Beyond tools: Reviewing AI's role in second language speaking skills and implications for classroom practice

Global Journal of Social Sciences Studies
Vol. 12, No. 2, 34-44, 2026.
e-ISSN: 2518-0614

Zhong Linling

School of Humanities and Foreign Languages, Zhejiang Shuren University, China.
Email: zhonglinling9@gmail.com

ABSTRACT

This review examines the transformative role of artificial intelligence (AI) in second language (L2) speaking instruction, analyzing its pedagogical potential and implementation challenges. It synthesizes findings from recent empirical studies to demonstrate how AI technologies, including automatic speech recognition (ASR), intelligent tutoring systems, and conversational agents, are reshaping L2 education. The analysis reveals that AI effectively addresses traditional limitations in speaking instruction, such as limited practice opportunities and delayed feedback, while also helping to reduce learners' speaking anxiety and improve fluency and pronunciation. However, the review also highlights critical considerations, including the variability in effectiveness across different AI applications and the need to balance AI-mediated support with human interaction. By exploring both the theoretical underpinnings and practical applications, this study provides evidence-based guidance for educators and researchers seeking to leverage AI to enhance L2 speaking competencies. It concludes that future efforts should focus on strategic integration to transition AI from a supplementary tool to an integral component of comprehensive language education.

Keywords: *Artificial intelligence, Automatic speech recognition, Conversational agents, Second language (L2), Speaking anxiety, Speaking instruction.*

DOI: 10.55284/gjss.v12i2.1921
Citation | Linling, Z. (2026). Beyond tools: Reviewing AI's role in second language speaking skills and implications for classroom practice. *Global Journal of Social Sciences Studies, 12*(2), 34–44.

Funding: This research is supported by Zhejiang Shuren University, Grant number (2025SK145).
Institutional Review Board Statement: Not applicable.
Transparency: The author confirms that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
Competing Interests: The author declares that there are no conflicts of interests regarding the publication of this paper.
History: Received: 18 June 2026/ Revised: 7 July 2026/ Accepted: 29 July 2026/ Published: 7 August 2026
Publisher: Online Science Publishing

Highlights of this paper

- This paper identifies key challenges in implementing AI-supported L2 speaking education.
- It positions AI as a complementary tool rather than a replacement for teachers or human interaction.
- The study provides implications for balanced AI integration in L2 speaking instruction.

1. INTRODUCTION

Second language (L2) education has undergone a profound transformation in recent decades, with artificial intelligence (AI) emerging as a pivotal force reshaping pedagogical practices and learning outcomes. While technology has long served as a supplementary tool in language classrooms (Kessler, 2018), the advent of sophisticated AI applications has fundamentally altered the paradigm of L2 instruction, particularly in the development of speaking skills. This shift represents more than mere technological enhancement; it signifies a paradigmatic evolution from viewing AI as external educational tools to recognizing their potential as integrated pedagogical agents capable of providing personalized, adaptive, and contextually responsive learning environments (Kayal, 2024).

Speaking is a complex language skill that requires learners to coordinate linguistic knowledge, communicative competence, and real-time cognitive processing (Sabnani & Renandya, 2019). Its development is often constrained in formal educational settings by limited opportunities for individual practice, insufficient exposure to authentic interaction, large class sizes, and delays in receiving feedback. These constraints have historically positioned speaking as perhaps the most challenging skill to develop and assess in formal educational environments (Aleksandrzak, 2011).

Contemporary AI technologies offer unprecedented solutions to these longstanding pedagogical challenges. AI is changing many aspects of education and is gradually being introduced to language education, with applications including natural language processing (NLP), automated writing evaluation (AWE), intelligent tutoring systems (ITSs), and automatic speech recognition (ASR). The integration of these technologies has created new possibilities for addressing the complexity of speaking skill development through personalized feedback, authentic interaction opportunities, and continuous assessment capabilities (Macías, Solano, Perero, Panimboza, & Lucas, 2024).

