

Enhancing English Speaking Skills through Artificial Intelligence: Opportunities, Tools, and Future Directions

Dr. Kedar S. Joshi

Associate Professor, Annasaheb Dange College of Engineering and Technology, Ashta, Maharashtra, India

ABSTRACT

Learners of English as a second or third language require ample opportunities to speak and read English to attain language proficiency. Such proficiency is crucial not only for academic success but also for social and emotional development. With the advent of Artificial Intelligence (AI), language learning has entered a transformative phase. AI now offers unprecedented tools that promote autonomous learning and provide learners with non-judgmental, personalized speaking practice. This paper explores current AI tools and applications that enhance English speaking skills among EFL and ESL learners. The study reviews the state-of-the-art technologies, discusses their pedagogical potential, and highlights real-world educational implementations. AI technologies such as speech recognition software, conversational AI, and VR-based language simulations show considerable promise in supporting speaking skills. Learners benefit from real-time feedback, adaptive learning models, and immersive scenarios. However, limitations exist in terms of human interaction, contextual nuance, and pronunciation recognition. Despite current limitations, the potential of AI to transform English language instruction is immense. Future development should focus on integrating emotional intelligence, improving speech recognition across accents, and enhancing cultural and contextual learning.

Keywords: AI, AI tools, EFL, ESL, Speaking, Language Learning

INTRODUCTION

Artificial Intelligence (AI) is now prominently utilized in five key areas of English Language Teaching (ELT): (1) speaking skill development, (2) writing enhancement, (3) reading improvement, (4) pedagogical support, and (5) promotion of learner self-regulation. While listening skills remain less supported, AI has shown notable success in fostering proficiency in the other core skills. This paper focuses specifically on the development of speaking skills through AI. It explores the impact of AI-based tools on English language teaching and learning, with a particular focus on their applications in EFL and ESL contexts.

METHODS AND MATERIALS

This paper is based on a review and synthesis of current literature and technological trends in AI-assisted language learning. The sources include peer-reviewed articles, organizational reports, and case studies that examine the integration of AI in language education. Tools and platforms discussed include speech recognition systems, AI-driven mobile apps, chatbots, VR simulations, and pronunciation analysis technologies. The analysis also incorporates examples of classroom applications, such as the study by Shivakumar et al. (2019), which highlights the effectiveness of AI coaches in higher education.

RESULTS

The integration of AI in English speaking instruction has significantly improved the way learners engage with language learning tasks. One of the primary benefits of AI is its ability to offer personalized, immediate feedback to learners, fostering a responsive and adaptive learning environment (Jiang et al., 2024). Among the most effective technologies are speech recognition systems, which enable learners to receive instant corrections on pronunciation, fluency, and intonation.

Tools like Google Speech Recognition and ELSA Speak are now widely adopted due to their ability to provide precise phonetic feedback and tailored recommendations for improvement (Xiaohong & Yanzheng, 2021).

Mobile language learning applications, such as Duolingo, Babbel, and Rosetta Stone, incorporate gamified speaking exercises that keep learners engaged while practicing everyday conversational English. These platforms often employ adaptive learning algorithms that adjust the difficulty and type of speaking tasks based on the user's performance, thereby encouraging consistent progress.

Conversational chatbots, including ChatGPT and Replika, have become valuable in simulating real-time dialogues. They help learners practice language in various social contexts without the anxiety of being judged by a human interlocutor (Shivakumar et al., 2019). Unlike traditional classroom environments, these AI systems are available on-demand, allowing learners to interact in both structured and informal settings at any time.

Another advancement includes virtual reality (VR)-based simulations offered by platforms like ImmerseMe, which immerse learners in realistic environments such as virtual cafes or airports where they can engage in context-specific conversations. This form of experiential learning enhances not only linguistic accuracy but also situational fluency, bridging the gap between academic knowledge and real-world communication (Xiaohong & Yanzheng, 2021).

In addition, AI features embedded in video conferencing platforms, such as Zoom's live transcription and translation, support multilingual interactions and aid learners in understanding and producing speech in real-time. Speech spectrograms and voice analysis tools, like those used by Speechling, visually represent voice elements like pitch and stress, enabling learners to adjust their articulation accordingly.

