INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN PROGRESSIVE EDUCATION & DEVELOPMENT

OPEN ACCESS JOURNAL

Vol 14, Issue 3, (2025) E-ISSN: 2226-6348

Evaluate the Effectiveness of Using AI in Spoken English: A Literature Review

Bin Xu, *Hanita Hanim Ismail

Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Malaysia

Email: p114323@siswa.ukm.edu.my

Corresponding Author Email: hanitahanim@ukm.edu.my

DOI Link: http://dx.doi.org/10.6007/IJARPED/v14-i3/26366

Published Online: 11 September 2025

Abstract

With the growing prevalence of artificial intelligence (AI) in education, spoken English learning has undergone significant transformation. This literature review explores the effectiveness of AI-powered tools—including speech recognition, chatbots, and intelligent tutoring systems—in improving pronunciation, fluency, and learner engagement. Framed by Computer-Assisted Language Learning (CALL) and Second Language Acquisition (SLA) theories, the review synthesizes findings from over a decade of research. Evidence suggests that AI facilitates individualized learning and provides instant feedback that enhances oral proficiency. However, limitations persist, particularly in providing contextual feedback, accommodating diverse accents, and mitigating overreliance on AI systems. The review concludes by highlighting future directions such as context-aware AI systems, inclusive speech recognition, hybrid learning models, and metacognitive training. These insights inform pedagogical design and call for balanced integration of AI and human-led instruction to foster effective spoken English acquisition.

Keywords: Artificial Intelligence (AI), Spoken English, CALL, Second Language Acquisition, Pronunciation, Speech Technology, Chatbot, Learner Engagement

Introduction

In recent years, artificial intelligence (AI) has rapidly transformed numerous domains, from healthcare to finance, and education is no exception. In the context of language learning, AI technologies are reshaping how spoken English is taught and acquired. Traditionally, classroom-based approaches have relied heavily on face-to-face instruction, repetitive drills, and teacher-led feedback (Munro & Derwing, 2019). While these methods have long been foundational, they often provide limited speaking opportunities and may exacerbate learner anxiety, constraining the development of oral proficiency (Aichhorn & Puck, 2017).

AI-driven tools—including speech recognition systems, conversational chatbots, and virtual tutors—offer a potential solution by providing personalized, immediate feedback and flexible practice opportunities that transcend classroom limitations (Godwin-Jones, 2023; Holozsai & József, 2024). These innovations promise to enhance pronunciation, fluency, and learner

engagement, yet their adoption raises critical questions about efficacy and equity. Empirical studies report measurable improvements in pronunciation accuracy (Li & Hegelheimer, 2013), fluency (Zuo et al., 2023), and learner motivation (Moybeka et al., 2023). At the same time, critics caution that AI applications may fall short in delivering context-sensitive feedback, fostering pragmatic competence, and ensuring equal access across diverse learner populations (Eragamreddy, 2025).

Given these mixed outcomes, it is essential to examine AI's role in spoken English learning through robust theoretical lenses. Computer-Assisted Language Learning (CALL) and interactionist Second Language Acquisition (SLA) frameworks provide insights into how input, interaction, and meaningful communication contribute to language development (Long, 1996; Warschauer & Healey, 1998).

Guided by these frameworks, this literature review addresses three central questions:

1. How can AI-driven tools effectively enhance the pronunciation, fluency, and participation of spoken English learners?

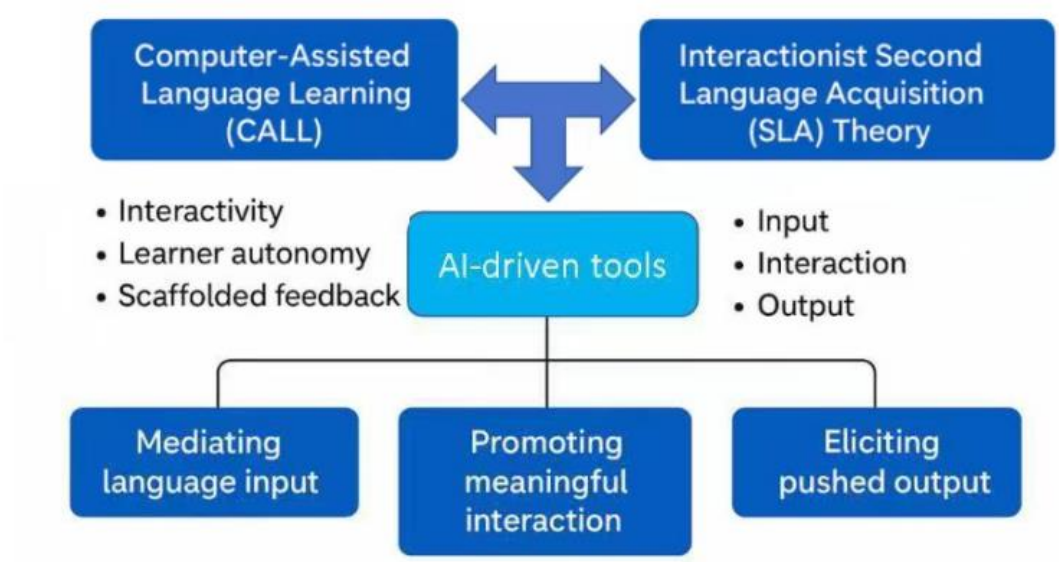
2. What theoretical principles best explain AI's influence on spoken English learning outcomes?

3. What persistent limitations exist, and how might future innovations address these challenges?

The review is organized as follows. Section 2 presents the theoretical framework, drawing on CALL and interactionist SLA perspectives. Section 3 synthesizes key empirical findings, focusing on pronunciation, fluency, and learner engagement. Section 4 critically examines existing limitations, including insufficient contextual feedback, technical and accessibility constraints, and potential over-reliance on AI at the expense of human interaction. Section 5 proposes directions for future research and practice, emphasizing context-aware AI systems, inclusive multilingual speech recognition, hybrid learning models, and the integration of autonomous learning with metacognitive training. Finally, Section 6 concludes by summarizing the insights derived from this review and their implications for spoken English education.

Theoretical Basis

The evaluation of the effectiveness of using AI in English speaking is grounded in two interrelated theoretical frameworks: Computer-Assisted Language Learning (CALL) and the interactionist theory of Second Language Acquisition (SLA). This review employs a dual-theoretical lens to investigate how AI-driven tools mediate language input, interaction, and output, as well as how they align with established principles of effective second language development.