Recent empirical investigations have demonstrated the transformative potential of AI in L2 speaking instruction. Recent studies suggest that AI chatbots can provide additional opportunities for speaking practice and may support learners in achieving course-related language learning objectives (Macías et al., 2024). Also, some benefits (e.g. alleviating speaking anxiety and improving speaking skills) have been consistently documented across multiple studies (Rusmiyanto et al., 2023).

ASR technology has received particular attention in research on pronunciation and speaking development. By converting speech into text and providing immediate automated feedback, ASR systems allow learners to monitor their oral production and repeat speaking tasks without continuous teacher intervention. This capacity may promote learner autonomy and increase the frequency of pronunciation practice. Meta-analytic evidence suggests that ASR-supported instruction can produce positive effects on pronunciation development in English as a second or foreign language contexts, although the magnitude of these effects varies across instructional conditions and system designs (Ngo, Chen, & Lai, 2024).

However, the integration of AI in L2 speaking instruction also presents significant challenges and considerations. Questions surrounding the balance between human interaction and AI support, cultural sensitivity in AI responses, and the pedagogical preparation required for effective implementation remain central concerns for educators and researchers. It is observed that ChatGPT has not yet been effective in improving speaking skills,

whereas other AI-based language learning applications like "Elsa Speak" have been shown to be beneficial in developing speaking skills, particularly pronunciation ability (Almutairi & Alghammas, 2025) highlighting the need for careful selection and implementation of AI technologies based on specific pedagogical objectives.

The current state of research reveals both promising developments and critical gaps in our understanding of AI's role in L2 speaking education. While numerous studies have documented the effectiveness of individual AI applications, comprehensive frameworks for strategic implementation remain underdeveloped. Furthermore, questions regarding long-term retention, cross-cultural applicability, and the optimal balance between AI-mediated and human-mediated instruction require sustained empirical investigation.

This review addresses these gaps by providing a comprehensive analysis of current research on AI applications in L2 speaking instruction, examining both their pedagogical potential and implementation challenges. Moving beyond the traditional conceptualization of AI as supplementary tools, this study investigates how AI technologies are fundamentally reshaping the theoretical and practical foundations of speaking skill development. The analysis synthesizes empirical findings from recent research to identify effective practices, persistent challenges, and emerging trends in AI-enhanced L2 speaking instruction.

The significance of this review extends beyond academic discourse to encompass practical implications for language educators, curriculum developers, and educational technology specialists. As educational institutions worldwide grapple with the integration of AI technologies, evidence-based guidance for implementation becomes increasingly critical. This review aims to provide such guidance while identifying areas requiring further research and development.

2. LITERATURE REVIEW

2.1. Theoretical Foundations of AI in Second Language (L2) Speaking

The integration of AI in L2 speaking instruction is grounded in several theoretical frameworks that have evolved from traditional computer-assisted language learning (CALL) paradigms. Theoretical frameworks, such as the cognitive theory of multimedia learning and sociocultural theories of language development, provide a basis for integrating CALL and AIALL (Bahari, Han, & Strzelecki, 2025). These foundational theories have been instrumental in shaping our current understanding of how AI technologies may support L2 speaking development.

The cognitive theory of multimedia learning, originally proposed by Moreno and Mayer (1999), provides crucial insights into how AI-enhanced speaking environments can optimize learning through the coordination of visual and auditory processing channels. This theory is particularly relevant to AI speaking applications that combine speech recognition, visual feedback, and multimodal interaction interfaces. When applied to speaking instruction, the theory suggests that AI systems should be designed to reduce unnecessary cognitive load while supporting meaningful engagement with linguistic input and output.

