

<https://doi.org/10.69639/arandu.v13i3.2518>

ChatGPT Voice and Role-Play Activities on English Oral Communicative Competence: A Pilot Study with Secondary Learners

ChatGPT Voice y actividades de juego de roles para fortalecer la competencia comunicativa oral en inglés: Un estudio piloto con estudiantes de educación secundaria

Gabriela Jacqueline Galeas Arboleda

ggaleas@uteq.edu.ec

<https://orcid.org/0000-0001-9759-5474>

Universidad Técnica Estatal de Quevedo
Quevedo – Ecuador

David Enrique Montenegro Montenegro

dmontenegro@yachaytech.edu.ec

<https://orcid.org/0009-0009-2549-7256>

Yachay Tech University
Urcuquí – Ecuador

Joel Sebastián Chevez Mendaño

<https://orcid.org/0009-0001-2907-1190>

jchevezm@uteq.edu.ec

Universidad Técnica Estatal de Quevedo
Quevedo – Ecuador

Artículo recibido: 10 julio 2026- Aceptado para publicación: 16 agosto 2026
Conflictos de intereses: Ninguno que declarar.

ABSTRACT

The rapid implementation of AI (Artificial Intelligence) technology has gained the attention of many educators to enhance complex EFL skills, such as speaking, especially in contexts with limited opportunities for language practice. This pilot study inspects the effects of the use of ChatGPT's voice feature to enhance oral presentations (role-plays). A group of secondary EFL learners in a public school participated in a four-week intervention, engaging in daily practice with ChatGPT voice. In each English session, participants engaged in a role-play practice with the AI feature. A speaking rubric with five sub-skills was designed and used to assess students' performance each week, and a semi-structured interview was conducted to collect specific participants' perceptions. ChatGPT was employed to evaluate student performance, with the exception of interaction/expressiveness, to get more objective results. To examine students' performance across the four weeks and to address the first research question regarding their progression in oral communicative competence, a repeated-measures analysis of variance (ANOVA) was conducted to observe statistically significant changes over time. Qualitative data were analyzed using thematic analysis. Quantitative data indicated that ChatGPT Voice effectively enhanced EFL students' oral skills in combination with role-plays; however, no

significant differences in gains were observed among subskills. On the other hand, qualitative analysis portrayed positive perceptions of the AI tool for speaking practice in alignment with previous research and major considerations of its usage. The study offers important insights for language teachers and researchers into the use of AI technologies and how they complement traditional instruction and assessment.

Keywords: AI, EFL oral communication, role-play, secondary education

RESUMEN

La rápida incorporación de tecnologías de inteligencia artificial (IA) ha despertado el interés de docentes por fortalecer habilidades complejas del inglés como lengua extranjera (EFL), especialmente la expresión oral, en contextos con oportunidades limitadas para practicar el idioma. Este estudio piloto examinó los efectos del uso de la función de voz de ChatGPT (ChatGPT Voice) para mejorar las presentaciones orales mediante actividades de juego de roles. Participó un grupo de estudiantes de educación secundaria de una institución pública en una intervención de cuatro semanas. Durante este periodo, los estudiantes practicaron diariamente con ChatGPT Voice y realizaron actividades de juego de roles en las sesiones de inglés. El desempeño oral se evaluó semanalmente mediante una rúbrica diseñada para el estudio, que incluyó cinco subhabilidades. Además, se realizaron entrevistas semiestructuradas para conocer las percepciones de algunos participantes. ChatGPT se empleó como herramienta de evaluación, excepto para la dimensión de interacción/expresividad, con el propósito de favorecer una valoración más objetiva. Para analizar los cambios en el desempeño oral a lo largo de las cuatro semanas, se aplicó un análisis de varianza de medidas repetidas (ANOVA), mientras que los datos cualitativos se examinaron mediante análisis temático. Los resultados cuantitativos mostraron que ChatGPT Voice, combinado con actividades de juego de roles, contribuyó significativamente a mejorar las habilidades orales de los estudiantes, aunque no se encontraron diferencias significativas entre las ganancias de las distintas subhabilidades. El análisis cualitativo reveló percepciones positivas y consideraciones relevantes sobre el uso de esta herramienta. El estudio aporta evidencia sobre el potencial de la IA para complementar la enseñanza y evaluación tradicionales de lenguas.

Palabras claves: inteligencia artificial, comunicación oral de Inglés como lengua extranjera, juego de roles, educación secundaria

Todo el contenido de la Revista Científica Internacional Arandu UTIC publicado en este sitio está disponible bajo licencia Creative Commons Attribution 4.0 International. 

INTRODUCTION

The acquisition of English as a foreign language remains a central focus in secondary education, where oral fluency is often considered one of the most demanding skills to develop. Many learners struggle to express themselves fluently and confidently in English due to limited opportunities for authentic communicative practice, which frequently leads to anxiety, low motivation, and reduced oral performance (Fernández-García & Fonseca-Mora, 2019). As highlighted by Sosa-López and Mora (2022), speaking anxiety has a direct impact on learners' fluency, accuracy, and overall communicative complexity, making it one of the most critical challenges in English language learning.

In response to speaking difficulties, artificial intelligence (AI) has emerged as a promising support for foreign language education. Recent advances in AI-powered tools, such as conversational chatbots, provide learners with interactive and adaptive environments that simulate real-life communication. According to Koç and Savaş (2024), the integration of AI chatbots in English language learning enhances motivation, increases learner engagement, and offers opportunities for individualized practice, thereby reinforcing speaking skills in more natural and dynamic ways.

Empirical studies indicate that AI-mediated interactions not only improve fluency but also contribute to a broader range of speaking competencies, including pronunciation, lexical development, grammatical accuracy, and willingness to communicate (Fathi et al., 2024; Nuñez et al., 2025). These findings suggest that AI tools, when integrated with pedagogical strategies such as role-playing, can significantly enhance oral language learning outcomes, provided appropriate pedagogical considerations are in place (Kim & Park, 2023).

There is a considerable body of studies on AI technologies used in EFL/ESL speaking contexts, predominantly in Asian countries such as China or Indonesia (Xing, 2025; Fauzi et al., 2025). For instance, a similar study was carried out by Huang (2024), who investigated how, by using voice prompts, students can leverage the potential of ChatGPT to gain meaningful feedback on their oral performance and strengthen their pronunciation skills. However, this research was developed in a culturally different background. Furthermore, little is known about AI technologies to enhance speaking skills in secondary learners for oral presentation, specifically role-plays. Most studies have been carried out in higher education (Deng & Jamaludin, 2026).

