



Volume 2 Issue 1 Year 2026 Pages 1-21

e-ISSN 3090-6245 | DOI: 10.70152

<https://journal.akademimerdeka.com/ojs/index.php/matcha/index>

Enhancing Oral Communication in Real Time: University Students' Self-Directed Use of AI-Powered Speech Recognition in English Classrooms

Masrurotul Ajiza^{1*}, Ali Said Al Matari², Muthmainnah²

¹ Institut Teknologi Nasional Malang, Indonesia

² A'Sharqiyah University, Oman

³ Universitas Al Asyariah Mandar Indonesia, Indonesia

*Corresponding author's email: masrurotul_ajiza@lecturer.itn.ac.id

DOI: <https://doi.org/10.70152/matcha.v2i1.332>

Abstract: The increasing availability of AI-powered speech recognition has created new possibilities for supporting oral communication in English classrooms. This study explores how university students use AI-powered speech recognition tools in a self-directed manner during real-time speaking activities and how they perceive the usefulness of these tools for enhancing oral communication. Adopting a convergent mixed-methods design, the study integrates qualitative data from classroom observations and semi-structured interviews with quantitative data from a descriptive Likert-scale questionnaire. The qualitative findings reveal that students use speech recognition strategically to monitor intelligibility, rehearse spoken output, and manage speaking-related anxiety during communicative tasks. Rather than replacing interaction, the tool functions as a flexible support that learners draw on selectively according to situational needs. The quantitative results indicate generally positive perceptions of the tool's usefulness, particularly in relation to confidence, fluency awareness, and ease of use. By combining observed practices with learners' reported perceptions, this study offers a classroom-grounded account of AI-powered speech recognition use in real-time speaking contexts. The findings contribute to ongoing discussions on AI-assisted language learning by highlighting the pedagogical potential of speech recognition as a learner-oriented resource for supporting oral communication.

Keywords: AI-powered speech recognition; English oral communication; mixed methods; real-time speaking; self-directed learning

INTRODUCTION

Oral communication remains a core yet challenging component of English language learning in higher education. University students are frequently expected to participate in discussions, deliver presentations, and engage in collaborative tasks that require spontaneous and accurate spoken interaction in English (Leeming et al., 2024; Oli et al.,

2023). Despite continued exposure to English instruction, many learners struggle with fluency, pronunciation, and confidence during real-time communication (Bashori et al., 2024; Chou, 2024; Manorom et al., 2024). These difficulties are often intensified by anxiety, fear of making errors, and limited opportunities for individualized feedback during speaking activities (Liao, 2025). Consequently, increasing attention has been directed toward pedagogical approaches and technological tools that can support learners during the speaking process itself rather than relying solely on post-task evaluation.

Recent developments in artificial intelligence have expanded the range of digital tools available for language learning. Among these, AI-powered speech recognition technologies have attracted considerable attention due to their ability to provide immediate automated transcription of spoken language (Broeren et al., 2025; Russell et al., 2024). By converting speech into text in real time, these tools allow learners to visually monitor their spoken output during communication (Zou et al., 2023). This feature enables learners to notice pronunciation inaccuracies, grammatical errors, and lexical limitations while speaking, which might otherwise go unnoticed in fast-paced classroom interactions (Mingyan et al., 2025; Qiao & Zhao, 2023). Compared with traditional feedback mechanisms, which are often delayed or limited by instructional time, speech recognition technologies can offer continuous and individualized support.

Speech recognition is now integrated into various platforms, including mobile applications, web-based learning environments, and AI-driven language tools. In English classrooms, these technologies are increasingly used not only for isolated pronunciation practice but also during communicative activities such as role plays, group discussions, presentation rehearsals, and spontaneous speaking tasks (Jakonen et al., 2025; Sun, 2023). However, as noted by Sangkawong et al. (2025), the pedagogical value of these technologies cannot be fully understood without examining how learners engage with them during real classroom interactions.

A notable characteristic of AI-powered speech recognition in educational settings is that its use is often initiated and regulated by learners themselves. Rather than being strictly embedded in instructional designs, these tools are frequently accessed at learners' discretion depending on their immediate communicative needs (Asratie et al., 2023; Mingyan et al., 2025; Qiao & Zhao, 2023; Zhang, 2024). Learners may use them to check pronunciation, monitor accuracy, rehearse spoken output, or reduce anxiety before speaking in front of peers (Dillon & Wells, 2023; Kusz & Pawliszko, 2025; Wardat & Akour, 2025). Such practices illustrate how AI tools are integrated into ongoing classroom activities (Fitriati & Williyon, 2025). Nevertheless, these naturally occurring practices remain underexplored.

Existing research on AI-assisted speaking has largely focused on technical accuracy, system performance, or learning outcomes measured through experimental or quasi-experimental designs (El Shazly, 2021; Wang et al., 2025). Although these studies provide valuable insights into the potential benefits of speech recognition technologies,

they often rely on short-term interventions or controlled tasks that may not reflect everyday classroom realities (Suzuki et al., 2025). Consequently, empirical evidence describing how university students independently use speech recognition tools during real-time classroom communication remains limited.

Students' perceptions of AI-powered speech recognition also represent an important but relatively underexplored dimension. Learners' perceptions of usefulness, ease of use, and emotional support strongly influence whether technologies are meaningfully adopted and sustained (Al-kfairy, 2024; Alshammari & Alkhwalidi, 2025; Li et al., 2025). In speaking contexts, where affective factors such as anxiety and confidence are particularly salient, understanding learners' experiences with speech recognition tools becomes especially important (Elov et al., 2025; Sun, 2023).

Another limitation concerns the methodological orientation of existing research. Many studies emphasize causal explanations or effectiveness claims, often overlooking descriptive mixed-methods approaches (Ding & Yusof, 2025; Nguy et al., 2025; Yang et al., 2025). Qualitative-dominant mixed-methods designs, which combine classroom observations and learner accounts with descriptive quantitative data such as Likert-scale questionnaires, are well suited for exploratory research in emerging technological contexts (Liu et al., 2024; Mizza et al., 2025).

