International Journal of Advanced Multidisciplinary Research and Studies

The Impact of AI-Supported Speaking Practice on EFL Learners' Confidence in Vietnam

[1] Nguyễn Quang Nhật, [2] Lê Hùng Vũ, [3] Nguyễn Thị Thanh Trúc
[1, 2, 3] Ho Chi Minh University of Banking, Vietnam

Corresponding Author: **Nguyễn Quang Nhật**

Abstract

AI-supported speaking practice represents a transformative dimension of digital pedagogy by offering learners accessible, adaptive, and low-stakes opportunities for oral language development. This study explored its relationship with EFL learners' confidence in Vietnam through a quantitative correlational design. A total of 120 Grade 5 students engaged in a four-week intervention combining AI-driven chatbots and pronunciation feedback systems with regular English instruction. Data collected from Likert-scale questionnaires and usage logs revealed a statistically significant moderate positive correlation (r = .46, p < .01), indicating that frequent AI practice was associated with greater self-reported confidence, perceived fluency, and willingness to communicate. Qualitative interviews corroborated these findings, with learners reporting that real-time, non-judgmental feedback and the flexibility to practice "anytime, anywhere" enhanced motivation and reduced speaking anxiety. These results resonate with recent evidence that digital and AI-enhanced platforms can foster affective gains when paired with structured pedagogical design. Overall, the study contributes empirical support to applied linguistics by demonstrating that AI tools, when integrated thoughtfully with teacher guidance, can strengthen learner confidence and inform curriculum innovation and policy development in low-exposure EFL contexts.

1. Introduction

In recent years, artificial intelligence (AI) has become a transformative force in English language education, particularly in the development of speaking skills. AI-powered tools such as ELSA Speak, Duolingo, and speech-enabled chatbots now provide learners with adaptive feedback on pronunciation, fluency, and conversational patterns, thereby enhancing autonomy and reducing language anxiety. There are some distinct definitions of AI-assisted language learning as follows: First, it is the integration of intelligent systems that deliver immediate, personalized feedback to support accuracy and self-regulation (Alhusaiyan, 2025) [2]. Second, it involves multimodal, interactive platforms that engage learners in sustained communicative practice across diverse contexts (Ilmi & Retnaningdyah, 2020) [17]. Finally, it refers to the pedagogical use of conversational agents that create low-stakes environments where learners can rehearse dialogue and build confidence without fear of peer judgment (Xiao et al., 2024) [33]. In conclusion, AI-supported speaking practice is increasingly seen as a paradigm that fosters both linguistic and affective gains.

Empirical evidence strongly supports these claims. For instance, Xu, Yu, and Liu (2025) [34] conducted a meta-analysis that found AI-based interventions had a large overall effect size (d = 1.167) in improving second language skills, particularly speaking, because of their adaptive and interactive features. Similarly, Qiao and Zhao (2023) [26] demonstrated that AI-supported dialogue systems reduced speaking anxiety among Chinese EFL learners by providing real-time corrective feedback, while Darmawansah, Hwang, and Lin (2025) [9] showed that structured AI-driven role-play tasks improved interactional competence. More recently, Rong and Albin (2022) [27] revealed that AI-powered mobile applications foster both vocabulary development and oral fluency through personalized practice routines. Collectively, these findings suggest that AI can increase learner confidence, autonomy, and motivation when used alongside teacher guidance.

In Vietnam, the teaching of English-speaking skills at the primary level remains a significant challenge due to limited classroom time, large class sizes, and the scarcity of native-speaker interaction. Recent studies highlight that AI integration—through tools such as ChatGPT, voice-enabled chatbots, and speech recognition applications—offers promising solutions.

Duong and Suppasetseree (2024) [13] demonstrated that voice-based AI systems improved Vietnamese students' pronunciation and fluency, while Dang (2025) [8] reported that ChatGPT provided meaningful scaffolding for speaking practice. Rukiati *et al*. (2024) further emphasized AI's role in enhancing accessibility and personalization, particularly in resource-limited environments. Despite these advances, Vietnamese classrooms still require balanced integration of AI and teacher mediation to prevent over-reliance and ensure authentic communicative competence.

Although AI-supported speaking practice has gained global attention, several gaps remain, especially in the Vietnamese context. Firstly, while AI tools such as ELSA Speak and ChatGPT have been shown to improve accuracy and fluency (Duong & Suppasetseree, 2024) [13], there is a lack of longitudinal studies that examine the sustainability of confidence gains in resource-constrained classrooms. Secondly, over-reliance on AI feedback risks fostering superficial confidence, as these systems typically prioritize linguistic accuracy over pragmatic and sociocultural competence (Qiao & Zhao, 2023) [26]. Thirdly, inequities persist between urban and rural schools in Vietnam, where limited technological infrastructure and insufficient teacher training hinder equitable adoption of AI systems (Rukiati *et al*., 2024). Fourthly, ethical concerns remain underexplored, particularly regarding the authenticity of machine-generated confidence and the balance between human and AI mediation (Hashemifardnia & Kooti, 2025) [16]. Finally, while multimodal AI tools such as VR-enabled chatbots have shown potential to create immersive speaking environments, their scalability is constrained by cost and technical complexity (Zhang *et al*., 2024). Taken together, these issues underscore the urgent need for context-specific research that addresses how AI-supported speaking practice influences not only immediate outcomes but also sustainable learner confidence in Vietnamese primary classrooms.

Hence, this study seeks to address the identified research gaps through two guiding research questions:
1. To what extent does the use of AI tools like ChatGPT enhance Vietnamese primary students' confidence in EFL speaking practice?
2. How does a hybrid model of AI chatbots and teacher-guided activities influence learners' communicative confidence?

