OPEN ACCESS

Technology in Language Teaching & Learning
ISSN 2652-1687
https://www.castledown.com/journals/tltl/

Technology in Language Teaching & Learning, 8, 104582 (2026)
https://doi.org/10.29140/tltl.2026.104582

AI-Assisted Speaking Practice for Elementary EFL Learners: An Action Research Study in a Rural Context

KUAN HUNG[a]
RUEI-TENG HUNG[b]
KATE TZU-CHING CHEN[c]

[a]*Department of Applied English, Chaoyang University of Technology, Taiwan R.O.C.*
s11122604@gm.cyut.edu.tw

[b]*Department of Applied English, Chaoyang University of Technology, Taiwan R.O.C.*
hung_poyi@o365.cyut.edu.tw

[c]*Department of Applied English, Chaoyang University of Technology, Taiwan R.O.C.*
katechen@o365.cyut.edu.tw

Abstract

Artificial intelligence (AI)-assisted technologies are increasingly used in language education to provide individualized practice and feedback. However, limited research has examined how AI-assisted tools support young English as a Foreign Language (EFL) learners' speaking practice and self-learning capacity, particularly in rural elementary contexts. This classroom action research study examined the implementation of Microsoft Reading Coach and NaturalReader in a 12-week AI-supported speaking program at a rural elementary school in Taiwan. Thirteen students from Grades 4 to 6 participated in weekly activities focusing on pronunciation, fluency, and autonomous practice. Data were collected through Cambridge A1 Movers speaking pre- and post-tests, weekly Microsoft Reading Coach records, classroom observations, and semi-structured interviews. The quantitative results showed a statistically significant increase in students' speaking scores after the instructional period. Older students appeared to engage more consistently with the AI-supported routines, while younger students required more scaffolding, modeling, and teacher guidance. The qualitative findings suggested that students perceived the activities positively and reported greater willingness to practice, more independent learning behaviors, and a perceived reduction in speaking-related nervousness during classroom activities. However, because the study used a single-group action research design, the findings should be interpreted as context-specific classroom evidence rather than causal evidence of AI tool effectiveness.

Keywords: AI-assisted learning, EFL speaking, elementary education, learner autonomy, educational technology

Introduction

The rapid advancement of artificial intelligence (AI) has reshaped educational technology, offering new possibilities for personalized, interactive, and learner-centered instruction (Du & Daniel, 2024). In language education, AI-assisted tools such as text-to-speech, speech recognition, and automated feedback systems have attracted growing attention for their potential to enhance learners' engagement and language development (Bibauw et al., 2019). In English as a Foreign Language (EFL) contexts, these technologies are increasingly viewed as promising resources for addressing long-standing challenges, including limited exposure, insufficient practice opportunities, and learner anxiety, particularly in speaking instruction (Fathi et al., 2024).

Speaking proficiency is widely recognized as one of the most essential yet challenging skills for EFL learners to develop (Brown & Yule, 2019). Unlike receptive skills, speaking requires the real-time integration of vocabulary, grammar, pronunciation, and pragmatic competence. For young EFL learners, these demands are often compounded by limited opportunities to engage in meaningful oral practice, especially in contexts where English is not used outside the classroom (Hwang et al., 2016). In Taiwan, despite English being a compulsory subject from elementary school onward, speaking instruction has traditionally received less emphasis than reading and listening, largely due to examination-oriented curricula and time constraints (Hwang et al., 2011). As a result, many students demonstrate relatively strong receptive skills but struggle to speak English with confidence and fluency.

In response to these challenges, Taiwan's Ministry of Education has promoted large-scale digital transformation through initiatives such as the Digital Learning Enhancement Plan and the Bilingual 2030 policy (Hwang & Tu, 2021). These policies aim to improve students' English proficiency while expanding access to digital learning resources, particularly in rural and under-resourced areas. Under such conditions, AI-assisted technologies offer an appealing solution, providing consistent input, individualized pacing, and immediate feedback without relying solely on teacher availability (Jeon, 2024). However, the educational value of AI-assisted tools depends not only on their technological capabilities but also on their pedagogical integration into real classroom contexts (Fan & Zhang, 2024).

Existing research has demonstrated that AI-assisted tools can support language learning outcomes, including pronunciation accuracy, fluency, and learner engagement (Dou et al., 2025). Nevertheless, several gaps remain. First, much of the existing literature focuses on adult or tertiary-level learners, leaving elementary EFL contexts underexplored (Aravantinos et al., 2024). Second, many studies examine AI tools in isolation or experimental settings, rather than as part of sustained classroom practice (Du & Daniel, 2024). Third, limited attention has been paid to how AI-assisted speaking practice influences young learners' self-learning ability, an increasingly important competence in digital education (de Ruig et al., 2023).

To address this gap, the present study examines the classroom implementation of AI-supported speaking activities using text-to-speech (TTS) and automatic speech recognition (ASR) tools in a rural elementary EFL context. Based on this purpose, the study addressed the following research questions:

1. How did students' speaking performance change after the 12-week AI-supported instructional period?
2. In what ways, if any, did students' self-directed learning behaviors change during formative speaking practice with AI-supported tools across the action research cycles?
3. How did students perceive the use of AI-supported tools for speaking practice?

Literature Review

Previous research has shown that EFL learners often face speaking difficulties, including limited vocabulary, pronunciation problems, low confidence, anxiety, and insufficient opportunities for meaningful oral practice (Derwing & Munro, 2015). These challenges may be more noticeable in rural elementary school contexts, where learners often have fewer chances to interact with English speakers or receive individualized feedback. In this regard, AI-supported tools such as TTS and ASR may provide useful support by offering pronunciation models, immediate feedback, and opportunities for repeated practice. These tools may also promote learner autonomy because students can listen, repeat, compare, and revise their speaking more independently (Reinders & White, 2016). However, existing studies have often examined EFL speaking challenges, AI-supported pronunciation tools, learner autonomy, and rural English education separately. Fewer studies have explored how TTS and ASR tools can be integrated into speaking instruction for rural elementary EFL learners through classroom-based action research. Therefore, the present literature review connects these areas to clarify the rationale and research gap of this study.

EFL Speaking Challenges in Elementary Education

Speaking is often regarded as one of the most complex skills in second and foreign language learning, as it requires learners to produce language spontaneously while simultaneously attending to accuracy, fluency, and appropriateness (Kormos, 2014). Unlike receptive skills, speaking places substantial cognitive demands on learners due to the need for real-time processing and coordination of multiple linguistic components (Skehan, 1998).

For EFL learners, particularly children, speaking difficulties are frequently associated with limited exposure to authentic input, insufficient opportunities for oral practice, and affective factors such as anxiety and fear of negative evaluation (Horwitz et al., 1986). Previous research has shown that young learners may hesitate to speak due to low confidence, limited linguistic resources, and concern about making mistakes in front of peers or teachers (Wang et al., 2022).

In Taiwan and other Asian EFL contexts, speaking instruction has traditionally received less emphasis than receptive skills (Butler, 2011). Examination-oriented curricula, large class sizes, and limited instructional time often constrain opportunities for sustained oral practice in classrooms (Hwang et al., 2011). As a result, students may progress through years of English instruction without developing adequate speaking proficiency. Studies have further indicated that insufficient speaking practice in the early stages of language learning can lead to persistent difficulties in fluency and pronunciation later on, underscoring the importance of early and systematic intervention (Kormos, 2014; Skehan, 1998).

AI-Assisted Technologies in EFL Learning

The integration of AI-assisted technologies into EFL education has expanded rapidly in recent years, driven by advances in natural language processing, speech recognition, and adaptive learning systems (Zawacki-Richter et al., 2019). AI-assisted tools can analyze learners' input, provide immediate feedback, and tailor learning experiences to individual needs (Tapingkae et al., 2020). Such features align closely with learner-centered pedagogies and have been shown to enhance motivation and engagement (Hwang et al., 2020).