Overall, the literature reveals several interconnected gaps. First, limited research examines how university students use AI-powered speech recognition tools in a self-directed manner during real-time classroom communication. Second, students' perceptions of the usefulness of these tools for supporting oral communication remain insufficiently documented. Third, there is a need for research designs that foreground classroom practices and learner experiences rather than focusing primarily on experimental outcomes. Addressing these gaps is essential for developing a grounded understanding of the role of AI-powered speech recognition in contemporary English language teaching. To address these issues, the present study investigates university students' self-directed use of AI-powered speech recognition in English classrooms through the following research questions:

1. How do university students use AI-powered speech recognition tools in a self-directed manner to support real-time oral communication in English classrooms?
2. What are students' perceptions of the usefulness of AI-powered speech recognition for enhancing real-time oral communication in English classrooms?

The study contributes to research on AI-assisted language learning by providing empirical insights into the real-time, self-directed use of speech recognition in authentic classroom interaction. It also highlights the experiential and affective dimensions of AI-supported speaking and demonstrates the value of descriptive mixed-methods approaches for studying emerging educational technologies. The findings may further inform teachers and curriculum designers considering the integration of AI-powered speech recognition into communicative speaking activities.

LITERATURE REVIEW

This section reviews theoretical and empirical studies on AI-powered speech recognition in English language classrooms. The first subsection presents the analytical lens of the study, while the second summarizes previous research and identifies key gaps.

Conceptual Framing of AI-Supported Real-Time Speaking

The present study adopts an analytical lens that combines three complementary perspectives: real-time oral communication in language learning, self-directed use of digital tools, and descriptive approaches to learner perceptions of educational technology. Rather than advancing a single theory, this lens provides a practical framework for examining how AI-powered speech recognition functions during authentic speaking activities and how learners experience its use.

Real-time oral communication is central to communicative language teaching and interaction-based learning (Alkilani et al., 2025). Speaking in real time requires learners to process input, formulate responses, and produce intelligible output under time pressure (Sok & Shin, 2025; Suzuki et al., 2025). Research shows that these cognitive and affective demands often limit learners' ability to notice linguistic errors while speaking (Çelik et al., 2025; Mora et al., 2024). From this perspective, tools that provide immediate feedback may support noticing during oral production.

AI-powered speech recognition offers such an affordance by converting spoken output into visible text in real time. This representation allows learners to monitor linguistic features such as word choice, grammatical structure, and pronunciation-related recognition errors (Bashori et al., 2024; Hirschi et al., 2025). In this study, speech recognition is viewed not as a corrective authority but as a mediating tool that learners may use to monitor and adjust their spoken output during communicative tasks.

A second component of the analytical lens concerns self-directed technology use. In many learning environments, students access AI tools through personal devices and widely available applications (Pesovski et al., 2024; Zhai & Nezakatgoo, 2025). Such tools are often used flexibly, allowing learners to decide when and how to employ them. Here, self-directed use refers to learners' voluntary engagement with speech recognition tools to address immediate communicative needs (Dai & Wu, 2025; Mingyan et al., 2025; Xiao, 2025).

The third element focuses on learner perceptions of educational technology. Research in technology-enhanced language learning shows that perceptions of usefulness, ease of use, and emotional impact influence technology adoption (Al-kfairy, 2024; Alshammari & Alkhwaldi, 2025; Li et al., 2025). In speaking contexts, affective factors such as confidence and anxiety are particularly relevant (Waluyo & Bakoko, 2022). A descriptive perspective on learner perceptions helps capture these experiential dimensions without making causal claims about learning outcomes. In this study, Likert-scale questionnaire data are used to identify general patterns in students' experiences with AI-powered speech recognition, complementing qualitative insights into their classroom practices.

Empirical Insights into Speech Recognition and AI-Assisted Speaking

The existing literature on AI-assisted speaking and speech recognition technologies spans several interconnected areas, including pronunciation training, communicative speaking

practice, learner motivation, and classroom integration of AI tools. These studies employ a variety of methodological approaches, ranging from experimental and quasi-experimental designs to survey-based and classroom-based investigations. While this body of research provides valuable insights into the potential of speech recognition technologies for language learning, differences in research focus and methodological orientation have led to varying emphases across studies. To clarify these patterns and highlight the research gaps addressed in the present study, a summary of representative studies is presented in Table 1.

Table 1. Summary of Previous Studies

Study	Focus of Study	Methodology	Key Findings	Identified Limitation
(Ahn et al., 2025; Loakes, 2024; McCrocklin & Levis, 2025)	Technical accuracy of speech recognition for pronunciation	Experimental / controlled tasks	Speech recognition tools can detect pronunciation errors and support segmental accuracy	Focus on controlled reading tasks rather than communicative speaking
(Farrús, 2023; Thi-Nhu Ngo et al., 2024)	Pronunciation training using speech recognition	Experimental / quasi-experimental	Improvement in pronunciation awareness and pronunciation accuracy	Limited exploration of classroom interaction
(Terzioğlu & Kurt, 2022; Zhang, 2024)	Speech recognition in communicative speaking tasks	Pre-test/post-test design	Increased speaking fluency and learner confidence	Focus on learning outcomes rather than learner practices
(Li & Kim, 2024)	Learners' motivation using AI speaking tools	Survey-based studies	Learners value private and immediate automated feedback	Limited observation of real-time classroom use
(Al-kfairy, 2024; Alshammari & Alkhwalidi,	Learners' perceptions of speech recognition tools	Questionnaires and interviews	Generally positive attitudes toward pronunciation support	Perception-based data with limited behavioral observation

2025; Li et al., 2025)

(Gottardi & Silveira, 2025; Sun, 2023)	Classroom integration of AI speaking tools	Classroom-based studies	Technology use varies depending on task and instructional context	Limited investigation of self-directed learner use
(Ballance, 2024; Hung & Yeh, 2023; Jiang et al., 2022; Zhao & Yang, 2023)	AI-assisted language learning effectiveness	Experimental / quasi-experimental	Evidence of potential learning benefits	Limited qualitative exploration of real-time technology use
Present Study	Self-directed use of AI-powered speech recognition during authentic speaking activities	Convergent mixed methods (observation + perception data)	Examines how learners use speech recognition tools and how they perceive their usefulness in real classroom interaction	Addresses gaps related to real-time classroom practices and learner agency

As shown in Table 1, previous studies have primarily focused on technical accuracy, pronunciation training, and learning outcomes measured through experimental or quasi-experimental designs. Although these studies demonstrate the potential benefits of speech recognition technologies, relatively few have examined how learners actually use these tools during real-time classroom communication. Similarly, studies investigating learner perceptions often rely on questionnaires or interviews without direct observation of classroom practices. Consequently, there remains limited empirical evidence on how university students integrate speech recognition tools into authentic speaking activities and how they experience these practices in situ. The present study addresses this gap by adopting a convergent mixed-methods approach that documents students' self-directed use of AI-powered speech recognition during classroom speaking tasks and explores their perceptions of its usefulness for supporting oral communication.