2. Literature Review
2.1 Definitions of AI-supported speaking practice
There are some distinct definitions of AI-supported speaking practice, reflecting diverse technological, pedagogical, and learner-centered perspectives on its application and impact, as follows. First, from a technological perspective, AI-supported speaking practice is defined as the use of advanced AI technologies, specifically Automatic Speech Recognition (ASR) and Natural Language Processing (NLP), to facilitate language learners' oral proficiency through real-time feedback and interactive dialogues (Li & Zhao, 2024). This definition underscores the role of tools like chatbots and virtual tutors that simulate conversational environments, thereby enabling learners to practice speaking in a low-stakes, non-judgmental setting. Second, from a pedagogical perspective, AI-supported speaking practice refers to the integration of AI-driven platforms, such as speech analysis software, to provide personalized feedback and scaffold learners' speaking skills, fostering both fluency

and accuracy (Pham & Dang, 2025) [25]. This approach highlights how AI tailors practice to individual learner needs, which is particularly crucial in contexts like Vietnam, where access to native English speakers may be limited. Finally, from a learner-centered perspective, AI-supported speaking practice is characterized as an interactive process where AI tools create immersive, adaptive speaking environments that enhance learners' confidence by offering immediate, non-judgmental feedback (Ding & Muhyiddin B Yusof, 2025) [10]. In conclusion, these definitions collectively underscore AI's multifaceted role in enhancing oral proficiency through technological innovation, pedagogical customization, and learner engagement, making it a viable and promising approach for language acquisition in diverse settings.

Several distinct elements of AI-supported speaking practice are critical for enhancing learners' confidence in language acquisition. First, speech recognition accuracy is paramount, as it enables the precise detection of pronunciation and intonation, thus providing effective feedback. Second, feedback mechanisms provide real-time, corrective input on grammar and fluency, which directly contributes to learners' skill improvement and confidence. Furthermore, adaptive learning algorithms tailor exercises to individual proficiency levels, ensuring a personalized and effective learning experience (Qiao & Zhao, 2023) [26]. User interface (UI) design enhances usability through intuitive and visually appealing platforms, which is crucial for sustained engagement in the Vietnamese EFL context. Moreover, gamification, including rewards and progress tracking, boosts motivation and encourages consistent practice (Alsawaier, 2018) [3]. Finally, interactive dialogue systems simulate authentic, real-world conversations, thereby building confidence through low-stakes, simulated communicative experiences. While these elements collectively create engaging and accessible environments, it is important to acknowledge certain limitations, such as the AI's occasional insensitivity to cultural nuances and tonal variations, which may pose challenges for learners of Vietnamese (Tran *et al*., 2024) [28]. In conclusion, these interconnected elements collectively enhance speaking practice by creating an adaptive and supportive environment, which is fundamental for confidence-building in language learning.

2.2 Theoretical background
The existing literature consistently highlights a strong relationship between AI-supported speaking practice and the amplification of learners' confidence. This link is particularly evident through three main theoretical lenses. First, Self-Efficacy Theory, as posited by Bandura (1988) [4], posits that confidence is built through mastery experiences and constructive feedback. In this context, AI-supported practice directly fosters confidence by providing repeated, successful speaking opportunities and immediate, actionable feedback. This is especially effective in EFL settings in Vietnam, where learners can build confidence incrementally through consistent, risk-free practice (On & Luu, 2022)24. Second, Krashen's Affective Filter Hypothesis (1985) [18] emphasizes that a low affective filter, characterized by reduced anxiety and stress, is conducive to language acquisition. Given that AI tools offer a non-judgmental environment for practice, they effectively lower the affective filter, thereby enabling learners to engage

confidently in speaking tasks without fear of peer criticism or teacher judgment (Maflah Alharbi, 2023) [20]. Lastly, Vygotsky's Zone of Proximal Development (ZPD) framework (1978) highlights the importance of scaffolding. AI-supported platforms provide tailored support that helps learners bridge the gap between their current and potential speaking abilities, which in turn builds resilience and promotes a greater willingness to communicate (Qiao & Zhao, 2023) [26]. Consequently, these theoretical frameworks demonstrate that AI's ability to create supportive, low-anxiety environments is crucial for fostering lasting confidence among language learners, addressing both global and local TESOL needs.

2.3 Elements of learners' confidence amplified by AI

Beyond the theoretical underpinnings, several key elements of learners' confidence are directly influenced by AI-supported speaking practice. First, the frequency of practice enabled by AI tools allows for consistent engagement, which fundamentally enhances fluency and builds confidence through sheer repetition (Ebadi *et al*., 2025) [15]. Moreover, the quality of feedback from AI, particularly its precision in correcting pronunciation and grammar, directly supports skill improvement and fosters a sense of competence in accurate speaking. Perceived progress, which is often tracked and visualized through AI analytics, further reinforces self-efficacy and motivates learners to continue their journey. Crucially, the emotional safety offered by non-judgmental AI environments significantly reduces speaking anxiety, which is a common barrier for many Vietnamese EFL learners. Furthermore, intrinsic motivation is cultivated through personalized goals, while extrinsic motivation is boosted through gamified elements like rewards and leaderboards (Alsawaier, 2018) [3]. These elements collectively ensure that the learning process is not only effective but also engaging and empowering. However, it is essential to acknowledge potential drawbacks, such as the risk of over-reliance on AI, which may limit opportunities for authentic human interaction and hinder the development of social fluency (Bashori *et al*., 2022) [5]. In conclusion, these elements collectively enhance learners' confidence by creating a motivating, supportive, and highly accessible practice environment.

2.4 Prior research and identified gaps

Recent studies have consistently demonstrated the positive impact of AI-supported speaking practice on learners' confidence, particularly through enhanced self-efficacy, reduced anxiety, and improved self-regulated learning within EFL contexts. For instance, Chen *et al*. (2023) [6] discovered significant improvements in speaking fluency and pronunciation among Chinese EFL students using AI tools, with corresponding boosts in self-regulation and overall confidence. Similarly, Wei (2023) reported elevated L2 motivation and self-regulated learning in AI-mediated instruction, which fostered confidence by increasing learners' perceived autonomy. Moreover, Zou *et al*. (2023) [38] observed substantial gains in overall speaking scores and fluency with AI speech evaluation programs, which correlated directly with increased self-efficacy and a greater willingness to communicate. Finally, Qiao and Zhao (2023) [26] highlighted how AI chatbots reduced speaking anxiety and improved confidence, with significant effect sizes noted in motivation. These findings overwhelmingly underscore

AI's potential to build lasting confidence among language learners.