According to Warschauer and Healey (1998), CALL thrives when technology-mediated language learning fosters interactivity, learner autonomy, and scaffolded feedback—features that are central to contemporary AI-based tools. As a field, CALL examines the intersection of technology and language education, emphasizing the role of digital tools in facilitating meaningful language use (Warschauer & Healey, 1998). Over time, CALL has evolved from drill-and-practice software to adaptive, AI-driven systems (Schmidt & Strasser, 2022), such as AI-powered speech recognition platforms that provide real-time pronunciation analysis, aligning with CALL's emphasis on immediate, form-focused feedback (Dennis, 2024). Similarly, adaptive tutoring systems employ machine learning to tailor practice to learners' proficiency levels, embodying CALL's learner-centered ethos (Gligorea et al., 2013).Three core CALL principles guide this review.First, interaction: CALL underscores the importance of dynamic, responsive language engagement. AI-driven tools, such as conversational agents, generate context-relevant prompts, adapt to learner input, and deliver immediate feedback, thereby simulating interactive communication. These affordances align with the "integrative CALL" vision of Warschauer and Healey (1998), in which technology mediates authentic communication.Second, autonomy: CALL promotes learner-centered, self-paced practice and individualized goal-setting (Benson, 2013). AI systems operationalize these principles through adaptive learning paths (e.g., difficulty adjustments based on performance) and self-monitoring tools (e.g., visualizations of pronunciation errors), enabling learners to take ownership of their progress (Li & Hegelheimer, 2013).Third, feedback: Effective CALL environments provide timely, specific, and actionable feedback (Engwall & Bälter, 2007). AI-powered speech recognition tools excel in this area by offering phoneme-level corrections (e.g., identifying mispronounced vowels) and grammatical feedback on spoken output—addressing a persistent challenge in traditional classrooms, where teachers often lack the time for individualized correction (Wang, 2024).Nevertheless, CALL scholars caution against overreliance on technology. Wiboolyasarin et al.(2025) note that AI-mediated systems often prioritize mechanical accuracy over communicative competence, potentially creating a disconnect between structured practice and real-world language use.

Based on the work of Long (1996) and Gass (2013), the interactionist theory of second language acquisition posits that language is acquired through meaningful interaction, particularly through negotiation of meaning. The core principles involved include comprehensible input, output and feedback, and social context. Studies have shown that when learners are exposed to input slightly above their current proficiency level (i + 1), comprehensible input is achieved through interactional modifications (Krashen, 1985). Regarding output and feedback, Swain and Lapkin (1995) emphasizes that producing language (output) pushes learners to notice the gaps in their knowledge, and that corrective feedback helps refine their interlanguage system. Furthermore, Canale and Swain (1980) stress that language learning occurs within social interaction, and that pragmatic and cultural norms play a critical role in communicative competence formation. Taken together, these principles underscore a dynamic interplay among interaction, cognitive processing, and social environment, thereby supporting AI-driven tools that simulate dialogues with virtual interlocutors engaging learners in negotiation of meaning. For instance, chatbotscan guide interactions on various topics ("Can you explain the present tense?") or ask learners to explain their input, creating opportunities for interactional modification akin to those in human dialogue (Alrajhi, 2024). Such low-anxiety environments encourage frequent output, which is a key driver in the development of oral fluency in English (Ding & Yusof, 2025). However, interactionist theory also highlights the limitations of AI. Long (1996) emphasizes that the "negotiation of meaning" in natural interaction is inherently collaborative and context-sensitive, involving nonverbal cues, emotional subtleties, and cultural knowledge—elements that current AI efforts attempt to replicate (AbuSahyon et al., 2023). For example, AI chatbots may fail to detect sarcasm or misunderstand culturally specific references, limiting their capacity to support pragmatic competence. Moreover, AI's ability to replicate human-like interaction remains limited. As Ziegler (2016) demonstrated, chatbots often fail to identify pragmatic errors (e.g., inappropriate politeness strategies), impeding learners' sociolinguistic development.

CALL (Computer-Assisted Language Learning) and interactionist theories of second language acquisition (SLA) together offer a comprehensive framework for assessing the role of artificial intelligence in English speaking learning. CALL elucidates how AI can enhance interactivity, learner autonomy, and feedback mechanisms. Conversely, interactionist SLA theory situates these technological affordances within a broader understanding of how language is acquired through use, emphasizing that AI-driven tools must foster not only accuracy but also meaningful communication. Moreover, both CALL and SLA theories highlight AI's potential to facilitate speaking practice by creating personalized, interactive, and feedback-rich environments. This theoretical foundation underpins the subsequent review of empirical studies that examine AI's effects on pronunciation, fluency, and learner engagement—and provides a basis for an analysis of the effectiveness of AI-mediated language learning experiences.

Key Findings in the Literature

Improvements in Pronunciation and Intonation

In English speaking proficiency, precise mastery of pronunciation and intonation is widely recognized as a core indicator of linguistic competence. AI technologies are driving transformative changes in this domain, offering opportunities that were previously unattainable in conventional classroom settings. A substantial body of empirical research

demonstrates that AI-powered speech recognition systems, underpinned by advanced algorithms, can transcend traditional pedagogical limitations by delivering real-time, detailed, and highly personalized pronunciation feedback, an advantage that has been extensively acknowledged in the literature (Li & Hegelheimer, 2013; Loor et al., 2024). These systems not only provide corrective feedback but also allow learners to engage in repeated practice without the constraints of teacher availability or classroom time, thereby fostering a self-directed learning environment.

The core functionality of these tools relies on sophisticated automatic speech evaluation algorithms, which can quantify multiple dimensions of learner speech, including segmental accuracy (vowels and consonants), stress patterns, intonation contours, and speech rhythm. This multi-dimensional analysis enhances learners' phonetic awareness and helps establish a metacognitive understanding of their own oral production. For example, in research, Saadia (2023) used a pretest-posttest control experimental group design to assess learners' pronunciation skills. A questionnaire was also conducted to evaluate participants' perceptions and attitudes towards the use of both tools. These findings suggest that TTS and ASR have the potential to facilitate different stages of pronunciation development and can aid learners in acquiring desired pronunciation features.

This finding highlights the capacity of AI to accelerate phonetic acquisition and provide learners with immediate, actionable insights into their pronunciation performance.Beyond segmental features, AI-driven tools demonstrate substantial value in training suprasegmental elements such as intonation, stress placement, and rhythm. Research by Bu et al. (2021) shows that systems offering visualized pitch contours and phoneme-level analysis enable learners to better perceive intonation patterns and the hierarchical structuring of stress in sentences. Such visual feedback not only makes abstract prosodic concepts tangible but also encourages learners to engage in reflective practice, comparing their speech output to model pronunciations and iteratively adjusting their production. This process fosters the development of self-monitoring skills and metacognitive awareness, allowing learners to transition from passive reception to active regulation of their oral output (El Kheir et al., 2023).

