

The Role of Artificial Intelligence in Enhancing English Speaking and Listening Skills in Higher Education

Aigerim A. Gilea¹, Katerina V. Melnova², Gyuzelya K. Temirgaliyeva³

^{1,2,3} Eurasian Humanities Institute named after A.K. Kussayinov, Kazakhstan



DOI : <https://doi.org/10.61796/icossh.v2i3.142>



Sections Info

Article history:

Submitted: April 15, 2025

Final Revised: May 20, 2025

Accepted: June 07, 2025

Published: June 14, 2025

Keywords:

Artificial Intelligence (AI)

EFL education

Speaking skills

Listening skills

Higher education

ABSTRACT

Objective: This study aims to analyze the role and effectiveness of Artificial Intelligence (AI) tools in enhancing English speaking and listening proficiency among EFL learners in higher education. **Method:** The research adopts a comprehensive literature review approach, synthesizing findings from recent empirical studies and scholarly discussions on AI applications such as Automated Speech Recognition (ASR), chatbots, Intelligent Tutoring Systems (ITS), and immersive technologies. **Results:** The analysis reveals that AI tools contribute significantly to improving pronunciation accuracy, speaking fluency, and listening comprehension, including intercultural and metacognitive dimensions. Moreover, these tools enhance learner confidence and engagement. Despite these benefits, limitations persist in the development of higher-order interactional competence and the consistent use of active vocabulary. **Novelty:** This study provides a novel contribution by concurrently examining the impact of AI on both speaking and listening skills, integrating recent advancements such as Generative AI, and offering critical insights for educators and researchers on optimizing AI-assisted language learning environments.

INTRODUCTION

Developing proficiency in English, particularly in speaking and listening skills, remains a crucial objective in higher education globally. However, many English as a Foreign Language (EFL) learners encounter significant obstacles in attaining adequate competence, especially in spoken communication. A major contributing factor lies in the limitations of traditional pedagogical approaches. These conventional classroom settings often fail to provide the necessary volume of individualized speaking practice required for substantial improvement [1]. Moreover, the delivery of timely, specific, and corrective feedback on learners' spoken output—critical for skill development—is frequently hampered by practical constraints such as large class sizes and limited instructional time. This persistent gap between the pedagogical need for extensive speaking and listening practice and the capacity of traditional methods to meet this demand highlights a critical challenge in language education [2].

In response to these challenges, Artificial Intelligence (AI) has emerged as a promising avenue for enhancing English language learning, particularly since 2020. The fundamental goal is to leverage AI to create learning environments that are more immersive, personalized, engaging, and ultimately more effective than those afforded by traditional methods [3]. A variety of AI technologies are being deployed, including Automated Speech Recognition (ASR), conversational agents (chatbots), Intelligent Tutoring Systems (ITS), and immersive platforms utilizing Virtual Reality (VR). The core potential of these tools lies in their ability to offer scalable, personalized practice environments. They provide learners with opportunities for repeated practice, often

alleviating the anxiety associated with live performance, and crucially, offer immediate, targeted feedback on diverse aspects of speech production and listening comprehension. By automating key elements of practice and feedback, AI directly addresses the limitation of insufficient individualized attention that is common in many language learning contexts [4].

The current landscape of AI in language education demonstrates both convergence—through the integration of multiple technologies into comprehensive platforms—and specialization, with tools focusing on specific sub-skills such as pronunciation (e.g., ELSA Speak) or contextual practice (e.g., EnglishCentral). Furthermore, the rapid advancement and increased accessibility of Generative AI (GenAI) since 2022–2023 represent a potentially transformative development. Unlike earlier, more scripted AI tools, GenAI enables flexible, dynamic, and open-ended interactions, potentially providing more nuanced feedback and closely simulating human conversation [5]. These advancements significantly enhance the potential for AI to impact a broad range of language skills. Research indicates that ASR-based tools are effective for improving pronunciation, while chatbots enhance fluency and reduce speaking anxiety. AI is also being used to improve intonation, rhythm, grammatical accuracy, and interactional skills—contributing not only to linguistic competence but also to learner confidence. Recent developments further address listening comprehension, including tone and emotion recognition, intercultural listening competence through exposure to diverse accents, support for metacognitive listening strategies, and authentic, performance-based assessment [6].

