ARTIFICIAL INTELLIGENCE IN EFL SPEAKING INSTRUCTION: PEDAGOGICAL OPPORTUNITIES AND CHALLENGES

Hoshimova Fotima Lutfullo qizi
Master's Student, Department of Linguistics (English Language)
Uzbekistan State University of World Languages
Email: fatimsaydullaeva@gmail.com

Abstract: This thesis examines the pedagogical role of artificial intelligence (AI) in English as a Foreign Language (EFL) speaking instruction. The topic is relevant because conventional classrooms often provide limited individual speaking time, delayed feedback and insufficiently safe conditions for repeated oral performance. A focused synthesis of six peer-reviewed studies published between 2023 and 2026 shows that AI can extend interaction, provide automated feedback, personalize practice and support engagement and affective readiness. However, speech-recognition limitations, uneven suitability across proficiency levels and inconsistent evidence for higher-order communicative competence prevent AI from being treated as a replacement for teachers or authentic human interaction. The thesis therefore conceptualizes AI as a design-sensitive instructional resource embedded in teacher-guided communicative tasks.

Keywords: artificial intelligence, EFL speaking, speaking instruction, automated feedback, adaptive learning.

Speaking is one of the most demanding dimensions of foreign-language competence because learners must coordinate linguistic knowledge, real-time processing, pronunciation, fluency, interactional responsiveness and affective control. In many EFL settings, meaningful oral practice is restricted by class size, lesson duration and the teacher's limited capacity to provide immediate individualized feedback. The resulting gap between the practice learners need and the practice classrooms can realistically provide makes speaking instruction a particularly relevant area for AI-supported intervention.

Recent developments in conversational AI, automatic speech recognition, intelligent tutoring and generative language models have created new possibilities for addressing this methodological constraint. AI-based systems can act as conversation partners, evaluate speech, generate prompts, adapt task difficulty and allow learners to repeat oral performance without the social pressure that may accompany public classroom speaking. A systematic review by Du and Daniel (2024), based on 24 empirical studies published between 2017 and 2023, identified recurring benefits

related to speaking anxiety, willingness to communicate, fluency, pronunciation, motivation and confidence.

Yet the rapid growth of AI-assisted language learning has also produced a conceptual problem: studies often report fluency, pronunciation, engagement, anxiety or willingness to communicate as separate outcomes. This can make the literature appear as a collection of tool-specific effects rather than a coherent explanation of how AI contributes to speaking methodology. Bhar's (2026) review of 36 empirical studies helps bridge this gap by showing that benefits are strongest when AI interaction, feedback and instructional input are pedagogically structured, whereas evidence for sustained higher-order communicative competence is less consistent.

The purpose of this thesis is to synthesize recent evidence from the perspective of EFL speaking methodology and identify the mechanisms through which AI may support oral development, as well as the conditions that limit these benefits. Methodologically, it uses a focused narrative synthesis of six peer-reviewed studies published from 2023 to 2026: two systematic reviews and four empirical studies. The analysis is organized around interactional practice, feedback, adaptive and affective support, and pedagogical mediation. Its central argument is that AI is most useful not as an autonomous teacher but as a mediated resource aligned with communicative goals, learner level and teacher-designed tasks.

1. AI as an interactional and adaptive speaking environment

The first pedagogical opportunity of AI lies in the expansion of oral interaction. Traditional classroom speaking often depends on pair work, group work or short teacher-led exchanges, which means that the actual speaking time available to each learner may be limited. Conversational AI changes the quantity and accessibility of practice by providing an interlocutor that is available beyond lesson time and can sustain repeated exchanges. Du and Daniel (2024) identify interaction, accessibility and psychological comfort among the recurring affordances of AI-powered chatbots. These affordances are pedagogically significant because speaking development requires not only exposure to language but also opportunities to formulate, produce and reformulate spoken output.

Additional practice alone does not explain AI's value; adaptive practice is a stronger methodological contribution. Xu et al. (2026) compared three AI-supported conditions with traditional computer-assisted language learning among 414 postgraduate EFL learners. Personalized learning pathways produced the largest gains: fluency increased by 3.21 points, error density fell by about 37.5%, syntactic complexity rose by 1.8 clauses per T-unit, metacognitive self-monitoring improved, and speaking anxiety decreased by about 44%. These results are context-specific, but they illustrate a key principle: AI can respond to learner performance and vary practice rather than giving all learners an identical task sequence.

