Ciencia Latina Revista Científica Multidisciplinar, Ciudad de México, México.
ISSN 2707-2207 / ISSN 2707-2215 (en línea), enero-febrero 2025,
Volumen 9, Número 1.

https://doi.org/10.37811/cl_rcm.v9i1

USING CHATGPT VOICE TO IMPROVE SPEAKING SKILLS IN ENGLISH LANGUAGE LEARNERS

USO DE CHATGPT VOICE PARA MEJORAR LAS HABILIDADES DE HABLA EN ESTUDIANTES DEL IDIOMA INGLÉS

Angel Alexander Carrera Nuñez
Universidad Estatal de Milagro, Ecuador

Marlene Stefania Carrera Nuñez
Universidad Estatal de Milagro, Ecuador

Jorge Francisco Zambrano Pachay
Universidad Estatal de Milagro, Ecuador

Angel Moises Carrera Bosquez
Unidad Educativa Cascales, Ecuador

DOI: https://doi.org/10.37811/cl_rcm.v9i1.16390

Using ChatGPT Voice to Improve Speaking Skills in English Language Learners

Angel Alexander Carrera Nuñez[1]
aalexcarrera9@gmail.com
https://orcid.org/0009-0003-6905-3637
Universidad Estatal de Milagro
Ecuador

Marlene Stefania Carrera Nuñez
marstefanicarrera@gmail.com
https://orcid.org/0009-0005-5852-9403
Universidad Estatal de Milagro
Ecuador

Jorge Francisco Zambrano Pachay
jzambranop10@unemi.edu.ec
https://orcid.org/0000-0001-9456-2765
Universidad Estatal de Milagro
Ecuador

Angel Moises Carrera Bosquez
moises.carrera@educacion.gob.ec
https://orcid.org/0009-0003-3347-6398
Unidad Educativa Cascales
Ecuador

ABSTRACT

This study examined the impact of ChatGPT Voice on improving speaking skills among Ecuadorian English as a Foreign Language (EFL) learners. Forty-nine third-year high school students from Unidad Educativa Cascales engaged in three weeks of AI-driven speaking practice focused on health-related topics. Using a quasi-experimental design with pre- and post-tests (assessed via Speechace), four parameters were analyzed: pronunciation, fluency, vocabulary, and grammar. Results demonstrated statistically significant improvements across all areas ($p < 0.001$), with large effect sizes ($r = 0.84–0.90$). Fluency and vocabulary showed the most notable gains (Δ +2.8 and +2.7 points, respectively). The findings indicate that ChatGPT Voice reduces language anxiety by offering a non-judgmental practice environment, thereby encouraging active participation. This research underscores artificial intelligence's potential to democratize access to authentic conversational practice in resource-limited educational contexts, positioning it as a valuable pedagogical complement for EFL instruction.

Keywords: ChatGPT Voice, speaking skills, learning, english, artificial intelligence

[1] Autor principal.
Correspondencia: aalexcarrera9@gmail.com

Uso De Chatgpt Voice para Mejorar las Habilidades de Habla en Estudiantes del Idioma Inglés

RESUMEN

Este estudio investigó el impacto de ChatGPT Voice en el desarrollo de las habilidades orales de estudiantes ecuatorianos de inglés como lengua extranjera (ILE). Participaron 49 estudiantes de tercer año de bachillerato de la Unidad Educativa Cascales, quienes utilizaron la herramienta durante tres semanas en conversaciones temáticas sobre salud. Mediante un diseño cuasi-experimental con pruebas pre y post intervención (evaluadas mediante Speechace), se analizaron cuatro parámetros: pronunciación, fluidez, vocabulario y gramática. Los resultados mostraron mejoras significativas en todas las áreas ($p < 0.001$), con tamaños de efecto grandes ($r = 0.84–0.90$). La fluidez y el vocabulario registraron los mayores avances (Δ +2.8 y +2.7 puntos, respectivamente). Los hallazgos sugieren que ChatGPT Voice reduce la ansiedad lingüística al proporcionar un entorno de práctica no crítico, fomentando la participación activa. Este estudio evidencia el potencial de la inteligencia artificial para democratizar el acceso a prácticas conversacionales auténticas en contextos con recursos educativos limitados, destacando su utilidad como complemento pedagógico en ILE.