Sociocultural theories of language development, rooted in Vygotsky and Cole (1978), and elaborated by scholars such as Lantolf (1994), emphasize the social nature of language learning and the importance of meaningful interaction in developing communicative competence. In AI-mediated speaking contexts, these theories highlight the potential for conversational agents to serve as mediating tools that provide scaffolded support within learners' zones of proximal development. The application of sociocultural theory to AI speaking instruction raises important questions about the authenticity of human-AI interaction and its role in developing communicative competence that can transfer to human interaction.

Activity theory has emerged as a particularly relevant framework for understanding AI chatbot integration in language learning contexts (Li, Zhou, Yin, & Chiu, 2025). Activity theory offers a framework of elements and

functions inside an activity system to accomplish desired objectives, providing a comprehensive lens through which to analyze the complex interactions between learners, AI systems, and instructional goals in contexts of L2 speaking development. This theoretical approach recognizes the dynamic, goal-oriented nature of language learning activities and emphasizes the importance of considering technological tools within broader activity systems rather than as isolated interventions.

The interaction hypothesis provides another relevant perspective is the interaction hypothesis for AI applications in speaking instruction (Wang, Zou, Du, & Wang, 2024). This hypothesis posits that meaningful interaction, negotiation of meaning, and feedback are essential components of successful L2 acquisition. Drawing on the interaction hypothesis, Wang et al. (2024) examined how different conversational AI chatbots influenced EFL learners' willingness to communicate, foreign language speaking anxiety, self-perceived communicative competence, and speaking performance. This theoretical grounding has important implications for designing AI systems that can facilitate meaningful negotiation of meaning and provide appropriate corrective feedback.

Control-value theory is also relevant to AI-enhanced speaking instruction because it explains how learners' perceptions of control and task value shape achievement emotions and engagement (Wu et al., 2021). This theory is particularly relevant given the emotional challenges often associated with speaking practice, including anxiety, frustration, and confidence issues that can either facilitate or hinder learning outcomes.

2.2. Current AI Technologies in L2 Speaking Development

The landscape of AI technologies supporting L2 speaking development encompasses a diverse range of applications, each offering unique affordances for addressing specific aspects of speaking competence. The literature commonly distinguishes several categories of AI technologies relevant to speaking instruction.

Conversational AI agents, commonly referred to as chatbots, represent perhaps the most visible and widely adopted AI technology in language learning contexts. Conversational agents are AI-enabled systems designed to engage users in text- or speech-based interaction. Earlier systems typically relied on rule-based responses, whereas more recent systems use neural language models to generate context-sensitive output. These systems range from rule-based chatbots that follow predetermined conversation paths to advanced neural language models capable of generating contextually appropriate responses to unexpected input.

ASR is widely used in AI-supported speaking instruction. ASR systems have advanced significantly in recent years, with improvements in accuracy, real-time processing capabilities, and integration with instructional platforms (Dhahbi, Saleem, Bourouis, Berrima, & Verdu, 2025). These systems enable automated analysis of learner speech, providing immediate feedback on pronunciation, fluency, and other speaking parameters. The integration of ASR with intelligent tutoring systems has created new possibilities for personalized speaking practice that adapts to individual learner needs and progress patterns.

Natural language processing (NLP) technologies underpin many AI applications in speaking instruction, enabling systems to process learners' linguistic input and generate context-sensitive responses. Advanced NLP capabilities include sentiment analysis, semantic understanding, may support functions such as sentiment detection, semantic analysis, and limited discourse-level processing. These technologies enable AI systems to provide more nuanced feedback on communicative effectiveness beyond simple accuracy measures.

Generative AI technologies, particularly large language models such as ChatGPT and similar systems, have introduced new possibilities for providing interactive conversational practice. These systems can generate contextually appropriate responses, maintain coherent dialogue across extended interactions, and adapt to different conversational styles and topics. However, findings regarding their effectiveness remain mixed. Some studies have

reported reductions in speaking anxiety and improvements in pronunciation, whereas others have found limited effects on broader dimensions of speaking performance.