Building on this perspective, the present study examines the effect of ChatGPT Voice Chat on the oral competence of secondary school students in role-play activities. The research is conducted in Quevedo, a coastal city characterized by public educational contexts in which opportunities for oral English practice are often limited and largely confined to the classroom. Within the broader context of Ecuador, English as a Foreign Language (EFL) instruction has gained increasing importance in recent years; however, persistent challenges remain, particularly

the limited exposure to authentic communicative situations. This underscores the need to explore innovative tools that can effectively support the development of linguistic competence.

Over a four-week intervention, students engaged in role-play activities guided by teacher-designed prompts that simulated real-life communicative scenarios (e.g., everyday interactions, academic situations, and social contexts). These activities were mediated through ChatGPT Voice Chat, enabling continuous interaction with an artificial intelligence system capable of providing immediate and adaptive feedback. This approach aimed not only to increase learners' exposure to the target language but also to create a more dynamic, accessible, and low-anxiety learning environment.

The primary purpose of this study is to evaluate the extent to which the use of ChatGPT Voice Chat + role-play contributes to improvements in key components of oral competence, including fluency, pronunciation, vocabulary use, grammatical accuracy, and speaking confidence. Additionally, the study explores students' perceptions of artificial intelligence as a supportive tool in their English learning process, considering dimensions such as motivation, learner autonomy, and the reduction of language anxiety.

It is important to note that, although there has been growing international interest in the use of artificial intelligence in language education, there is still limited empirical evidence in the Latin American context—and particularly in Ecuador—regarding the implementation of tools such as ChatGPT Voice for the development of oral skills in secondary education. Furthermore, Chat GPT Voice and role-plays contribute to an under-researched area of speaking with voice-enabled Gen-AI in secondary contexts (Lo et al, 2024).

Therefore, this study seeks to address this gap in the literature by providing context-specific evidence that contributes to a better understanding of the potential of these technologies in similar educational settings.

The purpose of this pilot study is to explore the potential of ChatGPT Voice as an innovative digital tool to enhance English oral proficiency among secondary school students within a short intervention period. To achieve this aim, the following question was set:

- To what extent does the use of ChatGPT Voice Chat improve the speaking skills of secondary students during a four-week intervention?

Secondary question:

- What measurable changes are observed in students' pronunciation, fluency, vocabulary, and grammar after practicing with ChatGPT Voice Chat?
- How do secondary students perceive the use of ChatGPT Voice for practicing role-plays?

LITERATURE REVIEW

Speaking Skills in EFL

Among the four core language skills—reading, writing, listening, and speaking—speaking is widely recognized as the most difficult to develop, despite being essential for effective

communication and real-life interaction (Rao, 2019). Learners often struggle to demonstrate their speaking abilities due to multiple challenges, including speaking anxiety, limited vocabulary and grammatical knowledge, pronunciation difficulties, and insufficient opportunities for oral practice (Du et al., 2025; Aziz & Kashinathan, 2021). These factors frequently result in low participation and hinder the development of communicative competence in EFL contexts.

A primary goal of learning a foreign language is the ability to communicate; however, many students lack confidence when attempting to improve their speaking skills (Fernández-García & Fonseca-Mora, 2019). In response to these challenges, recent research highlights the growing role of artificial intelligence in enhancing oral proficiency. AI-based tools help reduce foreign language anxiety (FLA), promote self-regulated learning, and provide adaptive instruction. By creating low-pressure conversational environments, these tools enable learners to practice speaking without fear of judgment, thereby fostering both confidence and fluency. Moreover, their flexible and explicit design allows students to monitor their progress through continuous oral feedback, which further supports the development of learner autonomy (Nhan et al., 2025). In this sense, the integration of technology into the development of oral skills facilitates the design of student-centered activities that actively promote autonomous learning (Cuevas-Montero, 2021).

According to the Common European Framework of Reference for Languages, speaking competence involves the ability to interact orally, express ideas, and engage effectively in conversations across proficiency levels (Council of Europe, 2020). Nevertheless, the development of speaking skills in EFL classrooms remains constrained by various factors, such as students' anxiety during interaction and difficulties in applying grammatical knowledge (Tapia & Vega, 2024). In the Ecuadorian context, these challenges are further compounded by systemic issues, including limited instructional time, an emphasis on grammar and writing over oral communication, a shortage of qualified English teachers, and low student motivation. Together, these barriers continue to limit opportunities for meaningful speaking practice in the classroom (Rodríguez & Baquerizo, 2025).

GenAI and EFL Speaking

Generative Artificial Intelligence (GenAI) refers to artificial intelligence systems capable of generating new content, such as text, images, or audio, by learning patterns from large datasets through advanced machine learning models (Lim, 2023). Furthermore, large language models such as ChatGPT represent a significant advancement in artificial intelligence because they can generate human-like responses and support interactive learning processes in educational contexts (Kasneci et al., 2023). Despite the benefits of generative AI in education, several challenges have been identified. For example, large language models may generate inaccurate information, raise ethical concerns, and encourage excessive reliance on technology if not used critically in educational contexts (Kasneci et al., 2023).

On the other hand, according to Deng and Jamaludin (2026), just a few studies focus on GenAI in speaking skills, even though AI systems offer students meaningful opportunities to engage in real-life language use while receiving prompt and constructive feedback (Mageira et al., 2022; Park, 2023). ChatGPT Voice proved highly effective in improving learners' oral proficiency, showing observable progress in pronunciation accuracy, fluency, vocabulary range, and grammatical precision. These improvements were largely associated with sustained AI-supported practice and the provision of immediate corrective feedback (Nuñez et al., 2025). Recently, a study found that AI-based dialogue systems fulfill an important communicative function in EFL learning at the university level, as students report that these systems reduce speaking anxiety, improve vocabulary retention, and provide instant feedback (Zhai & Wibowo, 2023).

The integration of AI-assisted tools has shown considerable potential for enhancing English language learning, particularly in the development of speaking skills. Evidence suggests that regular and sustained use—rather than simple access—plays a crucial role in achieving positive learning outcomes, especially in underserved or rural settings. In these contexts, AI technologies offer a scalable, adaptive, and learner-centered solution to support the development of oral fluency. Furthermore, such tools are both pedagogically sound and feasible for implementation in real classroom environments (Alenezi & Alenezi, 2025). Studies also indicate that learners make noticeable progress in key aspects of speaking, including intonation, stress, and fluency, while benefiting from increased opportunities for practice both within and beyond the classroom. In addition, AI chatbots promote self-regulated learning by providing adaptive support and personalized feedback, thereby enriching the overall learning experience.

