



Exploring ChatGPT's Voice Speech Recognition and Real-Time Feedback Features to Enhance EFL Speaking Skills: A Systematic Literature Review

Daniel Urslenton Wiser Sumang¹, Margana², Dyah Setyowati
Ciptaningrum³

^{1,2,3} Magister Pendidikan Bahasa Inggris, Universitas Negeri
Yogyakarta, Yogyakarta, Daerah Istimewa Yogyakarta

Article Info	Abstract
Received: 2026 - 06- 01 Revised: 2026 06-18 Accepted: 2026 06-30	<i>This study examines the integration of ChatGPT's voice recognition and real-time feedback features in enhancing EFL students' speaking skills through a systematic literature review that was conducted through identification, analysis, and synthesis. Ten peer-reviewed articles published between 2022 and 2025 were selected based on specific inclusion criteria, with a focus on AI-powered tools that support the development of EFL speech. Thematic analysis revealed that most of the research emphasized the effectiveness of AI-based voice-enabled platforms and feedback in improving learners' pronunciation, fluency, and engagement in English speaking skills. Some tools use machine learning algorithms to personalize feedback, while others promote autonomous speaking practices through interactive conversational interfaces. The findings also identify challenges such as recognition accuracy and pedagogical implementation gaps. These results show that ChatGPT and similar technologies strongly offer meaningful support for speaking instruction when carefully integrated into the EFL learning environment.</i>
Keywords: Code-mixing, Muysken's typology, intellectual digital identity, educational YouTube discourse, digital sociolinguistics	
DOI: 10.24256/ideasv14i1.11556	
Corresponding Author: Daniel Urslenton Wiser Sumang danielurslenton.2024@student.uny.ac.id Magister Pendidikan Bahasa Inggris, Universitas Negeri Yogyakarta, Yogyakarta, Daerah Istimewa Yogyakarta	

1. Introduction

ChatGPT is one of the products of Open AI that leverages natural language model processing (Roumeliotis & Tselikas, 2023). It stands for Generative Pre-Trained Transformers, a family of large language models (LLMs) that can understand and generate text in a natural language way. LLMs like (GPT -3) and (GPT-4) have made significant improvements in natural language processing and the models are basically trained on amounts of data to possibly generate human-like text, answer questions, and complete other related language tasks (Fryer et al., 2017). According to Chen et al. (2023) Machine learning (ML) is recognized as the main software of artificial intelligence (AI), which deals with numerous issues, such as speech recognition and natural language processing (NLP).

Systems that can learn and interact with human language, are built with the strong framework that Natural Language Processing (NLP) deals with algorithm voice speech recognition (Joshi & Taherdoost, 2025). It is fundamentally how machine learning comes with algorithm works and provides the variety of data information that is extracted to realize output. According to Lo et al. (2024), ChatGPT has delivered attention in the area of English language education, which is English as a second language (ESL) and English as a foreign language (EFL). ChatGPT is a promising tool for educational sectors such as educators and researchers (Karataş et al., 2024). In English language education, readiness is necessary for students to dive into modern technology use, covering ChatGPT's voice (Real Time Feedback). However, learners still face numerous issues in English language learning, especially conducting speaking skills competencies (Enesi et al., 2021).

Increasingly, 21st-century technology has been rapidly modern in assisting learners by utilizing oral abilities. According to Y. Wang, (2025) ChatGPT provides voice assistance which stands for speech recognition under the algorithm systems. This feature helps learners shape their critical thinking to boost oral communication under the algorithm of a language model processing system. According to Pantsar (2024), Tenorth & Beetz (2017), and Ciatto (2022) AI research encompasses various subfields, including theorem proving, knowledge representation, formal logic, and the extension to autonomous robots. Thus, diving into LLMs creates the development of English language Teaching instruments.

The process of speaking includes the conceptualization, formulation, articulation, and self-monitoring. Those are the key futures for navigating and building oral abilities that are derived from idea formation, language encoding, physical production, feedback, and correction (Hougham et al., 2024). The incorporation of ChatGPT and speaking skills are set to complete each other in an English language teaching field. Delivering voice or speech to the voice future of ChatGPT is a framework to enhance learners' speaking ability because that is the process of monitoring their habits. Nevertheless, learners of the English foreign language (EFL) have challenges recognizing the topic in speaking tasks because

they lack speaking components such as pronunciation, fluency, vocabulary, and comprehension (Zare-Behtash & Sarlak, 2017).

Furthermore, speaking skills need to be shaped to acquire the language input by being accustomed to speaking English with peers and teachers, but this practice is limited to the classroom. Therefore, by engaging EFL students with modern technology, they will be assisted in enhancing their speaking ability and technological skills, allowing them to innovate their learning style in speaking skills. Thus, students can practice speaking skills in formal or informal settings. Furthermore, this study aims to explore ChatGPT's voice speech algorithm in enhancing EFL speaking skills (Real-Time Feedback of Machine Learning). The following research question (RQ) was formulated.

RQ: 1. In what ways do ChatGPT's voice recognition and real-time feedback affect the speaking performance of EFL learners?

Literature review

ChatGPT Voice and EFL Speaking Skills

ChatGPT, an AI generative model with a voice interaction feature, has become a crucial innovation in learning English as a foreign language (EFL), particularly in enhancing students' speaking skills (Tai & Chen, 2024). By practicing English through conversational simulation and voice-based interactive exercises, students can flexibly and independently occupy ChatGPT's voice to practice their speaking skills without fear of judgment from their peers and teachers (Wu et al., 2025). According Tsang, (2025); Zheng et al., (2025) this approach creates a safe learning environment psychologically, where students are able to enhance their fluency, pronunciation, and their confidence in speaking skills.