***Palabras clave*:** ChatGPT Voice, habilidades orales, aprendizaje, inglés, inteligencia artificial

Artículo recibido 17 enero 2025
Aceptado para publicación: 22 febrero 2025

INTRODUCTION

Technology has significantly transformed language education by offering dynamic tools that facilitate interactive and personalized learning experiences. In particular, Artificial Intelligence (AI)-powered conversational agents have emerged as promising platforms for language practice, as they can simulate real-time discussions and offer instantaneous feedback. Despite these advancements, many learners of English as a foreign language (EFL) continue to face challenges in developing their speaking skills. Factors such as limited opportunities for authentic interaction, anxiety about making mistakes, and lack of immediate corrective feedback can impede oral proficiency.

One of the latest AI-driven solutions is ChatGPT Voice, which leverages natural language processing to engage students in immersive, low-stress dialogues. This technology allows learners to practice spoken English in realistic scenarios, replicating the give-and-take of natural conversations. While these features show great potential for addressing key barriers in oral language development, such as lack of interaction and fear of judgment—there is still insufficient empirical evidence regarding their effectiveness in enhancing speaking skills.

The present study seeks to fill this gap by investigating the benefits, challenges, and best practices of using Voice ChatGPT to bolster oral competence among English learners. Through a focused examination of learners' experiences and speaking performance before and after using Voice ChatGPT, this research aims to shed light on how AI-powered conversation can foster confidence, improve fluency, and reinforce accurate pronunciation and grammar in EFL settings. Ultimately, the findings will offer practical insights for instructors and students interested in integrating Voice ChatGPT into language curricula, promoting more engaging and effective pathways for developing communicative competence in English.

LITERATURE REVIEW

Speaking skills are fundamental in the process of language acquisition, serving as the primary means through which learners express themselves and engage in authentic communication. Unlike receptive skills such as listening and reading, speaking requires active language production, directly connecting learners to social interactions and allowing for immediate feedback and adjustment (Denston, 2021).

Proficiency in speaking not only builds learners' confidence but also enhances their ability to participate in meaningful conversations, a key component of communicative competence (Qasserras, 2023).

Traditional language teaching methods have employed various strategies to develop speaking skills, including guided classroom discussions, oral presentations, and the use of language laboratories. While these approaches aim to simulate conversational situations, they often fail to replicate the dynamic and unpredictable nature of real-life dialogues (Gutierrez-Heras & Villacrés, 2022). Moreover, classroom settings may inhibit learners due to fear of negative evaluation from peers and instructors, leading to reluctance in speaking and hindering progress (Yılmaz & De Jong, 2024). This underscores the need for environments that provide supportive and low-pressure opportunities for learners to practice speaking.

The advent of technology has introduced new possibilities for language learning, particularly in overcoming the limitations of traditional speaking practice. (Haleem et al., 2022) observes that digital platforms offer learners the flexibility to practice at their own pace, receive immediate feedback, and interact with simulated conversational partners. These tools align with the shift toward learner-centered approaches; (Gunawardena et al., 2024) highlights that they provide personalized learning experiences that cater to individual needs and preferences. Technology-enhanced language learning has been recognized for its potential to enhance engagement and motivation among learners, as (Hossain, 2024) acknowledges.

Recent advancements in natural language processing have led to the development of conversational AI tools, such as chatbots, that engage users in real-time, interactive dialogues. These tools provide several benefits for language learners. Firstly, they provide low-anxiety practice environments where learners can practice speaking without fear of judgment, reducing language anxiety and promoting willingness to communicate (Du & Ben Kei, 2024). Secondly, conversational AI can simulate authentic interactions, allowing learners to handle diverse topics and conversational flows, which enhances fluency and communicative competence (Wang et al., 2024) . Additionally, the accessibility and flexibility of AI tools enable learners to practice at any time and from any location, facilitating consistent practice and accommodating different learning schedules (Donghwa et al., 2023).

Within this context, specific AI tools have been developed to enhance speaking skills. ChatGPT Voice,

for example, is an AI-powered tool designed to facilitate natural and interactive speaking practice. It enables users to engage in conversations resembling interactions with a human interlocutor, making practice more enjoyable and practical. This tool is especially beneficial for language learners as it helps build confidence, refined pronunciation, and improve fluency without the pressure of a live audience. Its ability to adapt to individual needs and provide instant feedback creates a supportive environment for learners to strengthen their communication skills and apply them in real-world situations.

Integrating ChatGPT Voice into language learning offers several advantages beyond what traditional methods can provide. It offers immediate feedback and error correction, enabling learners to recognize and adjust their errors in real-time, reinforcing accurate pronunciation and language use (Cavalcanti et al., 2021). Furthermore, this tool creates a supportive learning environment by providing a non-judgmental space that encourages learners to practice speaking more freely, thereby reducing anxiety and increasing motivation (Annamalai et al., 2023). Additionally, the interactive and dynamic nature of AI tools can increase learner engagement and sustain motivation over time (Yuan & Liu, 2024).