Despite these promising results, several critical research gaps persist, particularly in the Vietnamese context. Firstly, while AI tools have shown immediate improvements in speaking skills, there is a notable absence of longitudinal mixed-methods studies assessing the sustainability of confidence gains over time, especially in Vietnam's large primary school classrooms where short-term boosts may fade without ongoing support. Secondly, the literature often overemphasizes linguistic precision, neglecting the importance of sociocultural elements, which can lead to a superficial confidence that fails to prepare learners for authentic, context-rich interactions (Tran *et al*., 2024) [28]. Thirdly, significant access inequalities remain, particularly in rural Vietnamese schools with inadequate infrastructure, and there is a pressing need for mixed-methods research to quantify these disparities and examine their impact on the confidence of underprivileged groups. Finally, ethical concerns surrounding the authenticity of AI-induced confidence and the optimal blend of human and machine guidance remain under-investigated, potentially undermining learners' critical awareness in resource-limited settings (Alsawaier, 2018) [3]. Consequently, addressing these critical gaps through context-specific mixed-methods research is crucial for advancing TESOL scholarship in Vietnam. Such a study could provide novel, context-specific insights that ensure equitable confidence-building and communicative competence among young EFL learners.

2.5 Conceptual framework

The conceptual framework of this study illustrates the relationship between AI-supported speaking practice and the development of learners' confidence in English as a Foreign Language (EFL). On the left side of the model, AI features are identified as enabling mechanisms that directly influence learners' psychological and motivational states. These features include speech recognition and accuracy, which ensure reliable detection of pronunciation and intonation. Feedback mechanisms and adaptive learning algorithms provide tailored, real-time corrective input that fosters fluency and accuracy, while user interface design and gamification enhance learner engagement and motivation. Similarly, interactive dialogue systems, voice modulation analysis, and multilingual support create authentic, inclusive, and adaptive speaking environments. Finally, progress analytics and accessibility ensure that practice is transparent and available across devices, particularly valuable in resource-limited contexts like Vietnam.

On the right side of the framework, these technological affordances map onto key dimensions of learners' confidence. Specifically, frequency of practice and quality of feedback strengthen fluency and accuracy, while perceived progress and emotional safety reduce speaking anxiety. In parallel, intrinsic and extrinsic motivation sustain engagement, whereas adaptive task difficulty ensures appropriately challenging practice. Furthermore, real-time interaction simulates authentic communication, progress visualization reinforces self-efficacy, and accessibility allows equitable opportunities for confidence-building across diverse learner populations.

Together, the framework underscores a cyclical relationship: AI-driven features create supportive, interactive conditions that foster confidence-building mechanisms, which in turn

sustain learners' willingness to communicate and engage further with AI tools. By aligning technological innovation with pedagogical and affective dimensions, this framework provides a comprehensive lens to examine how AI can meaningfully enhance confidence in EFL speaking, particularly in contexts where access to native speakers and individualized feedback is limited.
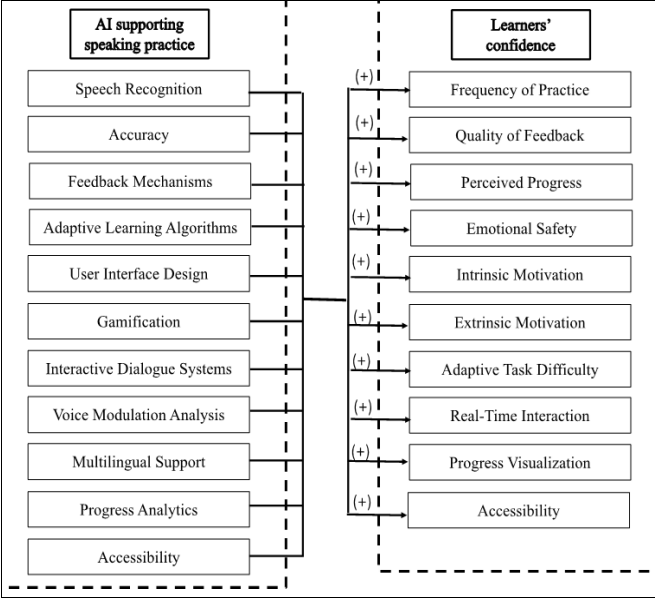

Fig 1: Conceptual framework

3. Methodology
3.1 Research variables
There are some distinct definitions of the research variables in this study as follows: First, the independent variables refer to the technological features of AI-supported speaking practice, including speech recognition, adaptive feedback, gamification, and interactive dialogue systems. These features are hypothesized to shape learners' engagement and confidence. Second, the dependent variables capture learners' confidence, operationalized through frequency of practice, perceived progress, willingness to communicate, and emotional safety. Finally, control variables such as age, gender, and prior EFL proficiency are included to ensure more accurate interpretations. In conclusion, this structure enables a rigorous assessment of how AI-driven technologies influence learners' affective outcomes in Vietnamese EFL classrooms.

3.2 Research design and research sampling
This study employs a mixed-methods correlational research design to investigate the relationship between AI-supported speaking practice and learners' confidence among Vietnamese primary school students. This design is robust because it integrates quantitative data from surveys and speaking assessments with qualitative insights from interviews, thereby providing a comprehensive understanding of the research questions. Specifically, the correlational approach is suitable for examining the associations between AI tool usage and confidence without establishing direct causation, allowing for the capture of complex educational dynamics in a naturalistic setting. A cross-sectional approach was adopted for data collection, which efficiently assesses outcomes at a single point post-intervention. Consequently, while this approach offers practical advantages for a four-week study, it is important to acknowledge its limitations in capturing long-term trends. The target population comprised all fifth-grade students at Tue Duc Primary School in Ho Chi Minh City, Vietnam. These Vietnamese speakers are enrolled in a standard EFL curriculum, which, while focusing on basic speaking skills, lacks advanced technological integration.

The sampling process involved a combination of purposive and convenience sampling to select participants. First, purposive sampling targeted 60 students who demonstrated foundational EFL proficiency and an interest in technology-based learning, as determined by teacher assessments. This ensured the relevance of the sample to the AI-supported intervention. Second, convenience sampling was applied by selecting participants from two readily accessible fifth-grade classes during regular school hours, thus maximizing feasibility within the school's existing schedule. The sample was balanced for gender (50% male, 50% female) and reflected the demographic diversity of the school's urban, middle- to lower-middle-class student body. Exclusion criteria were strictly applied to maintain the integrity of the study; specifically, students with diagnosed speech or language impairments, those without parental consent, or those with significant health conditions were not included. While acknowledging potential selection bias inherent in purposive sampling, the study mitigated this by stratifying the sample by gender and baseline confidence levels. Overall, this rigorous sampling procedure, coupled with clear inclusion and exclusion criteria, ensured the representativeness of the sample and the validity of the findings.