Moreover, this technology helps reduce language anxiety and speaking-related apprehension, creating supportive environments that encourage active participation and confidence building (Klímová & Seraj, 2023). Their ease of use and accessibility make them suitable for informal learning contexts, enabling students to engage in meaningful communication at any time without requiring advanced technical skills. Short, guided interactions are particularly effective for automated assessment, while strategies such as Socratic questioning contribute to the development of critical thinking skills in EFL learners. Importantly, students with lower levels of proficiency tend to benefit the most, demonstrating gains in confidence, vocabulary development, self-assessment, and oral presentation abilities. Additionally, research on tools such as ChatGPT emphasizes the relevance of personalized feedback and careful instructional design, offering valuable insights for optimizing oral communication practice in EFL contexts (Cha *et al.*, 2024).

ChatGPT Voice Chat

The adoption and integration of AI tools across various sectors have rapidly increased since the launch of ChatGPT in 2022. This trend has also reached the educational field, where

these technologies are becoming increasingly present in everyday teaching and learning practices (Bustamante Salinas *et al.*, 2025). Recently, more and more instructors are incorporating Chatbots in speaking instructions (Klímová & Seraj, 2023; Tai & Chen, 2024) as they offer opportunities for learners to engage in meaningful language input and productive output within a low-stress environment (Jeon *et al.*, 2023). Pau *et al.* (2026) emphasize that, despite technical challenges, ChatGPT stands out significantly, excelling in the ability to adapt and be evaluated in real-time, as well as in integration with virtual reality (VR) technologies. Its use has increased significantly since 2023 (Xing, 2025).

Conversational AI tools such as ChatGPT enable learners to interact with spoken dialogue systems, allowing them to practice the target language through simulated conversations. These tools provide opportunities for students to receive individualized feedback while communicating in a low-anxiety environment, which can support the development of their speaking skills (Ericsson & Johansson, 2023). AI-mediated speaking activities allow learners to practice oral communication skills in interactive learning environments. Through these technologies, students can engage in repeated speaking practice, which may contribute to improvements in fluency, pronunciation, and grammatical accuracy in second-language learning contexts (Fathi *et al.*, 2024). Studies also show that AI chatbots can enhance speaking outcomes by increasing learners' confidence, engagement, and motivation during oral language practice. By providing opportunities for constant interaction, these tools may encourage students to participate more actively in speaking activities (Du & Daniel, 2024).

METHODOLOGY

This pilot study adopted a mixed-method exploratory research design and aims to provide preliminary evidence of ChatGPT Voice for oral English development among secondary school learners. A mixed- method approach is particularly well-suited for language learning studies, as it enables a comprehensive understanding of both outcomes and learner experiences (Wei, 2022; Niglas *et al.*, 2021).

Participants

Given the exploratory nature of the intervention, fifteen high school students, aged 12–15, were initially nominated to participate in the study. The group included students of both genders, allowing for a more diverse range of experiences and perceptions within the learning process. Subsequently, an oral diagnostic interview was administered to the entire group to determine the final participants. This evaluation assessed key components of oral competence, including pronunciation, fluency, vocabulary use, and grammatical accuracy.

The total scores were converted into percentages, and seven students who achieved 70% or higher were selected to form the final sample. This selection was based on intentional sampling, considering their academic performance in English and their relevance to the analysis of oral

competence development. In this way, the process ensured alignment between the study’s objectives and the characteristics of the participants. The informed consents were approved by the legal guardians of participants who must be students under 18 years old. Pseudonyms were used to protect the identities of participants.

The study was conducted in a public educational setting characterized by limited resources and few opportunities for authentic English interaction outside of the classroom. These conditions make the development of oral skills a significant challenge for students, reinforcing the relevance of implementing technology-mediated pedagogical interventions. It should be mentioned that the students had previously experienced using ChatGPT, but not for oral English practice.

Instruments

The techniques used for data collection in this research are as follows:

- **A speaking rubric.** The speaking rubric was developed by the researchers and informed by established frameworks for assessing oral proficiency (Brown & Abeywickrama, 2019; Luoma, 2004). The instrument assessed five key dimensions of speaking performance: pronunciation, fluency, vocabulary, grammar, and interaction. To ensure content validity, the rubric was reviewed by an expert in English language teaching and assessment, who examined the relevance, clarity, and appropriateness of each criterion. The expert’s recommendations were incorporated into the final version of the instrument before its implementation. Additionally, a second evaluator independently assessed a subset of the speaking performances to enhance scoring reliability and reduce potential evaluator bias (McNamara, 1996). The rubric was employed to evaluate students’ performance in a role-play task based on a fashion show scenario, which served as the final project of the intervention. The primary purpose was to assess their speaking skills, considering all aspects involved in oral communication, and to monitor their progress over the weeks.

Tabla 1

Rubric for Evaluating Speaking in a Role Play

Criteria	2 pts (Excellent)	1 pt (Acceptable)	0 pts (Deficient)
Pronunciation	Clear and understandable pronunciation; few difficulties in comprehension.	Mostly understandable, though with some mistakes.	Very difficult to understand; frequent errors.
Fluency	Speaks naturally, continuously, with no	Some pauses or fillers, but meaning is maintained.	Many pauses, rigid reading, or lack of fluency.

	long pauses or overuse of notes.		
Vocabulary	Varied and appropriate vocabulary for the role play context.	Limited or repetitive vocabulary, but understandable.	Very basic vocabulary, frequent incorrect usage.
Grammar	Correct sentences, few errors that do not affect meaning.	Some grammatical errors, but overall understandable.	Many errors that make comprehension difficult.
Interaction / Expressiveness	Actively participates, responds naturally, uses gestures, and appropriate intonation.	Participates but with little expressiveness or minimal interaction.	Does not interact or depend fully on notes, without expressiveness.
Source: Own production			

***Note:** This table shows the criteria for evaluating a speaking role play. Scores range from 0 to 2 points: 2 = Excellent, 1 = Acceptable, 0 = Deficient.*

- **Semi-structured interview.** A semi-structured interview was employed to explore students' perceptions, experiences, and evaluations of the learning process while allowing flexibility in the interaction and the opportunity to probe emerging ideas (Kallio et al., 2016). This type of interview is particularly suitable for educational research because it enables the collection of rich and detailed data while maintaining a clear focus on the research objectives (Creswell & Creswell, 2018; Cohen et al., 2018). Following the four-week intervention, each participant completed a short semi-structured interview consisting of two open-ended questions. Although limited in number, the questions were carefully designed to elicit reflections on the use of ChatGPT Voice Chat as a learning support tool and its perceived influence on confidence and speaking development. The use of open-ended questions encouraged participants to provide spontaneous and detailed responses, generating insights into affective and cognitive aspects of language learning, such as motivation, anxiety, and self-perceived progress. The interview data complemented the quantitative findings obtained through the speaking rubric, contributing to a more comprehensive understanding of the intervention's impact. To facilitate participants' freedom of expression and ensure the accuracy of their responses, the interviews were conducted in Spanish, their native language.