Despite these advantages, AI-driven pronunciation training tools are not without limitations. Hincks (2005) points out that although these systems are highly effective at detecting mispronounced phonemes, they still struggle with assessing the naturalness and overall prosodic fluency of speech. Specifically, AI often fails to capture the subtle nuances of intonation and expressive cues that native speakers naturally convey during everyday communication. Moreover, the effectiveness of AI feedback depends heavily on the diversity and quality of the training datasets. Systems trained predominantly on standardized accents may produce biased evaluations when confronted with non-native or regionally accented speech, potentially leading to inaccurate feedback and, in some cases, negatively impacting learner motivation (El Kheir et al., 2023).Nevertheless, the consensus in the literature remains clear: AI offers a scalable, flexible, and engaging avenue for improving pronunciation skills. When integrated with teacher-guided instruction, AI-driven tools and human expertise can complement each other, creating a synergistic learning environment. AI provides individualized practice opportunities and immediate corrective feedback, while teachers contribute cultural and pragmatic insights, emotional support, and nuanced modeling of

speech in context (Borrego , 2025; El Kheir et al., 2023). This blended approach leverages the strengths of both AI and traditional instruction, optimizing learners' phonetic development while maintaining a holistic focus on communicative competence.

Development of Fluency and Learner Confidence

Improving oral fluency and establishing communicative confidence are interrelated and crucial objectives in English speaking education. The effectiveness of artificial intelligence (AI) tools in this domain has garnered significant attention. AI-driven tools, exemplified by chatbot systems and virtual tutors, offer learners frequent and normalized opportunities for speaking practice, playing an irreplaceable role in automating language use and reducing expressive anxiety (AbuSahyon et al., 2023;Liu et al., 2023).

Compared to traditional classroom settings, AI-mediated practice environments present distinct advantages. They allow learners to engage in unrestricted repetitive practice and private language rehearsal, effectively eliminating the psychological pressure of making mistakes in front of others. This reduction in emotional barriers lays a foundation for the gradual development of self-confidence (Wei, 2023).Research by Zou et al. (2023) provides compelling evidence supporting this. Their longitudinal study tracked learners interacting with AI-driven conversational agents, revealing measurable improvements in speaking rate and reduced frequency of pauses, directly reflecting enhanced temporal fluency. Additionally, learners' self-efficacy and perceived speaking proficiency significantly increased, aligning with the findings of Ding and Yusof (2025), thereby highlighting the motivational impact of AI interactions in a non-judgmental environment.Furthermore, these AI-driven tools often incorporate structured dialogue prompts and adaptive scaffolding based on individual proficiency levels, maintaining an optimal balance between learning challenges and support (Stockwell, 2010). This design philosophy closely aligns with Vygotsky's Zone of Proximal Development (ZPD) theory, which posits that learners achieve the most significant learning gains when supported in challenging tasks beyond their current abilities (Vygotsky, 1978). Modern AI systems, particularly those integrating natural language processing and machine learning technologies, can dynamically adjust conversation difficulty and turn-taking patterns, progressively increasing demands as learning advances, thereby continuously promoting fluency (Qin & Zhong, 2024).

However, researchers have also pointed out limitations in AI-mediated dialogues. AI conversations, often based on preset algorithms and data models, lack the unpredictability and rich cultural nuances inherent in real human interactions, potentially limiting learners' development of pragmatic fluency (Anh et al., 2025; Huang, 2022). For instance, AI systems may struggle to simulate language comprehension biases arising from cultural differences in real social contexts. Moreover, prolonged reliance on scripted or guided dialogues may lead to overuse of formulaic language, weakening learners' ability to engage in spontaneous and flexible communication (Chapelle, 2007).Therefore, while AI systems provide a valuable supplementary platform for enhancing oral fluency, their effectiveness is maximized when integrated with authentic communicative tasks and interpersonal interactions. Combining AI practice with real-life interaction scenarios leverages the convenience of AI to increase practice frequency while cultivating the ability to use language flexibly in authentic contexts.

Enhanced Learner Engagement and Motivation

Learner engagement and intrinsic motivation are pivotal factors influencing second language acquisition outcomes. Artificial Intelligence (AI) have demonstrated significant potential in fostering cognitive engagement and emotional investment in English speaking practice. Numerous studies indicate that AI systems, through the integration of gamified design, personalized learning pathways, and real-time feedback mechanisms, can substantially enhance learners' intrinsic motivation and sustain their engagement over extended periods (Dehghanzadeh et al., 2021;Wati et al., 2024).Mobile applications such as Duolingo and ELSA Speak exemplify this trend. These platforms adeptly utilize gamified AI techniques to establish reward-based learning cycles, incorporating features like points, badges, and level progression to effectively encourage learners to engage in repetitive practice and actively set learning goals (Akhmedova, 2025). A comparative study by Nykyporets et al. (2024) further corroborates this, revealing that students using AI-driven pronunciation applications reported higher learning motivation and invested significantly more time in tasks compared to those using traditional audio language materials.Additionally, the interactivity and multimodal characteristics of AI-driven tools—such as the integration of text, voice, images, and videos—cater to diverse learner preferences and learning styles, thereby enhancing accessibility and learner satisfaction (Burston, 2015). For instance, visual learners can deepen their understanding through video demonstrations, while auditory learners can reinforce memory via voice feedback.Moreover, AI-driven systems support self-paced and self-directed learning modes, granting learners greater autonomy. This aligns with the core principles of Deci and Ryan's (2000) Self-Determination Theory, which posits that autonomy, competence, and relatedness are fundamental to fostering intrinsic motivation. AI-driven tools provide personalized learning feedback and detailed goal tracking features, enabling learners to monitor their progress trajectories and frequently experience a sense of achievement. This positive feedback loop significantly strengthens their long-term commitment to language learning (Zhang & Miao, 2025).

However, the literature also highlights potential risks. Excessive gamification may lead to superficial engagement, with learners focusing more on obtaining rewards than on deeply understanding and meaningfully using the language (Chan & Lo, 2024). Furthermore, while highly personalized learning paths can effectively stimulate learning motivation, they may also isolate learners from collaborative learning environments, which are crucial for developing social cognition and teamwork skills (Philp & Duchesne, 2016).Therefore, while the role of AI systems in promoting learner engagement and motivation is undeniable, they should be judiciously integrated into broader educational frameworks. When combined with collaborative learning activities and reflective learning practices, AI-driven tools can leverage their advantages while avoiding over-reliance on technology, ultimately achieving comprehensive and effective language learning.