Despite these advancements, challenges persist. These include limitations in developing sophisticated interactional competence, inconsistent effects on active vocabulary use, the complexity of providing nuanced feedback for listening comprehension errors, and the difficulty of preparing learners for real-world, unpredictable communication scenarios. Ethical concerns related to data privacy, security, algorithmic bias, and inclusive design also remain paramount as these technologies become more embedded in educational contexts. Ongoing research continues to investigate the effectiveness of these tools, the transferability of skills to real-life communication, and the development of evidence-based implementation models [7].

Accordingly, the integration of AI holds immense potential to reshape language learning into a more interactive, strategic, and self-regulated process. This study distinguishes itself by focusing specifically on the synergistic potential and documented impact of AI tools designed to enhance both English speaking and listening skills concurrently within higher education settings, reflecting the integrated nature of language competence [8].

Therefore, the objective of this research is to comprehensively analyze the role and effectiveness of various Artificial Intelligence tools in enhancing English speaking and listening skills among higher education learners, examining specific skill improvements, learner experiences, and the implications of recent technological advancements such as Generative AI [9].

RESEARCH METHOD

To address the research objective concerning AI's impact on language acquisition, a methodological approach is employed that ensures a solid and sound analysis, yielding credible and potentially replicable findings. This section outlines the typical components of such an investigation within the higher education context [10].

Participants

Studies in this area typically involve English as a Foreign Language (EFL) learners enrolled in higher education institutions. Defining specific characteristics such as proficiency level, native language background, age range, and field of study is crucial for understanding the population to whom findings apply. Appropriate sampling methods, like convenience, purposive, or random assignment (for comparative studies), are selected based on the research questions and design.

Instruments and Procedures

Assessing the impact of AI tools commonly involves a combination of instruments and procedures.

Instruments typically include:

Specific AI applications identified as interventions, such as selected chatbots, ASR-based pronunciation trainers like ELSA Speak, ITS platforms, VR simulations, or AI-driven listening platforms.

Pre- and post-tests designed to measure changes in specific speaking skills (e.g., pronunciation accuracy, fluency metrics, grammatical accuracy, task completion) and listening skills (e.g., comprehension scores, ability to identify main ideas or specific details).

Qualitative data collection tools, including questionnaires or surveys, to gauge learner perceptions regarding motivation, confidence, and user experience.

Semi-structured interviews or focus groups to provide in-depth insights into learners' experiences, perceived benefits, and challenges encountered while using the AI tools.

System logs and learning analytics data generated by the AI tools, offering objective measures of usage patterns, interaction frequency, time-on-task, specific errors, and progress within the tool.

Observation protocols, potentially used in blended learning scenarios to observe the integration of AI tool usage with other learning activities.

Procedures generally entail:

A defined intervention period during which participants engage with the selected AI tool(s) for a specified duration and frequency.

Administration of pre-tests before the intervention and post-tests upon completion to measure gains.

Collection of qualitative data through surveys or interviews, conducted either during or after the intervention period.

Monitoring participant engagement and collecting usage data via AI platform analytics.

In comparative studies (e.g., AI intervention vs. traditional instruction), procedures involve assigning participants to conditions and ensuring comparable instruction time and content where relevant.

Data Analysis

The collected data requires analysis using methods appropriate for the data type and research questions.

Quantitative Data analysis involves statistical techniques. Paired-sample t-tests are suitable for pre/post-test comparisons within a single group, while independent-sample t-tests or ANOVAs are used when comparing different groups to determine statistically significant changes in proficiency metrics. Correlation analyses can explore relationships between usage patterns derived from logs and skill improvement.

Qualitative Data analysis typically employs thematic analysis. This approach is applied to interview transcripts, focus group data, and open-ended survey responses to identify recurring themes and patterns in learner perceptions, reported strategies, perceived benefits, and challenges associated with AI tool usage.

