289–312. https://doi.org/10.31098/ijmesh.v8i2.3229
32. Ruslin, R., Mashuri, S., Sarib, M., & Alhabsyi, F. (2022). Semi-structured Interview: A Methodological Reflection on the Development of a Qualitative Research Instrument in Educational Studies Ruslin. IOSR Journal of Research & Method in Education, 12(1), 22–29. https://doi.org/10.9790/7388-1201052229
33. Sahoo, R. (2022). Interview as a Tool for Data Collection in Educational Research. In Tools of Data Collection in Educational Research. Lucky International.
34. Tamayo, J., Sotto, J. C., Vicente, C., & Lim, M. (2025). The Impact of AI Tools on Student Engagement and Academic Performance. GEO Academic Journal, 6(1). https://doi.org/10.56738/issn29603986.geo2025.6.94
35. Traga Philippakos, Z. A., & Rocconi, L. (2025). AI Literacy: Elementary and Secondary Teachers' Use of AI-Tools, Reported Confidence, and Professional Development Needs. Education Sciences, 15(9), 1186. https://doi.org/10.3390/educsci15091186
36. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? International Journal of Educational Technology in Higher Education, 16. https://doi.org/10.1186/s41239-019-0171-0
37. Zou, B., Du, Y., Wang, Z., & Chen, J. (2023). An Investigation Into Artificial Intelligence Speech Evaluation Programs With Automatic Feedback for Developing EFL Learners' Speaking Skills. SAGE Open, 13(3). https://doi.org/10.1177/21582440231193818