Limitations and Challenges

Although artificial intelligence (AI) demonstrates numerous significant advantages in supporting English speaking learning, academic research has also clearly identified critical limitations and serious challenges. These limitations and challenges can broadly be categorized into three main areas: deficiencies in contextual understanding and pragmatic feedback, constraints arising from the technology itself and its accessibility, and risks associated with learners’ overreliance on AI systems. A thorough understanding of these

challenges is essential for setting reasonable expectations in practical teaching contexts, guiding the appropriate integration of AI into language learning environments, and fulfilling pedagogical responsibilities.

Lack of Contextual and Pragmatic Feedback

Current AI systems have increasingly strong capabilities in analyzing surface-level linguistic features, such as pronunciation accuracy and grammatical correctness. However, their ability to provide deeper feedback concerning discourse appropriateness, pragmatic rules, and socio-cultural context remains highly limited. This represents a significant shortcoming, as effective oral communication relies not only on accurate pronunciation but also on the ability to respond sensitively to specific contexts and employ language that aligns with cultural norms (Taguchi, 2011).Multiple studies have emphasized that contemporary AI models struggle to interpret subtle nuances in dialogue. For instance, AbuSahyon et al. (2023) observed that AI-driven tools often fail to detect behaviors that violate pragmatic norms, such as inappropriate speech acts in certain social situations or impolite tones that human interlocutors easily recognize. Similarly, Ziegler (2016) noted that most AI chatbots and virtual tutors lack mechanisms to ensure discourse coherence, often providing generic or contextually disconnected responses.

Moreover, an utterance may be linguistically correct yet socially inappropriate or inconsistent with the speaker's communicative intent, a scenario particularly common in cross-cultural interactions. As Kasper and Rose (2002) argued, the development of pragmatic competence requires exposure to rich, diverse contexts and authentic interactions, a process that current AI cannot fully replicate. This limitation may foster a suboptimal learning environment, overemphasizing linguistic form while neglecting communicative appropriateness, ultimately hindering learners' development of real-world speaking competence. For example, a learner may achieve correct grammar and pronunciation through AI-based practice but produce socially inappropriate remarks during actual interaction with foreign interlocutors due to unfamiliarity with cultural norms, potentially causing awkward or embarrassing situations (Borrego, 2025).

Technological and Accessibility Constraints

The application of artificial intelligence in English speaking instruction faces a persistent challenge: the digital divide, which is particularly pronounced in regions with limited technological infrastructure. AI-driven learning tools typically require stable internet connectivity, high-performance processing devices, and extensive training data to operate effectively (Burston, 2015). However, in resource-constrained educational environments or for learners in rural areas, these necessary conditions may be entirely unavailable, further exacerbating existing educational inequalities (Xalxo et al., 2025). For example, schools in remote mountainous areas often suffer from poor network signals and outdated computer equipment, rendering advanced AI-based speaking tools inaccessible. Compared to their urban counterparts, students in these areas face significantly fewer opportunities and conveniences for improving their English speaking skills (Ahmed, 2024).Moreover, speech recognition systems often exhibit performance biases when processing accented speech or non-standard English varieties. Vo(2016) found that learners with strong regional accents or diverse linguistic backgrounds frequently receive inconsistent, or even inaccurate, feedback from AI-based pronunciation systems. Such inaccurate feedback not only hinders learners'

progress and misguides their pronunciation practice but may also negatively impact their self-esteem and reduce their willingness to speak English. For instance, a learner with a regional accent may feel frustrated if the AI system repeatedly misjudges their pronunciation, gradually losing confidence in learning English(Nuryah, 2024).

Additionally, technical failures, software instability, and a lack of multilingual support further constrain the reliability of AI-driven tools across classrooms worldwide. As Godwin-Jones (2023) noted, AI systems are often trained on datasets that prioritize native-speaker norms, resulting in reduced performance when used by English speakers from other contexts. Without more inclusive training datasets and recognition algorithms capable of adapting to various accents and language varieties, AI-driven tools may inadvertently marginalize the very learners they are intended to support, leaving them unable to benefit from technological advances and facing additional learning obstacles due to technological limitations.

Overreliance on Technology and Reduced Interpersonal Interaction

An increasing number of studies have expressed concern that learners may overrely on AI-driven tools, thereby reducing interactions with real human interlocutors(Fathi et al., 2024). While AI systems are indeed effective in providing mechanical feedback and repetitive practice, they cannot facilitate the interpersonal negotiation and emotional co-construction that are intrinsic to natural language use (Lindgren, 2024). In authentic human communication, individuals continuously adjust their language and expression based on their interlocutor's responses, engaging in emotional exchange and mutual resonance—processes that AI cannot replicate(Mieczkowski et al., 2021). Some scholars argue that excessive use of AI-mediated dialogues, particularly scripted or context-independent interactions, may hinder learners' ability to transfer skills acquired in AI environments to dynamic, real-world conversations (Jawaid et al, 2025). For instance, Zuo et al. (2024) reported that learners who relied exclusively on AI tutors for speaking practice performed well in terms of linguistic accuracy during interviews with native speakers but displayed weaknesses in pragmatic adaptability. This finding highlights a disconnect between controlled practice environments and authentic spontaneous language use, echoing Long's (1996) early assertion that negotiation of meaning is central to successful second language acquisition.Furthermore, excessive automation may be counterproductive, adversely affecting learner autonomy. If learners overly depend on AI-generated prompts and corrections without developing intrinsic strategies to monitor and regulate their own language output, they may become passive recipients of feedback rather than active language users (Ziqi et al., 2017). Over time, this can undermine independent thinking and the ability to communicate effectively without AI assistance.

To mitigate this risk, some researchers advocate adopting blended learning models, positioning AI as a supportive tool rather than a replacement for human instruction (Yang et al., 2013). In summary, while AI-driven tools offer unprecedented opportunities for language practice, their current limitations underscore the need to maintain balance in implementation. Future developments must prioritize inclusivity, adaptability, and alignment with pedagogical approaches to ensure that AI enhances, rather than undermines, the overall language acquisition process.

Future Directions and Pedagogical Implications

Given the current limitations and potential of AI in spoken English learning, several future directions emerge for both research and classroom practice. These developments emphasize the need for a more nuanced, equitable, and learner-centered integration of AI technologies.