METHODS

This study employed a convergent mixed-methods approach to explore university students' self-directed use of AI-powered speech recognition during real-time oral communication in English classrooms and their perceptions of its usefulness. The methodological choices were guided by the exploratory nature of the research questions

and the aim to document naturally occurring classroom practices and learner experiences rather than to test causal relationships (Younas et al., 2023). This section describes the research design, participants, data collection procedures, and data analysis strategies.

Design of the Study

The study adopted a convergent mixed-methods design, in which qualitative and quantitative data were collected concurrently and integrated during interpretation (Creswell & Creswell, 2023). The design combined qualitative data from classroom observations and student interviews with quantitative data from a descriptive questionnaire using Likert-scale items. This approach allowed for an in-depth examination of how students used AI-powered speech recognition tools during real-time speaking activities while also identifying broader patterns in their perceptions of the technology.

Within this design, greater emphasis was placed on the qualitative strand, as it was central to addressing the first research question concerning students' self-directed use of speech recognition during classroom speaking activities. The qualitative data focused on students' observable practices and reported experiences, capturing how the technology was used in situ during oral communication tasks. The quantitative strand served a complementary role by providing descriptive evidence related to students' perceived usefulness of speech recognition for supporting oral communication.

Qualitative and quantitative data were collected during the same instructional period. Integration occurred at the interpretation stage, where findings from both strands were compared and synthesized to address the research questions. This convergent design is appropriate for exploratory research on emerging educational technologies, as it enables the examination of both situated practices and general perceptual trends without extending claims about effectiveness or causality.

Participants

The participants were 32 undergraduate students enrolled in English courses at a public university in Indonesia. The courses focused on general English communication and academic speaking activities, which regularly involved classroom discussions, role plays, and short presentations. The participants represented diverse academic majors and varied levels of English proficiency. Although no standardized proficiency test was administered specifically for the study, the students were generally considered to fall within the intermediate to upper-intermediate range, based on their placement in university English courses and their ability to participate in classroom speaking tasks conducted in English.

Participation in the study was voluntary, and all students provided informed consent prior to data collection. To capture authentic classroom practices, students were included based on their participation in speaking activities where AI-powered speech recognition tools were available for optional use. This inclusive approach allowed the study to document naturally occurring variations in students' self-directed use of speech recognition during classroom communication.

Data Collection

Data were collected from multiple sources to support methodological triangulation and enhance the trustworthiness of the findings. The primary qualitative data sources

consisted of classroom observations and semi-structured interviews, while quantitative data were obtained through a self-report questionnaire.

In the present study, the term *AI-powered speech recognition* refers to widely available speech-to-text functions embedded in commonly used digital platforms rather than a single specialized application. During classroom activities, students accessed speech recognition through mobile-based and web-based tools available on their personal devices, including built-in voice typing features in smartphone keyboards and speech-to-text functions integrated into widely used applications such as Google Docs and similar platforms. Because the study focused on naturally occurring classroom practices and self-directed technology use, students were not required to use a specific platform. Instead, they selected tools that were readily accessible to them. Despite variation in specific applications, the tools shared a common functionality: converting spoken English into real-time textual transcription, allowing learners to visually monitor their spoken output during communication.

Classroom observations were conducted during scheduled English-speaking sessions over a four-week instructional period, covering eight classroom meetings in which communicative speaking activities were implemented. Each session lasted approximately 90 minutes. The observations focused on students' self-directed use of AI-powered speech recognition during real-time oral communication tasks such as discussions, role plays, and presentation preparation. During the observations, the researcher adopted a non-participant observer role in order to minimize disruption to regular classroom interaction. Detailed field notes were taken to document when and how students activated speech-to-text functions, the purposes for which the tools were used, and how students interacted with the recognition output during speaking activities. Contextual factors such as task type, interaction format, and classroom dynamics were also recorded.

Following the observation phase, semi-structured interviews were conducted with a subset of participants to gain deeper insights into their experiences and decision-making processes. The interview protocol included open-ended questions addressing reasons for using or not using speech recognition tools, perceived benefits and challenges, and emotional responses during speaking activities. Interviews were conducted in a language comfortable for the participants to encourage rich and reflective responses. All interviews were audio-recorded with participants' consent and transcribed verbatim for analysis.

Quantitative data were collected using a questionnaire consisting of Likert-scale items designed to capture students' perceptions of the usefulness of AI-powered speech recognition for real-time oral communication. The questionnaire included items related to perceived support for pronunciation, fluency, confidence, and error awareness, as well as overall usefulness and ease of use. Responses were measured on a five-point Likert scale ranging from strongly disagree to strongly agree. The questionnaire was administered after participants had gained sufficient experience using speech recognition tools during classroom speaking activities.

Data Analysis

Qualitative and quantitative data were analysed separately and integrated during interpretation. For the qualitative data, an inductive thematic analysis approach was employed (Braun & Clarke, 2021). Classroom observation notes and interview transcripts were read repeatedly to achieve familiarization with the data. Initial codes were generated

to capture recurring patterns related to purposes of tool use, modes of interaction with speech recognition output, and situational factors influencing self-directed use.

The codes were subsequently examined and grouped into broader themes addressing the first research question. Constant comparison was applied throughout the analysis to refine themes and ensure consistency across participants and contexts. Analytical memos were maintained to document emerging interpretations and support reflexivity.

Quantitative questionnaire data were analysed using descriptive statistics (Hu & Xu, 2024). Mean scores, standard deviations, and frequency distributions were calculated for each item to identify general trends in students' perceptions of AI-powered speech recognition. The analysis was descriptive in nature and did not involve hypothesis testing or inferential statistics. These results were used to complement the qualitative findings by highlighting areas of convergence and divergence between observed practices and reported perceptions.