From a speaking-methodology perspective, personalization matters because oral development is rarely linear. Learners may display different combinations of strengths and weaknesses: one student may speak fluently but inaccurately, another may be grammatically accurate but hesitant, while another may avoid speaking because of anxiety. An adaptive system can, in principle, redistribute practice toward the learner's current needs. The pedagogical mechanism is therefore not simply "AI causes improvement"; rather, AI can increase the fit between task demands and individual performance, which may in turn create more productive practice. This interpretation also helps explain why Bhar (2026) describes AI as a design-sensitive instructional resource rather than an autonomous teaching agent.

2. Feedback, engagement and affective mediation

A second mechanism is the immediacy of feedback. In ordinary classroom conditions, teachers cannot continuously diagnose every learner's pronunciation, fluency or grammatical performance during every speaking attempt. AI-based speech evaluation systems can partially address this temporal limitation by analyzing spoken production and returning feedback during or immediately after practice. Such systems can offer immediate insights into speaking performance and support learners in identifying areas for improvement. Research has also indicated that AI-generated feedback may be perceived as useful by language learners and can be associated with measurable improvements in speaking performance over time. The methodological value of such systems is not that automated feedback is necessarily superior to teacher feedback, but that it can make formative feedback more frequent and available at the moment when learners are still attending to their own performance.

AI-assisted speaking also influences the learner's relationship with the task. Huang et al. (2025) examined 332 Chinese EFL learners and found that perceived AI affordances were significantly related to emotional, cognitive and behavioural engagement in spoken-English practice. In their structural model, AI affordances accounted for 54.7% of the variance in emotional engagement, 52.4% in cognitive engagement and 56.0% in behavioural engagement. These values should be interpreted as explanatory relationships within the model, not as proof that AI causally "improved" engagement by those percentages. Even so, the findings are important because sustained speaking development depends on whether learners actually invest attention, effort and repeated participation in oral tasks.

The affective dimension is equally significant. Speaking is a public and socially evaluative activity, and learners who fear mistakes may possess linguistic knowledge that they do not readily mobilize in interaction. Shi and Shakibaei (2025) used a quasi-experimental design to compare AI-integrated speaking instruction with traditional instruction and reported better speaking and social-emotional outcomes in the AI condition, together with significantly lower demotivation and shyness. Du and Daniel

(2024) likewise identify reduced speaking anxiety and increased confidence among recurrent findings in the chatbot literature, while Bhar (2026) concludes that affective benefits—particularly reduced anxiety and greater willingness to communicate—are among the more consistent patterns across the recent evidence base. The methodological implication is that AI may support speaking not only through linguistic input or correction, but also by modifying the emotional conditions under which practice occurs.

3. Pedagogical challenges and boundary conditions

The opportunities described above do not mean that AI interaction is pedagogically equivalent to human communication. Du and Daniel (2024) identify several recurring challenges: communication disengagement, limitations of speech recognition, insufficient fit across proficiency levels, and the possibility that chatbot content or task formats become monotonous. These problems are not merely technical. They affect the quality of the speaking activity itself. If a learner's speech is repeatedly misrecognized, feedback loses credibility; if interaction becomes predictable, the learner may produce language without genuinely negotiating meaning; and if the system is poorly calibrated to proficiency, practice may become either trivial or cognitively overwhelming. A deeper limitation concerns communicative authenticity. Speaking competence includes more than producing fluent and grammatically acceptable sentences. Real communication requires turn-taking, pragmatic interpretation, repair, sensitivity to interlocutor intentions, management of ambiguity and adjustment to social context. Bhar (2026) found that improvements in fluency, pronunciation and accuracy were frequently reported, particularly in structured task-based designs, but evidence for sustained development of higher-order communicative competence was more variable. This distinction is central to a rigorous interpretation of AI research. Improvement in measurable components of speaking should not automatically be equated with full communicative competence.