Development of Context-Aware AI Systems

In English speaking instruction, the cultivation of pragmatic competence and the enhancement of context-based communicative skills are crucial. However, current artificial intelligence (AI) technologies exhibit notable limitations in these areas. To address this issue, future AI applications should actively integrate advanced techniques in affective computing and contextual natural language processing (NLP). Deeply incorporating extensive socio-cultural databases and practical corpora into AI algorithms can significantly enhance the system's ability to assess and respond to learners' spoken English, thereby improving contextual accuracy (Devi & Sharma, 2024; Praveena & Anupama, 2025; Xue & Liu, 2025). For instance, when learners engage in dialogues simulating business negotiation scenarios, AI systems can evaluate whether their utterances are appropriate by referencing the socio-cultural norms of business etiquette.Research on intelligent agents capable of detecting emotional tone, conversational intent, and discourse cohesion further improves the realism and communicative value of AI-mediated interactions (Ismail & Alharkan, 2024; Zhang, 2024). In spoken English dialogues, variations in learners' intonation and speech rate often convey different emotions and intentions. If AI systems can accurately capture these signals, they can generate responses that are more contextually relevant. For example, when a learner expresses an opinion hesitantly, the system can detect uncertainty and provide encouraging feedback or additional guidance.

Furthermore, developing simulation-based role-playing dialogues with multiple branching paths, guided by learner responses, can better emulate real-life English conversational scenarios. In such dialogues, each learner utterance triggers different conversational pathways, allowing AI systems to provide nuanced feedback according to the learner's performance in various branches. For example, in a simulated restaurant ordering scenario, learners' different English expressions elicit corresponding responses from the AI-driven server. If a learner's phrasing is inappropriate or the tone is unsuitable, the AI can promptly highlight the issue and offer more appropriate expressions (Graesser et al., 2005).

Inclusive and Multilingual Speech Recognition

English speaking learners come from diverse linguistic backgrounds, producing a wide range of accents, which presents a major challenge for AI speech recognition. To ensure that AI-driven tools can be effectively used by a broad spectrum of learners, especially non-native speakers with varying accents, future AI systems must be trained on larger and more diverse datasets encompassing regional, ethnic, and learner-specific English pronunciation patterns (Anis, 2023; Danka, 2024). This is essential for AI systems to adapt to the unique pronunciation characteristics of different learners.Developers should prioritize the creation of automatic speech recognition (ASR) systems that mitigate accent bias to provide a fair and equitable English learning experience(Ngueajio & Washington, 2022). ASR systems should accurately recognize the pronunciations of learners with Chinese, Indian, Japanese, or other distinct accents, providing objective assessment and feedback rather than penalizing deviations from a so-called "standard" accent (Van Moere & Suzuki, 2017). For instance, common

pronunciation challenges among Chinese learners, such as the production of the “th” sound, should be correctly identified and addressed with targeted corrective suggestions rather than merely marked as incorrect(Ercan & Kunt, 2019).

Additionally, supporting code-switching and multilingual input is critical in English speaking practice. Many learners unconsciously mix their native language into English sentences. AI-driven tools that understand and accommodate this type of language use can more accurately reflect learners’ linguistic habits, thereby increasing engagement and fostering a sense of inclusion. For example, when a learner says, “I like eat apple (I like to eat apples),” an effective AI system can comprehend the intended meaning and guide the learner toward expressing it fully in English, rather than simply marking it as wrong (Olson, 2016; Stockwell, 2012).

Blended Learning Models

Although artificial intelligence (AI) offers significant advantages in scalability and personalized feedback for English speaking learning, it should ultimately serve as a complement to human interaction rather than a replacement. Future English speaking instructional models can benefit greatly from adopting a blended learning approach that integrates AI into the Communicative Language Teaching (CLT) framework (Oudghiri, 2025; Yaroslavova et al., 2020; Zhou et al., 2025).Specifically, AI can be effectively utilized for repetitive pronunciation practice. Due to its capacity to provide immediate and precise feedback, learners can repeatedly practice specific English phonemes, words, or sentences until achieving satisfactory accuracy. In contrast, peer interactions and teacher feedback can focus more on helping learners enhance pragmatic competence and fluency in real-world English communication(Fathi et al., 2024). For example, in classroom settings, learners might first use AI-driven tools for extensive phonetic exercises to master basic pronunciation skills. Subsequently, teachers can organize group discussions where learners communicate on a given topic in English, with teachers and peers concentrating on the appropriateness, fluency, and clarity of learners’ language use, providing targeted guidance (Borrego, 2025; Godwin-Jones, 2018).

Multiple studies indicate that human–AI collaborative teaching models—where teachers leverage data analysis generated by AI to diagnose learners’ specific needs and tailor individualized instruction—can effectively improve learning outcomes (Çiftçi, 2024; Shafiee Rad & Roohani, 2024). AI can analyze learners’ spoken English practice data to identify weaknesses, such as frequent grammatical errors or pragmatic missteps in particular scenarios, enabling teachers to design corresponding instructional activities to help learners overcome these challenges(Daniela, 2024).

Autonomous Learning and Metacognitive Training

To prevent learners from becoming overly reliant on AI-driven tools, it is essential to train them in the strategic use of these resources. Future research should explore instructional interventions that cultivate learners’ metacognitive awareness, teaching them how to critically interpret and evaluate AI-generated spoken English feedback (Benson, 2013; Flavell, 1979). This entails that learners should not blindly accept AI feedback but learn to analyze its validity and integrate it with their own learning goals and context to make informed adjustments.Integrating reflective prompts and self-assessment tools into AI platforms can help learners take greater control of their English speaking development(Mohebbi, 2025). For

instance, after providing pronunciation feedback, AI-driven tools can simultaneously pose reflective questions(Dai & Wu, 2023;Nuryah, 2024). Additionally, self-assessment tools enable learners to periodically evaluate their spoken English performance, compare it with AI-generated assessments, identify discrepancies, and gain a clearer understanding of their learning status, thereby facilitating more effective planning of English speaking practice (Goto Butler & Lee, 2010; Jamrus & Razali, 2019). Through such training, learners can gradually cultivate autonomous learning skills, enabling them to practice and improve their spoken English effectively even without AI assistance.

Conclusion

The integration of artificial intelligence in spoken English learning has ushered in a new era of personalized, accessible, and efficient language education. The research findings from the literature review demonstrate that artificial intelligence (AI) technology shows considerable promise in enhancing various aspects of English oral language acquisition, particularly in pronunciation, fluency, and learning motivation. However, the integration of AI into language learning is not without challenges. This discussion aims to synthesize key findings through the lens of theoretical frameworks (CALL and SLA), evaluate their pedagogical implications and limitations, and propose future research directions.