Intelligent tutoring systems (ITSs) specifically designed for speaking instruction represent a more comprehensive approach to AI integration. These systems combine multiple AI technologies to provide personalized instruction, adaptive practice opportunities, and comprehensive feedback on speaking performance. Recent developments have focused on creating more naturalistic interaction patterns, incorporating gamification elements, and providing culturally responsive content.

2.3. Empirical Evidence on AI-Supported L2 Speaking Development

The empirical evidence regarding AI effectiveness in L2 speaking skills development indicates both promising outcomes and persistent limitations. A comprehensive analysis of AI-driven chatbots in L2 education has revealed significant positive outcomes across multiple dimensions of language learning (Xiu-Yi, 2024). The review reported improvements in learners' productive skills, particularly speaking and writing, due to features such as real-time feedback, anxiety reduction, and increased practice opportunities. These findings suggest that AI technologies may help address traditional constraints in speaking instruction, such as limited practice opportunities and delayed feedback mechanisms.

Reduced speaking anxiety is one of the most frequently reported benefits of AI-supported speaking practice (Ding & Yusof, 2025). Research focusing specifically on conversation bots has demonstrated their potential for enhancing L2 speaking skills and reducing speaking anxiety through mixed-methods investigations that combine quantitative measures of speaking performance with qualitative analyses of learner experiences. The reduction in speaking anxiety is attributed to several factors, including the non-judgmental nature of AI interactions, the availability of unlimited practice opportunities, and the reduced social pressure compared to human-to-human interactions.

However, the relationship between AI interaction and speaking anxiety is not uniformly positive across all contexts and learner populations. Some studies have identified situations where AI interactions may actually increase anxiety levels. The limited agentic outcomes may be attributed to students' lack of motivation or interest in learning English, particularly among those with lower language proficiency who experience anxiety and nervousness when they are unable to understand chatbots. This finding highlights the importance of considering individual differences in language proficiency, technological familiarity, and learning preferences when implementing AI technologies.

Long-term retention and transfer effects represent areas where empirical evidence remains limited. Most studies have focused on immediate or short-term outcomes, with few investigations examining whether improvements gained through AI-mediated practice are maintained over time or transfer to non-AI speaking contexts. This limitation represents a significant gap in current understanding and highlights the need for longitudinal research designs.

The effectiveness of AI technologies also appears to be moderated by implementation factors, including the quality of pedagogical integration, teacher training and support, and the broader instructional context. Studies have consistently found that AI technologies are most effective when integrated into comprehensive instructional approaches rather than implemented as standalone interventions. Learners interact with others primarily through language, and conversation creates learning in the class and out of the class, emphasizing the importance of connecting AI-mediated practice with broader communicative contexts.

3. METHODOLOGICAL APPROACHES AND CURRENT APPLICATIONS OF AI IN L2 SPEAKING

3.1. AI-Driven Assessment and Feedback Mechanisms

The integration of AI in L2 speaking assessment represents a paradigmatic shift from traditional evaluation methods toward automated, real-time, and comprehensive feedback systems. Contemporary AI-driven assessment mechanisms leverage sophisticated technologies to provide learners with immediate, detailed, and personalized feedback on multiple dimensions of speaking performance, fundamentally transforming the assessment landscape in language education (Alshehri, 2025).

The sophistication of current AI assessment systems extends beyond simple pronunciation evaluation to encompass comprehensive speaking competence analysis. Modern AI-powered platforms are capable of evaluating multiple dimensions of speaking performance simultaneously, including pronunciation accuracy, fluency, prosody, vocabulary usage, and grammatical complexity. These systems utilize advanced algorithms to analyze acoustic features, temporal patterns, and linguistic content to provide holistic assessments of speaking ability.

Speech recognition sensors and AI automatic evaluation applications have been particularly effective in addressing the challenges of oral English correction systems. These systems provide automated assessment of pronunciation, intonation, and speaking fluency, enabling learners to receive consistent feedback without the constraints of human evaluator availability.