Empirical research has increasingly documented the potential of AI-assisted tools to support multiple dimensions of language learning, including vocabulary development, pronunciation accuracy, and speaking fluency (Kohnke et al., 2023; Okonkwo & Ade-Ibijola, 2021). For example, AI-powered platforms equipped with speech recognition have been found to help learners notice pronunciation errors and engage in repeated practice, leading to measurable improvements (Chen, 2022). In addition, AI-mediated interaction environments may reduce speaking anxiety by allowing learners to practice without the social pressure of face-to-face communication (Zhai & Wibowo, 2023).

However, much of this evidence is derived from short-term interventions or controlled experimental settings, raising concerns about the ecological validity of these findings. Reported improvements are often linked to structured practice and immediate system feedback, rather than to the development of sustained communicative competence (Benek, 2025). Consequently, while AI tools provide personalized and efficient learning support, their limitations, such as a lack of authentic interaction and contextual understanding, suggest that their broader pedagogical value depends heavily on effective integration into classroom instruction (Benek, 2025).

Despite these promising findings, the existing literature has largely focused on secondary and higher education contexts, with relatively limited attention to young learners in elementary EFL settings (Cizrelioğlu & Aydin, 2025). Furthermore, the predominance of short-term, laboratory-based designs limits our understanding of how AI-assisted learning operates in authentic classroom environments (Zhai & Wibowo, 2023). There is therefore a need for research examining the implementation of AI-assisted language learning within sustained instructional practice, particularly in resource-constrained contexts (Holmes & Tuomi, 2022).

Text-to-Speech and Speech Recognition for Speaking Development

TTS and ASR technologies constitute two core components of AI-assisted speaking practice. TTS systems provide learners with consistent and repeatable pronunciation models, while ASR enables immediate feedback on learners' spoken output, facilitating self-monitoring and pronunciation improvement (Evers & Chen, 2021). Empirical studies have shown that integrating ASR into learning environments significantly enhances learners' speaking performance and engagement (Tsai, 2023). Moreover, recent AI-driven speech recognition systems have demonstrated measurable improvements in learning outcomes and reduced language anxiety (McCrocklin, 2016). Advances in deep learning have further improved the accuracy, naturalness, and intelligibility of speech technologies, making them increasingly effective tools for language learning (Nassif et al., 2019).

Exposure to accurate pronunciation models is essential for developing speaking proficiency, particularly for young learners who may lack access to native or proficient English speakers (Saito & Plonsky, 2019). TTS tools allow learners to listen repeatedly, control playback speed, and focus on specific segments, thereby supporting individualized learning and self-paced practice (Uchihara et al., 2023). Research has shown that such repeated exposure can improve learners' pronunciation accuracy and strengthen the alignment between listening and speaking skills (Evers & Chen, 2021).

ASR technology, on the other hand, enables learners to receive feedback on their spoken output by converting speech into text and analyzing pronunciation features. When integrated into language learning platforms, ASR systems can provide immediate, automated feedback, enabling learners to identify errors and monitor their progress (Wu et al., 2023). Studies have shown that ASR-supported practice can enhance pronunciation accuracy, fluency, and learner confidence, particularly when combined with structured guidance and explicit corrective feedback (Thi-Nhu Ngo et al., 2024).

However, the effectiveness of TTS and ASR tools depends heavily on pedagogical design. Without appropriate scaffolding, learners may struggle to interpret automated feedback or become discouraged by recognition errors and low performance scores (McCrocklin, 2019). This highlights the importance of teacher mediation and structured support, as ASR and TTS technologies are most effective when integrated as complementary tools within guided instructional frameworks (Liakin et al., 2017). In particular, younger or less experienced learners benefit from gradual skill development and explicit guidance when using AI-assisted speaking tools.

AI-Assisted Learning and Learner Autonomy

Learner autonomy has become a central construct in contemporary educational research, particularly within digitally mediated language learning environments (Chang & Sun, 2024). It is commonly defined as learners' capacity to take control of their own learning processes, including goal setting, strategy use, and self-evaluation (Benson, 2011). Closely related to this concept, self-regulated learning (SRL) emphasizes learners' active role in monitoring and regulating their cognitive, motivational, and behavioral engagement (Zimmerman, 2002). In the context of English as a Foreign Language (EFL) learning, fostering autonomy is especially critical as language proficiency development requires sustained practice beyond formal classroom instruction (Reinders & Benson, 2017).

Recent advances in AI have introduced new possibilities for supporting learner autonomy. AI-assisted technologies, such as conversational agents and automated feedback systems, offer flexible, personalized learning opportunities that enable learners to engage in self-paced practice and receive immediate, adaptive feedback (Du, 2025). Empirical and review studies have demonstrated that such features can enhance learners' self-regulation skills, motivation, and sense of agency, thereby promoting more autonomous learning behaviors (Zhang, 2025). Moreover, AI-supported environments have been shown to facilitate iterative learning processes, enabling learners to plan, monitor, and evaluate their performance more effectively (Zhang et al., 2025).

However, the development of learner autonomy in AI-mediated contexts should not be assumed to occur automatically. Research indicates that learners' ability to benefit from AI tools is mediated by factors such as cognitive maturity, prior technological experience, and language proficiency (Zai & Zhou, 2026). This issue is particularly salient in elementary EFL settings, where young learners may lack the metacognitive awareness and strategic competence required for effective self-regulation. Consequently, scholars emphasized the importance of explicit instruction, guided practice, and ongoing scaffolding to support learners in using AI tools productively and developing sustainable self-learning habits (Chong & Reinders, 2025; Nguyen & Doan, 2025).

Method

Theoretical Framework

The study was informed by three connected concepts: AI-supported speaking practice, learner autonomy, and classroom action research. First, EFL speaking development requires repeated opportunities for oral production, pronunciation modeling, feedback, and revision. In this study, NaturalReader was used as a text-to-speech tool to provide pronunciation, rhythm, and intonation models, while Microsoft Reading Coach was used as an automatic speech recognition tool to support oral reading practice and feedback.

Second, the study was informed by Swain's Output Hypothesis (1985), which argues that language development is promoted when learners actively produce language, recognize knowledge gaps, and revise their output. NaturalReader and Microsoft Reading Coach provided opportunities for repeated oral production, pronunciation modeling, and immediate feedback, enabling students to refine their spoken English through repeated practice. This theoretical perspective guided the design and interpretation of the AI-supported speaking activities implemented in this classroom action research.

Third, the study was framed by classroom action research. Rather than testing the causal effectiveness of AI tools under controlled experimental conditions, the study focused on how AI-supported speaking activities were implemented, observed, and adjusted in one rural elementary EFL classroom. This framework connects technology-supported speaking practice, learner autonomy, and reflective classroom inquiry as the conceptual basis for the study.

Research Design

This study employed a single-group classroom action research design to investigate the implementation of AI-supported speaking activities in one local EFL classroom over a 12-week period. The study was designed to examine classroom practice and document observed changes in students' speaking performance within this specific instructional context, rather than to test the causal effectiveness of AI tools. Because no control or comparison group was included, the observed changes cannot be attributed solely to the AI-supported activities. Instead, the study aimed to generate context-sensitive evidence to inform local teaching practice and teacher reflection.

Action research is widely recognized as an appropriate methodology for educational settings because it enables teachers to systematically investigate and improve their instructional practices within authentic classroom contexts while responding to learners' emerging needs (Burns, 2010; Kemmis et al., 2014). Rather than implementing a fixed intervention under controlled experimental conditions, action research involves iterative cycles of planning, action, observation, and reflection that inform ongoing instructional refinement (Kemmis et al., 2014). Accordingly, the present study adopted a classroom-based action research design to examine how AI-supported speaking activities were implemented and refined over a 12-week instructional period. This design was intended to support context-sensitive pedagogical improvement rather than establish causal relationships between AI tools and learning outcomes.