Mixed Methods approaches integrate findings from both quantitative and qualitative data. Qualitative data might be used to explain or contextualize quantitative results, such as clarifying why confidence increased alongside fluency scores. Learning analytics data is analyzed to identify usage patterns correlated with learning outcomes.

RESULTS AND DISCUSSION

Results

The application of Artificial Intelligence tools in higher education, as analyzed in the literature, yields specific outcomes related to the enhancement of English speaking and listening skills. These findings, derived from studies employing methodologies similar to those outlined above, are presented below.

Speaking Skills Enhancement:

Pronunciation: Automated Speech Recognition (ASR)-based tools effectively improve pronunciation by providing immediate, detailed feedback on both individual sounds (segmental features) and aspects like intonation, stress, and rhythm (suprasegmental features). Specialized applications such as ELSA Speak target these features directly, while platforms like EnglishCentral integrate pronunciation practice within meaningful video contexts. Furthermore, pronunciation feedback is a common element in broader systems like Intelligent Tutoring Systems (ITS).

Fluency: Chatbots and virtual assistants are frequently linked to improvements in speaking fluency. They offer extensive opportunities for low-stakes, repetitive conversational practice, which aids in building automaticity and reducing hesitation. Notably, the non-judgmental nature of AI interaction often alleviates speaking anxiety. Studies involving tools such as LingoChamp and Smalltalk2Me have reported significant gains in fluency metrics.

Intonation and Rhythm: Beyond basic pronunciation, specialized AI tools like ELSA Speak focus specifically on suprasegmental aspects. They often employ visual aids

(e.g., pitch contours) and auditory models to assist learners in perceiving and replicating more natural-sounding speech patterns.

Vocabulary: The impact of current AI tools on the active use of vocabulary during spontaneous speech appears less consistently documented than gains in pronunciation or fluency. While general vocabulary knowledge may improve, the direct transfer to effective vocabulary deployment in real-time speaking tasks is often less clearly evidenced. Immersive technologies like VR and AR show potential in this area, but their effectiveness likely depends heavily on specific task design and integration.

Grammatical Accuracy: AI tools incorporating ASR and Natural Language Processing (NLP) capabilities can contribute to improving grammatical accuracy in spoken English. Some platforms are designed to detect grammatical errors in learners' speech and offer feedback on sentence structure. While ITS systems and specific grammar-focused apps have demonstrated improvements in accuracy, these gains are sometimes perceived as less substantial than those observed for fluency or pronunciation. Advanced chatbots and ITS might also provide explicit grammatical explanations upon request or error detection.

Interactional Skills: Chatbots and other conversational agents serve as platforms for practicing fundamental conversational mechanics, such as initiating turns, responding appropriately, and managing simple dialogue flow. Engaging with these tools may enhance learners' willingness to communicate and develop certain aspects of pragmatic competence. Additionally, VR simulations provide valuable, context-rich environments for practicing interactional skills needed in specific social or professional situations, while AI-enhanced Project-Based Learning (PBL) can also foster communicative competence through collaborative tasks.

Overall Proficiency and Confidence: Cumulatively, improvements achieved in various sub-skills, especially pronunciation and fluency, contribute to gains in learners' overall assessed speaking proficiency. Perhaps one of the most consistently reported positive outcomes is a significant increase in learner confidence. This boost seems strongly linked to the provision of ample, low-pressure practice opportunities and the reception of constructive, non-threatening feedback from AI systems. Enhanced confidence, in turn, actively encourages learners to take more communicative risks, experiment with language, and participate more readily in speaking activities.

Listening Skills Enhancement:

Nuanced Comprehension: Recent AI developments increasingly emphasize the interpretation of subtle nuances in communication, such as emotion, tone, and speaker intention. Advanced systems utilize paralinguistic feature analysis, enabling learners to practice recognizing cues like sarcasm or politeness levels through simulated, emotionally varied dialogues.

Intercultural Listening: AI systems trained on diverse linguistic corpora can expose learners to a wide range of accents, dialects, and discourse styles, including varieties of English from non-native speakers. This exposure fosters intercultural listening competence and better prepares learners for global communication realities.