From the perspective of Computer-Assisted Language Learning (CALL), AI technologies provide opportunities for immediate feedback, autonomous learning, and interactive engagement—all key elements for effective language acquisition as identified by Warschauer and Healey (1998). Tools such as speech recognition applications, conversational chatbots, and virtual pronunciation coaches have demonstrated their ability to help learners improve phonetic accuracy and build fluency through repetitive, low-pressure practice. Furthermore, consistent with the interactionist perspective of Second Language Acquisition (SLA) (Long, 1996), these tools simulate real-time conversations, enabling learners to engage in meaning negotiation and interaction sequences that approximate natural communication.

Despite these benefits, several limitations hinder the full effectiveness of AI in English oral proficiency development. A major concern is the lack of contextually rich and pragmatically accurate feedback. While AI can detect pronunciation errors, it often fails to assess the appropriateness of utterances within sociocultural contexts, which is crucial for pragmatic competence. As AbuSahyon et al. (2023) noted, the ability to respond appropriately to discourse nuances remains a limitation of current AI systems.Technical barriers also persist, particularly for learners in under-resourced environments. Many AI-powered tools require high-bandwidth internet and advanced hardware, creating an accessibility gap(Taguchi, 2011). Additionally, speech recognition systems frequently struggle with accented speech, leading to misinterpretations and erroneous feedback that may frustrate learners or impede progress. Another emerging challenge is the potential overreliance on AI-driven tools(Zhang & Miao, 2025). Although learners may gain confidence through AI-supported training, reduced opportunities for real-world interaction may weaken their ability to transfer these skills to authentic communicative environments(Kasper & Rose, 2002). As some researchers have cautioned (Borrego, 2025), the lack of peer or teacher interaction may result in superficial learning.

Given these limitations, future research should focus on developing more context-aware and inclusive AI systems. Emotionally intelligent and pragmatically sensitive AI capable of detecting tone, intention, and discourse cohesion could better simulate authentic communication and enhance learner engagement. For instance, adaptive role-play scenarios with branching dialogues based on learner input could provide more nuanced feedback and reflect real-life communication needs. Furthermore, to address linguistic diversity and equity issues, future Automatic Speech Recognition (ASR) systems must be trained on larger, more representative datasets that account for regional and non-native accents (Yuan et al., 2022). Incorporating multilingual and code-switching environments into AI tool design could also reflect learners' real-world language use, fostering inclusivity and relevance. Importantly, AI should complement rather than replace human interaction. Blended learning models that integrate AI into Communicative Language Teaching (CLT) can leverage the strengths of both approaches. For example, AI could provide individualized pronunciation training while teachers focus on feedback regarding pragmatic appropriateness and discourse coherence. Evidence suggests that such co-teaching models can enhance learning outcomes (Chen, 2021). Finally, to mitigate overreliance on AI, learners need training in metacognitive strategies for critical tool use. Incorporating reflective prompts, self-assessment checklists, and strategy guidance into AI platforms can foster learner autonomy (Benson, 2013). Teaching learners to interpret AI feedback judiciously and reflect on their progress aligns with broader goals of learner empowerment and lifelong learning..

This review contributes to the social sciences and educational technology literature by synthesizing empirical and theoretical insights on AI’s role in spoken English learning. Its novelty lies in integrating findings across pronunciation, fluency, and engagement domains while highlighting the interplay between AI tools and pedagogical frameworks such as CALL and interactionist SLA. By systematically identifying both opportunities and limitations, this study provides a foundation for future research and practical implementation, emphasizing inclusive, context-aware, and human-complementary AI interventions. In conclusion, AI serves as a valuable complement to traditional English oral language pedagogy, offering scalable and personalized learning opportunities. However, maximizing its effectiveness requires addressing contextual, technological, and pedagogical challenges. Future innovations must emphasize contextual awareness, inclusivity, and integration with human-mediated learning environments.

References

AbuSahyon, A. S. A. E., Alzyoud, A., Alshorman, O., & Al-Absi, B. (2023). AI-driven technology and chatbots as tools for enhancing English language learning in the context of second language acquisition: A review study. *International Journal of Membrane Science and Technology, 10*(1), 1209-1223.

Ahmed, F. (2024). The digital divide and AI in education: Addressing equity and accessibility. *AI EDIFY Journal, 1*(2), 12-23.

Aichhorn, N., & Puck, J. (2017). "I just don't feel comfortable speaking English": Foreign language anxiety as a catalyst for spoken-language barriers in MNCs. *International Business Review, 26*(4), 749-763.

Akhmedova, M. M. (2025, May). Gamification and Artificial Intelligence in Language Teaching: Innovative Ways to Enhance Effectiveness. In *International Conference on MODERN EDUCATION AND INNOVATIVE APPROACHES TO TEACHING FOREIGN LANGUAGES*" (pp. 77-81).

Alrajhi, A. S. (2024). Artificial intelligence pedagogical chatbots as L2 conversational agents. *Cogent Education, 11*(1), 2327789.

Anh, N. H., Baharloo, A., & Pasaribu, G. R. (2025). Integrating AI-Based Tools to Enhance Speaking Fluency in EFL Classrooms in Labuhanbatu Utara. *JELIM: Journal of Education, Language, Social and Management, 3*(1), 7-16.

Anis, M. (2023). Leveraging artificial intelligence for inclusive English language teaching: Strategies and implications for learner diversity. *Journal of Multidisciplinary Educational Research, 12*(6), 54-70.

Benson, P. (2013). *Teaching and researching: Autonomy in language learning.* Routledge.

Borrego, M. M. (2025). Artificial Intelligence-Assisted Error Analysis: Developing Pragmatic Competence in EFL Learners. In *Application of AI in the Teaching and Learning of English as a Foreign Language (EFL)* (pp. 99-148). IGI Global Scientific Publishing.

Bu, Y., Ma, T., Li, W., Zhou, H., Jia, J., Chen, S., ... & Liu, Z. (2021, May). PTeacher: A computer-aided personalized pronunciation training system with exaggerated audio-visual corrective feedback. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-14).

Burston, J. (2015). Twenty years of MALL project implementation: A meta-analysis of learning outcomes. *ReCALL, 27*(1), 4-20.

Canale, M., & Swain, M. (1980). Theoretical bases of commuicative approaches to second language teaching and testing.*Applied Linguistics, 1*(1), 1–47.

Chan, S., & Lo, N. (2024, August). Enhancing EFL/ESL instruction through gamification: a comprehensive review of empirical evidence. In *Frontiers in Education* (Vol. 9, p. 1395155). Frontiers Media SA.