Action Research Cycles and Instructional Adjustments

To provide clearer evidence of the action research process, the 12-week instruction was organized into three four-week cycles of planning, action, observation, and reflection, following the action research framework proposed by Burns (2010) and Kemmis et al. (2014). In the planning stage, the teacher-researcher identified students' speaking difficulties, including limited oral confidence, pronunciation problems, and insufficient opportunities for repeated practice. Based on these needs, two AI-supported tools were integrated into classroom instruction: NaturalReader, a TTS tool that provides pronunciation models, and Reading Coach, an ASR tool that supports students' oral reading practice and provides feedback on their spoken output.

Cycle 1 (Weeks 1–4). Students used NaturalReader to listen to model readings of target words, sentences, and short passages. They were guided to notice pronunciation, stress, rhythm, and intonation before practicing aloud. Weekly classroom observations and field notes collected during Weeks 1–4 indicated that some students listened passively or imitated the audio without fully understanding the text. Based on these observations, the teacher modified the instruction for the next cycle by incorporating vocabulary explanation, sentence-level comprehension checks, guided repetition, and pair practice before independent speaking activities.

Cycle 2 (Weeks 5–8). Students used Reading Coach to read aloud and receive feedback on pronunciation and oral reading performance. Observation records from this cycle showed that several learners focused primarily on whether the system correctly recognized their speech, while others became anxious when their pronunciation was not recognized accurately. These observations informed the instructional adjustments implemented in Cycle 3, where students were reminded that Reading Coach was intended as a formative practice tool rather than a test.

Cycle 3 (Weeks 9–12). Students were encouraged to combine NaturalReader and Reading Coach more independently. They first listened to pronunciation models, practiced with peers, read aloud using Reading Coach, reviewed the feedback, and revised their pronunciation through repeated practice. Based on classroom observations and students' learning responses, the teacher provided additional scaffolding for lower-proficiency learners, including sentence frames, vocabulary support, slower-paced listening models, repeated modeling, and shorter reading segments. After each weekly lesson, the teacher-researcher documented classroom observations and instructional reflections in field notes, which were reviewed before the subsequent lesson to guide instructional adjustments across the action research cycles. These iterative refinements reflected the action research purpose of improving local classroom practice rather than implementing a fixed intervention.

Context and Participants

The study was conducted in a rural elementary school in Taiwan, where access to supplementary English learning resources was limited. English is taught as a compulsory subject from Grade 3, with students receiving one to two general English classes per week, supplemented by a weekly session taught by a native English-speaking instructor. None of the participants attended private English tutoring or cram schools outside regular school hours.

Table 1. *Demographic Characteristics of the Participants*

Grade	Participants	Age	Pre-test mean	Boys	Girls
Grade 4	5	9–10	7.60	2	3
Grade 5	5	10–11	7.60	4	1
Grade 6	3	11–12	9.67	3	0

Thirteen students from Grades 4 to 6 participated in the study. The mixed-grade design reflected the actual instructional arrangement of the school's English-speaking program. To ensure anonymity, participants were assigned codes from P1 to P13. Based on grade level, participants were grouped as follows: Grade 6 (P1–P3), Grade 5 (P4–P8), and Grade 4 (P9–P13). Because each grade-level subgroup included only three to five students, grade-level patterns were interpreted cautiously as tentative and illustrative observations rather than as developmental findings or generalizable grade-based differences.

Baseline characteristics, including students' age, gender, grade level, and pre-test speaking scores, were summarized to provide contextual information about the participants and to check whether the grade groups differed noticeably before the instructional period. Given the small subgroup sizes, no inferential comparison was conducted by grade level. Instead, baseline information was used descriptively to support cautious interpretation of students' learning patterns across the mixed-grade classroom.

Prior to data collection, written informed consent was obtained from parents or legal guardians, and verbal assent was obtained from all participating students. Students were informed that participation was voluntary and that they could withdraw from the study at any time without penalty. The study complied with ethical standards for research involving minors, including confidentiality, anonymity, secure data handling, and the right to withdraw from the study without penalty. Ethical approval was obtained from the university review board. The datasets generated and analyzed during the current study are not publicly available because they contain identifiable information from minor participants, but are available from the corresponding author upon reasonable request. The authors declare no conflict of interest.

AI-Assisted Tools

Two AI-assisted platforms were selected for the intervention: Microsoft Reading Coach and NaturalReader. They were selected because they provided age-appropriate support for speaking practice through two complementary functions: ASR and TTS. Microsoft Reading Coach served as the output and feedback tool. Students read aloud selected texts, and the platform provided ASR-based feedback on pronunciation, pacing, and accuracy. It also highlights mispronounced words, allowing learners to focus on specific problem areas. In this study, Reading Coach was used as a formative practice tool rather than as an independent proficiency assessment.

NaturalReader served as the input and modeling tool. It converted written texts into natural-sounding speech and allowed students to replay the audio and adjust the playback speed. Students used NaturalReader before speaking practice to listen to models of pronunciation, rhythm, stress, and intonation. Together, NaturalReader and Microsoft Reading Coach supported a learning sequence in which students first received pronunciation input and then practiced oral output with feedback.

Instructional Procedure

The intervention lasted 12 weeks, with one 50-minute AI-assisted speaking session conducted each week. Each session followed a consistent instructional structure designed to promote listening–speaking integration and self-directed practice. The AI-supported speaking sessions supplemented students' regular English curriculum, which consisted of one to two English classes per week and one additional weekly lesson taught by a native English-speaking instructor. Accordingly, the AI intervention should be regarded as one component of students' overall English learning experience rather than as an isolated instructional treatment. The difficulty of the AI-generated passages was gradually increased to align with students' developing English proficiency throughout the intervention. Consequently, weekly Reading Coach scores were interpreted as formative indicators of students engagement and progress in classroom practice rather than as standardized measures of speaking proficiency.

At the beginning of each session, the researcher introduced the speaking task and demonstrated how to use the two AI-assisted tools when necessary. Students were then assigned two speaking passages: one teacher-selected text aligned with the textbook content and one AI-generated text created through Microsoft Reading Coach at a designated difficulty level. The AI-generated texts allowed for variation while maintaining linguistic appropriateness.

Students first listened to the passages using NaturalReader, adjusting the playback speed as needed. They were encouraged to listen repeatedly and focus on pronunciation, stress, and intonation. After sufficient listening practice, students read the passages aloud using Microsoft Reading Coach, which provided immediate automated feedback and a numerical performance score. Students were allowed to repeat the task multiple times before submitting their final scores.

Throughout the intervention, the researcher provided scaffolded support, particularly during the initial weeks. This support included technical assistance, pronunciation guidance, and encouragement strategies to reduce anxiety. As students became more familiar with the tools, teacher intervention was gradually reduced to promote independent practice.

Instruments and Data Collection

Multiple data sources were used to examine the impact of AI-assisted tools on speaking proficiency, self-learning ability, and learner perceptions.

Speaking Proficiency Tests

Students' speaking performance was assessed through pre- and post-tests adapted from the Cambridge English A1 Movers Speaking test format. The test was selected because its task types were age-appropriate for elementary EFL learners and aligned with the students' instructional level. However, the present study did not use the official Cambridge reporting system, which is normally presented in shields. Instead, students' responses were evaluated using a researcher-developed analytic speaking rubric designed for classroom-based assessment.