Intervention

During the initial phase of the intervention, a mechanical practice based on pre-written dialogues was implemented to familiarize students with the dynamics of the activity. In this sense, during the first week, participants worked with dialogues previously designed and provided by the researchers, corresponding to a role-playing game that included two presenters (Angela and

Deymar). Students were required to memorize their interventions and practice with ChatGPT, focused solely on checking whether they could reproduce the dialogue correctly. At this stage, the chatbot played a fundamentally mechanical role: listening, comparing, and validating the accuracy of the memorized lines.

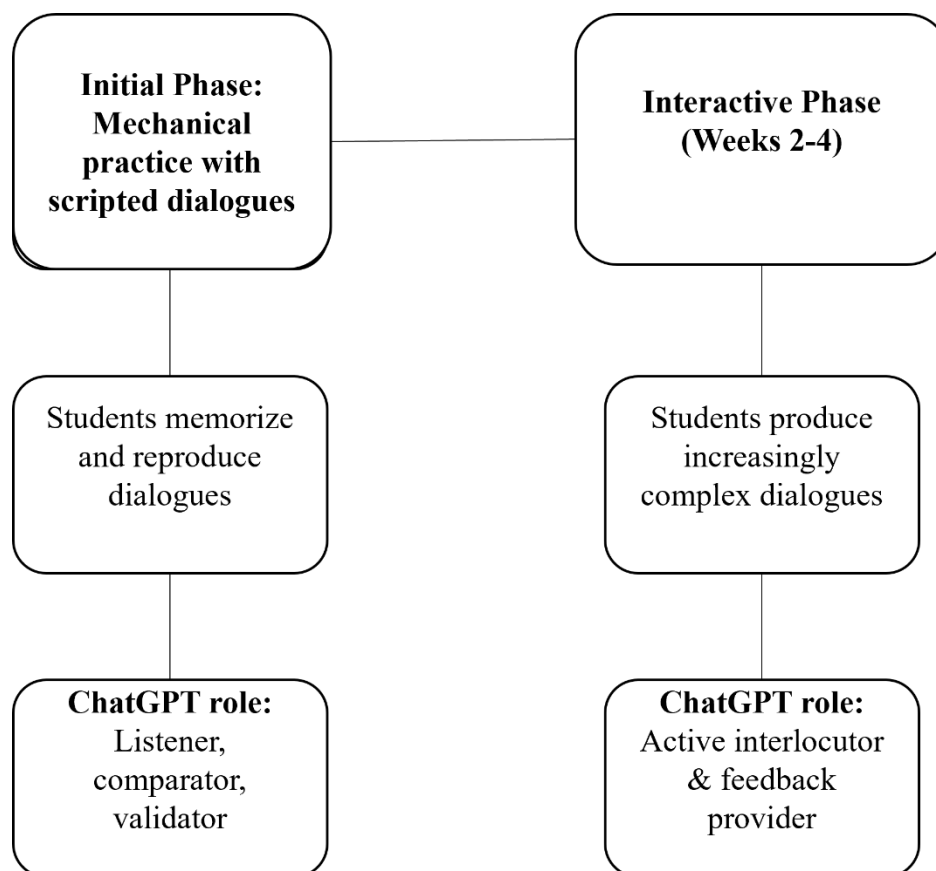
At a later stage of the intervention, a progressive practice based on oral interaction was promoted through the voice chat function. From the second week to the fourth, students began interacting with ChatGPT, producing increasingly complex dialogues. The purpose of this stage was to observe gradual development in key aspects such as grammar, vocabulary, fluency, and pronunciation. In this context, ChatGPT took on the role of an active interlocutor, allowing for a more authentic practice and providing feedback that facilitated the identification of progress in participants' oral competence. Figure 1 shows the two main stages and the virtual assistance role.

In order to structure the interaction during practice activities clearly and consistently, the researchers designed a specific prompt that guided the use of the tool. The statement provided was as follows: "First of all, let's have a conversation using your voice chat feature with this RPG. You will say all the sentences of the presenter, both Angela and Deymar. Let's do the whole conversation at the same time. We started with voice chat, and after that, I'll ask you for ratings and feedback." This prompt made it possible to establish an organized dynamic, ensuring that the interaction developed in a coherent way and aligned with the objectives of the activity.

This instruction also made it easier for ChatGPT to take an active role in the process, simultaneously representing both dialogue presenters. In this way, the system not only functioned as an interlocutor but also as a linguistic model that guided the structure of the conversation, maintaining continuity of exchange and offering students a clear reference for language use. This mediation contributed to a more controlled and structured practice environment, in which participants could focus on oral production. The design of the prompt also incorporated a later evaluation phase, in which the tool was required to provide ratings and feedback based on student performance. This component made it possible to immediately integrate the formative evaluation within the same interaction, favoring processes of self-adjustment and reflection on one's own linguistic performance.

Figura 1

Intervention Process



Source: Own production

It is important to note that the original presenters (the characters of the role-playing game) were not included as participants in the study because they only read the scripted lines. They were not involved in the process of learning or improvement. Their role was purely performative. During the actual practice sessions, ChatGPT acted as a presenter, involving each student.

Performance evaluation and data analysis

Performance was assessed at two levels. The first level was assessed by Automatic Evaluation by ChatGPT. The virtual assistant provided feedback on grammar, vocabulary, fluency, and pronunciation. However, the researchers did not use the chatbot's automatic scoring for interaction or expressiveness, because the system produced random values and could not accurately detect emotions, acting skills, or expressiveness. The second level was human evaluation (done by researchers). The researchers evaluated the criteria related to interaction and expressiveness.

Data analysis retrieved from the criteria was carried out using the SPSS Statistics program. Descriptive statistics (means and standard deviations) were calculated to examine students' performance across the four weeks and to address the first research question regarding their progression in oral communicative competence. To determine whether the observed changes over time were statistically significant, a repeated-measures analysis of variance (ANOVA) was

conducted. Qualitative data obtained from the open-ended interview questions were analyzed using thematic analysis. Responses were coded, categorized into themes, and supported with representative excerpts from participants' comments. Pseudonyms were employed to protect participants' identities.