The personalization capabilities of AI assessment systems represent a significant advancement over traditional evaluation methods. AI-powered pronunciation and fluency assessment platforms utilize machine learning algorithms trained specifically for non-native speaker evaluation, focusing on intelligibility, naturalness, and linguistic richness rather than native-like pronunciation alone. This approach recognizes the diversity of acceptable pronunciation patterns while still providing meaningful feedback for improvement.

Real-time feedback delivery is another crucial advantage of AI-driven assessment systems. Unlike traditional assessment approaches that require delayed evaluation and feedback provision, AI systems can provide immediate analysis of speaking performance, enabling learners to adjust their production patterns during practice sessions. This immediate feedback loop has been shown to enhance learning efficiency and motivation, as learners can observe the direct impact of their adjustments on assessment scores.

The integration of AI assessment systems with learning analytics platforms has enabled the collection and analysis of detailed performance data over time. These systems can track learner progress across multiple sessions, identify patterns in error production, and provide recommendations for targeted practice activities. The data-driven insights generated by these systems support both learner self-regulation and instructor decision-making regarding curriculum and intervention strategies.

3.2. Personalized Learning Pathways and Adaptive Technologies

The development of personalized learning pathways through AI technologies represents a fundamental transformation in how L2 speaking instruction is conceptualized and delivered. AI-powered adaptive technologies create individualized learning experiences that respond dynamically to learner needs, preferences, and progress patterns, moving beyond traditional one-size-fits-all approaches to embrace truly differentiated instruction.

AI-based instruction offers learners personalized and adaptive learning experiences, facilitating the analysis of their performance and enabling the identification of areas for improvement, as well as the provision of tailored feedback and practice materials. This personalization is achieved through sophisticated algorithms that analyze

multiple data points, including learner performance history, error patterns, learning pace, and engagement levels, to create customized instructional sequences that optimize individual learning outcomes.

The theoretical foundation for AI-driven personalization in language learning draws from multiple educational frameworks, including adaptive learning theory and intelligent tutoring systems design. AI offers transformative potential in language learning by providing personalized, adaptive instruction via tools like ITS and NLP, enabling biometric feedback and adaptive reading for improved comprehension and reduced anxiety. This multifaceted approach recognizes that effective personalization requires attention to cognitive, affective, and behavioral dimensions of learning.

Contemporary AI systems demonstrate remarkable sophistication in their ability to adapt content, pacing, and interaction styles to individual learner characteristics. An AI-driven teaching approach differentiates between adaptive L2 learning applications, which personalize content, and interactive applications, which encourage conversational practice and real-time feedback. This differentiation allows for targeted interventions that address specific aspects of speaking development while maintaining coherent overall learning progression.

The implementation of AI-powered interactive scaffolding has shown particularly promising results in supporting speaking performance development. Recent research examining the effects of AI-based interactive scaffolding on secondary students' speaking performance, goal setting, self-evaluation, and motivation revealed significant improvements across multiple outcome measures. The study's AI conversational agent provided personalized support through structured learning activities, including guided practice, contextual application exercises, and adaptive feedback mechanisms.

Personalized learning pathways enabled by AI technologies also address individual differences in learning preferences and anxiety levels. Learners engaging with interactive AI platforms receive instant feedback, comments, and alternative sentences, fostering continuous improvement and instilling confidence in their language abilities. This real-time support creates a supportive and encouraging atmosphere, promoting positive learning experiences that are particularly beneficial for learners with high speaking anxiety.

The adaptive capabilities of AI systems extend to content selection and sequencing, ensuring that learners encounter appropriately challenging material that promotes optimal learning without overwhelming cognitive resources. AI-powered chatbots have gained prominence in language education by offering personalized, adaptive learning experiences and continuous practice opportunities. These systems can adjust topic complexity, interaction patterns, and support levels based on ongoing assessment of learner performance and engagement.