The rubric included three dimensions: vocabulary use, pronunciation, and interaction. Each dimension was scored on a 5-point scale, yielding a maximum score of 15. The rubric was developed based on the speaking descriptors of the Cambridge A1 Movers test and reviewed by two experienced EFL educators to establish content validity. Equal weighting was adopted because the study aimed to provide a balanced classroom-based assessment of overall speaking performance. Vocabulary use was assessed to determine students' ability to use appropriate words and expressions in response to the speaking tasks. Pronunciation assessed the clarity and comprehensibility of spoken output. Interaction assessed students' ability to respond to prompts, maintain simple communication, and participate in the speaking exchange.

All speaking tests were conducted individually by the researcher. To ensure scoring reliability, two experienced English language educators independently rated the recorded performances using the same analytic rubric. The final speaking score for each participant was calculated as the average of the two raters' scores, and these averaged scores were used in all subsequent statistical analyses. Inter-rater reliability was examined using intraclass correlation coefficients (ICCs) based on a two-way random-effects, absolute-agreement, average-measures model. Because the researcher also served as the instructor, test administrator, and classroom observer, potential expectancy and observer bias should be acknowledged. To reduce this risk, students' speaking performances were audio-recorded and independently evaluated by the two external raters. The qualitative findings were triangulated using multiple data sources, including student interviews, teacher field notes, weekly Reading Coach records, and pre- and post-test results. Representative student quotations and field-note excerpts were included to enhance the transparency of the analysis. Although these procedures strengthened the trustworthiness of the findings, the researcher's multiple roles remain a limitation of the study.

Weekly Performance Records

Weekly performance data were collected from Microsoft Reading Coach. Each student submitted two speaking scores per week, corresponding to the teacher-selected and AI-generated passages. Students were permitted to repeat each speaking task multiple times before submitting their final score. Accordingly, the weekly Reading Coach scores were treated as formative process data documenting students' engagement with AI-supported speaking practice rather than standardized measures of speaking proficiency. Because the passages varied across weeks and the number of practice attempts was not recorded, the weekly scores may also have been influenced by repeated practice, task familiarity, and differences in text difficulty.

Classroom Observations and Field Notes

Classroom observations were conducted throughout the intervention using an observer-as-participant approach. The researcher documented students' engagement, use of learning strategies, interactions with the AI tools, and behavioral changes related to autonomy and confidence. Field notes were recorded during and immediately after each session to capture instructional adjustments and notable learning behaviors.

Semi-Structured Interviews

Semi-structured interviews were conducted during the final week of the intervention. Interviews were organized by grade level in Mandarin Chinese to encourage comfortable peer interaction and age-appropriate responses. The interview protocol consisted of open-ended questions exploring students' experiences using the AI-assisted tools, perceived changes in speaking ability, self-directed learning behaviors, and overall attitudes toward AI-supported learning. All interviews were audio-recorded and transcribed verbatim. Selected quotations included in the manuscript were translated into English by the researcher and reviewed by a bilingual English language educator to ensure translation accuracy. Although the group interview format encouraged discussion among students of similar ages, it may also have introduced peer-conformity effects, as some participants may have been influenced by their classmates' responses. This limitation was taken into account when interpreting the qualitative findings.

Data Analysis

Quantitative data from the pre- and post-tests were analyzed using paired-samples t-tests to examine observed changes in speaking proficiency. Given the small sample size, the t-test was interpreted as supportive evidence and supplemented by a Wilcoxon signed-rank test as a robustness check. Weekly Reading Coach scores were analyzed descriptively as formative process data. Qualitative data from student interviews and teacher field notes were analyzed using reflexive thematic analysis (Braun & Clarke, 2006). The researchers repeatedly reviewed the transcripts and field notes to become familiar with the data before generating initial codes. Related codes were grouped into broader themes, which were subsequently reviewed, refined, and interpreted in relation to the research questions.

To enhance trustworthiness, several strategies were used. First, data triangulation was applied by comparing student interview responses, teacher field notes, weekly Microsoft Reading Coach process data, and Cambridge A1 Movers speaking test results. Second, representative student quotations and field-note excerpts were included to support the interpretation of the qualitative findings. Third, the teacher-researcher reviewed the coding results and classroom records repeatedly to check whether the themes were consistent with the observed classroom process. Finally, the findings were interpreted cautiously as context-specific evidence from a local action research study rather than as broadly generalizable results.

Results

This section presents the findings related to the three research questions, integrating quantitative and qualitative data to illustrate how AI-assisted tools influenced (1) speaking proficiency, (2) self-learning ability, and (3) learner perceptions. Rather than reporting outcomes in isolation, the results are interpreted through theoretical lenses of input–output interaction, feedback-mediated learning, and self-regulated learning development.

Changes in Students' Speaking Performance

To examine changes in students' speaking performance, a paired-samples *t*-test was conducted using Cambridge A1 Movers speaking test scores collected before and after the 12-week instructional period. The Shapiro-Wilk test indicated that the difference scores did not significantly deviate from normality, $W = 0.897$, $p = .122$, suggesting that the assumption of normality for the paired-samples *t*-test was met. As shown in Table 2, students' speaking scores increased significantly from pre-test ($M = 8.08$, $SD = 1.73$) to post-test ($M = 9.77$, $SD = 2.14$), $t\ (12) = -7.138$, $p < .001$. The mean difference was 1.69 points, with an *SD* of the differences of 0.855 and a 95% confidence interval of [1.18, 2.21]. The effect size was Cohen's $d = 1.98$, representing a large within-subject effect according to Cohen's (1988) guidelines.

Table 2. *Paired Sample t-test Analysis of Pre- and Post-Test Speaking Scores*

Variable	Pre-test M (SD)	Post-test M (SD)	Mean Difference (SD)	95% CI	*t* (12)	*p*
Speaking score	8.08 (1.73)	9.77 (2.14)	1.69 (0.86)	[1.18, 2.21]	−7.14	< .001

Table 3. *Intraclass Correlation Coefficient (ICC) of Both Raters*

Test	ICC	Lower 95% CI	Upper 95% CI
Pre-test	0.733	0.329	0.910
Post-test	0.750	0.361	0.916

**Note.* 13 subjects and 2 raters.

The final speaking score for each participant was calculated as the average of the two external raters' scores, and these averaged scores were used in the paired-samples *t*-test. Inter-rater reliability was examined using intraclass correlation coefficients (ICCs) calculated with a two-way random-effects, absolute-agreement, average-measures model. As shown in Table 3, the ICCs indicated moderate agreement between raters for both the pre-test (ICC = .733, 95% CI [.329, .910]) and the post-test (ICC = .750, 95% CI [.361, .916]). Although the reliability estimates suggested reasonable consistency, the relatively wide 95% confidence intervals reflect the uncertainty associated with the small sample size. Consequently, while the observed mean improvement of 1.69 points provides evidence of positive change during the instructional period, it should be interpreted cautiously, as the reliability estimates are limited.

In addition to pre–post comparisons, weekly speaking scores derived from Microsoft Reading Coach provided longitudinal evidence of learners' performance development. Descriptive analysis indicated a clear upward trajectory over the 12-week period, with relatively rapid improvement during the initial four weeks, followed by a phase of stabilization and gradual growth.