Integration of qualitative and quantitative findings occurred at the interpretation stage. Findings from both strands were compared to develop a cohesive understanding of how students used speech recognition tools and how they perceived their usefulness for real-time oral communication. This integrative approach enabled the study to address both research questions comprehensively and to generate empirically grounded implications.

FINDINGS

This section presents the findings of the study in relation to the two research questions. The findings for RQ1 are based on qualitative data from classroom observations and semi-structured interviews, while the findings for RQ2 draw on quantitative data from a Likert-scale questionnaire. The results are organized according to each research question to ensure clarity and alignment with the study objectives.

Students' Self-Directed Use of AI-Powered Speech Recognition

The first research question examined how university students used AI-powered speech recognition tools in a self-directed manner to support real-time oral communication in English classrooms. Analysis of classroom observations and interview data revealed several recurring patterns of use. These patterns were grouped into four major themes: real-time monitoring of spoken output, rehearsal and preparation for oral tasks, support for pronunciation and lexical accuracy, and affective regulation during speaking activities. Table 2 presents an overview of the themes and sub-themes identified from the qualitative data.

Table 2.

Themes Identified from Classroom Observations and Interviews

Theme	Description	Data Sources
Real-time monitoring of spoken output	Using speech-to-text to check accuracy while speaking	Observations, Interviews
Rehearsal and preparation	Practicing utterances before speaking publicly	Observations, Interviews

Pronunciation and lexical support	Identifying unclear pronunciation and word choice	Interviews
Affective regulation	Reducing anxiety and increasing confidence	Interviews

The four themes presented in Table 1 were derived through an inductive thematic analysis of observation notes and interview transcripts. During the coding process, recurring patterns related to the purposes and contexts of speech recognition use were identified and gradually grouped into broader analytical categories. The final themes were selected because they captured the most consistent and meaningful ways in which students integrated speech recognition tools into their speaking practices. Together, these themes illustrate how the technology functioned as linguistic support, preparation assistance, and affective scaffolding during real-time communication.

Real-Time Monitoring of Spoken Output

One prominent pattern observed across multiple speaking activities was students' use of speech recognition tools to monitor their spoken output in real time. During pair and group discussions, several students activated speech-to-text functions on their devices while speaking, occasionally glancing at the transcription to check whether their utterances were accurately recognized. This practice appeared to help students notice grammatical omissions, word order issues, or unclear pronunciation that resulted in incorrect transcription. An interview participant explained:

“When I speak and see the text, I can immediately know if my sentence is correct or not. If the words are wrong on the screen, I usually try to say it again in a clearer way.”

This finding suggests that speech recognition served as a form of immediate feedback, allowing students to adjust their speech while maintaining engagement in communicative tasks. Rather than interrupting interaction entirely, students selectively used the tool to support accuracy without fully disengaging from the conversation.

Rehearsal and Preparation for Oral Tasks

Another frequently observed use of AI-powered speech recognition involved rehearsal prior to speaking in front of the class. During presentation preparation and role-play activities, students were observed practicing their utterances individually by speaking into their devices and reviewing the transcribed output. This self-directed rehearsal allowed students to refine their language before delivering it publicly. One student noted:

“Before I speak in front of the class, I usually practice with speech-to-text. If the text looks okay, I feel more ready to speak.”

This pattern indicates that speech recognition tools functioned as a private rehearsal space where students could test and refine their spoken language. The tool was not used as a substitute for speaking, but rather as a preparatory resource to support oral performance.

Pronunciation and Lexical Support

Interview data further revealed that students used speech recognition to identify pronunciation and vocabulary-related issues. When the transcription displayed unexpected words or failed to recognize certain utterances, students interpreted this as an

indication that their pronunciation might be unclear or that their word choice was inappropriate. Some students reported repeating specific words multiple times to achieve accurate recognition. As one participant stated:

“If the app cannot catch my word, I think my pronunciation is not clear. I try to change it until the text is correct.”

This practice reflects students’ reliance on speech recognition as a diagnostic tool for pronunciation clarity and lexical accuracy. Rather than focusing on formal phonetic rules, students used recognition accuracy as a practical indicator of intelligibility.

Affective Regulation During Speaking

In addition to linguistic support, speech recognition tools played a role in regulating students’ emotional responses to speaking tasks. Several interview participants reported that using speech recognition reduced anxiety and increased confidence, particularly when preparing for public speaking. One student explained:

“I feel less nervous because I already check my speaking with the app. Even if I make mistakes, at least I know what I want to say.”

This finding highlights the affective dimension of self-directed speech recognition use, suggesting that the tool provided emotional reassurance alongside linguistic support.

Students’ Perceptions of the Usefulness of AI-Powered Speech Recognition

The second research question examined students’ perceptions of the usefulness of AI-powered speech recognition for enhancing real-time oral communication. Quantitative data from the Likert-scale questionnaire were analyzed using descriptive statistics, including mean, median, and standard deviation values. Overall, the results indicated generally positive perceptions of speech recognition as a supportive tool for speaking activities. Table 3 summarizes the descriptive statistics for selected questionnaire items.

Table 3.

Descriptive Statistics of Students’ Perceptions of AI-Powered Speech Recognition

Item	Mean	Median	SD
Helps me notice speaking errors	4.12	4.00	0.68
Improves my pronunciation awareness	4.05	4.00	0.71
Increases my confidence when speaking	3.98	4.00	0.75
Useful for real-time speaking activities	4.15	4.00	0.64
Easy to use during class	4.22	4.00	0.59

The highest mean score was reported for the item related to ease of use during class, indicating that students generally found speech recognition tools accessible and manageable in real-time classroom settings. Similarly, high mean scores for usefulness and error awareness suggest that students perceived the tools as beneficial for monitoring their spoken language.

Items related to confidence also received relatively high ratings, although with slightly greater variability. This variation suggests that while many students felt more confident using speech recognition, the degree of affective support differed across individuals.

The quantitative findings complement the qualitative results by demonstrating that students not only used speech recognition tools in self-directed ways, but also perceived these tools as useful for supporting real-time oral communication. The convergence between observed practices and reported perceptions strengthens the credibility of the findings and provides a comprehensive account of AI-powered speech recognition use in English classrooms.

DISCUSSION

This study set out to explore university students' self-directed use of AI-powered speech recognition during real-time oral communication in English classrooms and their perceptions of its usefulness. Rather than reiterating the findings, this discussion interprets the results by situating them within existing literature, highlighting their conceptual and pedagogical significance, and considering their implications for research and practice in AI-assisted language learning.