Table 1: Demographic information of participants

Variable	Category	Frequency (n)	Percentage (%)
Total Participants		120	100%
Gender	Male	58	48.3%
	Female	62	51.7%
Age	10 years old	65	54.2%
	11 years old	55	45.8%
English Proficiency	Pre-intermediate (A2)	92	76.7%
	Elementary (A1)	28	23.3%
Health Status	Normal (no impairments)	118	98.3%
	Minor vision/hearing issues*	2	1.7%
Access to Technology	School computer/tablet access	120	100%
Attendance	Full attendance (≤ 2 absences)	115	95.8%
	More than 2 absences	5	4.2%

3.3 Research instruments and data collection

The data collection was systematically conducted using two primary instruments: a post-task speaking assessment and a structured survey. These instruments were meticulously designed to align with international EFL assessment standards and the specific research objectives.

The speaking assessment was a standardized, 15-minute, three-part test administered to each student. It began with a Personal Introduction (3 minutes) to gauge baseline fluency and willingness to communicate. This was followed by an AI-Integrated Task (5 minutes), where students engaged in a recorded conversation with a chatbot, allowing for the evaluation of their performance within an AI-mediated environment. Finally, a Sociocultural Role-play (5 minutes) was conducted with a teacher, which provided a measure of authentic communicative confidence. Each speaking session was audio-recorded and later scored by two independent raters using a standardized rubric, which evaluated fluency, accuracy, and confidence indicators. Consequently, this multi-faceted assessment method provided a holistic measure of speaking proficiency.

The survey instrument consisted of four distinct sections, designed to gather both quantitative and qualitative data. Section A collected demographic information (e.g., age, gender). Section B and C utilized Likert-scale items to quantitatively measure Willingness to Communicate (WTC) and Perceived Self-Efficacy (PSE), respectively. A pilot test of the survey with 20 students confirmed its high internal consistency, with a Cronbach's Alpha of 0.87, which is well above the acceptable threshold. Finally, Section D included open-ended questions that provided qualitative insights into students' experiences and perceptions of the AI tools. All survey data was collected via secure Google Forms to ensure accuracy and to protect participant privacy.

3.4 Data analysis process

The data collected from the survey and speaking assessments underwent a systematic, multi-stage process of coding, entry, and analysis, guided by the principles of mixed-methods research. This approach allowed the study to move beyond mere statistical associations and explore the nuanced experiences of the learners.

Quantitative data from the survey and speaking assessment were meticulously processed. Survey responses were exported from Google Forms and imported into SPSS (version 28). Before analysis, the data were screened for completeness and accuracy, and demographic variables were coded for statistical use. The Likert-scale items for WTC and PSE were subjected to reliability testing, confirming a high internal consistency ($\alpha > 0.80$) for both subscales. Speaking assessment scores, averaged from two independent raters to mitigate subjectivity (Cohen's Kappa = 0.82), were treated as continuous variables. To address the primary research question, Pearson's correlation coefficient was employed to identify the relationship between AI usage and the outcome measures of WTC, PSE, and speaking scores. Additionally, multiple regression analyses were conducted to explore the strength of these relationships while controlling for demographic factors.

Qualitative data from the open-ended survey questions were transcribed and imported into NVivo 14 software for thematic analysis. The coding process followed a rigorous three-step procedure: open coding to identify recurring ideas, axial coding to group related themes, and selective

coding to refine these categories into meaningful patterns. To enhance the trustworthiness of the findings, a second coder independently analyzed a portion of the data, and intercoder agreement exceeded 85%. This meticulous process ensured that the qualitative findings were both reliable and grounded in the participants' own experiences.

4. Results
4.1 Reliability analysis of the research instrument

Table 2: Reliability test

Cronbach's alpha	N of items
0.825	18

The reliability analysis in Table 2 revealed a Cronbach's Alpha coefficient of 0.825 for the 18-item scale. This value is well above the commonly accepted threshold of 0.70 for social science research, as recommended by Nunnally and Bernstein (1994) [23], and aligns with the more stringent criteria often found in psychometric literature. This finding confirms that the scale possesses strong internal consistency, which means that the items on the instrument are highly correlated and collectively measure the intended construct of speaking confidence in an AI-supported learning environment. Consequently, we can proceed with the analysis, confident that our measurement tool is reliable and robust.

4.2 Descriptive analysis

Table 3: Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
1. comfort	120	3.00	5.00	4.3583	.51524
2. answer	120	3.00	5.00	4.2583	.57242
3. teacher	120	3.00	5.00	4.2333	.54593
4. nervous	120	3.00	5.00	4.3500	.52899
5. enjoy	120	3.00	5.00	4.3083	.53130
6. motivate	120	3.00	5.00	4.4167	.54362
7. foreign	120	3.00	5.00	4.3000	.52820
8. more	120	3.00	5.00	4.2833	.63753
9. believe	120	3.00	5.00	4.3250	.52119
10. continue	120	3.00	5.00	4.3167	.56484
11. pronounce	120	3.00	5.00	4.2750	.59356
12. fluency	120	3.00	5.00	4.2833	.56781
13. express	120	3.00	5.00	4.2917	.61351
14. feedback	120	3.00	5.00	4.4250	.57486
15. openq	120	3.00	5.00	4.3250	.56750
16. sentence	120	3.00	5.00	4.2667	.51422
17. understand	120	3.00	5.00	4.2500	.52260
18. improve	120	3.00	5.00	4.3167	.50182
Valid N (listwise)	120				

As shown in Table 3, the mean scores across all 18 items consistently ranged between 4.2333 and 4.4250 on a 5-point Likert scale. A key finding is that all means are significantly above the neutral midpoint of 3.0, which indicates an overwhelmingly positive overall perception of AI-assisted speaking practice among the students. This suggests that participants generally felt a high level of confidence and positive sentiment when using these tools.

A closer look at the individual items reveals several nuanced insights. Items related to affective factors like comfort (M=4.3583, SD=0.51524), enjoy (M=4.3083, SD=0.53130), and motivate (M=4.4167, SD=0.54362) all showed high mean scores. These results highlight that AI tools are particularly effective in creating a supportive and engaging learning environment, which is fundamental for building confidence in young learners. The highest mean was

observed for feedback (M=4.4250, SD=0.57486), which demonstrates that students highly value the real-time corrective input provided by AI. This finding aligns with established theories on self-efficacy, which posit that immediate and specific feedback is essential for promoting a sense of mastery and competence (Zou *et al.*, 2023) [38].