Moreover, the feature of voice-enabled dialogue on ChatGPT gives experienced communication in real human interaction that can help students become accustomed to utilizing rhythm, intonation, and the structure of sentences in the English language (Ahmed et al., 2024; Rezai et al., 2024). According to Zhang et al., (2025) This technology crucially supports personalized learning activities, where students are able to repeat their training and receive feedback directly, which is non-judgmental. Some studies also reveal that the users of ChatGPT engage in learner autonomy and extend the practice time to enhance their speaking skills outside the classroom (Hong & Guo, 2025; Li et al., 2025).

In the context of speaking development, ChatGPT is not only a tool to assist the speaking practice but also becomes an adaptive co-training for EFL students (Khaliq et al., 2025; Reza et al., 2025). Its ability to respond to various languages, including idioms, open-ended questions, and informal expressions, necessarily gives students exposure to a variety of languages that are very large rather than the traditional method (Pawar & Augenstein, 2025). Therefore, integrating ChatGPT into EFL learning opens new opportunities for the communicative approach in which students as the learning center and technology-based (Darmawansah et al.,

2025; Esfandiari, 2024).

Speech Recognition Technology in EFL Speaking Development

Speech Recognition Technology (SRT) has been bringing significant changes in teaching speaking skills in the context of English as a foreign language (EFL) (Tzu & Chen, 2022). With its ability to change speech into texts in real-time, this technology can make students receive a direct evaluation in the aspects of pronunciation, intonation, and fluency (Thomas & Sun, 2023; Vu et al., 2025). Thus, these benefits significantly give more advantages in pedagogical aspects of teaching and learning, because the students can directly identify their mistakes and make corrections independently (Lin et al., 2025; Zou et al., 2024).

In practice, this technology is often combined with AI-based applications, which include conversation assistance or language-learning software that automatically analyzes speech and provides real-time feedback (Zou et al., 2023). This shows that automatic speech recognition (ASR) can enhance students' confidence and give them support in developing accurate phonology (Ngo et al., 2024). Moreover, according to Jiang & Chai, (2022) This technology leverages opportunities for students to practice English outside the formal classroom, with access to training in speaking skills that is more flexible.

Nevertheless, the effective technology of speech recognition is dependent on the accuracy of systems in recognizing various accents of non-native speakers, the level of noise, and the quality of the device that is used for practicing (Dar, 2025; Polspoel et al., 2025). Nonetheless, Orosoo et al., (2025), emphasized when it is used strategically in a learning environment, speech recognition has the potential to support a learning approach that is more communicative, personalized, and data-driven.

Real-Time Feedback in EFL Speaking Practice

Real-time feedback is a crucial element in speaking skill practice in the context of English as a foreign language (EFL) because it makes learners aware of areas to improve, such as pronunciation, intonation, and ways of speaking (Sa et al., 2025; Shadiev et al., 2024). In this situation, students can get feedback directly after asking numerous questions by clicking the voice mode. According to Liang & Wu, (2024), this voice feature will help EFL student shape and provoke their critical thinking in breaking down various topics in speaking English Activities.

In Practice, real-time feedback consists of various aspects, such as pronunciation, fluency, grammar, and intonation (Ping & Tao, 2025a). Jackson et al., (2025), the introduction of technology that is embedded with voice speech recognition has been integrated with artificial intelligence like ChatGPT voice which gives instant feedback about what students have said to the voice system. For example, systems serve certain information, whether the pronounced words have been well-uttered or suggest the right structure of sentences when students experience incorrect grammar (Sa et al., 2025; Vu et al., 2025).

Machine Learning in Language Learning and Speaking Assessment

Machine learning is one of the branches of artificial intelligence (AI), playing a significant role in developing systems of adaptive and personal language learning, particularly in speaking assessment (Zechner & Evanini, 2019). In the context of learning English as a foreign language, Zechner & Evanini, (2019) machine learning is used to analyze the students' data performance continuously to give significant feedback automatically and adapt the learning material to the individual's needs.

According to Ping & Tao, (2025) a system based on machine learning can identify the linguistic patterns in students' way of pronouncing, including clear pronunciation, fluency, sentence structure, and intonation. Through compiled data from students' interactions, this algorithm learns to recognize general mistakes and automatically gives suggestions to solve the problem in real-time (Trabelsi et al., 2023). This technology has already been integrated into various applications such as Duolingo, Elsa Speak, and systems that are based on ChatGPT which uses machine learning to enhance interaction quality with students (OpenAI, 2025; Sa et al., 2025; Yaşar & Murat, 2025).

Some studies identified the effectiveness of machine learning in supporting speaking assessment. For instance, the research by M. Chen & Zechner, (2011) shows that automatic-based machine learning can reveal the evaluation of speaking ability which is the same as human assessors in the context of the TOEFL test. Similarly, Zou et al., (2023) show that a learning algorithm can predict the students' achievement in speaking with high accuracy based on verbal interaction.

In addition, to the efficiency and assessment, the approach also encourages learning personalization. Adaptive systems powered by machine learning can recommend specific pieces of training, based on an individual's weaknesses, as well as track their progress from time to time (Dhananjaya et al., 2024; Human-hendricks, 2023). Furthermore, Human-hendricks, (2023) emphasized this study supports the process of learning more comprehensible and well-navigated as well as students' responses to learning activities. Therefore, integrating machine learning in the teaching and learning speaking assessment, not only makes the evaluation faster but also develops learning quality through responsive approaches and driven data (Jared et al., 2025; Liu et al., 2025).