Chapelle, C. A. (2007). Technology and second language acquisition. *Annual review of applied linguistics, 27*, 98-114.

Chen, C. (2021). *A bilingual/L2 hybrid intervention model: combining human and machine intelligences* (Doctoral dissertation, University of Illinois at Urbana-Champaign).

Çiftçi, A. (2024). *AI Assisted Teaching: Practices and Perspectives of Instructors on Using AI Tools in ELT* (Master's thesis, Maltepe University (Turkey)).

Dai, Y., & Wu, Z. (2023). Mobile-assisted pronunciation learning with feedback from peers and/or automatic speech recognition: A mixed-methods study. *Computer Assisted Language Learning, 36*(5-6), 861-884.

Daniela, M. (2024). The application of Artificial Intelligence in EFL learning and teaching: considerations, challenges and practical examples for educators. *Juan C. Figuereo Benítez*.

Danka, S. (2024). *The effectiveness of a Computer-Assisted Pronunciation Training Approach for the production and perception of linking by English L2 learners* (Doctoral dissertation, University of Leicester).

Deci, E. L., & Ryan, R. M. (2000). The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry, 11*(4), 227-268.

Dehghanzadeh, H., Fardanesh, H., Hatami, J., Talaee, E., & Noroozi, O. (2021). Using gamification to support learning English as a second language: a systematic review. *Computer Assisted Language Learning, 34*(7), 934-957.

Dennis, N. K. (2024). Using AI-Powered Speech Recognition Technology to Improve English Pronunciation and Speaking Skills. *IAFOR Journal of Education, 12*(2), 107-126.

Devi, S., & Sharma, V. (2024). The AI-supported Language Library Model: It's Not About the CALL Tools, It's About the Language Library. Library of Progress-Library Science, *Information Technology & Computer, 44*(3).

Ding, D., & Yusof, M. B. (2025). Investigating the role of AI-powered conversation bots in enhancing L2 speaking skills and reducing speaking anxiety: a mixed methods study. *Humanities and Social Sciences Communications, 12*(1), 1-16.

Engwall, O., & Bälter, O. (2007). Pronunciation feedback from real and virtual language teachers. *Computer Assisted Language Learning, 20*(3), 235-262.

Eragamreddy, N. (2025). The impact of AI on pragmatic competence. *Journal of Teaching English for Specific and Academic Purposes*, 169-189.

Ercan, H., & Kunt, N. (2019). Is it really difficult to produce correct th (/ð/and/θ/) sounds in English for EFL Learners? A case study in Northern Cyprus. *Journal of Language and Linguistic Studies, 15*(3), 926-936.

Fathi, J., Rahimi, M., & Derakhshan, A. (2024). Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System, 121*, 103254.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist, 34*(10), 906.

Gass, S. M. (2013). *Input, interaction, and the second language learner*. Routledge.

Gligorea, I., Cioca, M., Oancea, R., Gorski, A. T., Gorski, H., & Tudorache, P. (2023). Adaptive learning using artificial intelligence in e-learning: A literature review. *Education Sciences, 13*(12), 1216.

Godwin-Jones, R. (2018). Using mobile technology to develop language skills and cultural understanding. *Language Learning & Technology, 22*(3), 3-17.

Godwin-Jones, R. (2023). Emerging spaces for language learning: AI bots, ambient intelligence, and the metaverse. *Language Learning & Technology, 27*(2), 6-27.

Godwin-Jones, R. (2023). Presence and agency in real and virtual spaces: The promise of extended reality for language learning.

Goto Butler, Y., & Lee, J. (2010). The effects of self-assessment among young learners of English. *Language testing, 27*(1), 5-31.

Graesser, A. C., Chipman, P., Haynes, B. C., & Olney, A. (2005). AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education, 48*(4), 612-618.

Hincks, R. (2005). *Computer support for learners of spoken English* (Doctoral dissertation, KTH).

Holozsai, J., & József, H. (2024). Aspects of digital learning among foreign language learners in secondary education.

Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of computer assisted learning, 38*(1), 237-257.

Ismail, S., & Alharkan, A. (2024). EFL learners' positive emotions in the era of technology: Unpacking the effects of artificial intelligence on learning enjoyment, self-efficacy, and resilience. *Computer-Assisted Language Learning Electronic Journal, 25*(4), 526-551.

Jamrus, M. H. M., & Razali, A. B. (2019). Using self-assessment as a tool for English language learning. *English Language Teaching, 12*(11), 64-73.

Jawaid, A., Batool, M., Arshad, W., ul Haq, M. I., Kaur, P., & Sanaullah, S. (2025). AI and English language learning outcomes. *Contemporary Journal of Social Science Review, 3*(1), 927-935.

Kasper, G., & Rose, K. R. (2002). Pragmatic development in a second language. *Language learning*.

Kheir, Y. E., Ali, A., & Chowdhury, S. A. (2023). Automatic Pronunciation Assessment--A Review. *arXiv preprint arXiv:2310.13974.*

Krashen, S. D. (1985). The input hypothesis: Issues and implications. *(No Title)*.

Li, Z., & Hegelheimer, V. (2013). Mobile-assisted grammar exercises: Effects on self-editing in L2 writing.

Lindgren, H. (2024). Emerging roles and relationships among humans and interactive AI systems. *International Journal of Human–Computer Interaction*, 1-23.

Liu, X. J., Wang, J., & Zou, B. (2025). Evaluating an AI speaking assessment tool: Score accuracy, perceived validity, and oral peer feedback as feedback enhancement. *Journal of English for Academic Purposes, 75*, 101505.

Long, M. (1996). The role of the linguistic environment in second language acquisition. *Handbook of second language acquisition*.

Loor, M. A. M., Solorzano, D. M. A., Katherine, A., & Moreira, V. (2024). Integration of artificial intelligence in english teaching. *Journal of Cleaner Producfion, 289*, 125834.

Mieczkowski, H., Hancock, J. T., Naaman, M., Jung, M., & Hohenstein, J. (2021). AI-mediated communication: Language use and interpersonal effects in a referential communication task. *Proceedings of the ACM on Human-Computer Interaction, 5*(CSCW1), 1-14.

Mohebbi, A. (2025). Enabling learner independence and self-regulation in language education using AI tools: a systematic review. *Cogent Education, 12*(1), 2433814.