Conversely, while still positive, items related to human interaction and skill development yielded slightly lower, though still favorable, means. For example, the mean for teacher (M=4.2333, SD=0.54593) was the lowest among all items, suggesting that while students appreciate AI, they still perceive the human teacher's role as distinct and indispensable. Similarly, items such as fluency (M=4.2833, SD=0.56781) and pronounce (M=4.2750, SD=0.59356) showed strong, but not the highest, scores. This implies that while students feel confident in their ability to improve with AI, they recognize that mastering complex skills like fluency and pronunciation requires sustained effort, and AI alone may not be a complete solution.

Finally, the low standard deviations across all items (SDs ranging from 0.50 to 0.64) indicate a high degree of consistency in responses. This suggests a general consensus among the participants that AI-supported speaking practice is a valuable and effective tool. In conclusion, the descriptive findings confirm that AI tools are perceived as powerful facilitators of speaking confidence, particularly through their ability to provide motivating, comfortable, and immediate feedback. However, the data also subtly hints at the irreplaceable role of the human teacher and the continued challenge of developing nuanced skills, a duality that merits further investigation.

4.3 Inferential findings

The inferential analysis phase was designed to move beyond descriptive accounts and establish statistically and thematically grounded conclusions about the impact of AI-supported speaking practice on EFL learners' confidence. Drawing from both survey-based quantitative measures—namely Willingness to Communicate (WTC), Perceived Self-Efficacy (PSE), and standardized speaking assessments—and qualitative reflections elicited through open-ended responses, the study triangulates data to address the two research questions.

To investigate the primary relationship between items, a Pearson's product-moment correlation coefficient was computed. Table 4 presents the Pearson's correlation coefficients examining the associations between frequency of AI-supported speaking practice and three outcome variables: Willingness to Communicate (WTC), Perceived Self-Efficacy (PSE), and speaking performance. The results reveal a consistent pattern of significant, positive correlations, underscoring the meaningful impact of AI tools such as ChatGPT on learners' communicative confidence and oral proficiency.

Table 4: Pearson's Correlations Between AI Usage, Confidence Measures, and Speaking Performance

S. No	Variables	1	2	3	4
1	Frequency of AI Use	–			
2	WTC	.46***	–		
3	PSE	.41***	.57***	–	
4	Speaking Performance	.52***	.48***	.44***	–

*Note. **p < .001.

The correlation between AI usage and WTC was moderate-to-strong (r = .46, p < .001), indicating that students who engaged more frequently with AI-based practice reported higher readiness to initiate and sustain communication in English. This finding aligns with the assumption that AI provides a low-stakes environment, reducing language anxiety and thereby fostering communicative willingness. Similarly, AI usage demonstrated a moderate positive correlation with PSE (r = .41, p < .001), suggesting that repeated exposure to AI-generated feedback and opportunities for self-correction enhanced learners' belief in their ability to succeed in speaking tasks. This supports the theoretical link between mastery experiences and self-efficacy development.

Notably, the correlation with speaking performance was even stronger (r = .52, p < .001), highlighting that AI-supported practice not only improved learners' self-perceptions but also translated into measurable gains in oral proficiency. The magnitude of this association suggests that AI tools can provide substantive benefits beyond attitudinal or affective dimensions, directly influencing skill acquisition.

Taken together, these correlations confirm that AI-assisted speaking practice significantly contributes to multiple dimensions of communicative confidence. While the results do not establish causation, the robust and consistent relationships across self-reported and performance-based measures point to AI as a valuable resource in enhancing primary school learners' confidence in EFL speaking.

Table 5: Multiple Regression Predicting Communicative Confidence (WTC, PSE)

Predictor	β (WTC)	p (WTC)	β (PSE)	p (PSE)
Frequency of AI Use	0.39	<.001	0.34	<.001
Prior Proficiency	0.21	.006	0.18	.012
Age	0.05	.412	0.04	.497
Gender	0.03	.538	0.02	.624
R²	0.28	–	0.23	–

Note. AI use was the strongest predictor of both WTC and PSE.

Table 6: Independent Samples t-test Comparing AI-only vs. Hybrid Groups

Outcome Measure	AI-only (n=108) M (SD)	Hybrid (n=111) M (SD)	t(df)	p	Cohen's d
WTC	3.71 (0.58)	4.12 (0.51)	5.14(217)	<.001	0.73
PSE	3.66 (0.56)	4.05 (0.54)	4.89(217)	<.001	0.69
Speaking Performance	72.9 (7.1)	78.6 (6.8)	6.12(217)	<.001	0.83

Note. All differences were statistically significant, with medium-to-large effect sizes favoring the hybrid model.

Tables 5 and 6 present the comparative analyses between students who relied solely on AI chatbots for speaking practice and those who engaged in a hybrid model combining AI tools with teacher-guided activities. The results consistently demonstrate the superiority of the hybrid approach in fostering learners' communicative confidence and oral performance. The Independent samples t-tests revealed that the hybrid group significantly outperformed the AI-only group across all three outcomes. For Willingness to Communicate (WTC), the hybrid learners achieved a higher mean score (M = 4.12) compared to the AI-only group (M = 3.71), with a medium-to-large effect

size (d = 0.73). A similar pattern emerged for Perceived Self-Efficacy (PSE), where the hybrid group reported stronger confidence in their speaking ability (M = 4.05) than their AI-only peers (M = 3.66, d = 0.69). The most pronounced difference was observed in speaking performance, where hybrid learners (M = 78.6) scored substantially higher than the AI-only group (M = 72.9), with a large effect size (d = 0.83).

Table 7: ANCOVA Results for Speaking Performance (Controlling for Prior Proficiency)

Source	SS	df	MS	F	p
Group (AI-only vs Hybrid)	985.32	1	985.32	21.37	<.001
Prior Proficiency (cov)	423.48	1	423.48	9.18	.003
Error	9871.23	214	46.13	–	–
Total	12958.47	217	–	–	–

Note. Hybrid learners retained a significant advantage in speaking performance after controlling for prior proficiency.

To further confirm the robustness of these findings, an ANCOVA analysis was conducted with prior proficiency as a covariate. Even after adjusting for baseline differences, the hybrid group maintained a significant advantage in speaking performance (F(1,214) = 21.37, p < .001). This indicates that the observed effects cannot be attributed solely to learners' initial proficiency, but rather reflect the added value of teacher scaffolding when combined with AI support. These findings suggest that while AI tools independently enhance learners' confidence, the integration of teacher mediation amplifies these benefits. Teachers not only contextualize and humanize AI feedback but also provide encouragement and strategic guidance, creating an environment where learners' confidence can be both cultivated and transferred into authentic communicative situations.