2. Method

Search Strategy and Literature Retrieval

This research adheres to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines by (Moher et al., 2009). The data analyzed comes from the last five years, namely from 2022 to 2025 (Snyder, 2019). The selection of this time range is based on the development of artificial intelligence systems, specifically the use of ChatGPT, which has begun to be widely used in recent years (Kasneci et al., 2023; OpenAI, 2022). Thus, Kasneci et al., (2023) the analysis is focused on the latest publications that are relevant to the application of ChatGPT in the context of language learning. The search strategy was organized

into three main categories: (1) AI-Powered Tools / Systems, (2) Technology and Feedback Methods, and (3) EFL Speaking Skills Context. These terms were combined into a comprehensive search string to ensure the inclusion of studies that integrate AI-powered tools/systems, Technology, and Feedback Methods, as well as EFL Speaking Skills Context. The search string is presented as follows: ("ChatGPT" AND "Speaking Skills"), ("Machine Learning" OR "Voice Speech Recognition" AND "EFL Speaking Skill"), ("Language Learning Algorithm" AND "Speaking Skill" AND "ChatGPT Voice" OR "Voice Speech Recognition"), and ("ChatGPT" OR "Real Time-Time Feedback" AND "Speaking Skills").

Table 1. Database and Search String

Databases	Keywords	Boolean Operator
Google Scholar	"ChatGPT" AND "Speaking Skills"	AND
Elsevier	"Machine Learning" AND "Voice Speech Recognition" AND "EFL Speaking Skills"	AND, AND
Routledge, Taylor & Francis, Nature Communications	("Language Learning Algorithm" OR "Machine Learning") AND "Speaking Skills" AND ("ChatGPT Voice" OR "Voice Speech Recognition")	OR, AND, AND, OR
IEEE dan Sustainability	("ChatGPT" OR "Real-Time Feedback") AND "Speaking Skills"	OR, AND
Association for Computational Linguistics		

Study Selection Procedure and Inclusion/Exclusion Criteria

The PRISMA flowchart (Figure 2) illustrates that there is a study selection process, which includes inclusion and exclusion criteria. There were 142 data successfully identified between January 2022 and June 2025, consisting of 128 data from databases and 14 from registers. Before entering the screening stage, there were 30 data deleted: 22 due to duplication, 5 because they were deemed unsuitable by automated tools, and 3 were deleted for specific reasons. So, there are 112 data that are filtered in the next stage. At the screening stage, there were 70 data released, and 42 reports were sought for review. Of these, 6 reports were not successfully obtained, so that there were 36 reports that were assessed as feasible in screening. Of the 36 reports assessed, 26 were issued with details: 12 were irrelevant to ChatGPT or AI-based tools, 8 did not have a real-time voice feedback component, and there were 6 that did not target English as a foreign language (EFL) speaking skills. In the final session, there were as many as 10

studies included in the review. The PRISMA flowchart below clearly shows each stage in the process of identification, screening, and inclusion of the study. The illustrations of the PRISMA flow diagram of the study selection, are below:

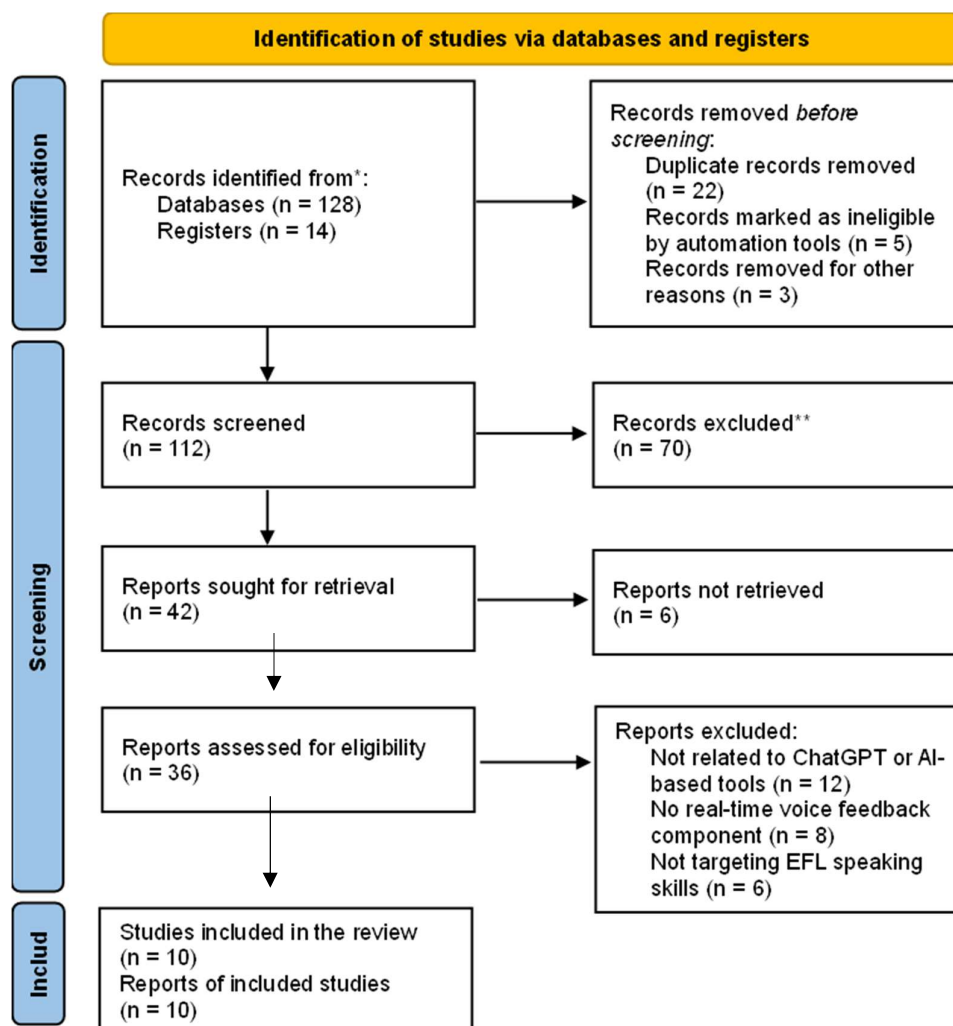


Figure 1. PRISMA flow diagram of the study selection

In the inclusion criteria, the researcher decided to collect the relevant data for the Identification of studies via databases and registers regarding the use of Artificial Intelligence tools, such as ChatGPT, the algorithm of speech recognition technology, machine learning, and real-time feedback systems in enhancing EFL learners' speaking skills. This step ensures that this can be aligned with the focus of the research by improving the credibility of inclusive databases. The objective is to maintain the relevance, credibility, and quality of the literature review which supports the analysis of how Artificial Intelligence fundamentally influences speaking ability in the context of a foreign language.