A methodological risk is over-attributing outcomes to technology itself. Huang et al. (2025) examined associations rather than direct causal effects, while Xu et al. (2026) studied a specific advanced learner population. These designs provide useful evidence but do not justify a universal claim that AI improves speaking independently of context. A more defensible conclusion is conditional: AI supports speaking when task quality, feedback, learner readiness and instructional integration are appropriate. The teacher therefore remains central in setting communicative goals, interpreting feedback and deciding when human interaction is necessary.

4. An integrated pedagogical model for AI-assisted EFL speaking

The reviewed evidence can be synthesized into a four-stage pedagogical model. First, AI expands the interactional space by increasing access to repeated speaking practice. Second, speech recognition and evaluation tools add a diagnostic layer

through immediate or near-immediate feedback. Third, personalization and lower social pressure can support engagement, self-monitoring and affective readiness. Fourth, these processes may contribute to linguistic outcomes such as fluency, pronunciation, accuracy and complexity, but only when they are embedded in purposeful instructional design. The model therefore treats learning outcomes as mediated rather than automatic.

The research gap lies in connecting outcomes that are usually measured separately. The reviewed evidence suggests a chain in which AI affordances expand interaction, automated feedback supports noticing and repair, personalization improves task–learner fit, and psychological safety helps sustain engagement. This chain can break when recognition is inaccurate, tasks lack communicative purpose, interaction becomes repetitive, or AI displaces rather than prepares learners for human communication.

For practical EFL instruction, this means that AI-supported activities should be designed as part of a progression rather than as isolated technology use. Learners may first rehearse with AI, receive feedback, repeat and self-correct, and then transfer the same linguistic and communicative resources into pair work, group discussion, presentations or other human-interaction tasks. Such sequencing preserves the advantages of individualized rehearsal while ensuring that speaking remains fundamentally social. In this model, AI extends the teacher's pedagogical reach but does not replace the teacher's role in task design, interpretation, interaction and evaluation.

Conclusion

The pedagogical value of AI in EFL speaking instruction lies less in technological novelty than in the learning conditions it can create. Across recent studies, four recurring opportunities emerge: expanded interaction, immediate feedback, adaptive practice and affective support. Evidence indicates potential gains in fluency, accuracy, complexity and speaking performance, while engagement, confidence, willingness to communicate, shyness and anxiety also matter. AI is therefore especially relevant where conventional classrooms cannot provide sufficient individualized speaking time or immediate feedback.

The evidence does not support treating AI as a substitute for teachers or authentic human communication. Speech-recognition limitations, uneven suitability across proficiency levels and less consistent evidence for higher-order communicative competence require caution. The strongest conclusion is conditional: AI can support EFL speaking when used as a mediated resource within structured, communicative and teacher-guided tasks.

In practice, AI should bridge private rehearsal and socially meaningful communication. It can increase practice and formative feedback, while classroom

interaction remains essential for negotiating meaning and developing pragmatic competence. Future research should examine long-term transfer to human interaction across proficiency levels and contexts.

References:

1. 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

2. Du, J., & Daniel, B. K. (2024). Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. Computers and Education: Artificial Intelligence, 6, 100230. https://doi.org/10.1016/j.caeai.2024.100230

3. Huang, F., Peng, D., & Teo, T. (2025). AI affordances and EFL learners' speaking engagement: The moderating roles of gender and learner type. European Journal of Education, 60(1), e70041. https://doi.org/10.1111/ejed.70041

4. Shi, W., & Shakibaei, G. (2025). Insights into the effectiveness of artificial intelligence-integrated speaking instruction in enhancing speaking skills and social–emotional competence as well as reducing demotivation and shyness. European Journal of Education, 60(3), e70174. https://doi.org/10.1111/ejed.70174

5. Xu, J., Ding, K., Yang, D., Wen, H., & Zou, T. (2026). Artificial intelligence as a catalyst for nonlinear skill development: Evaluating adaptive and collaborative AI tools for transforming academic speaking proficiency among EFL learners. Journal of Educational Computing Research, 64(5), 1187–1216. https://doi.org/10.1177/07356331261425496