Similarly, Speechace is an advanced tool specifically designed to streamline the evaluation of oral tests and generate effective reports. Utilizing cutting-edge voice recognition technology, Speechace accurately transcribes students' responses, providing educators with a fast and reliable method to record and analyze oral performance (Ningsih, 2024). It organizes results into clear, detailed reports, highlighting students' strengths and areas for improvement. This platform not only saves time in grading but also ensures a more objective and standardized evaluation, enhancing the effectiveness of teaching and learning processes.

The integration of conversational AI into language learning represents a significant advancement in the development of speaking skills. These technologies have the potential to transform traditional language education methods. This theoretical framework underscores the transformative role that technology, particularly AI, can play in enhancing speaking proficiency among English language learners.

METHODS

Research Design

A total of 95 third-year high school students from Unidad Educativa Cascales, a public institution located in Cascales, province of Sucumbíos, in the Ecuadorian Amazon, formed the initial population.

The students were distributed across three academic parallels (A, B, and C) and specialized in the Science track, emphasizing biology, chemistry, and environmental sciences as part of Ecuador's standardized high school curriculum. This specialization aligns with the study's focus on health-related topics during the AI-driven speaking practice. To determine the sample size, Yamane's formula for finite populations was applied at a 95% confidence level and a 10% margin of error. This formula, proposed by Taro Yamane in *"Statistics: An Introductory Analysis"*, is particularly suited for smaller populations and ensures proportional representation while balancing statistical rigor (Ahmed, 2024). The calculation is as follows:

$$n = \frac{N}{1 + N.e^2} = \frac{95}{1 + 95.(0.10)^2} \approx 48.72$$

Rounding to the nearest whole number yielded a sample size of 49 students. To ensure proportional representation across parallels, a stratified sampling procedure was employed. Quotas were aligned with each parallel's share of the total population (A: 33%, B: 34%, C: 33%), and participants were randomly selected within their respective strata. Eligibility criteria included: (1) current enrollment in the third year of high school, (2) participation in English as a foreign language classes, and (3) commitment to the full duration of the study.

The selection process produced a final sample of 49 students, all aged 16 to 18 and specializing in science. This approach maintained a homogeneous academic background while preserving proportional representation across all parallels. The use of Yamane's formula reinforced the statistical validity of the sample size, ensuring it aligned with the study's finite population context.

Instruments

This study relies on three main instruments to measure participants' speaking skills and assess the effectiveness of ChatGPT Voice as an intervention:

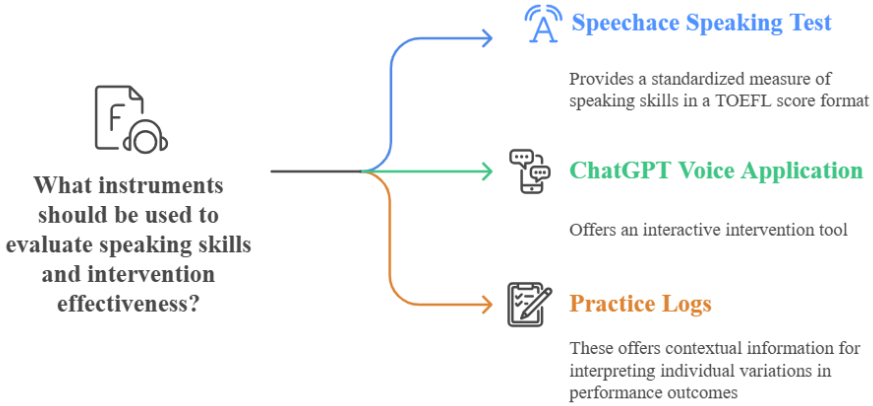
The Speechace Speaking Test provides an automated assessment of English oral proficiency aligned with the TOEFL scoring scale (0–30 points). It evaluates four key parameters: pronunciation, fluency, vocabulary, and grammar. Speechace analyzes spoken responses to standardized prompts and assigns numerical scores based on TOEFL criteria. It focuses on quantifying performance through standardized metrics. The test was administered before (pre-test) and after (post-test) the intervention to measure

improvements in each parameter and the total score.

ChatGPT Voice serves as the primary intervention tool. As an AI-powered conversational agent, it simulates authentic dialogues in a supportive, low-pressure environment. For three weeks, participants engage in regular practice sessions using ChatGPT Voice, following guidelines set by the study.

Each participant maintains a personal practice log to document engagement with ChatGPT Voice. These logs record each session's date, duration, and main topics, helping monitor adherence to the prescribed practice schedule. In addition, the practice logs offer contextual information for interpreting individual variations in performance outcomes.