In the exclusion criteria, the researcher decided to eliminate studies that have

no related focus on EFL students and do not really employ the use of Artificial Intelligence as the main tool for enhancing students' speaking skills by using ChatGPT voice for instance, ChatGPT, the algorithm of speech recognition technology, machine learning, and real-time feedback systems. The published literature in languages, specifically the literature that has weak findings, abstracts, and other opinions, statements, and those lacking empirical data, was also excluded. Therefore, selecting the published peer-reviewed literature should be done carefully and relevant to the needs of the databases and registers that cover the use of AI-powered tools to enhance EFL students' speaking skills.

Coding Framework and Analysis Procedures

The data analysis method in this study uses thematic analysis to identify, analyze, and synthesize the patterns of themes related to the literature review. Articles relevant to the study were chosen according to specific inclusion and exclusion criteria. Information from each article was systematically arranged in a table containing details such as authors, publication year, research methodology, key findings, and their relevance to the research topic. Subsequently, the data was analyzed by categorizing the main findings into pertinent themes, specifically focusing on the contributions of ChatGPT's voice in speaking skills, particularly, in the English Language Teaching (ELT) context. This analytical process was performed iteratively to ensure that each theme accurately represented the logical connections within the reviewed literature about the research objectives. The outcomes of this analysis were then presented in tables to clearly illustrate (Table 1) the Components of the Framework, Definitions, and Sample Studies of the role of ChatGPT's voice in speaking skills in ELT.

Table 2. Database and Search String

Components	N	Definitions and examples
Voice speech recognition Algorithm	7	Technology that allows students to interact with the system using voice input (Young & Donggil, 2021). For example, using speech APIs for speaking practice.
Real-time feedback	10	Automatic responses from the AI system to students' speech, such as pronunciation correction (Rogti, 2025).
Machine Learning	1	An adaptive learning system that adjusts feedback based on students' performance data (Casteleijn & Franzsen, 2024).

The table explains that there are three main components of AI technology in EFL speech learning, namely voice recognition algorithms, real-time feedback, and machine learning. Researchers illustrated the speech recognition component found

in 7 related studies and allowed students to practice speaking through voice input, such as using APIs as speech recognition (Young & Donggil, 2021). Meanwhile, real-time feedback appeared in 10 related studies and provided direct automatic corrections, such as pronunciation, which helped students learn independently and be more confident (Rogti, 2025).

The component, machine-based learning, was only found in one study (Casteleijn & Franzsen, 2024) , but it strongly offers a system that can make feedback adjustments based on students’ performance in learning to speak English. This shows the potential to create more personalized and adaptive learning, although its use is still limited in the context of EFL speakers because this feature is still very new and will be a brilliant innovation in the development of new research in speaking English that is basically already systematized with algorithms and machine learning.

Table 3. AI-Powered Speaking Tools: Voice Recognition and Real-Time Feedback in Language Learning

Code	Theme	Categ	Supporting Quotes
ory			
S1, S3, S5, S7, S10	Voice Recognition Algorithm	Speech Speaking Practice	Technology that allows students to interact with the system using voice input.” (Young & Donggil, 2021). For example, students practiced speaking through speech APIs integrated in apps
S2, S4, S5, S6, S7, S8, S9, S10	Real-time feedback	Pronunciati on Support	Automatic responses from the AI system to students’ speech, such as pronunciation correction.” (Rogti, 2025). Students received instant feedback after speaking to improve their fluency
S9	Machine Learning		An adaptive learning system that adjusts feedback based on students’ performance data.” (Casteleijn & Franzsen, 2024). The system adapted correction types based on learners’ progress

The first theme is the Speech Recognition Algorithm exemplified by the tool in the data code according to the inclusion criteria, namely S1, S3, S5, S6, S7, S8, and S10. These AI-powered speaking tools can encourage students to practice speaking through voice input, creating a more natural and interactive learning environment

for developing English language skills. As Young & Donggil, (2021) inferred, The students engage with the system orally, supported by an API that is speech integrated into the app, which can make authentic pronunciation and the development of speaking fluency, especially in English. This interactive mode certainly reflects real-life conversation scenarios and strengthens students' confidence in using the target language automatically and effectively.

The second theme is Real-Time Feedback which basically appears in the tools represented by S2, S4, S5, S6, S7, S8, S9, and S10. Researchers show that this system provides an instant response to students' speech, especially correcting pronunciation and fluency in improving English speaking. According to Rogti, (2025), direct feedback is very helpful for students to adjust their speech patterns during the learning process, resulting in faster improvement because, in their natural system, students can develop the way they talk to friends or direct feedback. Thus, this real-time correction greatly supports self-monitoring and motivates students to improve their oral output continuously, making the learning process more efficient and responsive, especially in developing their English language skills consistently.

The third theme, Machine Learning which is integrated with algorithms and although only represented in S9, underlines the adaptive capabilities of modern language learning systems. Casteleijn & Franzsen, (2024) Explains how machine learning enables these systems to analyze learners' performance data adjust feedback accordingly, and help them to develop their English-speaking skills. This ensures that the type and complexity of corrections evolve in response to the student's progress. This ability is certainly adaptable such as improving personalized learning, making the experience more effective and supporting the diverse needs of students in honing English naturally.