Adaptive Learning: Although still in early stages, the integration of technologies like eye-tracking and biometric feedback offers possibilities for real-time adjustment of listening content based on detected cognitive load or emotional responses, potentially slowing or simplifying input when confusion arises.

Metacognitive Development: AI is proving valuable in supporting the development of metacognitive listening strategies, such as predicting content, monitoring understanding, and evaluating success. Intelligent tutoring systems prompt learners to reflect on performance, select strategies, and track progress. Dialogic self-assessment tools further guide learners through post-listening reflections.

Learning Analytics: The analysis of user behavior within AI platforms – tracking playback patterns, pauses, and error trends – provides valuable insights into specific listening weaknesses. These analytics, often visualized through intuitive dashboards, empower both learners and educators to set informed goals and monitor progress across various dimensions of listening competence.

Performance-Based Assessment: AI facilitates a shift away from traditional discrete-item testing towards more authentic performance-based listening assessments. Learners can engage in open-ended tasks like summarizing spoken input or completing related conversational turns, with AI evaluating responses for comprehension depth and accuracy, thereby better reflecting real-world communication and integrating skill practice.

Context-Aware & Task-Based Listening: Emerging context-aware AI listening companions can curate customized content aligned with a learner's environment, interests, or ongoing language goals, significantly increasing ecological validity. Task-based listening embeds comprehension within meaningful activities (e.g., listening to instructions to complete a task), evaluating functional understanding rather than just abstract comprehension.

Collaborative Listening: AI moderation can now support shared listening experiences. In these environments, AI might facilitate turn-taking, highlight gaps in understanding among participants, provide differentiated scaffolding, and thus promote peer interaction and the cooperative development of listening strategies.

Affective Personalization & Diagnostics: AI tools are beginning to assess affective states like listening anxiety or cognitive load (potentially using facial recognition, voice analysis, or biometrics in research settings) to adapt task difficulty or support strategies. Furthermore, AI can perform language diagnostics in listening, identifying systematic comprehension breakdowns (e.g., issues with cohesive devices or inference) to inform targeted remedial interventions.

Multimodal Support: AI systems increasingly offer multimodal listening support, synchronizing audio with features like subtitles, gesture recognition in videos, visual reinforcement of vocabulary, or even AI-generated summaries and visual mind maps, catering to diverse learning styles and needs.

LMS Integration & Pedagogical Support: AI-driven listening tools are progressively being integrated into Learning Management Systems (LMS), enabling

instructors to track data, assign tasks, and align activities with curriculum outcomes, particularly in blended and hybrid models. AI-informed pedagogical dashboards can also offer teachers class-wide insights and suggest differentiated instructional strategies.

Discussion

The presented findings indicate that AI tools offer significant potential for enhancing specific components of English speaking and listening skills in higher education, although their documented effectiveness varies across different skill sets. Interpreting these results reveals several key insights [11]. Firstly, AI demonstrates considerable strength in improving the more mechanical aspects of language production, particularly pronunciation and fluency. This efficacy likely stems from the capacity of tools like ASR and chatbots to provide immediate, repetitive, and low-stakes practice opportunities, coupled with targeted feedback. The consistently reported boost in learner confidence stands out as a crucial outcome, potentially acting as a catalyst for increased participation and communicative risk-taking, both vital for overall language development [12].

However, the results simultaneously highlight certain limitations. The impact of current AI on higher-order skills—such as the fluid, active deployment of vocabulary in spontaneous speech or the development of sophisticated, nuanced interactional and strategic competence—appears less pronounced or consistently documented. This suggests that while AI excels at structured practice and feedback on discrete language elements, replicating the complexity and unpredictability inherent in authentic human interaction remains a significant challenge. Similarly, while advancements in AI for listening show promise, especially regarding nuance detection, intercultural exposure, and metacognitive support, the sophistication of feedback mechanisms explaining listening errors still lags behind those available for speaking or writing. Furthermore, the reliance of many current systems on controlled acoustic or interactional conditions may limit the effective transfer of acquired skills to noisy, unpredictable real-world communication scenarios [13].