Figure 1 Instruments to evaluate speaking skills and intervention effectiveness



Procedure

The study was carried out in three distinct phases: Pre-Test Phase, Intervention Phase, and Post-Test Phase. Each phase is designed to systematically assess the impact of AI-driven speaking practice on students' speaking proficiency.

Pre-Test Phase

Prior to the intervention, a baseline assessment will be conducted to measure students' initial speaking proficiency. The Speechace Speaking Test will be used as the primary evaluation tool, which provides an automated scoring system aligned with the TOEFL scale (0–30 points). This assessment will evaluate four key components of speaking skills:

1. Pronunciation: Accuracy in phoneme articulation and intelligibility.
2. Fluency: Naturalness and coherence in speech production.
3. Vocabulary: Appropriateness and range of lexical choices.

4. Grammar: Correctness in syntactic structures.

The test will be administered individually in the students' free time, allowing them to complete the assessment at their own convenience. This approach aims to reduce external pressures and scheduling conflicts, ensuring that students can focus entirely on their performance. The selected topic for the test, "Health", aligns with the interdisciplinary project of the current trimester. This alignment ensures that students have a foundational understanding of the vocabulary and concepts related to the topic.

The pre-test scores will serve as a baseline for comparison with post-test results, providing quantitative data on students' initial speaking proficiency before the implementation of the intervention.

Intervention Phase

Following the pre-test, a three-week structured speaking practice program will be implemented, where students will engage in interactive conversations using ChatGPT's voice tool. Each week will focus on a specific health-related theme relevant to their curriculum, ensuring meaningful engagement and vocabulary reinforcement.

Students will be required to: engage in at least 30 minutes of speaking practice per week, distributed across multiple sessions, use guided prompts related to the week's topic to encourage structured yet spontaneous conversation and maintain a personal log recording of each session's date, duration, and key takeaways.

Each week's topic and suggested prompts are detailed below:

Table 1 Weekly Speaking Practice Topics and Guided Prompts

Week	Topic	Suggested Prompts
1	Common Health Problems in Ecuador	What are the most common health problems in Ecuador? How can they be prevented?
2	Daily Health Care	What are the most important personal health habits? How do Ecuadorian traditional remedies compare to modern medicine?
3	Pros and Cons of Free Healthcare	What are the benefits and challenges of free healthcare in Ecuador? What improvements would you suggest?

Throughout the intervention, students will engage in at least three speaking sessions per week, ensuring

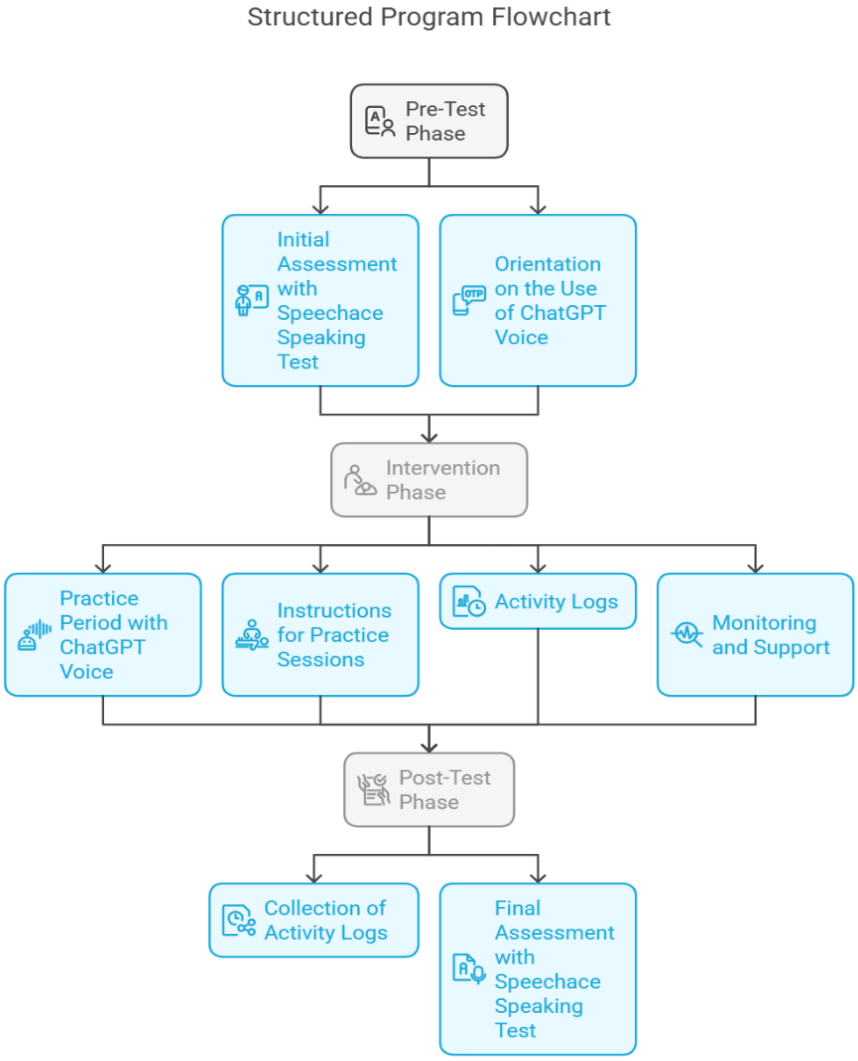
consistent practice.