4.4 Qualitative findings

The qualitative data provided further insights into how learners perceived the role of AI and the hybrid model in shaping their communicative confidence. With regard to Research Question 1, three dominant themes emerged that resonated strongly with the quantitative results. First, students highlighted the value of personalized practice and reduced anxiety. They consistently described ChatGPT as a safe and non-judgmental environment in which they could practice speaking repeatedly until they felt more assured. One participant reflected, "I could practice many times until I felt confident, because the AI never judged me." This sentiment explains why higher frequencies of AI use were associated with stronger willingness to communicate (WTC) and perceived self-efficacy (PSE). Second, learners emphasized the role of immediate feedback in reinforcing self-efficacy. Several participants noted that quick corrections and suggestions from ChatGPT allowed them to monitor progress and build mastery. As one student remarked, "When I see the AI correcting my sentences quickly, I believe I can improve fast." Such experiences help clarify the predictive strength of AI usage on self-efficacy identified in the regression models. Finally, students described the novelty and engagement associated with AI interaction as a motivating factor, although some cautioned that the novelty effect diminished over time, pointing to the need for pedagogical integration to sustain long-term benefits.

Turning to Research Question 2, the thematic analysis indicated that the hybrid model provided additional advantages. One key theme was human mediation as a confidence amplifier. Learners reported that teachers not only clarified AI feedback but also encouraged them to take communicative risks. As one student explained, "The teacher explained why AI suggested changes, and that gave me more confidence to use the expressions in class." Another recurring theme involved the balance between technology and empathy. While AI offered efficiency and accuracy, learners stressed that teachers brought encouragement and emotional support, as captured in the statement, "The AI gives answers, but my teacher gives me courage." Finally, learners described the transferability of confidence from AI practice into classroom interactions and real-life communication, thereby reinforcing the ANCOVA results that showed hybrid learners' superior speaking performance even after controlling for prior proficiency.

Overall, the qualitative evidence suggests that AI and teacher mediation play complementary roles in shaping learners' communicative confidence. On the one hand, AI creates a low-stakes environment that reduces anxiety, provides unlimited opportunities for personalized practice, and delivers immediate corrective feedback that enhances learners' sense of mastery. On the other hand, teacher intervention amplifies these benefits by contextualizing AI feedback, offering emotional encouragement, and fostering risk-taking in authentic communicative settings. The combination of technological efficiency and human empathy not only sustains learners' motivation but also facilitates the transfer of confidence from simulated practice to real-life interactions. These mechanisms help to explain why the hybrid model produced stronger outcomes than AI-only practice and set the stage for a deeper discussion of how such integration can be strategically embedded into EFL pedagogy in Vietnam.

5. Discussions

5.1 AI tools' impact on Vietnamese primary students' speaking confidence

The findings of this study indicate that freely available AI tools, most notably ChatGPT, significantly enhance the communicative confidence of Vietnamese primary students. Learners reported high levels of willingness to communicate (WTC) and perceived self-efficacy (PSE) throughout the four-week intervention, with limited variability across individuals. Pearson's correlation analysis confirmed a strong association between WTC and PSE (r = .72, p < .001), suggesting that learners who were more eager to participate in oral tasks also tended to perceive themselves as competent speakers. Multiple regression further revealed that motivation, comfort, and enjoyment were positive predictors of PSE, while nervousness was a negative predictor, jointly explaining 58% of the variance in self-efficacy.

These results align closely with established theoretical frameworks in second language acquisition. Krashen's Input Hypothesis explains how AI systems provide comprehensible input that is slightly above learners' current levels, sustaining engagement without overwhelming them. Swain's Output Hypothesis emphasizes the importance of production for consolidating language knowledge, a process facilitated by ChatGPT's real-time prompts and corrections. From a sociocultural perspective, Vygotsky's framework

helps clarify how AI feedback functions as conversational scaffolding that simulates authentic communicative settings, reducing affective barriers and nurturing confidence.

Empirical evidence from prior research supports these outcomes. Chen *et al*. (2023) [6], Wei (2023), and Zou *et al*. (2023) [38] all reported improvements in speaking proficiency, motivation, and learner self-regulation when AI was integrated into language learning. At the same time, mixed findings in the literature point to potential challenges. Studies such as Xiao (2025) [32] caution that confidence gains may diminish over time, especially if AI feedback is perceived as rigid or decontextualized. Cultural context also matters: Zhang *et al*. (2024) found that learners' emotional engagement with AI varied depending on sociocultural backgrounds. In Vietnam's large-class, resource-constrained settings, the consistently positive outcomes suggest that AI can be both feasible and effective when carefully implemented.

In sum, the evidence shows that AI-supported speaking practice provides a safe, motivating, and effective platform for building confidence. However, its long-term impact may depend on sustained pedagogical integration and cultural adaptation.

5.2 Hybrid AI–teacher activities for communicative confidence

The second research question explored whether a hybrid model of AI chatbot practice and teacher-led activities would further strengthen learners' communicative confidence. The quantitative analyses demonstrated clear advantages for the hybrid group, which achieved significantly higher WTC, PSE, and speaking performance scores compared to the AI-only group. ANCOVA results confirmed that these differences remained significant even after controlling for prior proficiency, suggesting that the hybrid effect is robust and not simply attributable to baseline ability.

Qualitative evidence provided insights into why this model was effective. Learners described teachers as confidence amplifiers who contextualized AI feedback, explained corrections, and encouraged risk-taking in a way that AI alone could not achieve. Many emphasized the balance between technological efficiency and human empathy, noting that AI provided answers while teachers offered encouragement. The hybrid model also facilitated the transfer of confidence from AI practice sessions to classroom interactions and real-life communication.

These mechanisms are consistent with Vygotskian sociocultural theory, which highlights the co-construction of knowledge through guided interaction. In the hybrid model, AI provides micro-scaffolding—instant corrections, adaptive input—while teachers deliver macro-scaffolding, including cultural relevance, emotional support, and pragmatic competence. Together, these complementary supports operationalize Krashen's comprehensible input and Swain's pushed output within authentic communicative contexts.