The findings related to RQ1 demonstrate that students did not use AI-powered speech recognition in a uniform or mechanical manner. Instead, their use was adaptive, situational, and closely tied to immediate communicative demands. This pattern reinforces the view that emerging AI tools function less as instructional replacements and more as flexible resources that learners appropriate according to context (Asratie et al., 2023; Fitriati & Williyani, 2025; Mingyan et al., 2025). The observed practices of real-time monitoring, rehearsal, and selective repetition suggest that students treated speech recognition as a support for intelligibility and clarity rather than as an authoritative evaluator of correctness, echoing classroom-based observations reported by Sun (2023) and Jakonen et al. (2025).

This interpretation aligns with prior research emphasizing the importance of noticing in spoken language development (Çelik et al., 2025; Mora et al., 2024). However, unlike traditional noticing mechanisms that rely on teacher feedback or post-task reflection, speech recognition enabled noticing to occur concurrently with speech production. The visual transcription of spoken output appeared to externalize language in a way that made gaps and inconsistencies more salient without fully disrupting communication, a process also noted in studies on AI-mediated feedback and intelligibility awareness (Bashori et al., 2024; Hirschi et al., 2025). This supports the argument that AI-powered tools can extend learners' attentional capacity during real-time speaking, particularly in contexts where teacher feedback is necessarily limited (Liao, 2025; Suzuki et al., 2025).

The use of speech recognition for rehearsal and preparation further suggests that students valued opportunities to privately test and refine their language before public performance. This finding resonates with studies highlighting the role of low-risk practice environments in reducing speaking anxiety and increasing readiness to communicate (Dillon & Wells, 2023; El Shazly, 2021; Wardat & Akour, 2025). Importantly, speech recognition did not replace oral communication but functioned as a bridge between internal formulation and external performance. This challenges concerns that AI tools might discourage spontaneous speaking (Ali et al., 2024; Kuhail et al., 2023), indicating

instead that they may scaffold participation by increasing learners' preparedness and confidence.

Students' reliance on speech recognition accuracy as an indicator of pronunciation and lexical clarity points to a pragmatic orientation toward spoken language. Rather than focusing on abstract linguistic rules, students interpreted recognition success or failure as feedback on intelligibility, a perspective consistent with communicative approaches to oral proficiency (Leeming et al., 2024; Sok & Shin, 2025). This reflects a shift from form-focused accuracy toward listener-oriented clarity, which is central to effective oral communication. While speech recognition systems are not flawless and may misinterpret speech for reasons unrelated to pronunciation (Russell et al., 2024; Loakes, 2024), students appeared to negotiate these limitations through repeated attempts and contextual judgment, as also reported in empirical studies on ASR-supported speaking practice (McCrocklin & Levis, 2025; Thi-Nhu Ngo et al., 2024).

This finding contributes to ongoing discussions about the pedagogical value of imperfect AI systems. Rather than viewing recognition errors solely as technological shortcomings, the results suggest that learners can engage productively with these inaccuracies (Farrús, 2023; Gottardi & Silveira, 2025). The process of adjusting speech to achieve clearer recognition may encourage greater attention to stress, articulation, and word choice. However, this also underscores the need for pedagogical guidance to help students interpret recognition output critically and avoid over-reliance on transcription accuracy as a definitive measure of spoken competence (Al-kfairy, 2024; Alshammari & Alkhwaldi, 2025).

One of the most salient interpretive insights from the study concerns the affective role of speech recognition in speaking activities. The qualitative findings indicate that students experienced reduced anxiety and increased confidence when using speech recognition, particularly in preparation for public speaking. This aligns with research emphasizing the emotional barriers associated with oral communication in a second or foreign language (Mora et al., 2024; Waluyo & Bakoko, 2022). The presence of an AI tool that provides private, nonjudgmental feedback appears to create a sense of reassurance, enabling learners to approach speaking tasks with greater confidence, as similarly observed in AI-assisted speaking studies (L. Li & Kim, 2024; Sun, 2023). The quantitative findings reinforce this interpretation by showing generally positive perceptions related to confidence and ease of use. While variability in responses suggests that affective benefits are not uniform across all learners, the overall trend indicates that speech recognition can function as an emotional scaffold (Elov et al., 2025; Wardat & Akour, 2025). This highlights the importance of considering affective outcomes alongside linguistic ones when evaluating AI-assisted speaking tools.

The convergent mixed-methods design of the study allows for a deeper interpretation of how observed practices and reported perceptions intersect. The alignment between qualitative evidence of self-directed use and quantitative trends indicating perceived usefulness strengthens the credibility of the findings (Creswell & Creswell, 2023; Younas et al., 2023). Students not only reported that speech recognition was useful, but also demonstrated purposeful engagement with the tool during classroom activities. This convergence suggests that learners' positive perceptions were grounded in actual experience rather than abstract attitudes toward technology, a point emphasized in mixed-methods research on AI-mediated language learning (Liu et al., 2024; Zhai &

Nezakatgoo, 2025). At the same time, the findings caution against simplistic assumptions that positive perceptions automatically translate into effective use. The qualitative data reveal nuanced patterns, including selective and situational engagement with speech recognition, underscoring the value of combining observational and self-report data (Hu & Xu, 2024).

The findings of this study have several implications for English language teaching in technology-rich classrooms. First, they suggest that AI-powered speech recognition can be meaningfully integrated into communicative speaking activities without undermining interaction (Jakonen et al., 2025; Sun, 2023). Teachers may consider allowing or encouraging optional use of speech recognition during rehearsal phases, pair work, or preparation stages, rather than restricting it to isolated pronunciation exercises. Second, the self-directed nature of tool use highlights the importance of learner autonomy in technology integration (Pesovski et al., 2024; Xiao, 2025). Rather than prescribing rigid procedures, teachers can support students by modeling possible uses of speech recognition and discussing its limitations. Explicit discussion about how to interpret transcription output may help learners use the tool more critically and effectively.

Third, the affective benefits observed in the study suggest that speech recognition may be particularly valuable for learners who experience high levels of speaking anxiety (El Shazly, 2021; Wardat & Akour, 2025). Incorporating AI tools as supportive resources rather than evaluative mechanisms may contribute to more inclusive and psychologically safe speaking environments (Elov et al., 2025).