The integration of learning analytics with personalized AI instruction enables continuous refinement of adaptive algorithms based on empirical evidence of effectiveness. AI chatbots, for example, can provide individualized instruction and feedback based on the needs and progress of each learner, revolutionizing the concept of personalized learning through data-driven optimization. This iterative improvement process ensures that personalization strategies become increasingly effective over time.

3.3. Interactive Conversational Agents and Immersive Learning Environments

The development of interactive conversational agents and immersive learning environments represents the cutting edge of AI applications in L2 speaking instruction. These sophisticated systems create authentic communicative contexts that simulate real-world interaction while providing the safety and support necessary for effective language learning. The evolution from simple chatbots to complex conversational partners has fundamentally transformed the possibilities for speaking practice and communicative competence development.

Contemporary conversational AI agents demonstrate remarkable capabilities in maintaining coherent, contextually appropriate dialogues across extended interactions. These systems utilize advanced natural language processing algorithms to understand learner input, generate appropriate responses, and maintain conversational coherence while adapting to individual learner characteristics and preferences. The sophistication of current conversational agents enables them to engage learners in meaningful exchanges that closely approximate authentic human interaction.

The pedagogical effectiveness of conversational agents in language learning contexts has been extensively documented through empirical research. AI-driven chatbots in L2 education have shown significant positive impacts on learners' productive and receptive skills, especially in speaking and writing, through features such as real-time feedback, anxiety reduction, and increased practice opportunities. These benefits are attributed to the unique affordances of conversational agents, including their availability for unlimited practice, non-judgmental interaction style, and ability to provide immediate corrective feedback.

The design of effective conversational agents for language learning requires careful consideration of multiple pedagogical and technological factors. Successful implementations incorporate principles from L2 acquisition theory, human-computer interaction design, and educational technology research. The agents must balance authenticity with pedagogical support, providing realistic conversational experiences while maintaining focus on learning objectives and providing appropriate scaffolding for skill development.

Immersive learning environments that incorporate conversational agents create rich contexts for authentic communication practice. These environments can simulate various real-world scenarios, from casual social interactions to professional communications, enabling learners to practice speaking in diverse contexts without the logistical challenges of arranging authentic interactions. The controlled nature of these environments allows for systematic practice of specific communication functions while maintaining the motivational benefits of meaningful interaction.

The integration of conversational agents with other AI technologies creates comprehensive learning ecosystems that address multiple aspects of speaking development simultaneously. For example, systems that combine conversational agents with ASR can provide immediate feedback on pronunciation while maintaining conversational flow, creating seamless learning experiences that address both accuracy and fluency development.
Recent advances in generative AI have significantly enhanced the capabilities of conversational agents in educational contexts. Large language models can generate contextually appropriate responses to unexpected learner input, maintain coherent dialogues across extended interactions, and adapt their communication style to match learner proficiency levels and preferences. However, research has revealed mixed results regarding the effectiveness of different AI systems for specific speaking skills development, highlighting the importance of careful system selection and implementation.

The potential of immersive learning environments extends beyond individual conversational practice to include collaborative learning experiences. AI-mediated environments can facilitate group interactions, peer collaboration, and community-building activities that enhance the social dimensions of language learning. These environments can create opportunities for learners to practice speaking in various social contexts while receiving AI-powered support and feedback.

The development of culturally responsive conversational agents represents an important frontier in AI applications for language learning. These systems must navigate complex issues of cultural representation, pragmatic appropriateness, and linguistic variation while maintaining pedagogical effectiveness. Recent research

has begun to address these challenges through the development of culturally adaptive AI systems that can adjust their interaction patterns to match learner cultural backgrounds and preferences.

The assessment capabilities integrated into conversational learning environments enable comprehensive evaluation of communicative competence beyond traditional accuracy measures. These systems can analyze discourse patterns, pragmatic appropriateness, conversational management skills, and other complex aspects of communicative ability that are difficult to assess through traditional methods. This comprehensive assessment capability supports both formative feedback during learning and summative evaluation of speaking development.