RESULTS

Quantitative Results

Over the four weeks of implementation, students demonstrated progressive improvement in their oral English performance. Table 1 shows that in Week 1, during the initial diagnosis and first practices, most participants showed willingness to participate and attempted to answer in English; however, their performance was weak, with very short answers, frequent pauses, and limited vocabulary, resulting in an average score of ($M=4,59$). By Week 2, the first improvements became evident as students began producing longer answers and using simple connectors such as “because” and “and”. Despite this progress, many pauses, grammar errors, and limited expressiveness persisted, with an average score of ($M=6,5$). During Week 3, students showed steady progress, delivering more natural speech, organizing ideas more clearly, and incorporating more connectors, which allowed them to give more detailed answers. Although some grammar and pronunciation mistakes remained, comprehension was much clearer, and scores improved to ($M=7,44$). Finally, in Week 4, during the last practices and the role-play presentation, students reached their best performance, providing fluent, detailed, and confident responses with clear and coherent communication. Only minor errors were observed, and their average scores increased to ($M=8,67$), marking a significant improvement compared to the first week. Students' weekly scores over four weeks and their total progress. Total progress is calculated as the difference between the Week 1 and Week 4 scores (see Table 1).

Tabla 2

Students' Progress Over Four Weeks

	N	Mean	Std. Deviation	Minimum	Maximum
Week 1	7	4.59	.20	4.30	4.80
Week 2	7	6.50	.22	6.20	6.80
Week 3	7	7.44	.17	7.20	7.70
Week 4	7	8.67	.11	8.50	8.80

To determine whether students' speaking performance improved over time, a Friedman test was performed on the weekly rubric scores. The analysis revealed a statistically significant difference among the four measurement points ($p < .001$). This finding demonstrated that students' speaking performance improved significantly throughout the intervention. The progressive

increase in scores across the four weeks suggests that regular practice with ChatGPT Voice may have contributed positively to the development of learners' oral skills.

Table 2 illustrates the weekly progress of the entire group in relation to each assessment criterion. The table shows that participants predominantly outperformed in vocabulary from ($M=0,83$) in week 1 to ($M=1,86$). Followed by fluency with ($M=0,67$) in the initial week to ($M=1,67$) in the final week of practice, and pronunciation with ($M=0,77$) in the first week and ($M=1,71$). Then the grammar criterion went from ($M=0,97$) to ($M=1,64$). Finally, interaction/expressiveness started at ($M=1,30$) and reached ($M=1,79$) at the end of the intervention.

Tabla 3

Progression over the weeks by each criterion

	Week 1	Week 2	Week 3	Week 4
Criteria	Mean	Mean	Mean	Mean
Pronunciation	0.77	1.29	1.86	1.71
Fluency	0.67	1.14	1.43	1.67
Vocabulary	0.83	1.46	1.43	1.86
Grammar	0.97	1.14	1.49	1.64
Interaction/ Expressiveness	1.30	1.47	1.24	1.79
Total	4.8	6.7	7.6	8.8

Note: The data presents the Median scores of each criterion.

A repeated-measures ANOVA in Table 3 revealed a statistically significant effect of Week on students' speaking performance, $F(3, 18) = 4157.33$, $p < .001$, $\eta^2 = .999$, indicating a substantial and consistent improvement across the four-week intervention. In contrast, no significant main effect of Subskill was found, $F(4, 24) = 0.62$, $p = .656$, $\eta^2 = .09$, suggesting that students performed similarly across pronunciation, fluency, vocabulary, grammar, and interaction. The Week $\times$ Subskill interaction was also not significant, $F(12, 72) = 1.00$, $p = .457$, $\eta^2 = .14$, indicating that all speaking subskills improved at a comparable rate over time.

Tabla 3

Repeated-measures ANOVA for speaking performance

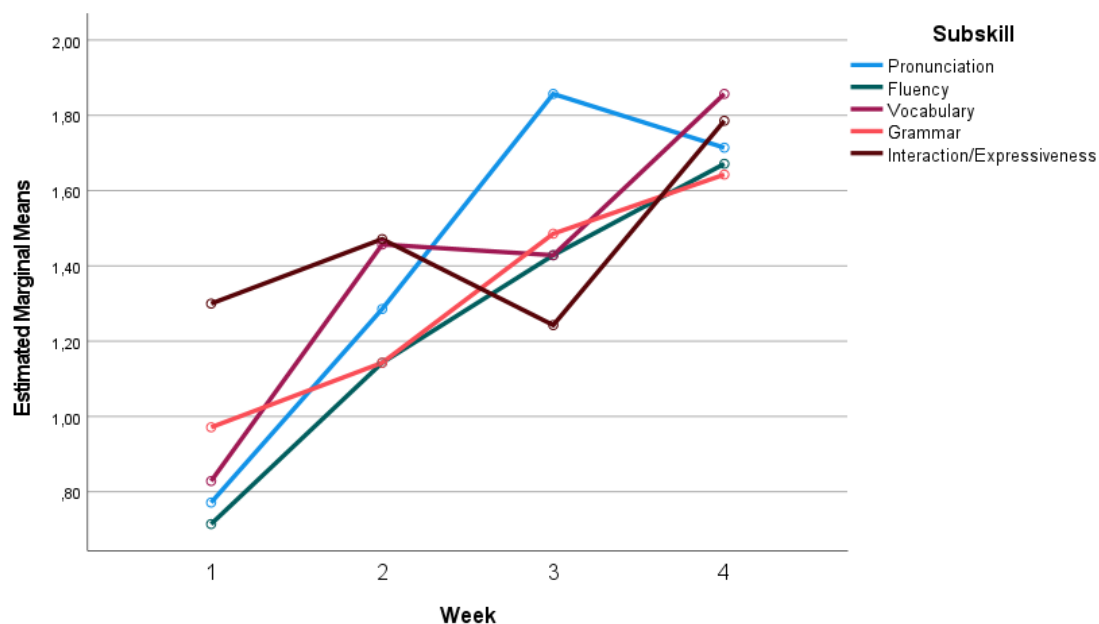
Effect	F	df	p	Partial η^2
Week	4157.33	3. 18	< .001	.999
Subskill	0.62	4. 24	.656	.093
Week $\times$ Subskill	1.00	12. 72	.457	.143

Figure 2 illustrates the progression of students' speaking subskills across the four-week intervention. Overall, all subskills exhibited an upward trend, indicating consistent improvement

in oral performance. Fluency showed the most stable linear development, while pronunciation demonstrated rapid initial gains followed by slight stabilization in Week 4. Vocabulary and grammar showed steady improvement with minor fluctuations, whereas interaction/expressiveness displayed a more variable pattern, with noticeable improvement emerging in the final week. These findings suggest that while all speaking components benefited from the intervention, each subskill developed at a different rate and with distinct trajectories.