From a research perspective, this study contributes to the literature by foregrounding real-time, classroom-based use of AI-powered speech recognition. While much existing research focuses on outcomes or controlled interventions (Ballance, 2024; Nguy et al., 2025), this study provides descriptive insights into how learners integrate AI tools into everyday speaking practices. The findings extend understanding of AI-assisted speaking beyond questions of effectiveness to include issues of use, perception, and experience. Methodologically, the study demonstrates the value of convergent mixed-methods designs for exploring emerging educational technologies (Ding & Yusof, 2025; Mizza et al., 2025).

Although the study provides valuable insights, certain limitations should be acknowledged. The descriptive nature of the quantitative data limits the ability to draw conclusions about learning outcomes over time. Future research may build on these findings by examining longitudinal patterns of speech recognition use or exploring how different instructional designs influence self-directed engagement (Wang et al., 2025; Yang et al., 2025). Additionally, further research could investigate how teacher mediation and classroom norms shape students' use of AI-powered speech recognition. Comparative studies across task types or proficiency levels may also yield deeper understanding of when and for whom such tools are most beneficial.

This study highlights the potential of AI-powered speech recognition as a flexible, learner-oriented support for real-time oral communication in English classrooms. By interpreting students' self-directed practices and perceptions, the discussion underscores the importance of viewing AI tools not as deterministic solutions but as resources whose value emerges through use (Fitriati & Williyani, 2025; Liu et al., 2024). These insights

contribute to ongoing conversations about the role of AI in language education and point toward more nuanced, practice-oriented approaches to technology integration.

CONCLUSION

This study explored university students' self-directed use of AI-powered speech recognition to support real-time oral communication in English classrooms, as well as their perceptions of its usefulness. The findings suggest that speech recognition functions as a flexible and learner-oriented support rather than a prescriptive instructional tool. Students appropriated the technology selectively to monitor intelligibility, rehearse spoken output, and manage affective challenges associated with speaking in English. Their practices indicate that AI-powered speech recognition can mediate spoken language production in ways that are immediate, contextualized, and responsive to learners' communicative needs. At the same time, students' generally positive perceptions highlight the perceived value of the technology in enhancing confidence, fluency awareness, and engagement during speaking activities. Together, these findings contribute to a more nuanced understanding of how AI tools are integrated into everyday classroom practices, moving beyond effectiveness claims to emphasize use, experience, and perception.

Despite its contributions, this study also points to several directions for future research. Further studies could adopt longitudinal designs to examine how sustained use of speech recognition influences spoken language development over time, particularly in relation to pronunciation, fluency, and willingness to communicate. Future research may also explore how instructional design, teacher mediation, and task types shape students' self-directed engagement with AI-powered tools. Comparative studies across proficiency levels or educational contexts could provide deeper insights into differential patterns of use and perception. By extending inquiry in these directions, future research can continue to clarify the pedagogical potential and limitations of AI-powered speech recognition in supporting oral communication in English language education.

REFERENCES

- Ahn, T., Hong, Y., Im, Y., Kim, D. H., Kang, D., Jeong, J. W., Kim, J. W., Kim, M. J., Cho, A.-R., Nam, H., & Jang, D.-H. (2025). Automatic speech recognition (ASR) for the diagnosis of pronunciation of speech sound disorders in Korean children. *Clinical Linguistics & Phonetics*, 39(10), 913–926. <https://doi.org/10.1080/02699206.2024.2387609>
- Al-kfairy, M. (2024). Factors impacting the adoption and acceptance of ChatGPT in educational settings: A narrative review of empirical studies. *Applied System Innovation*, 7(6), 110. <https://doi.org/10.3390/asi7060110>
- Ali, O., Murray, P. A., Momin, M., Dwivedi, Y. K., & Malik, T. (2024). The effects of artificial intelligence applications in educational settings: Challenges and strategies. *Technological Forecasting and Social Change*, 199. <https://doi.org/10.1016/j.techfore.2023.123076>
- Alkilani, R., Jaganathan, P., & Abd Rahim, S. (2025). Teachers' and Students' Perceptions of Oral Communicative Activities in a Libyan EFL Classroom. *Mextesol Journal*, 49(2), 1–15. <https://doi.org/10.61871/mj.v49n2-4>

- Alshammari, S. H., & Alkhwaldi, A. F. (2025). An integrated approach using social support theory and technology acceptance model to investigate the sustainable use of digital learning technologies. *Scientific Reports*, 15(1), 342. <https://doi.org/10.1038/s41598-024-83450-z>
- Asratie, M. G., Wale, B. D., & Aylet, Y. T. (2023). Effects of using educational technology tools to enhance EFL students' speaking performance. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-022-11562-y>
- Ballance, O. J. (2024). Sampling and randomisation in experimental and quasi-experimental CALL studies: Issues and recommendations for design, reporting, review, and interpretation. *ReCALL*, 36(1), 58–71. <https://doi.org/10.1017/S0958344023000162>
- Bashori, M., van Hout, R., Strik, H., & Cucchiarini, C. (2024). I can speak: Improving English pronunciation through automatic speech recognition-based language learning systems. *Innovation in Language Learning and Teaching*, 18(5), 443–461. <https://doi.org/10.1080/17501229.2024.2315101>
- Braun, V., & Clarke, V. (2021). Thematic analysis: A practical guide. In *Sage Publications* (p. 320). <https://books.google.se/books?id=Hr11DwAAQBAJ&hl=sv>
- Broeren, M., Gu, Y., Pitt, M., & Tompkins, V. (2025). A script and tutorial for using Rev AI's automatic speech transcription. *Infant and Child Development*, 34(2). <https://doi.org/10.1002/icd.70007>
- Çelik, F., Baturay, M. H., Yastibaş, A. E., Namaziandost, E., & Ersanlı, C. Y. (2025). Modelling the complex dynamic effect of L2 speaking anxiety in metaverse on students' academic enjoyment, pronunciation anxiety, speaking performance, and task load. *British Educational Research Journal*. <https://doi.org/10.1002/berj.70062>
- Chou, M.-H. (2024). Communication strategies, difficulties, and speaking tasks in foreign language learning. *Sage Open*, 14(3). <https://doi.org/10.1177/21582440241266324>
- Creswell, J. W., & Creswell, J. D. (2023). Research Design: Qualitative, quantitative and mixed methods approaches. In *SAGE Publications, Inc.: Vol. Sixth Edit* (Issue 1). SAGE Publications. <https://medium.com/@arifwicaksanaa/pengertian-use-case-a7e576e1b6bf>
- Dai, L., & Wu, F. (2025). An AI-powered conversational system for college students learning English as a second language. *Education and Information Technologies*, 30(16), 23393–23417. <https://doi.org/10.1007/s10639-025-13640-3>
- Dillon, T., & Wells, D. (2023). Effects of pronunciation training using automatic speech recognition on pronunciation accuracy of Korean English language learners. *English Teaching*, 78(1), 3–23. <https://doi.org/10.15858/engtea.78.1.202303.3>
- Ding, D., & Yusof, A. 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), 1223.