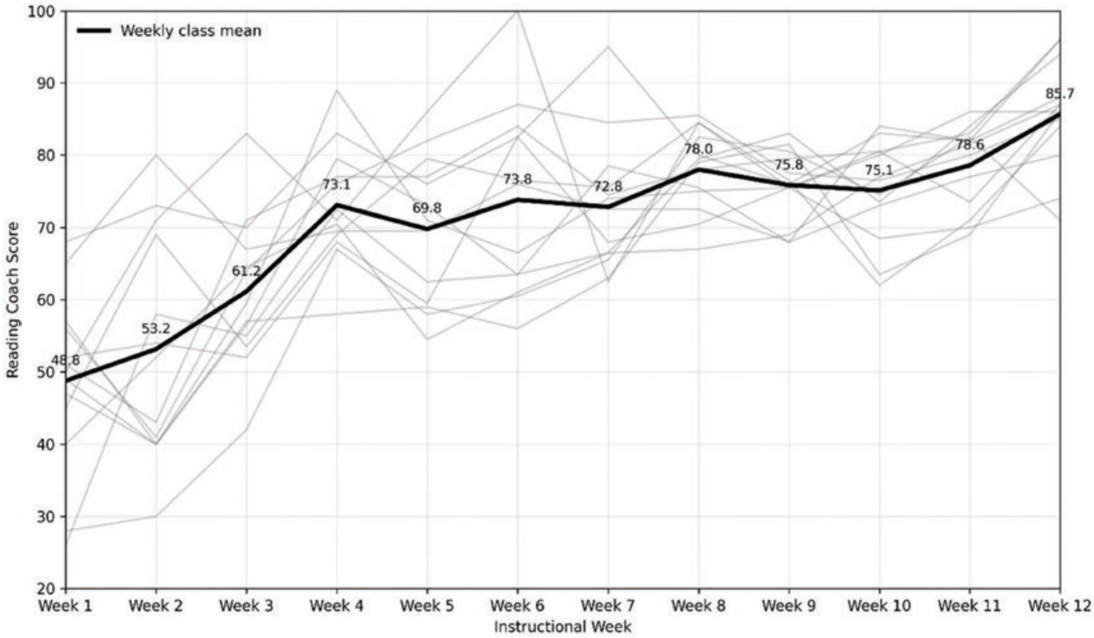

Figure 1 *Weekly Mean Reading Coach Scores Across the 12-Week Intervention*

Students' weekly Reading Coach scores were collected as formative process data throughout the 12-week intervention. Figure 1 presents the individual participants' final submitted Reading Coach scores for the two weekly speaking passages, together with the class means. Because students were permitted multiple attempts before submitting their final scores and the passages varied in difficulty across weeks, these scores were interpreted as indicators of students' engagement with AI-supported speaking practice rather than standardized measures of speaking proficiency. As shown in Figure 1, although many students demonstrated higher scores over the instructional period, substantial variation was evident across individual trajectories. Accordingly, the figure is presented as descriptive evidence of students' participation in AI-assisted speaking practice rather than as direct evidence of growth in speaking proficiency.

This pattern suggests that students became more familiar with the oral reading routines, Reading Coach feedback, and repeated pronunciation practice over time. However, because Reading Coach scores may be influenced by text difficulty, task familiarity, pronunciation recognition accuracy, background noise, and repeated practice, they should be interpreted as formative evidence of students' practice process rather than standardized proficiency outcomes.

Student interviews and teacher field notes further supported these process-oriented findings. One student stated, "At first, I felt nervous when I had to read English aloud. But after practicing many times with NaturalReader and Reading Coach, I felt less afraid to speak." Another student explained, "I liked that I could practice by myself. If I made a mistake, I could listen again, read again, and check my score." Field notes also showed changes in participation. In the early stage, several students were hesitant to read aloud and needed teacher modeling. By Week 9, students followed the routine more independently by listening to NaturalReader, practicing, and reading aloud in Reading Coach with less teacher prompting.

Classroom observations indicated that students varied in the extent to which they required teacher support during the intervention. Some students became familiar with the listen–practice–record–review routine more quickly than others, whereas some continued to require additional modeling, vocabulary support, and repeated practice. Given the small sample size and the single-group action research design, these observations should be interpreted as context-specific classroom observations rather than evidence of systematic differences among learner groups.

Development of Self-Learning Ability

Classroom observations and field notes indicated gradual yet consistent changes in students' learning behaviors throughout the intervention. During the initial phase, many students, particularly those in lower grade levels, relied heavily on teacher support for both technical operation and pronunciation guidance. Over time, however, learners increasingly engaged in independent problem-solving behaviors, such as replaying TTS input, adjusting playback speed, and repeating speaking tasks without external prompting.

These emerging behaviors can be interpreted through the lens of SRL, which conceptualizes learning as an active process involving planning, monitoring, and self-evaluation. In this study, AI-supported tools appeared to function as external regulatory supports that scaffolded learners' engagement with these processes. Through repeated interaction, learners gradually assumed greater control over their practice routines, suggesting a shift from externally regulated to more internally regulated learning.

By the latter half of the intervention, most students were able to complete speaking tasks independently with minimal teacher intervention. This transition reflects a movement toward learner-regulated practice, facilitated by the structured feedback and consistent task design embedded within the AI tools. This finding is consistent with recent research suggesting that AI-supported language learning environments can provide personalized feedback, support learner autonomy and competence, and encourage more active engagement in language practice (Griche & Bennis, 2026; Tran, 2026).

Interview data further supported students' self-directed speaking practice. Students reported actively adjusting playback speed, repeating challenging segments, and producing multiple recordings to improve their performance. These observations indicate that students actively used AI-supported tools to revise and refine their oral production during speaking practice. Although such behaviors may be associated with self-regulated learning, the present study did not directly measure metacognitive awareness or learner autonomy. Therefore, these findings should be interpreted as evidence of observed practice behaviors rather than changes in underlying self-regulatory processes. Similar patterns of repeated practice and AI-supported feedback have been reported in previous studies of technology-enhanced language learning.

Notably, differences in self-learning development were observed across grade levels. Grade 4 students demonstrated emerging autonomy, often exploring tool functionalities but still requiring occasional guidance. Grade 5 students exhibited more stable self-regulatory behaviors and were able to articulate strategies for independent tool use. In contrast, Grade 6 students demonstrated higher technical proficiency but tended to adopt a more outcome-oriented approach, frequently discontinuing practice once a satisfactory score was achieved. These variations suggest that developmental factors, including cognitive maturity and prior learning experience, mediate learners' engagement with AI-assisted environments.

Learner Perceptions of AI-Assisted Speaking Practice

Across all grade levels, students generally reported positive perceptions of AI-assisted speaking practice. A recurring theme in student responses was increased confidence in speaking English, which many attributed to the opportunity to practice privately and to repeat tasks without fear of making mistakes in front of classmates or the teacher. As one Grade 5 student explained, “I liked that I could practice on my own and didn’t have to worry about making mistakes while I was learning.” This finding suggests that the AI-supported environment may have reduced affective barriers to participation and encouraged greater willingness to engage in speaking activities.

Another commonly reported theme was that students viewed mistakes as part of the learning process rather than as failures. One Grade 4 student commented, “If I made a mistake, I listened again and tried again until I got it right.” The AI tools appeared to create a low-stakes learning environment in which errors were perceived as part of an iterative learning process rather than as instances of public failure. This feature was particularly beneficial for younger learners, who demonstrated sustained engagement even when their initial performance was relatively low. The ability to control the pace and frequency of practice allowed students to gradually build confidence while remaining actively engaged.

Despite these generally positive perceptions, differences emerged across grade levels. Older students, particularly Grade 6 students, expressed more critical evaluations of the AI tools. While they acknowledged improvements in fluency and pronunciation, they also noted limitations in the system’s ability to provide detailed explanations or nuanced feedback comparable to that of a human teacher. This pattern suggests that as learners’ cognitive and linguistic demands increase, their expectations for instructional support become more sophisticated.

From a pedagogical perspective, these findings underscore the complementary role of AI-assisted tools within language instruction. Rather than serving as a replacement for teacher-led instruction, AI-supported practice appears to function most effectively as a supplementary resource that enhances opportunities for individualized and repeated practice, a finding consistent with teachers’ perceptions of AI as a complement rather than a replacement for classroom instruction (Chung, 2025). The findings further indicate that integrating AI tools should be accompanied by teacher-mediated explanation and guidance, particularly for more advanced learners who require deeper conceptual support.