Figure 2

Changes in Learners' Speaking Performance by Subskill



Qualitative Results

After completing four weeks of practice, two open-ended questions were asked of each student to understand their perceptions and feelings about the process and the use of ChatGPT as a support tool. The qualitative findings reveal overwhelmingly positive perceptions of ChatGPT Voice Chat as a supportive tool for enhancing English-speaking skills and confidence. Participants consistently described emotional shifts from anxiety and nervousness to confidence and comfort throughout the intervention.

Several students emphasized the reduction of anxiety in speaking, which coincides with the affective filter hypothesis in the acquisition of second languages (Krashen, 1982). For example, Ana reported feeling "more confident practicing without pressure", highlighting improvements in both pronunciation and fluency, which suggests that an environment without immediate assessment favors oral production. In contrast, Sofia mentioned that although she felt insecure at first, the constant repetition of interactions allowed her to gain confidence progressively, evidencing a process of adaptation. Similarly, Elena noted that the tool made her feel supported and calm, making it easier to organize her ideas before responding. However, other

students indicated that sometimes the lack of explicit corrective feedback limited the identification of specific errors. Taken together, these reflections show that the flexible and non-critical nature of vocal interaction contributes significantly to reducing anxiety, although it also raises the need to balance this environment with more targeted feedback instances to enhance learning.

A recurring theme was the gradual transformation of self-perception. Sofia described moving from being shy to feeling confident and excited about speaking in English. Julia and Camila also reported transitioning from nervousness and anxiety to feeling prepared, comfortable, and fluent. These emotional changes indicate that repeated exposure to voice practice fostered both linguistic development and self-assurance.

Additionally, students highlighted gains in expressiveness and vocabulary. Isabela mentioned feeling more relaxed and natural when speaking, while Valeria expressed pride in her improvement, particularly in vocabulary and pronunciation. The sense of achievement and enjoyment further reinforced their motivation to continue practicing. The findings suggest that ChatGPT Voice Chat is perceived not only as a language practice tool but also as an emotionally supportive environment that promotes fluency, pronunciation development, and increased speaking confidence.

DISCUSSION

Speaking gains (Overall and subskills)

The noticeable progress in students' ability to express themselves confidently in English represents one of the most significant findings of this study. Participants demonstrated improvements in pronunciation, vocabulary use, and overall communicative fluency throughout the intervention. These findings are consistent with previous research on AI-mediated language learning, which has reported significant gains in speaking performance through the use of ChatGPT Voice, AI conversational agents, and chatbots, often yielding medium to large effect sizes (Nuñez et al., 2025; Guerrero, 2026). In particular, the results of the present study closely resemble those reported by Nuñez et al. (2025), especially regarding vocabulary development. Notably, comparable improvements were observed despite the relatively short duration of the intervention, which lasted only four weeks. Progress was evident in the final role-playing practices, where students showed longer, more coherent, and natural responses compared to their initial performances. These results are consistent with recent research on the use of artificial intelligence tools in oral skills development, which shows significant improvements in fluency, discursive coherence, and confidence in communication (Yan & Singh, 2026).

Speaking anxiety, engagement, and autonomy

It has also been noted that AI-mediated practice environments favor greater language production and reduce anxiety by offering free-judgment interactions, facilitating more frequent speaking, which resonates with Tan and Ismail's (2025) and Nhan et al. (2025) results,

respectively. These results are also consistent with broader evidence suggesting that AI-driven conversational practice can strengthen fluency, lexical development, and learners' willingness to communicate (Fathi *et al.*, 2024), highlighting the importance of interaction in technology-supported oral language development. However, other studies found that anxiety was not reduced or even slightly intensified, despite performance gains or task-specific reductions only (Huang, 2026). For instance, Shazly (2021) found that the use of conversationally enhanced AI chatbots slightly intensified learners' FLA speech-related anxieties, which were not reduced. The reduction of unpleasant emotions, such as anxiety, when interacting with AI technologies makes a straight connection with a previous study done by Susoy (2026), who found a significant reduction in speaking anxiety before AI-facilitated examinations compared to human-facilitated or neutralized the relationship between anxiety and performance.

Student feedback further supports quantitative results, as learners reported that interacting with ChatGPT Voice *Chat* provided them with consistent opportunities to practice without the pressure commonly experienced in traditional classroom settings. This aligns with recent studies emphasizing that repeated practice and immediate feedback are key to enhancing communicative competence in second language learning (Sosa-López & Mora, 2022). The use of AI allowed students to engage in meaningful and personalized tasks, enabling them to progress at their own pace while focusing on individual weaknesses (Koç & Savaş, 2024).

Students highlighted that this approach simulated real-life communicative scenarios, which not only reinforced their linguistic skills but also boosted their self-confidence when speaking in public. However, while AI provides a low-pressure environment, it may not eliminate performance-related apprehension for all learners (Huu *et al.*, 2026). Similarly, researchers like Wei (2023) found that AI-mediated instruction positively impacts learners' motivation, self-regulation, and learning outcomes, while Zhang *et al.* (2024) provided empirical validation for AI-driven technology to increase students' willingness to communicate and reduce foreign language anxiety, especially if working in pairs on work and presentation tasks (Huang & Kakham, 2026). Alenezi and Alenezi (2025) also reported that students expressed high levels of satisfaction, motivation, and confidence when using the chatbot, with the overall attitude score averaging 4.35 out of 5. Students frequently report positive attitudes, higher confidence, and a "safe space" for experimentation with positively matches Fathi *et al.* (2024) and Tan and Ismail (2025).

A central procedure in this study involved integrating an AI-based automated evaluation system alongside human assessment to ensure fairness and balance in measuring students' performance. This approach aligns with the findings of Susoy (2026), whose participants showed a strong preference for human facilitation, attributing this preference to the empathy and interpersonal connection provided by human evaluators. Although this may seem contradictory, the evidence suggests that the most effective assessment models may integrate both AI-driven

and human-mediated processes. Thereby capitalizing on the complementary advantages of each modality, rather than over-relying on AI, which may undermine autonomy and critical evaluation of alternative options due to automation (Ebadi et al., 2025).

CONCLUSION

This study showed that the systematic use of ChatGPT Voice Chat had significant positive effects on students' oral English proficiency in high school. Through a phased implementation over several weeks, the comparison between initial results and final evaluation based on the role play showed clear and measurable improvements in key components of communication competence. These developments suggest that consistent practice in interactive and simulated environments contributes effectively to strengthening oral skills in an English as a foreign language learning context. In addition, qualitative analysis provided a broader understanding of the students' experience, who perceived ChatGPT Voice Chat as an accessible, innovative, and motivating tool. In particular, they highlighted its usefulness in reducing the nervousness associated with oral production (or level of anxiety). This environment allowed students to experience the language more spontaneously and naturally, promoting a greater willingness to participate and take linguistic risks. This study positions ChatGPT voice + role-play design as a valuable contribution to enhance oral skills among secondary learners in a context where rich and authentic exposure is limited. However, teachers may use the tool to provide safe, non-judgmental rehearsal spaces, rather than over-reliance, a trade-off that can lead to a reduction of learner autonomy, lower ability to communicate without AI help, and decreased motivation to generate ideas.