3. Result

After selecting the data, articles are categorized based on inclusion and exclusion criteria. Each relevant article is then arranged in a table to facilitate analysis, and the results of the codification are organized under the umbrella of thematic analysis. The following is a summary of the results of the review of the 10 selected articles:

Theme	Description	N =	Supporting Studies
		10	
Voice Speech Recognition for Speaking Practice	Use of speech APIs and voice-enabled apps to allow learners to practice pronunciation and speaking	7/10	(Shadiev et al., 2024), (Hsu et al., 2023), (Zheng et al., 2025), (Mohamed, 2024), (Zou et al., 2024), (Dar, 2025), (Tzu & Chen, 2022).

Real-Time Feedback Fluency Improvement	fluency. Immediate responses from AI tools that correct pronunciation, grammar, or fluency to help learners self-correct.	8/10	(Qiao & Zhao, 2023), (Wu et al., 2025), (Zheng et al., 2025), (Mohamed, 2024), (Zou et al., 2024), (Hsu et al., 2023), (Jeon et al., 2023), (Dar, 2025).
Machine Learning Algorithms in EFL Learning	Adaptive feedback systems using ML models (e.g., SVM, CNN) to tailor support based on learner performance.	3/10	(Shadieiev et al., 2024), (Qiao & Zhao, 2023), (Dar, 2025).
Learner Motivation and Engagement	Increased student motivation and interaction due to interactive, responsive AI systems in speaking tasks.	6/10	(Qiao & Zhao, 2023), (Hsu et al., 2023), (Wu et al., 2025), (Zheng et al., 2025), (Mohamed, 2024), (Jeon et al., 2023).
Pedagogical and Technical Challenges	Issues such as limited algorithm transparency, ASR inaccuracies, or a lack of pedagogical integration in classrooms.	4/10	(Hsu et al., 2023), (Mohamed, 2024), (Jeon et al., 2023), (Tzu & Chen, 2022).

Thematic analysis of the 10 selected studies revealed that Speech Recognition for Speaking Practice and Real-Time Feedback for improved Fluency are the most prominent trends in AI-powered tools for structured and effective EFL speaking instruction. Seven out of ten studies underscore the use of speech APIs and voice-enabled platforms, such as Alexa or mobile ASR apps, that facilitate authentic spoken language training. This technology allows learners to interact naturally through voice input by strengthening pronunciation and fluency. Meanwhile, eight other studies emphasize how real-time feedback mechanisms, often embedded in AI tools, provide direct corrections to learners' verbal output. This instant response

certainly plays an important role in helping students correct themselves and develop more accurate use of language, especially in pronunciation and grammar.

Although slightly less rare, Machine Learning algorithms have appeal in the context of EFL, with three studies applying models such as Support Vector Machines (SVM) or Convolutional Neural Networks (CNN) to analyze learners' speech and generate adaptive feedback. This intelligent system certainly adjusts the type and frequency of corrections based on the performance of individual learners, offering a more personalized and data-driven learning experience. In addition, six studies reported increased motivation and student engagement when using AI-based speaking tools. The students can feel more motivated to talk and interact, driven by the responsiveness and gamification elements of the technology involved.

Despite these benefits, there are some studies also reporting pedagogical and technical challenges, found in four out of ten articles. These include the limitations of ASR accuracy, a lack of transparency in algorithm design, and inadequate integration of AI tools into structured classroom instruction. The issues highlight the need for more educator training, enhanced AI design, and pedagogical alignment to ensure these technologies are effective and sustainable in EFL speech development. Overall, the results suggest that while AI offers promising support for speaking skills, its implementation must be strategically designed to maximize learning outcomes.

4. Discussion

The findings of this review demonstrate the growing integration of artificial intelligence, particularly voice speech recognition and real-time feedback systems, into English as a Foreign Language (EFL) speaking instruction. The prevalence of tools utilizing speech APIs, such as mobile ASR applications or AI chatbots, shows that learners now have access to more interactive, autonomous, and authentic speaking environments. These technologies help bridge the gap between classroom practice and real-life communication by enabling learners to rehearse pronunciation and fluency in a low-anxiety setting. When students are able to speak to a system and receive immediate responses or corrections, they become more engaged and develop a better awareness of their spoken language accuracy.

The role of real-time feedback emerges as particularly crucial in promoting effective speaking development. Feedback delivered instantly allows learners to recognize and self-correct errors at the moment of speech, a feature that has been traditionally difficult to achieve in typical classroom settings with limited teacher time. AI tools, as reported in the reviewed studies, offer pronunciation correction, grammatical feedback, and fluency monitoring through automated systems. These benefits align with communicative language teaching principles, where meaningful interaction and corrective feedback are central to skill development. Additionally, the integration of machine learning algorithms has opened new possibilities for

adapting instruction to individual learner needs, enhancing the personalization of language learning.

However, several studies also caution against overreliance on these technologies without addressing their pedagogical and technical limitations. Issues such as low ASR accuracy, misinterpretation of learner intent, limited algorithm transparency, and insufficient teacher involvement can reduce the effectiveness of AI-based systems. Moreover, the lack of structured integration into curricula may hinder their full pedagogical impact. To address these challenges, it is essential for educators, developers, and researchers to collaborate in designing AI tools that are both technically robust and pedagogically sound. This includes ensuring that AI feedback aligns with learning objectives, complements human instruction, and respects learner diversity. Ultimately, the discussion reveals that while AI-powered speaking tools hold strong potential, thoughtful implementation is key to ensuring their success in EFL learning contexts.

5. Conclusion

This research concludes that AI-powered tools, particularly those utilizing voice speech recognition and real-time feedback, have a significant impact on enhancing EFL learners' speaking performance. The integration of these technologies allows learners to practice pronunciation and fluency in interactive environments that simulate real-life communication. Real-time feedback, in particular, helps learners self-correct their speech immediately, making learning more responsive and effective.

Moreover, the use of machine learning algorithms in some tools contributes to personalized feedback by adapting to individual learners' progress. This adaptive support fosters greater learner engagement and motivation. However, several studies also highlight technical and pedagogical challenges, such as ASR inaccuracies and the need for better integration into classroom instruction. Therefore, while AI tools like ChatGPT show strong potential in developing speaking skills, their implementation must be carefully aligned with pedagogical goals to maximize their effectiveness in EFL contexts.