Discussion

Designing AI-Assisted Speaking Practice as a Feedback-Rich Learning Environment

One of the central findings of this study is that sustained engagement with AI-assisted speaking tools was associated with measurable improvements in learners’ speaking proficiency. From a theoretical perspective, these findings suggest that AI-assisted speaking practice facilitated oral language development by providing repeated opportunities for language production supported by immediate feedback. Consistent with output-based theories of second language acquisition, such feedback likely promoted learners’ noticing of linguistic gaps and subsequent interlanguage restructuring. The integration of TTS input and ASR-based feedback created a feedback-rich environment that enabled learners to access pronunciation models, monitor their performance, and make immediate corrections. This interpretation is consistent with recent empirical research demonstrating that AI-assisted feedback can improve EFL learners’ pronunciation and speaking performance by facilitating timely error detection and correction (Nguyen Huu Lam & Tran Thanh, 2026; Zou et al., 2023).

Compared with traditional classroom contexts, where feedback is often delayed, AI-supported environments facilitate rapid practice–revision cycles, thereby supporting more efficient learning (Liu et al., 2025). Most importantly, the findings extend previous studies by highlighting the role of feedback loop design. The integration of input and output within a single system enabled iterative practice, which is critical for developing speaking proficiency (Aryanti & Santosa, 2024). In addition, learner control over pacing and repetition supported individualized learning and may have reduced cognitive load (Du, 2025).

Scaffolded Design for Young Learners: Gradual Release of Control

Another important design implication concerns the role of scaffolding in AI-assisted learning environments. Although such tools are often positioned as supporting self-directed learning, this study's findings indicate that young learners require structured guidance during initial engagement. In the early stages of the intervention, students, particularly those in lower grades, relied on teacher support to interpret automated feedback, regulate affective responses, and develop effective practice strategies.

From a design perspective, this underscores the importance of incorporating progressive scaffolding into AI-supported learning environments. Rather than assuming immediate learner autonomy, effective implementations should adopt a gradual release of control, enabling learners to transition from guided participation to independent operation. This aligns with recent research demonstrating that AI systems can function as external scaffolds that support the development of self-regulated learning, particularly when guidance is adaptively reduced over time (Du, 2025). Similarly, studies in K–12 EFL contexts have shown that scaffolded AI interactions can enhance learner autonomy, but only when combined with structured instructional support and opportunities for strategy internalization (Li & Wilson, 2025).

The present findings extend this line of research by highlighting how consistent task structures and repeated exposure to tool functionalities can facilitate the internalization of learning strategies. Over time, learners demonstrated increased independence, suggesting that scaffolding embedded within both instructional design and technological affordances can effectively support the transition toward autonomous learning. This process reflects a dynamic interplay between external regulation and internal control, rather than an immediate shift to self-directed learning.

These findings caution against conceptualizing AI-assisted tools as "plug-and-play" solutions. Instead, their effectiveness depends on deliberate instructional integration and ongoing human mediation, particularly for younger learners. This is consistent with prior work emphasizing that teacher scaffolding remains essential in AI-enhanced environments to enable learners to meaningfully engage with feedback and develop sustainable learning strategies (Zhang & Derakhshan, 2025). For educational technology designers, these results highlight the need to consider developmental appropriateness, structured onboarding, and scaffolded interaction when implementing AI tools in elementary language learning contexts.

Designing for Learner Autonomy: Tools as Cognitive Supports

The findings further indicate that AI-assisted tools can support the development of learner autonomy when conceptualized as cognitive supports rather than standalone instructional agents. Over the course of the intervention, students increasingly engaged in self-monitoring behaviors, including replaying audio input, adjusting speech rate, and repeating tasks to refine their performance. These behaviors reflect core components of SRL, particularly in terms of monitoring and strategic control of learning processes.

From a theoretical perspective, this pattern suggests that AI tools function as external regulators that scaffold learners' engagement with SRL processes. Rather than directly promoting autonomy, the tools provided structured feedback and interactional affordances that learners gradually appropriated for their own learning purposes. This interpretation is consistent with recent research indicating that AI systems can enhance learner autonomy by supporting metacognitive regulation and strategy use, particularly when learners actively engage with feedback (Xu et al., 2026). Similarly, empirical studies have shown that AI-supported environments can foster self-regulated behaviors, but the extent of autonomy development varies across learners depending on individual characteristics such as proficiency and prior experience (Du, 2025).

The present findings highlight variability in the emergence of autonomy. Learners differed in the extent to which they engaged in self-regulatory behaviors, suggesting that autonomy is not a uniform outcome of AI use. This aligns with meta-analytic evidence demonstrating significant variation in SRL outcomes in AI-mediated learning environments, influenced by factors such as age, proficiency level, and technological familiarity (Shafiee Rad, 2025). From a design perspective, these findings underscore the need for flexible autonomy in AI-assisted systems. Rather than maximizing learner independence from the outset, effective design should accommodate different levels of learner control. For instance, younger or lower-proficiency learners may benefit from more structured interfaces and explicit prompts, while more advanced learners may require greater flexibility and challenge.

Differentiated Design Needs Across Learner Levels

A notable finding of this study is the variation in learner perceptions across grade levels. While younger students expressed higher levels of engagement and motivation, older learners demonstrated more critical attitudes toward AI-assisted tools, particularly regarding the limited depth of explanation compared to teacher-provided instruction. This divergence suggests that learners' expectations of instructional support evolve alongside their cognitive and linguistic development. Such findings align with prior research indicating that although AI-assisted tools are effective in supporting lower-level language skills such as pronunciation and fluency, they are less capable of addressing higher-order cognitive and metalinguistic needs without additional pedagogical support (Zou et al., 2023). This limitation highlights an important gap between the technological affordances of AI tools and the evolving learning needs of more proficient users.

From a design perspective, these findings underscore the limitations of one-size-fits-all approaches in educational technology. Instead, they point to the need for adaptive or layered AI systems that can provide differentiated support based on learner characteristics such as age, proficiency, and learning goals. Emerging research suggests that incorporating multiple levels of feedback—ranging from surface-level corrective feedback to more elaborated explanations—can enhance both engagement and learning outcomes across diverse learner groups (Shafiee Rad, 2025; Xu et al., 2026). Future research should incorporate multiple levels of feedback, ranging from surface-level corrective feedback to more elaborate explanations, to enhance both engagement and learning outcomes across diverse learner groups.

Context-Sensitive Design: AI in Rural and Under-Resourced Settings

The rural context of this study highlights the importance of context-sensitive design in AI-assisted language learning. In this setting, AI tools provided learners with access to pronunciation models and immediate feedback that might otherwise be limited, thereby supporting more equitable learning opportunities. This aligns with prior research suggesting that AI-assisted technologies can help mitigate disparities in instructional resources by offering consistent and individualized support (Shafiee Rad, 2025).

At the same time, practical constraints such as unstable internet connectivity and limited device availability occasionally disrupted the learning process. These findings align with the existing literature, which indicates that AI's effectiveness in education depends not only on pedagogical design but also on infrastructural conditions, particularly in under-resourced contexts (Xu et al., 2026). From a design perspective, these challenges underscore the need for resilient and adaptable systems. Features such as lightweight platforms, offline functionality, and tolerance for connectivity interruptions are especially important in rural settings.

Contributions to Educational Technology Research

From an educational technology perspective, this study contributes empirical evidence on how AI-assisted tools function as designed learning environments in elementary EFL contexts. Rather than evaluating technology as an isolated variable, the findings highlight the importance of examining the dynamic interaction between design, use, and learner engagement over time. The action research approach enabled close observation of how learners appropriated technological affordances in authentic classroom settings, revealing developmental patterns that may be overlooked in short-term or controlled experimental studies (Yang, 2026).