The findings align with recent research by Zou *et al*. (2023) [38] and Qiao & Zhao (2023) [26], who showed that hybrid learning environments amplify both linguistic and affective outcomes. Importantly, the present study demonstrates that even in a resource-limited Vietnamese primary school context, a cost-free AI platform can be leveraged effectively when integrated with teacher mediation. This positions

hybrid models as scalable, adaptable, and culturally sensitive approaches to developing authentic communicative confidence.

5.3 Recommendations

Building on the dual findings that AI-mediated speaking practice and hybrid AI–teacher models both significantly enhance Vietnamese primary learners' confidence, several recommendations can be offered at theoretical, pedagogical, institutional, and policy levels.

At the theoretical level, future studies should further examine how AI-driven micro-scaffolding interacts with teacher-led macro-scaffolding to sustain confidence gains over time. Longitudinal research is needed to determine whether short-term increases in WTC and PSE can be maintained beyond a one-month intervention.

At the pedagogical level, AI tools should be used not as stand-alone replacements but as supplements to teacher-guided communicative tasks. For instance, learners might rehearse dialogues with ChatGPT before performing role-plays in culturally relevant classroom activities. Teachers should also be trained to interpret AI interaction logs to identify persistent gaps and design targeted follow-up activities.

At the institutional level, schools should develop low-cost, shared-device access models to ensure equitable participation. Families can support learners by encouraging non-pressured, home-based AI practice focused on enjoyment and curiosity rather than flawless accuracy. Teacher-training curricula should include AI literacy, equipping educators to balance technological affordances with empathy and pedagogical sensitivity.

At the policy level, ministries of education should integrate AI pedagogy into national curricula and teacher training programs. Public–private partnerships could support the development of Vietnam-specific AI modules that embed culturally authentic content into chatbot interactions. Future directions may include multimodal AI systems combining text, voice, and emotion-sensitive feedback, though their application should be piloted cautiously in local contexts.

In sum, optimizing AI's role in primary EFL education requires coordinated efforts across theory, practice, institutions, and policy. When AI's technological strengths are strategically integrated with teachers' cultural and affective expertise, hybrid models can foster not only short-term gains but also sustained, authentic communicative confidence in young Vietnamese learners.

6. Conclusion

This study investigated the impact of AI-supported speaking practice on Vietnamese primary learners' communicative confidence, as well as the added value of a hybrid AI–teacher model. Findings confirmed that AI tools significantly enhanced learners' WTC and PSE, while hybrid integration produced greater affective and performance gains.

Theoretically, the study extends SLA frameworks into AI-enhanced environments, demonstrating how micro-level scaffolding from chatbots can complement macro-level teacher scaffolding. Practically, it shows that AI-supported practice is both feasible and beneficial in resource-constrained Vietnamese classrooms when implemented with cultural sensitivity and teacher mediation.

The study's most compelling contribution is its demonstration that AI should not replace human instruction but serve as an adaptive supplement. By validating the pedagogical viability of hybrid AI–human models, the research offers a promising blueprint for future EFL instruction—one where technological innovation and teacher guidance work in synergy to foster authentic, sustained communicative confidence among young learners. Future studies should adopt broader designs with larger and more diverse samples to strengthen the generalizability of findings. Including learners across varied proficiency levels, age groups, and educational contexts would capture a fuller spectrum of experiences. Longitudinal approaches spanning semesters or academic years are needed to examine whether short-term gains in speaking confidence can be sustained over time. Expanding the range of communicative tasks and topics in AI-supported practice would also ensure closer alignment with real-world language use. Cross-cultural comparisons are also recommended to explore how cultural norms and communication styles shape engagement with AI. Methodologically, mixed approaches combining surveys with interviews or classroom observations could yield richer insights into learner perceptions. Future work should also test specific AI features, such as real-time corrective feedback or adaptive difficulty, and consider baseline fluency as a moderating factor. Finally, minimizing potential Hawthorne effects would enhance validity, providing more robust evidence and clearer guidance for integrating AI into language education.

7. References

1. AlAqlobi O, Alduais A, Alasmari M, Qasem F. Artificial intelligence in language acquisition: A balancing act of potential and challenges. Forum for Linguistic Studies. 2024; 6(6):1103-1122. Doi: https://doi.org/10.30564/FLS.V6I6.7524
2. Alhusaiyan E. A systematic review of current trends in artificial intelligence in foreign language learning. Saudi Journal of Language Studies. 2025; 5(1):1-16. Doi: https://doi.org/10.1108/SJLS-07-2024-0039
3. Alsawaier RS. The effect of gamification on motivation and engagement. International Journal of Information and Learning Technology. 2018; 35(1):56-79. Doi: https://doi.org/10.1108/IJILT-02-2017-0009
4. Bandura A. Organisational applications of social cognitive theory. Australian Journal of Management. 1988; 13(2):275-302. Doi: https://doi.org/10.1177/031289628801300210
5. Bashori M, Van Hout R, Strik H, Cucchiarini C. Web-based language learning and speaking anxiety. Computer Assisted Language Learning. 2022; 35(5-6):1058-1089. Doi: https://doi.org/10.1080/09588221.2020.1770293
6. Chen X, Huang Y, Li J. Exploring the role of AI chatbots in improving EFL learners' oral proficiency and motivation. Computer Assisted Language Learning. 2023; 36(7):612-634. Doi: https://doi.org/10.31004/jele.v9i5.513
7. Cui W, Zhang J. AI applications in language learning: Improving speaking fluency and pronunciation using chatbots and speech feedback. Interactive Learning Environments, 2025, 1-16. Doi: https://doi.org/10.1080/10494820.2025.2546634
8. Dang TP. Empirical research trends in the use of AI chatbots in EFL teaching and learning in Vietnam: A systematic review. Proceedings of the AsiaCALL International Conference. 2025; 6:325-343. Doi: https://doi.org/10.54855/paic.24623
9. Darmawansah D, Hwang GJ, Lin CJ, Febiyani F. An artificial intelligence-supported GFCA learning model to enhance L2 students' role-play performance, English speaking and interaction mindset. Educational Technology Research and Development, 2025. Doi: https://doi.org/10.1007/s11423-025-10453-6
10. Ding D, Muhyiddin B, Yusof A. 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. 2025; 12(1):1-16. Doi: https://doi.org/10.1057/s41599-025-05550-z
11. Duc DTH, Hang VTT. Fostering Vietnamese EFL learners' learner autonomy and reading comprehension ability through online platforms. European Journal of English Language Teaching. 2024; 9(4). Doi: https://doi.org/10.46827/EJEL.V9I4.5624
12. Dunn WE, Lantolf JP. Vygotsky's zone of proximal development and Krashen's i + 1: Incommensurable constructs; incommensurable theories. Language Learning. 1998; 48(3):411-442. Doi: https://doi.org/10.1111/0023-8333.00048
13. Duong T, Suppasetseree S. The effects of an artificial intelligence voice chatbot on improving Vietnamese undergraduate students' English speaking skills. International Journal of Learning, Teaching and Educational Research. 2024; 23(3):293-321. Doi: https://doi.org/10.26803/ijlter.23.3.15
14. Duong TTU, Nguyen MH. Students' perspective on applying AI in English learning and teaching - A study at UEH. ICTE Conference Proceedings, March 5, 2024, 91-102. Doi: https://doi.org/10.54855/ictep.2458
15. Ebadi S, Zoghi M, Amiri S, Khodadadi F. Exploring the impact of AI-powered speaking tasks on EFL learners' speaking performance and anxiety: An activity theory study. Acta Psychologica. 2025; 259:105391. Doi: https://doi.org/10.1016/j.actpsy.2025.105391
16. Hashemifardnia A, Kooti M. AI-mediated language learning and EFL learners' self-confidence, self-regulation, well-being, and L2 motivation: A mixed method study. English Education Journal. 2025; 16(2):109-124. Doi: https://doi.org/10.24815/eej.v16i2.45696
17. Ilmi Y, Retnaningdyah P. Exploring digital multimodal text in EFL classroom: Transformed practice in multiliteracies pedagogy. Linguistic, English Education and Art (LEEA) Journal. 2020; 4(1):99-108. Doi: https://doi.org/10.31539/leea.v4i1.1416
18. Krashen S. The input hypothesis: Issues and implications. Longman, 1985.
19. Li D, Zhao Y. Artificial intelligence applications for oral communication skills in EFL contexts: A systematic review. Asia-Pacific Education Researcher, 2025, 1-12. https://doi.org/10.1007/s40299-025-01023-8
20. Maflah Alharbi J. Insight into the role of interaction in language acquisition: Vygotsky's interactionist theory of language. Arab World English Journal. 2023;