Moybeka, A. M., Syariatin, N., Tatipang, D. P., Mushthoza, D. A., Dewi, N. P. J. L., & Tineh, S. (2023). Artificial Intelligence and English classroom: the implications of AI toward EFL students' motivation. *Edumaspul: Jurnal Pendidikan, 7*(2), 2444-2454.

Munro, M. J., & Derwing, T. M. (2019). Collecting data in L2 pronunciation research. *Pronunciation in Second Language Learning and Teaching Proceedings, 11*(1).

Ngueajio, M. K., & Washington, G. (2022, June). Hey ASR system! Why aren't you more inclusive? Automatic speech recognition systems' bias and proposed bias mitigation techniques. A literature review. In *International conference on human-computer interaction* (pp. 421-440). Cham: Springer Nature Switzerland.

Nuryah, A. (2024). *The Comparison Between Lecturer and AI-Powered Platform'Feedback in Assessing English Pronunciation* (Bachelor's thesis, FITK UIN Syarif Hidayatullah Jakarta).

Nykyporets, S. S., Melnyk, O. D., & Medvedieva, S. O. (2024). A comparative analysis of digital technology utilization for English language learning among technical students in Ukraine

and the European. *Цифровізація як інструмент забезпечення якості надання освітніх послуг з урахуванням європейського досвіду. Розд. 2.2: 112-161.*

Olson, C. L. (2016). *The priming effect of English subject-predicate on Chinese topic-comment in English to Chinese translation*. University of Hawai'i at Manoa.

Oudghiri, A. (2025). Optimizing Language Learning in Blended Environments: Enhancing Engagement and. *Proceedings of the E-Learning and Smart Engineering Systems (ELSES 2024), 36*, 414.

Philp, J., & Duchesne, S. (2016). Exploring engagement in tasks in the language classroom. *Annual Review of Applied Linguistics, 36*, 50-72.

Praveena, T., & Anupama, K. (2025). Machine Learning Meets Language Learning: The Transformative Potential of Artificial Intelligence in English Language Instruction. *Human Research in Rehabilitation, 15*(1).

Qin, L., & Zhong, W. (2024). Adaptive system of English-speaking learning based on artificial intelligence. *Journal of Electrical Systems, 20*(6s), 267-275.

Saadia, K. H. (2023). Assessing the effectiveness of text-to-speech and automatic speech recognition in improving EFL learner's pronunciation of regular past-ed.

Schmidt, T., & Strasser, T. (2022). Artificial intelligence in foreign language learning and teaching: a CALL for intelligent practice. *Anglistik: International Journal of English Studies, 33*(1), 165-184.

Shafiee Rad, H., & Roohani, A. (2024). Fostering L2 learners' pronunciation and motivation via affordances of artificial intelligence. *Computers in the Schools*, 1-22.

Stockwell, G. (2010). Using mobile phones for vocabulary activities: Examining the effect of platform.

Stockwell, G. (Ed.). (2012). *Computer-assisted language learning.* Cambridge University Press.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied linguistics, 16*(3), 371-391.

Taguchi, N. (2011). Teaching pragmatics: Trends and issues. *Annual review of applied linguistics, 31*, 289-310.

Van Moere, A., & Suzuki, M. (2017). Using speech processing technology in assessing pronunciation. In *Assessment in second language pronunciation* (pp. 137-152). Routledge.

Vo, S. (2016). Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research. *Pronunciation in Second Language Learning and Teaching Proceedings, 7*(1).

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (Vol. 86). Harvard university press.

Wang, D. (2024). Teacher-versus AI-generated (Poe application) corrective feedback and language learners' writing anxiety, complexity, fluency, and accuracy. *International Review of Research in Open and Distributed Learning, 25*(3), 37-56.

Warschauer, M., & Healey, D. (1998). Computers and language learning: An overview. *Language teaching, 31*(2), 57-71.

Wati, S. A., Ningsih, F., & Setiani, K. (2024). The Role of Mobile Apps in Increasing Student Engagement and Motivation in English Language Learning. ICOERESS, 1(1), 161-176.

Wei, L. (2023). Artificial intelligence in language instruction: impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in psychology, 14,* 1261955.

Wiboolyasarin, W., Wiboolyasarin, K., Tiranant, P., Jinowat, N., & Boonyakitanont, P. (2025). AI-driven chatbots in second language education: A systematic review of their efficacy and pedagogical implications. *Ampersand, 14*, 100224.

Xalxo, P. V., Kindo, J., & Kachhap, P. (2025). Online Education Ecosystem—Exploring the Challenges and Opportunities. *European Journal of Education, 60*(2), e70085.

Xue, H., & Liu, W. (2025). Bibliometric Analysis of Natural Language Processing Technology in Education: Hot Topics, Frontier Evolution, and Future Prospects. S*AGE Open, 15*(1), 21582440251319891.

Yang, Y. T. C., Chuang, Y. C., Li, L. Y., & Tseng, S. S. (2013). A blended learning environment for individualized English listening and speaking integrating critical thinking. Computers & Education, 63, 285-305.

Yaroslavova, E. N., Kolegova, I. A., & Stavtseva, I. V. (2020). Flipped classroom blended learning model for the development of students' foreign language communicative competence. *Перспективы науки и образования*, (1), 399-412.

Zhang, M. (2024). The Impact of Artificial Intelligence Virtual Oral Tutoring APPs on Chinese Youth's Anxiety in Oral English learning——Interview Research Based on Users of Artificial Intelligence Speaking Tutoring APPs. *Journal of Language, Culture and Education, 1*(1), 31-38.

Zhang, Y., & Miao, Z. (2025). Enhancing EFL Learners' Engagement and Motivation Through Immersive Technologies: The Role of Artificial Intelligence, Augmented Reality, Virtual Reality, and Mobile Applications. *European Journal of Education, 60*(2), e70128.

Zhou, Q., Hashim, H., & Sulaiman, N. A. (2025). Supporting english speaking practice in higher education: the impact of AI chatbot-integrated mobile-assisted blended learning framework. *Education and Information Technologies*, 1-32.

Ziegler, N. (2016). Synchronous computer-mediated communication and interaction: A meta-analysis. *Studies in Second Language Acquisition, 38*(3), 553-586.

Ziqi, C., Xinhua, Z., Qi, L., & Wei, W. (2024). L2 students' barriers in engaging with form and content-focused AI-generated feedback in revising their compositions. *Computer Assisted Language Learning*, 1-21.

Zou, B., Du, Y., Wang, Z., Chen, J., & Zhang, W. (2023). An investigation into artificial intelligence speech evaluation programs with automatic feedback for developing EFL learners' speaking skills. *Sage Open, 13*(3), 21582440231193818.

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*, 103497.