The advent of Generative AI holds considerable potential to address some of these limitations. Its capacity for more dynamic, flexible, and context-aware interactions could lead to improved feedback quality and more realistic simulations of authentic dialogue. Moreover, the trend towards context-aware, task-based, collaborative, and multimodal AI systems signifies an important pedagogical shift towards more holistic, engaging, and ecologically valid learning experiences. The integration of affective personalization and diagnostic capabilities further enhances the potential for deeply tailored learning support [14].

Interpreting these results necessitates a critical acknowledgment of the ethical considerations surrounding AI in education. Issues concerning data privacy, learner security, algorithmic bias (particularly regarding the equitable treatment of diverse accents and linguistic backgrounds), and the need for operational transparency and inclusive design are paramount. Ensuring equitable access and consistent effectiveness

across diverse learner populations remains a critical priority for development and implementation.

Looking forward, the findings suggest several directions for future research. More rigorous investigation is required into the long-term transfer of skills developed using AI tools to authentic, real-world communication contexts. There is also a pressing need for standardized evaluation methods to move beyond anecdotal evidence and establish robust, evidence-based models for assessing the effectiveness and scalability of different AI interventions [15]. Further exploration should examine how varying types of AI scaffolding impact the transfer of listening skills, how adaptive listening tools affect learner retention and motivation in online or blended environments, and the potential of AI to foster long-term listening autonomy through enhanced metalinguistic awareness. Critically, understanding how best to integrate these powerful tools within broader pedagogical frameworks that genuinely maximize learner growth, equity, and engagement is essential. Continued research must also focus on improving AI's ability to handle interactional nuance, provide more sophisticated and explanatory listening feedback, and proactively address the complex ethical challenges inherent in its deployment.

CONCLUSION

Fundamental Finding : The integration of Artificial Intelligence stands as a pivotal development in addressing the long-standing challenges of enhancing English speaking and listening skills within higher education. This research synthesis confirms that AI tools demonstrably improve discrete language components, particularly pronunciation accuracy and speaking fluency, while significantly boosting learner confidence through personalized, low-stakes practice. This core finding underscores AI's transformative potential in overcoming the limitations of traditional pedagogy, offering scalable solutions that cater to individual learner needs and giving the field a powerful new direction. **Implication :** The capabilities afforded by AI necessitate a significant re-evaluation of language teaching methodologies, signaling a move towards more dynamic, interactive, and learner-centered educational paradigms. The successful adoption of these technologies implies a shift in the educator's role towards facilitating learning with AI, curating appropriate tools, and interpreting nuanced performance data. For learners, AI offers unprecedented access to practice opportunities and feedback, fostering greater autonomy and self-regulation in their language development journey, thus providing a more complete and potentially more effective learning ecosystem. **Limitation :** Despite its promise, the current generation of AI tools is not without significant limitations that must be acknowledged for a complete understanding. Their impact on higher-order skills, such as nuanced interactional competence and the spontaneous use of diverse vocabulary, remains less certain. Furthermore, the sophistication of feedback mechanisms, particularly for diagnosing and explaining listening comprehension errors, requires considerable advancement. Concerns regarding the transferability of skills learned in controlled AI environments to unpredictable real-

world situations, alongside critical ethical considerations encompassing data privacy, algorithmic bias, and equitable access, temper unchecked enthusiasm and demand careful consideration. **Future Research** : To fully realize AI's potential in language education and provide a conclusive assessment, focused future research is imperative. Establishing standardized evaluation metrics is crucial to move beyond anecdotal reports towards evidence-based assessments of tool effectiveness and scalability. Priority should be given to investigating the long-term transfer of AI-facilitated skills to authentic communicative contexts and exploring how AI can better foster sophisticated interactional abilities and nuanced listening comprehension. Continued innovation is needed to refine feedback systems, enhance AI's adaptability to real-world conditions, and proactively address ethical challenges through transparent and inclusive design, ultimately shaping AI's role into one that holistically supports learner growth, equity, and meaningful engagement in language learning.