Acknowledgement

The author expresses his sincere gratitude to all parties who have provided moral and academic support during the completion of this systematic literature review. This research was conducted independently and was fully funded by the authors without external sponsorship or financial support. Appreciation is also given to those whose careers are critically reviewed and analyzed in this research, which makes a very significant contribution to the insights and findings presented in literature works related to the purpose of the research.

6. References

- Ahmed, N., Kumar, A., Al, A., Rahman, J., Mridha, M. F., & Kabir, M. (2024). Deep learning-based natural language processing in human – agent interaction : Applications , advancements and challenges. *Natural Language Processing Journal*, 9(October), 100112. <https://doi.org/10.1016/j.nlp.2024.100112>
- Bani, M., & Masruddin, M. (2021). Development of Android-based harmonic oscillation pocket book for senior high school students. *JOTSE: Journal of Technology and Science Education*, 11(1), 93-103.
- Casteleijn, D., & Franzsen, D. (2024). Heliyon Personalized adaptive learning in higher education : A scoping review of key characteristics and impact on academic performance and engagement. *Heliyon*, 10(21), e39630. <https://doi.org/10.1016/j.heliyon.2024.e39630>
- Chen, H. H. J., Yang, C. T. Y., & Lai, K. K. W. (2023). Investigating college EFL learners' perceptions toward the use of Google Assistant for foreign language learning. *Interactive Learning Environments*, 31(3), 1335–1350. <https://doi.org/10.1080/10494820.2020.1833043>
- Chen, M., & Zechner, K. (2011). *Computing and Evaluating Syntactic Complexity Features for Automated Scoring of Spontaneous Non-Native Speech*. 722–731.
- Ciatto, G. (2022). *On the role of Computational Logic in Data Science: representing, learning, reasoning, and explaining knowledge*.
- Dar, M. A. (2025). Machine Learning and Deep Learning Approaches for Accent Recognition : A Review. *IEEE Access*, 13(February), 51527–51550. <https://doi.org/10.1109/ACCESS.2025.3552935>
- Darmawansah, D., Rachman, D., & Febiyani, F. (2025). *ChatGPT-supported collaborative argumentation : Integrating collaboration script and argument mapping to enhance EFL students ' argumentation skills*. 3803–3827.
- Dhananjaya, G. M., Goudar, R. H., & Kulkarni, A. A. (2024). A Digital Recommendation System for Personalized Learning to Enhance Online Education : A Review. *IEEE Access*, 12(January), 34019–34041. <https://doi.org/10.1109/ACCESS.2024.3369901>
- Enesi, M., Vrapı, F., & Trifoni, A. (2021). Challenges of teaching and learning English language for ESP courses. *Journal of Educational and Social Research*, 11(4), 213–226. <https://doi.org/10.36941/jesr-2021-0090>
- Esfandiari, R. (2024). *The Role of ChatGPT-based Instruction and Flipped Language Learning in Metadiscourse Use in EFL Learners ' Argumentative Writing and their Perceptions of the two Instructional Methods*. 43(3), 27–51. <https://doi.org/10.22099/tesl.2024.49975.3277>
- Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners. *Computers in Human Behavior*, 75, 461–468. <https://doi.org/10.1016/j.chb.2017.05.045>
- Furwana, D., Muin, F. R., Zainuddin, A. A., & Mulyani, A. G. (2024). Unlocking the

- Potential: Exploring the Impact of Online Assessment in English Language Teaching. *IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature*, 12(1), 653-662.
- Hong, X., & Guo, L. (2025). Effects of AI-enhanced multi-display language teaching systems on learning motivation , cognitive load management , and learner autonomy. In *Education and Information Technologies* (Issue 0123456789). Springer US. <https://doi.org/10.1007/s10639-025-13472-1>
- Hougham, D., Clenton, J., Uchihara, T., & Higginbotham, G. (2024). The Impact of Lexical Bundle Length on L2 Oral Proficiency. *Languages*, 9(7), 1–30. <https://doi.org/10.3390/languages9070232>
- Hsu, H., Chen, H. H., Todd, A. G., Chen, H. H., Todd, A. G., & Chen, H. H. (2023). *Investigating the impact of the Amazon Alexa on the development of L2 listening and speaking skills Investigating the impact of the Amazon Alexa on the.* 4820. <https://doi.org/10.1080/10494820.2021.2016864>
- Human-hendricks, N. E. (2023). Personalized Adaptive Learning Technologies Based on Machine Learning Techniques to Identify Learning Styles : A Systematic Literature Review. *IEEE Access*, 11(May), 48392–48409. <https://doi.org/10.1109/ACCESS.2023.3276439>
- Jackson, L., Sidorkin, A. M., Jandrić, P., Costello, E., Heybach, A., Greenhalgh-spencer, H., Hytten, K., Gourlay, L., Buchanan, R., Tesar, M., Jackson, L., Sidorkin, A. M., Jandrić, P., Costello, E., Heybach, A., Greenhalgh-spencer, H., Hytten, K., Gourlay, L., & Buchanan, R. (2025). The voice of artificial intelligence : Philosophical and educational reflections The voice of artificial intelligence : Philosophical and. *Educational Philosophy and Theory*, 0(0), 1–12. <https://doi.org/10.1080/00131857.2025.2481588>
- Jared, X., Wang, J., & Zou, B. (2025). Journal of English for Academic Purposes Evaluating an AI speaking assessment tool : Score accuracy , perceived validity , and oral peer feedback as feedback enhancement. *Journal of English for Academic Purposes*, 75(March), 101505. <https://doi.org/10.1016/j.jeap.2025.101505>
- Jeon, J., Lee, S., & Choe, H. (2023). Computers & Education Beyond ChatGPT : A conceptual framework and systematic review of speech-recognition chatbots for language learning. *Computers & Education*, 206(July), 104898. <https://doi.org/10.1016/j.compedu.2023.104898>
- Jiang, M. Y., & Chai, C. (2022). *Exploring the effects of automatic speech recognition technology on oral accuracy and fluency in a flipped classroom.* August. <https://doi.org/10.1111/jcal.12732>
- Joshi, H. S., & Taherdoost, H. (2025). Developing Natural Language Processing Algorithms to Fact-Check Speech or Text. *Transportation Research Procedia*, 84, 291–298. <https://doi.org/10.1016/j.trpro.2025.03.075>
- Karataş, F., Abedi, F. Y., Ozek Gunyel, F., Karadeniz, D., & Kuzgun, Y. (2024). Incorporating AI in foreign language education: An investigation into