The study also contributes to ongoing discussions regarding the role of AI in education by demonstrating that technological effectiveness is not inherent to the tools themselves but emerges through pedagogical design, scaffolding, and contextual alignment. This perspective is consistent with recent research emphasizing that AI-supported learning outcomes depend on how feedback, interaction, and learner control are structured within the learning environment (Hou & Zhou, 2025).

Limitations and Threats to Validity

Several limitations should be acknowledged when interpreting the findings of this study. First, the study employed a single-group classroom action research design without a control or comparison group. Therefore, although students' speaking performance showed positive changes after the 12-week instructional period, these gains cannot be attributed solely to the AI-supported activities. Other possible explanations may have contributed to the observed improvement.

One possible threat to validity is maturation. Because the study took place over a 12-week period, students may have naturally developed greater speaking confidence, language awareness, or classroom familiarity. In addition, repeated speaking practice may have helped students improve their oral performance regardless of whether AI tools were used. The pretest-posttest format may also have introduced testing effects, as students may have become more familiar with the speaking task, assessment criteria, or test format by the time they completed the posttest.

Another limitation concerns the teacher-researcher role. Since the teacher also served as the researcher, the instructional process may have involved increased attention, encouragement, feedback, or support during the intervention period. These factors may have influenced students' motivation and performance. As a result, the findings should be interpreted as context-specific evidence from a local action research project rather than causal evidence proving the independent effectiveness of AI tools. Future studies could strengthen the research design by including a comparison group, collecting data across multiple classrooms, or adopting a quasi-experimental design to better examine the relationship between AI-supported instruction and speaking development.

Conclusion

This classroom action research examined the implementation of AI-supported speaking activities using NaturalReader and Microsoft Reading Coach in a rural elementary EFL classroom. The findings showed improvements in students' speaking scores, together with positive classroom observations and interview responses, indicating that the AI-supported activities encouraged repeated speaking practice, pronunciation awareness, and active engagement during the instructional period. Given the single-group action research design, these findings should be interpreted as context-specific evidence of classroom implementation rather than causal evidence of the independent effectiveness of AI-assisted tools.

The study contributes to educational technology research by demonstrating how TTS and ASR tools can be integrated into elementary EFL instruction as feedback-rich practice resources. Rather than replacing teacher-led instruction, AI-supported speaking activities appear to be most effective when combined with appropriate teacher guidance and classroom scaffolding.

Several limitations should be acknowledged. The study involved a small sample from a single rural elementary school and employed a single-group design without a comparison group. In addition, the AI-supported activities were implemented alongside students' regular English classes, and the weekly Reading Coach scores were collected as formative process data rather than standardized measures of speaking proficiency. Future research should include larger samples, comparison groups, and longer intervention periods to further examine the effectiveness of AI-supported speaking instruction across different educational contexts.

References

Aravantinos, S., Lavidas, K., Voulgari, I., Papadakis, S., Karalis, T., & Komis, V. (2024). Educational approaches with AI in primary school settings: A systematic review of the literature available in Scopus. *Education Sciences*, *14*(7), 744. https://doi.org/10.3390/educsci14070744

Aryanti, R. D., & Santosa, M. H. (2024). A systematic review on artificial intelligence applications for enhancing EFL students' pronunciation skill. *The Art of Teaching English as a Foreign Language*, *5*(1), 102–113. https://doi.org/10.36663/tatefl.v5i1.718

Benek, K. (2025). EFL learners' and teachers' perceptions of AI-powered language learning technologies: Benefits and challenges. *International Journal of Instruction*, *18*(2), 103–120. https://doi.org/10.29333/iji.2025.1827a

Benson, P. (2011). *Teaching and researching autonomy* (2nd ed.). Routledge.

Bibauw, S., François, T., & Desmet, P. (2019). Discussing with a computer to practice a foreign language: Research synthesis and conceptual framework of dialogue-based CALL. *Computer Assisted Language Learning*, *32*(8), 827–877. https://doi.org/10.1080/09588221.2018.1535508

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brown, G., & Yule, G. (2019). *Teaching the spoken language*. Cambridge University Press.

Burns, A. (2010). *Doing action research in English language teaching: A guide for practitioners*. Routledge.

Butler, Y. G. (2011). The implementation of communicative and task-based language teaching in the Asia-Pacific region. *Annual Review of Applied Linguistics*, *31*, 36–57. https://doi.org/10.1017/S0267190511000122

Chang, W.-L., & Sun, J. C.-Y. (2024). Evaluating AI's impact on self-regulated language learning: A systematic review. *System*, *126*, 103484. https://doi.org/10.1016/j.system.2024.103484

Chen, K. T. C. (2022). Speech-to-text recognition in university English as a foreign language learning. *Education and Information Technologies*, *27*(7), 9857–9875. https://doi.org/10.1007/s10639-022-11016-5

Chong, S. W., & Reinders, H. (2025). Autonomy of English language learners: A scoping review of research and practice. *Language Teaching Research*, *29*(2), 607–632. https://doi.org/10.1177/13621688221075812

Chung, A. Q. (2025). AI integration in pronunciation instruction: Insights from secondary school EFL teachers. *Technology in Language Teaching & Learning, 7*(4), 103014. https://doi.org/10.29140/tltl.v7n4.103014

Cizrelioğlu, İ., & Aydin, S. (2025). Using artificial intelligence in EFL learning among multilingual young learners. In M. Sanmugam, B. I. Edwards, N. M. Mohd Barkhaya, & Z. Khlaif (Eds.), *Fostering inclusive education with AI and emerging technologies* (pp. 303–322). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3693-7255-5.ch013

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

de Ruig, N. J., de Jong, P. F., & Zee, M. (2023). Stimulating elementary school students' self-regulated learning through high-quality interactions and relationships: A narrative review. *Educational Psychology Review*, *35*(3), 71. https://doi.org/10.1007/s10648-023-09795-5

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins.

Dou, A., Xu, W., Li, X., Zhang, S., & Zhang, J. (2025). Artificial intelligence in language learning. *International Journal of Distance Education Technologies*, *23*(1), 1–24. https://doi.org/10.4018/IJDET.385045

Du, J., & Daniel, B. K. (2024). Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. *Computers and Education: Artificial Intelligence*, *6*, 100230. https://doi.org/10.1016/j.caeai.2024.100230

Du, Q. (2025). How artificially intelligent conversational agents influence EFL learners' self-regulated learning and retention. *Education and Information Technologies*, *30*(15), 21635–21701. https://doi.org/10.1007/s10639-025-13602-9

Evers, K., & Chen, S. (2021). Effects of automatic speech recognition software on pronunciation for adults with different learning styles. *Journal of Educational Computing Research*, *59*(4), 669–685. https://doi.org/10.1177/0735633120972011

Fan, J., & Zhang, Q. (2024). From literacy to learning: The sequential mediation of attitudes and enjoyment in AI-assisted EFL education. *Heliyon*, *10*(17), e37158. https://doi.org/10.1016/j.heliyon.2024.e37158

Fathi, J., Rahimi, M., & Derakhshan, A. (2024). Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System*, *121*, 103254. https://doi.org/10.1016/j.system.2024.103254

Griche, M. S., & Bennis, M. (2026). Generative artificial intelligence integration in foreign language education in higher education: A systematic literature review. *Technology in Language Teaching & Learning*, *8*, 102827. https://doi.org/10.29140/tltl.2026.102827

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, *57*(4), 542–570. https://doi.org/10.1111/ejed.12533

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, *70*(2), 125–132. https://doi.org/10.1111/j.1540-4781.1986.tb05256.x

Hou, F., & Zhou, C. (2025). How AI shapes self-regulated learning in foreign language education. *Discover Computing*, *28*(1), 344. https://doi.org/10.1007/s10791-025-09801-0

Hwang, G.-J., & Tu, Y.-F. (2021). Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review. *Mathematics*, *9*(6), 584. https://doi.org/10.3390/math9060584

Hwang, G.-J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *1*, 100001. https://doi.org/10.1016/j.caeai.2020.100001

Hwang, W.-Y., Shadiev, R., Hsu, J.-L., Huang, Y.-M., Hsu, G.-L., & Lin, Y.-C. (2016). Effects of storytelling to facilitate EFL speaking using web-based multimedia system. *Computer Assisted Language Learning*, *29*(2), 215–241. https://doi.org/10.1080/09588221.2014.927367

Hwang, W.-Y., Shadiev, R., & Huang, S.-M. (2011). A study of a multimedia web annotation system and its effect on the EFL writing and speaking performance of junior high school students. *ReCALL*, *23*(2), 160–180. https://doi.org/10.1017/S0958344011000061

Jeon, J. (2024). Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives. *Computer Assisted Language Learning*, *37*(1-2), 1–26. https://doi.org/10.1080/09588221.2021.2021241

Kemmis, S., McTaggart, R., & Nixon, R. (2014). *The action research planner: Doing critical participatory action research*. Springer.