<https://doi.org/10.1057/s41599-025-05550-z>

- El Shazly, R. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems*, 38(3). <https://doi.org/10.1111/exsy.12667>
- Elov, B., Abdullayeva, I., Raupova, L., Kholikov, A., & Mirkasimova, M. (2025). Speaking exams with less anxiety in Intelligent Computer-Assisted Language Assessment (ICALA): mirroring EFL learners' foreign language anxiety, shyness, autonomy, and enjoyment. *Language Testing in Asia*, 15(1), 3. <https://doi.org/10.1186/s40468-024-00340-x>
- Farrús, M. (2023). Automatic speech recognition in L2 learning: A review based on PRISMA methodology. *Languages*, 8(4), 242. <https://doi.org/10.3390/languages8040242>
- Fitriati, S. W., & Williyana, A. (2025). AI-enhanced self-regulated learning: EFL learners' prioritization and utilization in presentation skills development. *Journal of Pedagogical Research*, 9(2), 22–37. <https://doi.org/https://doi.org/10.33902/JPR.202530647>
- Gottardi, W., & Silveira, R. (2025). A theoretical assessment of ASR-based pronunciation activities for classroom learning. *Texto Livre*, 18. <https://doi.org/10.1590/1983-3652.2025.56636>
- Hirschi, K., Kang, O., Yang, M., Hansen, J. H. L., & Beloin, K. (2025). Artificial intelligence-generated feedback for second language intelligibility: An exploratory intervention study on effects and perceptions. *Language Learning*, 75(S1), 204–241. <https://doi.org/10.1111/lang.12719>
- Hu, P., & Xu, H. (2024). Student teachers' perceptions and application of questionnaire survey in an EFL context. *Sage Open*, 14(2). <https://doi.org/10.1177/21582440241245621>
- Hung, H., & Yeh, H. (2023). Augmented-reality-enhanced game-based learning in flipped English classrooms: Effects on students' creative thinking and vocabulary acquisition. *Journal of Computer Assisted Learning*, 39(6), 1786–1800. <https://doi.org/10.1111/jcal.12839>
- Jakonen, T., Duran, D., & Peltonen, P. (2025). Situated L2 pronunciation instruction during small-group robot-assisted language learning activities. *Language Teaching Research*. <https://doi.org/10.1177/13621688251367852>
- Jiang, M. Y.-C., Jong, M. S.-Y., Wu, N., Shen, B., Chai, C.-S., Lau, W. W.-F., & Huang, B. (2022). Integrating automatic speech recognition technology into vocabulary learning in a flipped English class for Chinese college students. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.902429>
- Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information*

- Technologies*, 28(1), 973–1018. <https://doi.org/10.1007/s10639-022-11177-3>
- Kusz, E., & Pawliszko, J. (2025). Efficacy of artificial intelligence (AI) voice cloning in phonetic self-imitation for L2 pronunciation training. *International Journal of Applied Linguistics*. <https://doi.org/10.1111/ijal.70014>
- Leeming, P., Vitta, J. P., Hiver, P., Hicks, D., McLean, S., & Nicklin, C. (2024). Willingness to communicate, speaking self-efficacy, and perceived communicative competence as predictors of second language spoken task production. *Language Learning*, 74(4), 917–949. <https://doi.org/10.1111/lang.12640>
- Li, J., Jin, M., & Chen, X. (2025). Understanding continued use of smart learning platforms: Psychological wellbeing in an extended TAM-ISCM model. *Frontiers in Psychology*, 16. <https://doi.org/10.3389/fpsyg.2025.1521174>
- Li, L., & Kim, M. (2024). It is like a friend to me: Critical usage of automated feedback systems by self-regulating English learners in higher education. *Australasian Journal of Educational Technology*. <https://doi.org/10.14742/ajet.8821>
- Liao, M.-H. (2025). Cultivating proficient and efficacious L2 English speakers via VoiceThread-mediated self- and peer assessments. *Humanities and Social Sciences Communications*, 12(1), 1277. <https://doi.org/10.1057/s41599-025-05674-2>
- Liu, G. L., Darvin, R., & Ma, C. (2024). Exploring AI-mediated informal digital learning of English (AI-IDLE): A mixed-method investigation of Chinese EFL learners' AI adoption and experiences. *Computer Assisted Language Learning*, 0(0), 1–29. <https://doi.org/10.1080/09588221.2024.2310288>
- Loakes, D. (2024). Automatic speech recognition and the transcription of indistinct forensic audio: How do the new generation of systems fare? *Frontiers in Communication*, 9. <https://doi.org/10.3389/fcomm.2024.1281407>
- Manorom, P., Hunsapun, N., & Chansanam, W. (2024). An investigation of English speaking problems of Chinese EFL students in Thailand. *International Journal of English Language and Literature Studies*, 13(2), 270–286. <https://doi.org/10.55493/5019.v13i2.5046>
- McCrocklin, S., & Levis, J. (2025). Automatic speech recognition and pronunciation learning. *Language Teaching*, 1–17. <https://doi.org/10.1017/S0261444825100852>
- Mingyan, M., Noordin, N., & Razali, A. B. (2025). Improving EFL speaking performance among undergraduate students with an AI-powered mobile app in after-class assignments: An empirical investigation. *Humanities and Social Sciences Communications*, 12(1), 370. <https://doi.org/10.1057/s41599-025-04688-0>
- Mizza, D., Reese, M., & Malouche, D. (2025). Flipped classroom evaluation and blended learning potential: A case study of engagement and inclusion in quantitative education. *Smart Learning Environments*, 12(1), 56. <https://doi.org/10.1186/s40561-025-00412-2>