- ChatGPT's effect on foreign language learners. *Education and Information Technologies*, 29(15), 19343–19366. <https://doi.org/10.1007/s10639-024-12574-6>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., & Dementieva, D. (2023). *ChatGPT for good? On opportunities and challenges of large language models for education*. <https://www.doi.org/https://doi.org/10.1016/j.lindif.2023.102274>
- Khalique, A., Honggang, W., & Ali, M. (2025). Computers in Human Behavior Reports Empowering creativity and engagement : The impact of generative artificial intelligence usage on Chines EFL students ' language learning experience. *Computers in Human Behavior Reports*, 18(February), 100627. <https://doi.org/10.1016/j.chbr.2025.100627>
- Li, Y., Sadiq, G., Qambar, G., & Zheng, P. (2025). *The impact of students ' use of ChatGPT on their research skills: The mediating effects of autonomous motivation , engagement , and self-directed learning*.
- Liang, W., & Wu, Y. (2024). *Exploring the use of ChatGPT to foster EFL learners ' critical thinking skills from a post-humanist perspective*. 54(April).
- Lin, C., Wang, W., & Lee, H. (2025). *Interventions in STEM Education Through Speech Recognition-Based Learning Analysis*. 63(1), 311–335. <https://doi.org/10.1177/07356331241307904>
- Liu, T., Zhang, M., Li, Z., Dou, H., Zhang, W., Yang, J., Wu, P., Li, D., & Mu, X. (2025). Machine learning-assisted wearable sensing systems for speech recognition and interaction. *Nature Communications*. <https://doi.org/10.1038/s41467-025-57629-5>
- Lo, C. K., Yu, P. L. H., Xu, S., Ng, D. T. K., & Jong, M. S. yung. (2024). Exploring the application of ChatGPT in ESL/EFL education and related research issues: a systematic review of empirical studies. *Smart Learning Environments*, 11(1). <https://doi.org/10.1186/s40561-024-00342-5>
- Mohamed, A. M. (2024). *Exploring the potential of an AI-based Chatbot (ChatGPT) in enhancing English as a Foreign Language (EFL) teaching : perceptions of EFL Faculty Members*. 3195–3217.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, T. P. (2009). *Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement*. 6(7). <https://doi.org/10.1371/journal.pmed.1000097>
- Ngo, T. T., Chen, H. H., & Lai, K. K. (2024). *The effectiveness of automatic speech recognition in ESL / EFL pronunciation: A meta-analysis*. 36, 4–21. <https://doi.org/10.1017/S0958344023000113>
- OpenAI. (2022). OpenAI. *ChatGPT: Optimizing Language Models for Dialogue*. <https://openai.com/blog/chatgpt>
- OpenAI. (2025). *OpenAI*. 2015. <https://openai.com/chatgpt/overview/>
- Orosoo, M., Raash, N., Treve, M., Fareed, H., Lahza, M., Alshammry, N., Venkata, J., Ramesh, N., & Rengarajan, M. (2025). Transforming English language learning :

- Advanced speech recognition with MLP-LSTM for personalized education. *Alexandria Engineering Journal*, 111(October 2024), 21–32. <https://doi.org/10.1016/j.aej.2024.10.065>
- Pantsar, M. (2024). *Theorem proving in artificial neural networks : new frontiers in mathematical AI*. 1–22.
- Pawar, S., & Augenstein, I. (2025). *Survey of Cultural Awareness in Language Models : Text and Beyond*. 1–96.
- Ping, L., & Tao, N. (2025a). *Innovative approaches to English pronunciation instruction in ESL contexts : integration of multi-sensor detection and advanced algorithmic feedback*. January. <https://doi.org/10.3389/fpsyg.2024.1484630>
- Ping, L., & Tao, N. (2025b). Innovative approaches to English pronunciation instruction in ESL environments : integration of multi-sensor detection and advanced algorithmic feedback Innovative approaches to English pronunciation instruction in ESL environments: integration of multi-. *Interactive Learning Environments*, 4820, 1–21. <https://doi.org/10.1080/10494820.2025.2476718>
- Polspoel, S., Moore, D. R., Swanepoel, D. W., & Kramer, S. E. (2025). *Automatic development of speech- in-noise hearing tests using machine learning Framework for Auditory Discrimination Experiments*. 1–20.
- Qiao, H., & Zhao, A. (2023). *Artificial intelligence-based language learning : illuminating the impact on speaking skills and self-regulation in Chinese EFL context*. November. <https://doi.org/10.3389/fpsyg.2023.1255594>
- Reza, A., Sheykhholeslami, M., & Mahmoudi, A. (2025). Computers in Human Behavior Reports Uncovering personalized L2 motivation and self-regulation in ChatGPT-assisted language learning : A hybrid PLS-SEM-ANN approach. *Computers in Human Behavior Reports*, 17(December 2024), 100539. <https://doi.org/10.1016/j.chbr.2024.100539>
- Rezai, A., Namaziandost, E., & Hwang, G. (2024). Computers in Human Behavior Reports How can ChatGPT open promising avenues for L2 development ? A phenomenological study involving EFL university students in Iran. *Computers in Human Behavior Reports*, 16(July), 100510. <https://doi.org/10.1016/j.chbr.2024.100510>
- Rogti, M. (2025). Automatic pronunciation evaluation feedback and peer correction for shaping English pronunciation accuracy and interpersonal communication ARTICLE HISTORY. *Innovation in Language Learning and Teaching*, 1229, 1–17. <https://doi.org/10.1080/17501229.2025.2475931>
- Roumeliotis, K. I., & Tselikas, N. D. (2023). ChatGPT and Open-AI Models: A Preliminary Review. *Future Internet*, 15(6). <https://doi.org/10.3390/fi15060192>
- Sa, N. T., Thi, N., Thi, D., Ngoc, H., Nhu, Q., Thi, D., & Hieu, K. (2025). *How AI-Powered Voice Recognition Has Supported Pronunciation Competence among EFL University Learners*. 26(3), 64–83.