Taken together, the study's findings not only support the pedagogical value of integrating artificial-intelligence technologies into medium-system language teaching, but they also emphasize the importance of designing learning experiences that combine structured practice with opportunities for meaningful interaction (role-play). Therefore, it is recommended to consider the use of tools such as ChatGPT Voice Chat as a strategic complement to traditional English language teaching practices (Koç & Savaş, 2024), both in the classroom and outside it, capable of enriching teaching processes, learning, and practice while responding to the communicative needs of students in educational contexts where English language learning is not immersive.

While the present study offers positive and relevant results, it is important to recognize some methodological limitations that do not invalidate the findings but guide their interpretation. First, the intervention time was four weeks, which is a relatively short period. However, this lapse was sufficient to show significant improvements in oral production, highlighting the potential of the tool even in short-term interventions. Secondly, the sample consisted of a small number of students, all women. Although these characteristics limit the generalization of results to larger and more diverse populations, it does not affect the validity of findings within the study group.

Finally, the participants belonged to the same educational context, specifically a public institution. While this homogeneity may restrict extrapolation to other environments, it also allowed for greater control of contextual variables. In this sense, these limitations do not diminish the relevance of the results obtained, but open opportunities for future research.

REFERENCES

- Alenezi, A., & Alenezi, A. (2025). Evaluating the Effectiveness of Chatbot-Assisted Learning in Enhancing English Conversational Skills Among Secondary School Students. *Education Sciences*, 15(9), 1136. <https://doi.org/10.3390/educsci15091136>
- Aziz, A. A., & Kashinathan, S. (2021). ESL learners' challenges in speaking English in Malaysian classroom. *Development*, 10(2), 983-991. <https://doi.org/10.6007/ijarped/v10-i2/10355>
- Brown, H. D., & Abeywickrama, P. (2019). *Language assessment: Principles and classroom practices* (3rd ed.). Pearson.
- Bustamante Salinas, P. M., Rodríguez Maida, A. A., & Quisbert Pinedo, L. (2026). Integración de recursos de inteligencia artificial en los procesos educativos. *Revista Tecnología, Ciencia Y Educación*, (33), 139–167. <https://doi.org/10.51302/tce.2026.24177>
- Cha, J., Han, J., Yoo, H., & Oh, A. (2024). CHOP: Integrating ChatGPT into EFL Oral Presentation Practice. *arXiv preprint arXiv:2407.07393*. <https://doi.org/10.48550/arXiv.2407.07393>
- Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th ed.). Routledge.
- Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment—Companion volume*. Council of Europe Publishing. <https://www.coe.int/en/web/common-european-framework-reference-languages>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
- Cuevas-Montero, R. (2021). Desarrollo e integración de las destrezas orales en la educación a distancia a través de podcasts colaborativos. En J. Martín Párraga & J. M. Garrido Anguita (Eds.), *De Platón al "Homo Tecnologicus": Las humanidades en el siglo XXI* (pp. 75–82). Tirant Humanidades.
- Deng, L., & Jamaludin, K. A. (2026). Roles of Generative Artificial Intelligence (GENAI) in English as a Foreign Language (EFL) instruction: A systematic literature review. *SAGE Open*, 16(1). <https://doi.org/10.1177/21582440261418315>
- Du, T. T., Duc, L. V., The, N. H., & Dung, D. T. B. (2025). Challenges faced by EFL students at a university in Vietnam in developing English speaking skills and pedagogical strategies used by instructors. *Multidisciplinary Science Journal*, 8(2), 2026021. <https://doi.org/10.31893/multiscience.2026021>
- Ebadi, S., Velayati, S., Ramezanzadeh, A., & Salman, A. R. (2025). Exploring the impact of AI-powered speaking tasks on EFL learners' speaking performance and anxiety: An activity theory study. *Acta psychologica*, 259, 105391. <https://doi.org/10.1016/j.actpsy.2025.105391>

- Ericsson, E., & Johansson, S. (2023). English speaking practice with conversational AI: Lower secondary students' educational experiences over time. *Computers and Education Artificial Intelligence*, 5, 100164. <https://doi.org/10.1016/j.caeai.2023.100164>
- 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>
- Fauzi, I., Hartono, R., Rukmini, D., & Pratama, H. (2025). AI Applications for EFL Learners: Enhancing Speaking Performance and Reducing Anxiety with Gender-Based Analysis. *Forum for Linguistic Studies*, 7(9). <https://doi.org/10.30564/fls.v7i9.10192>
- Fernández-García, A., & Fonseca-Mora, M. C. (2019). EFL learners' speaking proficiency and its connection to emotional understanding, willingness to communicate and musical experience. *Language Teaching Research*, 26(1), 124–140. <https://doi.org/10.1177/1362168819891868>
- Guerrero, F. S. C., Guillén, L. A. P., Tenesaca, J. R. B., & Herrera, D. C. E. (2026). The Use of Artificial Intelligence to Improve Speaking Fluency in English as a Foreign Language (EFL) Learning. *Arandu UTIC*. <https://doi.org/10.69639/arandu.v13i1.1950>
- Huang, J. (2024). Enhancing EFL speaking feedback with ChatGPT's voice prompts. *International Journal of TESOL Studies*, 6(3). <https://doi.org/10.58304/ijts.20240302>
- Huang, Y., & Kakham, P. (2026). The impacts of AI conversational agents on EFL learners' oral proficiency and foreign language speaking anxiety. *Frontiers in Education*. <https://doi.org/10.3389/feduc.2026.1799269>
- Huu, N., Du, T. T., & Hoang, T. T. (2026). Exploring the Impact of Artificial Intelligence on English Majors' Speaking Skills: A Case Study at a Vietnamese University. *Arab World English Journal*, 3, 166–179. <https://doi.org/10.24093/awej/ai3.11>
- 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>
- Kallio, H., Pietilä, A. M., Johnson, M., & Kangasniemi, M. (2016). Systematic methodological review: Developing a framework for a qualitative semi-structured interview guide. *Journal of Advanced Nursing*, 72(12), 2954–2965. <https://doi.org/10.1111/jan.13031>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., . . . Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>