Despite these advancements, some challenges remain. Non-native accents are often misinterpreted by speech recognition systems, which are typically trained on native speaker data sets. Furthermore, current tools lack the ability to provide emotionally and culturally nuanced feedback, which is crucial for mastering pragmatic aspects of communication.

DISCUSSION

The results underscore the transformative role of AI in enhancing English speaking proficiency, particularly through interactive, adaptive, and autonomous learning mechanisms. However, as beneficial as these tools are, their limitations must be critically examined.

Firstly, one of the most pressing issues is the inability of AI systems to fully interpret human emotions, social cues, and context-specific language use (Sindermann et al., 2021). While chatbots and speech recognition technologies are advancing rapidly, they still struggle to replicate the depth of human interaction found in natural conversations. For instance, AI systems may fail to detect sarcasm, humor, or cultural references, all of which are vital components of communicative competence.

Secondly, accent bias remains a major obstacle. AI models are predominantly trained on native English datasets, resulting in misrecognition or miscorrection of speech by learners with diverse linguistic backgrounds (Jiang et al., 2024). This can lead to frustration and reduced motivation among learners, particularly when the system incorrectly flags correct utterances as incorrect.

Additionally, while automated feedback provides immediacy, it often lacks the pedagogical nuance of human instruction. AI tools may identify errors but cannot always explain why something is incorrect in a learner-friendly manner. This limits deeper learning, especially for learners who need detailed grammatical or phonological guidance (Shivakumar et al., 2019). From an ethical standpoint, concerns about data privacy and surveillance are also rising. Many AI-powered learning tools collect large amounts of learner data to function effectively. However, the transparency of how this data is used, stored, and shared remains limited (UNESCO, 2022).

Nonetheless, there are promising directions for future development. One such direction is the integration of emotional intelligence into AI systems. Advances in affective computing could allow AI to detect learner emotions and adjust the difficulty or tone of conversation accordingly, thereby improving engagement and reducing anxiety. Additionally, developing accent-inclusive models can improve accessibility and equity, making AI tools more inclusive for global learners.

The incorporation of cultural awareness into AI tools could also enhance pragmatic competence by exposing learners to different conversational norms and values. As educators and developers collaborate to address these limitations, AI is likely to evolve from a supplementary resource into a pedagogical co-teacher, complementing human instruction and enriching the learning experience.

CONCLUSION

AI-based English speaking tools are reshaping language learning, offering highly interactive and customized learning experiences. While current tools are effective, further innovations are required in contextual understanding, emotional intelligence, and cultural sensitivity. AI's evolving role as a learning partner rather than a passive tool holds immense promise. As AI continues to grow more autonomous and anthropomorphic, its role in human communication and education will become increasingly complex and impactful.

REFERENCES

- [1] Sindermann, C., Sha, P., Zhou, M. et al. Assessing the Attitude towards Artificial Intelligence: Introduction of a Short Measure in German, Chinese, and English Language. *Künstl Intell* 35, 109–118 (2021). <https://doi.org/10.1007/s13218-020-00689-0>
- [2] Xiaohong, W. & Yanzheng, W. (2021, February). The application of artificial intelligence in modern foreign language learning. In *Proceedings of the 2021 4th International Conference on Big Data and Education* (pp. 34–37). London, United Kingdom: ACM. <https://doi.org/10.1145/3451400.3451406>
- [3] UNESCO Recommendation on the ethics of artificial intelligence (2022). https://unesdoc.unesco.org/notice?id=p::usmarcdef_0000381137
- [4] Shivakumar, A., Shukla, S., Vasoya, M., Kasrani, I. M., & Pei, Y. (2019). AI-enabled language speaking coaching for dual language learners. *IADIS International Journal on WWW/Internet*, 17(1).
- [5] Jiang, T., Sun, Z., Fu, S., & Lv, Y. Human-AI interaction research agenda: A user-centered perspective. *Data and Information Management*, 8(4), 100078 (2024). <https://doi.org/10.1016/j.dim.2024.100078>