4. CONCLUSION

This review examined the emerging role of artificial intelligence (AI) in L2 speaking instruction by synthesizing recent research on major AI applications, including automatic speech recognition (ASR), conversational agents, generative AI systems, and intelligent tutoring systems. The findings indicate that AI technologies have expanded the possibilities for speaking instruction by addressing several persistent limitations of traditional learning environments, particularly restricted opportunities for oral practice, limited access to individualized feedback, and learners' concerns about negative evaluation during speaking activities. Through immediate feedback, repeated interaction opportunities, and personalized learning support, AI-mediated environments can provide additional resources for developing different dimensions of speaking competence, including pronunciation, fluency, and communicative confidence.

Nevertheless, the review also highlights that the pedagogical value of AI should not be understood as technology-driven improvement alone. The effectiveness of AI-supported speaking instruction varies considerably depending on the characteristics of the technology, instructional design, learner factors, and the context in which AI is implemented. Specialized applications designed for specific language learning purposes, such as pronunciation training, may offer more targeted support, whereas general-purpose generative AI systems may provide broader interaction opportunities but face limitations in areas such as systematic feedback, accuracy, and pedagogical alignment. Therefore, the question is not whether AI is effective in general, but rather under what conditions particular AI affordances can meaningfully support specific learning objectives.

The review further suggests that AI should be conceptualized as a complementary component within broader L2 speaking ecosystems rather than as a replacement for teachers or human interaction. Although AI systems can create accessible and low-pressure environments for speaking practice, human communication remains essential for developing pragmatic competence, interactional awareness, and the ability to negotiate meaning in authentic social contexts. Effective AI integration therefore requires a balanced approach in which AI-mediated practice supplements, rather than substitutes for, teacher guidance and human-to-human communication.

From a pedagogical perspective, educators should consider the alignment between technological affordances and instructional goals when selecting AI tools. AI applications should be incorporated through carefully designed tasks that encourage meaningful language production, reflection, and transfer beyond the AI environment. Teachers also require sufficient AI literacy to evaluate system capabilities, interpret AI-generated feedback critically, and guide learners in using AI responsibly and effectively. For educational institutions, successful AI implementation requires not only access to technology but also appropriate teacher training, curriculum integration, and consideration of ethical issues related to data privacy, algorithmic bias, and unequal access.

Several directions for future research emerge from this review. First, more longitudinal studies are needed to examine whether improvements gained through AI-mediated speaking practice can be maintained over time and transferred to human communication contexts. Second, future research should pay greater attention to individual

differences, including learners' proficiency levels, motivation, technological familiarity, and affective factors, which may influence how learners engage with AI systems. Third, cross-cultural investigations are needed to determine how learners from different educational backgrounds perceive and utilize AI-mediated speaking environments. Finally, future development of AI technologies should focus on improving contextual awareness, cultural responsiveness, and pedagogical adaptability.

Overall, AI represents a promising but complex development in L2 speaking education. Its contribution lies not simply in automating existing instructional practices, but in creating new forms of interaction, feedback, and learner engagement. Future progress will depend on moving beyond a technology-centered perspective toward a pedagogically informed model in which AI, teachers, and learners interact as complementary elements of a comprehensive language learning environment.