- Kim, S., & Park, S. (2023). Young Korean EFL Learners' Perception of Role-Playing Scripts: ChatGPT vs. Textbooks. *Korean Journal of English Language and Linguistics*, 23, 1136–1153. <https://doi.org/10.15738/kjell.23..202312.1136>
- Klímová, B., & Seraj, P. M. I. (2023). The use of chatbots in university EFL settings: Research trends and pedagogical implications. *Frontiers in Psychology*, 14, 1131506. <https://doi.org/10.3389/fpsyg.2023.1131506>
- Koç, F. Ş., & Savaş, P. (2024). The use of artificially intelligent chatbots in English language learning: A systematic meta-synthesis study of articles published between 2010 and 2024. *ReCALL*, 37(1), 4–21. <https://doi.org/10.1017/s0958344024000168>
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. L., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, 21(2), 100790. <https://doi.org/10.1016/j.ijme.2023.100790>
- Lo, C. K., Yu, P., Xu, S., Ng, D., & Jong, M. S. (2024). Exploring the application of ChatGPT in ESL/EFL education and related research issues: A systematic review of empirical studies. *Smart Learning Environments*, 11, Article 49. <https://doi.org/10.1186/s40561-024-00342-5>
- Luoma, S. (2004). *Assessing speaking*. Cambridge University Press.
- Mageira, K., Pittou, D., Papasalouros, A., Kotis, K., Zangogianni, P., & Daradoumis, A. (2022). Educational AI chatbots for content and language integrated learning. *Applied Sciences*, 12(7), 3239. <https://doi.org/10.3390/app12073239>
- McNamara, T. F. (1996). *Measuring second language performance*. Longman.
- Nguyen Thi Minh Nguyet, & Pham Thi Thanh Nhan. (2024). *Voice mode of ChatGPT: A new support tool for teaching English in university programs*. *EPRA International Journal of Research & Development (IJRD)*, 9(7). <https://eprajournals.com/IJSR/article/13628/abstract>
- Nhan, L. K., Hoa, N. T. M., & Quang, L. V. N. (2025). Revolutionizing speaking skills improvement: AI's role in personalized language learning. *International Journal of Innovative Research and Scientific Studies*, 8(2), 4637–4649. <https://doi.org/10.53894/ijirss.v8i2.6408>
- Niglas, K., Teddlie, C., & Tashakkori, A. (2021). The use of mixed methods in educational research: A review. *Quality & Quantity*, 55(3), 1007–1025. <https://doi.org/10.1007/s11135-021-01218-3>
- Núñez, A. A. C., Nuñez, M. S. C., Pachay, J. F. Z. P. Z., & Bosquez, A. M. C. B. C. (2025). Using ChatGPT Voice to improve speaking skills in English language learners. *Ciencia Latina*

- Park, H. (2023). Application of ChatGPT for an English learning platform. *STEM Journal*, 24(3), 30–48. <https://doi.org/10.16875/stem.2023.24.3.30>
- Pau, L. C., Yunus, M. M., & Lun, C. W. W. (2026). The impact of popular artificial intelligence tools on English language teaching and learning: a systematic literature review (2021–2025). *International Journal of Research and Innovation in Social Science*, IX(XII), 3804–3827. <https://doi.org/10.47772/ijriss.2025.91200296>
- Rao, P. S. (2019). The importance of speaking skills in English classrooms. *Alford Council of International English & Literature Journal*, 2(2). https://www.researchgate.net/publication/334283040_THE_IMPORTANCE_OF_SPEAKING_SKILLS_IN_ENGLISH_CLASSROOMS
- Rodriguez, S. E. G., & Baquerizo, A. S. M. (2025). Ecuadorian EFL teachers' experiences in fostering students' English-speaking skills: insights into strategies and challenges in public and private schools. *UNESUM - Ciencias Revista Científica Multidisciplinaria*, 9(2), 124–136. <https://doi.org/10.47230/unesum-ciencias.v9.n2.2025.124-136>
- Sosa-López, G., & Mora, J. (2022). The role of speaking anxiety on L2 English speaking fluency, accuracy, and complexity. In J. Levis & A. Guskaroska (Eds.), *Proceedings of the 12th Pronunciation in Second Language Learning and Teaching Conference* (held June 2021, Brock University, St. Catharines, ON). <https://doi.org/10.31274/psllt.13362>
- Susoy, Z. (2026). Reducing anxiety and enhancing performance: the impact of AI chatbots versus human facilitation on EFL speaking assessment outcomes. *Frontiers in Psychology*, 16, 1745942. <https://doi.org/10.3389/fpsyg.2025.1745942>
- Tai, T., & Chen, H. H. (2024). Improving elementary EFL speaking skills with generative AI chatbots: Exploring individual and paired interactions. *Computers & Education*, 220, 105112. <https://doi.org/10.1016/j.compedu.2024.105112>
- Tan, W., & Ismail, H. H. (2025). Malaysian ESL learners' experiences with AI-driven voice practice using ChatGPT and traditional role-play. *International Journal of Education Psychology and Counseling*. <https://doi.org/10.35631/ijepc.1061086>
- Tan, X., Cheng, G., & Ling, M. H. (2024). Artificial intelligence in teaching and teacher professional development: A systematic review. *Computers and Education Artificial Intelligence*, 8, 100355. <https://doi.org/10.1016/j.caeai.2024.100355>
- Tapia, L., & Vega, M. (2024). The impact of Content Learning Integrated Language (CLIL) on the speaking skill in the EFL classroom at the secondary level. *Religación*, 9(40), e2401220. <https://doi.org/10.46652/rgn.v9i40.1220>

- Wei, L. (2023). Artificial intelligence in language instruction: impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology*, 14, 1261955. <https://doi.org/10.3389/fpsyg.2023.1261955>
- Wei, X. (2022). Mixed methods research in language education: Applications and trends. *Language Education and Research*, 5(2), 14–23. <https://ojs.piscomed.com/index.php/LE/article/view/3103>
- Xing, C. (2025). A Systematic Review on Artificial intelligence (AI) Technologies in ESL/EFL speaking Skills. *International Journal of TESOL Studies*. <https://doi.org/10.58304/ijts.250908>
- Yan, H., & Singh, M. K. S. (2026). The impact of AI-mediated instruction on speaking proficiency, enjoyment, anxiety, and emotional engagement: A mixed-methods approach. *Humanities & Social Sciences Communications*. <https://doi.org/10.1057/s41599-026-06705-2>
- 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, C., Meng, Y., & Ma, X. (2024). Artificial intelligence in EFL speaking: Impact on enjoyment, anxiety, and willingness to communicate. *System*, 121, 103259. <https://doi.org/10.1016/j.system.2024.103259>
- Shazly, R. E. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems*, 38. <https://doi.org/10.1111/exsy.12667>