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, *54*(2), 537–550. https://doi.org/10.1177/00336882231162

Kormos, J. (2014). *Speech production and second language acquisition*. Routledge.

Li, M., & Wilson, J. (2025). AI-integrated scaffolding to enhance agency and creativity in K-12 English language learners: A systematic review. *Information*, *16*(7), 519. https://doi.org/10.3390/info16070519

Liakin, D., Cardoso, W., & Liakina, N. (2017). Mobilizing instruction in a second-language context: Learners' perceptions of two speech technologies. *Languages*, *2*(3), 11. https://doi.org/10.3390/languages2030011

Liu, C.-C., Hwang, G.-J., Yu, P., Tu, Y.-F., & Wang, Y. (2025). Effects of an automated corrective feedback-based peer assessment approach on students' learning achievement, motivation, and self-regulated learning conceptions in foreign language pronunciation. *Educational Technology Research and Development*, *73*(4), 2403–2424. https://doi.org/10.1007/s11423-025-10484-z

McCrocklin, S. (2019). Learners' feedback regarding ASR-based dictation practice for pronunciation learning. *CALICO Journal*, *36*(2), 119–137. https://doi.org/10.1558/cj.34738

McCrocklin, S. M. (2016). Pronunciation learner autonomy: The potential of automatic speech recognition. *System*, *57*, 25–42. https://doi.org/10.1016/j.system.2015.12.013

Nassif, A. B., Shahin, I., Attili, I., Azzeh, M., & Shaalan, K. (2019). Speech recognition using deep neural networks: A systematic review. *IEEE Access*, *7*, 19143–19165. https://doi.org/10.1109/ACCESS.2019.2896880

Nguyen Huu Lam, K., & Tran Thanh, T. (2026). Segmental and suprasegmental pronunciation training via AI: An exploration of university students' perceptions and attitudes. *Technology in Language Teaching & Learning, 8*, 103341. https://doi.org/10.29140/tltl.2026.103341

Nguyen, Q. N., & Doan, D. T. H. (2025). ChatGPT's effects on EFL learners' autonomy and self-efficacy in self-regulated learning contexts. *JALT CALL Journal*, *21*(2), 102968. https://doi.org/10.29140/jaltcall.v21n2.102968

Okonkwo, C. W., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence*, *2*, 100033. https://doi.org/10.1016/j.caeai.2021.100033

Reinders, H., & Benson, P. (2017). Research agenda: Language learning beyond the classroom. *Language Teaching*, *50*(4), 561–578. https://doi.org/10.1017/S0261444817000192

Reinders, H., & White, C. (2016). 20 years of autonomy and technology: How far have we come and where to next? *Language Learning & Technology, 20*(2), 143–154.

Saito, K., & Plonsky, L. (2019). Effects of second language pronunciation teaching revisited: A proposed measurement framework and meta-analysis. *Language Learning*, *69*(3), 652–708. https://doi.org/10.1111/lang.12345

Shafiee Rad, H. (2025). Reinforcing L2 reading comprehension through artificial intelligence intervention: Refining engagement to foster self-regulated learning. *Smart Learning Environments*, *12*(1), 23. https://doi.org/10.1186/s40561-025-00377-2

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Tapingkae, P., Panjaburee, P., Hwang, G.-J., & Srisawasdi, N. (2020). Effects of a formative assessment-based contextual gaming approach on students' digital citizenship behaviours, learning motivations, and perceptions. *Computers & Education*, *159*, 103998. https://doi.org/10.1016/j.compedu.2020.103998

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

Tran, H. (2026). How does AI work as a "game changer"? Self-determination and influencing factors in EFL students' use of AI for academic writing. *Technology in Language Teaching & Learning*, *8*, 103723. https://doi.org/10.29140/tltl.2026.103723

Tsai, S.-C. (2023). Learning with mobile augmented reality- and automatic speech recognition-based materials for English listening and speaking skills: Effectiveness and perceptions of non-English major English as a foreign language students. *Journal of Educational Computing Research*, *61*(2), 444–465. https://doi.org/10.1177/07356331221111203

Uchihara, T., Webb, S., Saito, K., & Trofimovich, P. (2023). Frequency of exposure influences accentedness and comprehensibility in learners' pronunciation of second language words. *Language Learning*, *73*(1), 84–125. https://doi.org/10.1111/lang.12517

Wang, M., Wang, H., & Shi, Y. (2022). The role of English as a foreign language learners' grit and foreign language anxiety in their willingness to communicate: Theoretical perspectives. *Frontiers in Psychology*, *13*, 1002562. https://doi.org/10.3389/fpsyg.2022.1002562

Wu, X., Zhang, Y., & Zhu, W. (2023). Study on an English speaking practice system based on automatic speech recognition technology. *Journal of Education and Educational Research*, *4*(1), 143–146. https://doi.org/10.54097/jeer.v4i1.10273

Xu, J., Luo, Y., Wang, C., Wang, M., & Wu, Y. (2026). AI support in self-regulated learning: A decade of technological evolution and meta-analysis. *British Journal of Educational Technology*, *57*, 1191–1220. https://doi.org/10.1111/bjet.70058

Yang, L. (2026). Empowering the autonomous learner: How AI-assisted language learning environments shape self-regulation, autonomy, and self-directed behaviors. *Language Teaching Research*, *0*(0). https://doi.org/10.1177/13621688261422129

Zai, F., & Zhou, X. (2026). The impact of AI-assisted learning on the agency of foreign language learners: A meta-analysis. *Behavioral Sciences*, *16*(3), 379. https://doi.org/10.3390/bs16030379

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, *16*(1), 39. https://doi.org/10.1186/s41239-019-0171-0

Zhai, C., & Wibowo, S. (2023). A systematic review on artificial intelligence dialogue systems for enhancing English as foreign language students' interactional competence in the university. *Computers and Education: Artificial Intelligence*, *4*, 100134. https://doi.org/10.1016/j.caeai.2023.100134

Zhang, J., & Derakhshan, A. (2025). Integrating 'GPT-4o' as an AI tool into K-12 contexts: Assessing its long-term effects on students' self-regulated learning (SRL) and task engagement through latent growth curve modeling. *European Journal of Education*, *60*(4), e70329. https://doi.org/10.1111/ejed.70329

Zhang, Z. (2025). The role of artificial intelligence tools on Chinese EFL learners' self-regulation, resilience, and autonomy. *European Journal of Education*, *60*(2), e70127. https://doi.org/10.1111/ejed.70127

Zhang, Z., Zhou, S., & Bai, B. (2025). A systematic scoping review of self-regulated learning with AI in language education. *International Journal of Applied Linguistics*, *0,* 1–15. https://doi.org/10.1111/ijal.70030

Zimmerman, B. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, *41*, 64–70. https://doi.org/10.1207/s15430421tip4102_2

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