REFERENCES

Aleksandrzak, M. (2011). Problems and challenges in teaching and learning speaking at advanced level. *Glottodidactica. An International Journal of Applied Linguistics*, *37*(1), 37-48. https://doi.org/10.14746/gl.2011.37.3

Almutairi, R., & Alghammas, A. (2025). The effectiveness of using ELSA app on improving Saudi students' English-speaking skills. *International Journal of English Linguistics*, *15*(1), 140. https://doi.org/10.5539/ijel.v15n1p140

Alshehri, A. (2025). AI's effectiveness in language testing and feedback provision. *Social Sciences & Humanities Open*, *12*, 101892. https://doi.org/10.1016/j.ssaho.2025.101892

Bahari, A., Han, F., & Strzelecki, A. (2025). Integrating CALL and AIALL for an interactive pedagogical model of language learning. *Education and Information Technologies*, *30*(10), 14305-14333. https://doi.org/10.1007/s10639-025-13388-w

Dhahbi, S., Saleem, N., Bourouis, S., Berrima, M., & Verdu, E. (2025). End-to-end neural automatic speech recognition system for low resource languages. *Egyptian Informatics Journal*, *29*, 100615. https://doi.org/10.1016/j.eij.2025.100615

Ding, D., & Yusof, 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*(1), 1223. https://doi.org/10.1057/s41599-025-05550-z

Kayal, A. (2024). *Transformative pedagogy: A comprehensive framework for AI integration in education in explainable AI for education: Recent trends and challenges* Cham: Springer Nature Switzerland.

Kessler, G. (2018). Technology and the future of language teaching. *Foreign Language Annals*, *51*(1), 205-218. https://doi.org/10.1111/flan.12318

Lantolf, J. P. (1994). Sociocultural theory and second language learning: Introduction to the special issue. *The Modern Language Journal*, *78*(4), 418-420. https://doi.org/10.2307/328580

Li, Y., Zhou, X., Yin, H.-b., & Chiu, T. K. (2025). Design language learning with artificial intelligence (AI) chatbots based on activity theory from a systematic review. *Smart Learning Environments*, *12*(1), 24. https://doi.org/10.1186/s40561-025-00379-0

Macías, A. M. D., Solano, D. O. A., Perero, L. M. P., Panimboza, J. A. R., & Lucas, C. (2024). The potential of artificial intelligence to improve speaking skills in a second language (English) fluently. *Ciencia Latina: Revista Multidisciplinar*, *8*(3), 3826-3836. https://doi.org/10.37811/cl_rcm.v8i3.11592

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, *91*(2), 358-368. https://psycnet.apa.org/doi/10.1037/0022-0663.91.2.358

Ngo, T. T.-N., Chen, H. H.-J., & Lai, K. K.-W. (2024). The effectiveness of automatic speech recognition in ESL/EFL pronunciation: A meta-analysis. *ReCALL*, *36*(1), 4-21. https://doi.org/10.1017/S0958344023000113

Rusmiyanto, R., Huriati, N., Fitriani, N., Tyas, N. K., Rofi'i, A., & Sari, M. N. (2023). The role of artificial intelligence (AI) in developing English language learner's communication skills. *Journal on Education*, *6*(1), 750–757.

Sabnani, R. L., & Renandya, W. A. (2019). A comprehensive approach to developing L2 speaking competence. *English Language Teaching and Research*, *1*(1), 16-25.

Vygotsky, L. S., & Cole, M. (1978). *Mind in society: Development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Wang, C., Zou, B., Du, Y., & Wang, Z. (2024). The impact of different conversational generative AI chatbots on EFL learners: An analysis of willingness to communicate, foreign language speaking anxiety, and self-perceived communicative competence. *System*, *127*, 103533. https://doi.org/10.1016/j.system.2024.103533

Wu, C., Gong, X., Luo, L., Zhao, Q., Hu, S., Mou, Y., & Jing, B. (2021). Applying control-value theory and unified theory of acceptance and use of technology to explore pre-service teachers' academic emotions and learning satisfaction. *Frontiers in Psychology*, *12*, 738959. https://doi.org/10.3389/fpsyg.2021.738959

Xiu-Yi, W. (2024). Artificial intelligence in L2 learning: A meta-analysis of contextual, instructional, and social-emotional moderators. *System*, *126*, 103498. https://doi.org/10.1016/j.system.2024.103498