- Mora, J. C., Mora-Plaza, I., & Bermejo Miranda, G. (2024). Speaking anxiety and task complexity effects on second language speech. *International Journal of Applied Linguistics*, 34(1), 292–315. <https://doi.org/10.1111/ijal.12494>
- Nguy, V. T., Thi Hoai Ly, N., & Thi Ha Le, D. (2025). Exploring the effectiveness of generative AI tools in enhancing English speaking skills among non-English major students at Nghe an university. *The International Journal of Language Studies (ISSN : 3078 - 2244)*, 2(2), 55–60. <https://doi.org/10.60087/ijls.v2.n2.005>
- Oli, L., Manickam, R., Rajesh, K., Ponmuthuramalingam, M., & Aravind, B. R. (2023). Probing the determinant factors affecting students' English speaking skills. *World Journal of English Language*, 13(7), 232. <https://doi.org/10.5430/wjel.v13n7p232>
- Pesovski, I., Santos, R., Henriques, R., & Trajkovik, V. (2024). Generative AI for customizable learning experiences. *Sustainability*, 16(7), 3034. <https://doi.org/10.3390/su16073034>
- Qiao, H., & Zhao, A. (2023). Artificial intelligence-based language learning: Illuminating the impact on speaking skills and self-regulation in Chinese EFL context. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1255594>
- Russell, S. O., Gessinger, I., Krason, A., Vigliocco, G., & Harte, N. (2024). What automatic speech recognition can and cannot do for conversational speech transcription. *Research Methods in Applied Linguistics*, 3(3), 100163. <https://doi.org/10.1016/j.rmal.2024.100163>
- Sangkawong, N., Leal Bucol, J., & Luciano, R. (2025). Bridging language gaps in real-time: Investigating university students' self-initiated use of speech-to-text translation in English language classrooms. *Teaching English with Technology*, 2025(2). <https://doi.org/10.56297/vaca6841/ZDDD4208/KTKA2748>
- Sok, S., & Shin, H. W. (2025). Do interactions with ChatGPT influence L2 learners' oral speaking ability, summarization ability, and perceptions of generative AI tasks? *TESOL Quarterly*, 59(S1). <https://doi.org/10.1002/tesq.70001>
- Sun, W. (2023). The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: A mixed methods investigation. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1210187>
- Suzuki, S., Takatsu, H., Matsuura, R., Koyama, M., Saeki, M., & Matsuyama, Y. (2025). Feedforwarding diagnostic language assessment: Artificial intelligence- (AI-) driven weakness identification and contextualised feedback for second language speaking. *Language Testing*, 42(4), 476–507. <https://doi.org/10.1177/02655322251348725>
- Terzioğlu, Y., & Kurt, M. (2022). Elevating English language learners' speaking fluency and listening skill through a learning management system. *Sage Open*, 12(2). <https://doi.org/10.1177/21582440221099937>
- Thi-Nhu Ngo, T., Hao-Jan Chen, H., & Kuo-Wei Lai, K. (2024). The effectiveness of

- automatic speech recognition in ESL/EFL pronunciation: A meta-analysis. *ReCALL*, 36(1), 4–21. <https://doi.org/10.1017/S0958344023000113>
- Waluyo, B., & Bakoko, R. (2022). Effects of affective variables and willingness to communicate on students' English-speaking performance in Thailand. *Studies in English Language and Education*, 9(1), 45–61. <https://doi.org/10.24815/siele.v9i1.21090>
- Wang, F., Zhou, X., Li, K., Cheung, A. C. K., & Tian, M. (2025). The effects of artificial intelligence-based interactive scaffolding on secondary students' speaking performance, goal setting, self-evaluation, and motivation in informal digital learning of English. *Interactive Learning Environments*, 33(7), 4633–4652. <https://doi.org/10.1080/10494820.2025.2470319>
- Wardat, S., & Akour, M. (2025). AI-powered language learning tools and their impact on EFL students' speaking anxiety in Jordanian universities. *Journal of Posthumanism*, 5(3). <https://doi.org/10.63332/joph.v5i3.785>
- Xiao, Y. (2025). The impact of AI-driven speech recognition on EFL listening comprehension, flow experience, and anxiety: A randomized controlled trial. *Humanities and Social Sciences Communications*, 12(1), 425. <https://doi.org/10.1057/s41599-025-04672-8>
- Yang, G., Wang, Y., Zhang, Y., Yang, M., Zeng, Q., & Song, Z. (2025). An empirical study of AI-supported interleaved training strategy to improve EFL students' English impromptu speaking performance, learning engagement, technology acceptance and epistemic network structure. *The Asia-Pacific Education Researcher*, 34(4), 1519–1540. <https://doi.org/10.1007/s40299-024-00962-y>
- Younas, A., Fàbregues, S., & Creswell, J. W. (2023). Generating metainferences in mixed methods research: A worked example in convergent mixed methods designs. *Methodological Innovations*, 16(3), 276–291. <https://doi.org/10.1177/20597991231188121>
- Zhai, Y., & Nezakatgoo, B. (2025). Evaluating AI-powered applications for enhancing undergraduate students' metacognitive strategies, self-determined motivation, and social learning in English language education. *Scientific Reports*, 15(1), 35199. <https://doi.org/10.1038/s41598-025-19118-z>
- Zhang, M. (2024). Enhancing self-regulation and learner engagement in L2 speaking: Exploring the potential of intelligent personal assistants within a learning-oriented feedback framework. *BMC Psychology*, 12(1), 421. <https://doi.org/10.1186/s40359-024-01917-0>
- Zhao, J., & Yang, Q. (2023). Promoting international high-school students' Chinese language learning achievements and perceptions: A mind mapping-based spherical video-based virtual reality learning system in Chinese language courses. *Journal of Computer Assisted Learning*, 39(3), 1002–1016. <https://doi.org/10.1111/jcal.12782>

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). <https://doi.org/10.1177/21582440231193818>

Zou, B., Guan, X., Shao, Y., & Chen, P. (2023). Supporting speaking practice by social network-based interaction in artificial intelligence (AI)-assisted language learning. *Sustainability*, 15(4), 2872. <https://doi.org/10.3390/su15042872>