- <https://doi.org/https://doi.org/10.54855/callej.252634>
- Shadiev, R., Feng, Y., Zhussupova, R., & Altinay, F. (2024). Effects of speech-enabled corrective feedback technology on EFL speaking skills , anxiety and confidence. *Computer Assisted Language Learning*, 0(0), 1–37. <https://doi.org/10.1080/09588221.2024.2430761>
- Snyder, H. (2019). Literature review as a research methodology : An overview and guidelines. *Journal of Business Research*, 104(August), 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Tai, T., & Chen, H. H. (2024). Computers & Education Improving elementary EFL speaking skills with generative AI chatbots : Exploring individual and paired interactions. *Computers & Education*, 220(July), 105112. <https://doi.org/10.1016/j.compedu.2024.105112>
- Tenorth, M., & Beetz, M. (2017). Representations for robot knowledge in the KNOWROB framework. *Artificial Intelligence*, 247, 151–169. <https://doi.org/10.1016/j.artint.2015.05.010>
- Thomas, M., & Sun, W. (2023). *The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners : a mixed methods investigation*. August. <https://doi.org/10.3389/fpsyg.2023.1210187>
- Trabelsi, Z., Alnajjar, F., Mohan, M., Parambil, A., & Gochoo, M. (2023). *Real-Time Attention Monitoring System for Classroom : A Deep Learning Approach for Student 's Behavior Recognition*. 1–17.
- Tsang, A. (2025). *The relationships between EFL learners ' anxiety in oral presentations , self-perceived pronunciation , and speaking proficiency*. <https://doi.org/10.1177/13621688221102522>
- Tzu, K., & Chen, C. (2022). *Speech-to-text recognition in University English as a Foreign Language Learning*. 9857–9875.
- Vu, N. N., Thach, V. T., Vu, T., & Thuy, D. (2025). *AI-Driven Pronunciation Assessment : The Impact of SpeechAce on EFL Learners ' Pronunciation Competency*. 26(3), 84–106.
- Wang, Y. (2025). A Study on the Efficacy of ChatGPT-4 in Enhancing Students' English Communication Skills. *SAGE Open*, 15(1), 1–17. <https://doi.org/10.1177/21582440241310644>
- Wu, T., Hapsari, I. P., & Huang, Y. (2025). Effects of incorporating AI chatbots into think – pair – share activities on EFL speaking anxiety , language enjoyment , and speaking performance think – pair – share activities on EFL speaking anxiety ,. *Computer Assisted Language Learning*, 0(0), 1–39. <https://doi.org/10.1080/09588221.2025.2478271>
- Yaşar, İ., & Murat, K. (2025). *Digital language learning with Duolingo : assessing its impact on listening , speaking , reading , and writing skills*. 0123456789. <https://doi.org/10.1007/s40692-025-00355-0>
- Young, E., & Donggil, O. (2021). Developmental research on an interactive

- application for language speaking practice using speech recognition technology. *Educational Technology Research and Development*, 69(2), 861–884. <https://doi.org/10.1007/s11423-020-09910-1>
- Zare-Behtash, E., & Sarlak, T. (2017). The Effect of Project Based Learning (PBL) on the Components of Speaking Ability of Iranian EFL Beginner Learners. *Journal of Applied Linguistics and Language Research*, 4(3), 119–130. www.jallr.com
- Zechner, K., & Evanini, K. (2019). *Automated Speaking Assessment; Using Language Technologies to Score Spontaneous Speech; First Edition*. Routledge. https://www.routledge.com/Automated-Speaking-Assessment-Using-Language-Technologies-to-Score-Spontaneous-Speech/Zechner-Evanini/p/book/9781138056879?srsId=AfmBOooKR3ayY3jjSmNFllwernNd668jc5DsF_fELiVAohtTzt6a2hMm
- Zhang, L., Jiang, Q., Xiong, W., & Zhao, W. (2025). *Effects of ChatGPT-Based Human – Computer Dialogic Interaction Programming Activities on Student Engagement* (Vol. 0, Issue 2555). <https://doi.org/10.1177/07356331251333874>
- Zheng, Y., Zhou, Y., Chen, X., & Ye, X. (2025). The influence of large language models as collaborative dialogue partners on EFL English oral proficiency and foreign language anxiety. *Computer Assisted Language Learning*, 0(0), 1–27. <https://doi.org/10.1080/09588221.2025.2453191>
- Zou, B., Guan, X., & Shao, Y. (2023). *Supporting Speaking Practice by Social Network-Based Interaction in Artificial Intelligence (AI) -Assisted Language Learning*.
- Zou, B., Liviero, S., Ma, Q., Zhang, W., Du, Y., & Xing, P. (2024). Exploring EFL learners' perceived promise and limitations of using an artificial intelligence speech evaluation system for speaking practice. *System*, 126(March), 103497. <https://doi.org/10.1016/j.system.2024.103497>