REFERENCES

- [1] C. M. Goh, "Metacognitive instruction for second language listening development: Theory, practice and research implications," *RELC Journal*, vol. 39, no. 2, pp. 188–213, 2008. doi: 10.1177/0033688208092184
- [2] C. M. Goh and L. Vandergrift, *Teaching and Learning Second Language Listening: Metacognition in Action*. New York: Routledge, 2012
- [3] M. Özdere, "The integration of artificial intelligence in English education: Opportunities and challenges," *Language Education & Technology (LET Journal)*, vol. 3, no. 2, pp. 137–172, 2023.
- [4] J. E. Stefaniak and S. L. Moore, "The use of generative AI to support inclusivity and design deliberation for online instruction," *Online Learning*, vol. 28, no. 3, pp. 181–205, 2024. doi: 10.24059/olj.v28i3.4458.
- [5] W. Y. Leong and J. B. Zhang, "Ethical design of AI for education and learning systems," *ASM Science Journal*, vol. 20, no. 1, 2025. doi: 10.32802/asmscj.2025.1917.
- [6] P. Cui and B. Alias, "Opportunities and challenges in higher education arising from AI: A systematic literature review (2020–2024)," *Journal of Infrastructure, Policy and Development*, vol. 8, no. 11, art. 8390, 2024. doi: 10.24294/jipd.v8i11.8390.
- [7] A. Tierney, "Student perceptions on the impact of AI on their teaching and learning experiences in higher education," *Research and Practice in Technology Enhanced Learning*, vol. 20, art. e20005, 2025. doi: 10.58459/rptel.2025.20005.
- [8] W. Ling, "Artificial intelligence in language instruction: impact on English learning achievement, L2 motivation, and self-regulated learning," *Frontiers in Psychology*, vol. 14, 2023
- [9] B. Peña-Acuña and R. Corga Fernandes Durão, "Learning English as a second language with artificial intelligence for prospective teachers: a systematic review," *Frontiers in Education*, vol. 9, 2024
- [10] V. Quan, T. Bui, and C. Thanh, "Enhancing English speaking skills among English major students: the role of artificial intelligence-based," in *Proc. Conf. Nurturing Open Minds, Shaping Inclusive Futures in Language Education*, Ho Chi Minh City Open University, 2024.

- [11] S. M. Ruan, "Exploring the effectiveness of AI-powered chatbots in improving speaking confidence of EFL learners," **Computer Assisted Language Learning**, vol. 36, no. 4, pp. 345–364, 2023, doi: 10.1080/09588221.2022.2090038.
- [12] K. R. Chiu and M. L. Lin, "Automatic speech recognition technology and pronunciation training in English learning: A meta-analysis," **ReCALL**, vol. 35, no. 1, pp. 25–44, 2023, doi: 10.1017/S0958344022000078.
- [13] T. F. Kang and J. Wu, "Evaluating the impact of intelligent tutoring systems on English listening comprehension for university students," **International Journal of Artificial Intelligence in Education**, vol. 33, no. 2, pp. 213–230, 2023, doi: 10.1007/s40593-022-00291-3.
- [14] L. Zheng and H. Yang, "Using AI-driven immersive VR to enhance English listening and speaking proficiency: A quasi-experimental study," **British Journal of Educational Technology**, vol. 55, no. 1, pp. 80–98, 2024, doi: 10.1111/bjet.13357.
- [15] J. Lee, "Generative AI for EFL speaking tasks: A case study of GPT integration in university-level English classes," **AsiaCALL Online Journal**, vol. 15, no. 1, pp. 49–67, 2024, doi: 10.54855/asiacall.v15i1.832.

***Aigerim A. Gilea (Corresponding Author)**

Eurasian Humanities Institute named after A.K. Kussayinov, Kazakhstan

Email: Aigerimgilea@gmail.com

Katerina V. Melnova

Eurasian Humanities Institute named after A.K. Kussayinov, Kazakhstan

Email: melkatharinenova@gmail.com

Gyuzelya K. Temirgaliyeva

Eurasian Humanities Institute named after A.K. Kussayinov, Kazakhstan

Email: gyuzel3001@gmail.com