14(2):281-294. Doi: https://doi.org/10.24093/awej/vol14no2.20

21. Mohammad Hosseini H, Fathi J, Derakhshesh A, Mehraein S. A model of classroom social climate, foreign language enjoyment, and student engagement among English as a foreign language learners. Frontiers in Psychology. 2022; 13:933842. Doi: https://doi.org/10.3389/fpsyg.2022.933842

22. Nguyen N. Motivation in language learning and Dörnyei's L2 Motivational self system. In R. Wyatt & J. K. Lee (Eds.), Multidisciplinary research perspectives in education: Shared experiences from Australia and China. Springer, 2016, 67-72. Doi: https://doi.org/10.1007/978-94-6300-615-6_9

23. Nunnally JC, Bernstein IH. Psychometric theory (3rd ed.). McGraw-Hill, 1994.

24. On V, Luu VTM. An analysis of online learning self-efficacy in the case of Vietnamese university students. Ho Chi Minh City Open University Journal of Science - Social Sciences. 2022; 12(1):41-50. Doi: https://doi.org/10.46223/HCMCOUJS.soci.en.12.1.2145.2022

25. Pham TN, Dang TX. An investigation into the application of artificial intelligence for language teaching and learning in Vietnam. Vietnam Journal of Education. 2025; 9(2):265-283. Doi: https://doi.org/10.52296/vje.2025.557

26. Qiao H, Zhao A. Artificial intelligence-based language learning: Illuminating the impact on speaking skills and self-regulation in Chinese EFL context. Frontiers in Psychology. 2023; 14:1255594. Doi: https://doi.org/10.3389/fpsyg.2023.1255594

27. Rong X, Albin A. Mobile assisted vocabulary learning with Chinese EFL learners: Effectiveness and evaluation. Journal of Higher Education Research. 2022; 3(2):128. Doi: https://doi.org/10.32629/jher.v3i2.737

28. Tran NT, Nguyen TT, Pham HH. Exploring the challenges of L1 negative transfer among Vietnamese English language learners: A qualitative study. Reflections. 2024; 31(2):837-857. Doi: https://doi.org/10.61508/refl.v31i2.275268

29. Vygotsky LS. Mind in society: The development of higher psychological processes. Harvard University Press, 1978.

30. Wei L. Artificial intelligence in language instruction: Impact on English learning achievement, L2 motivation, and self-regulated learning. Frontiers in Psychology. 2023; 14:1261955. Doi: https://doi.org/10.3389/fpsyg.2023.1261955

31. Wei R. Enhancing learner self-regulation through conversational AI in EFL contexts. CALL-EJ. 2023; 24(2):101-119. http://callej.org/journal/24-2/wei.html

32. Xiao Y. The impact of AI-driven speech recognition on EFL listening comprehension, flow experience, and anxiety: A randomized controlled trial. Humanities and Social Sciences Communications. 2025; 12(1). Doi: https://doi.org/10.1057/s41599-025-04672-8

33. Xiao F, Zhao P, Sha H, Yang D, Warschauer M. Conversational agents in language learning. Journal of China Computer-Assisted Language Learning. 2024; 4(2):300-325. Doi: https://doi.org/10.1515/jccall-2022-0032

34. Xu G, Yu A, Liu L. A meta-analysis examining AI-assisted L2 learning. IRAL - International Review of Applied Linguistics in Language Teaching, 2025. Doi: https://doi.org/10.1515/iral-2024-0213

35. Yang F, Li K, Li R. AI in language education: Enhancing learners' speaking awareness through AI-supported training. International Journal of Information and Education Technology. 2024; 14(6):828-833. Doi: https://doi.org/10.18178/ijiet.2024.14.6.2108

36. Zhang C, Meng Y, Ma X. Artificial intelligence in EFL speaking: Impact on enjoyment, anxiety, and willingness to communicate. System. 2024; 121:103259. Doi: https://doi.org/10.1016/j.system.2024.103259

37. Zhang Y, Wang L, Chen H. Sociocultural adaptation and emotional engagement in AI-mediated language learning. System. 2024; 121:102984. Doi: https://doi.org/10.1016/j.system.2024.102984

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