Post-Test Phase

After completing the three-week intervention, students will undergo a post-test assessment using the same Speechace Speaking Test as in the pre-test. This test will be administered under the same controlled conditions to maintain consistency and validity in the results. The post-test aims to measure improvements in speaking proficiency, specifically in: pronunciation accuracy, fluency in speech production, expansion of vocabulary usage and enhanced grammatical accuracy.

The pre-test and post-test scores will be statistically analyzed to determine the effectiveness of the AI-driven speaking intervention. A paired sample t-test will be conducted to assess whether the observed differences in mean scores are statistically significant.

Figure 2 Enhancing Language Skills: A Structured Approach



Data Analysis

The data collected from the pre-test and post-test were processed using the Real Statistics add-in for Excel. Initially, the Shapiro-Wilk test was performed to determine whether each variable followed a normal distribution. For those variables that met the normality assumption, a paired Student's t-test was conducted, accompanied by Cohen's *d* to gauge effect size. In cases where the data did not meet the normality assumption, the Wilcoxon signed-rank test was employed as a nonparametric alternative, reporting *Z*-scores and the corresponding effect size (*r*). All statistical analyses were performed at a 95% confidence level ($\alpha = 0.05$). This dual focus on statistical significance and effect size provided a comprehensive evaluation of whether, and to what extent, the intervention influenced students' speaking proficiency.

RESULTS AND DISCUSSION

The study involved 49 students (aged 16–18) specializing in science. Participants' speaking scores, measured by the Speechace Speaking Test, were obtained before and after the three-week intervention with ChatGPT Voice. Table 1 presents the descriptive statistics (mean, standard deviation, median, and range) for pronunciation, fluency, vocabulary, grammar, and total scores at both pre-test and post-test stages.

Table 2 Descriptive Statistics of Pre-Test and Post-Test Scores for Each Parameter

Parameter	Pre-test	Post-test
Pronunciation	16.2 ± 4.76, Median 17 (Range 18)	18.7 ± 4.48, Median 19 (Range 17)
Fluency	16.4 ± 4.62, Median 16 (Range 18)	19.2 ± 4.29, Median 19 (Range 16)
Vocabulary	18.3 ± 5.17, Median 19 (Range 16)	21.0 ± 4.63, Median 22 (Range 15)
Grammar	16.4 ± 5.46, Median 15 (Range 24)	18.7 ± 4.88, Median 18 (Range 22)
Total	16.8 ± 3.44, Median 17 (Range 14)	19.4 ± 3.16, Median 19 (Range 14)

Before selecting appropriate inferential tests, the distribution of difference scores (post-test minus pre-test) was examined using the Shapiro-Wilk test. Results indicated that pronunciation, fluency,

vocabulary, and grammar difference scores did not meet normality assumptions ($p < 0.05$), whereas the total score difference demonstrated no significant deviation from normality ($p > 0.05$). Consequently, Wilcoxon Signed-Rank tests were conducted for the non-normal parameters, and a paired t-test was applied to the total score.

Table 3 Normality Assessment Results

Parameter	W-statistic	p-value	Normality	Inferential Test
Pronunciation	0.6992	1.02×10^{-8}	No	Wilcoxon
Fluency	0.7971	9.15×10^{-7}	No	Wilcoxon
Vocabulary	0.8822	0.000152	No	Wilcoxon
Grammar	0.6717	3.39×10^{-9}	No	Wilcoxon
Total	0.9666	0.1771	Yes	Paired t-test

Since the difference scores for pronunciation, fluency, vocabulary, and grammar did not follow a normal distribution (as determined by the Shapiro-Wilk test), the Wilcoxon Signed-Rank test was applied to compare pre-test and post-test scores for these parameters. Table 3 presents the Wilcoxon results, including Z-scores, p-values, and effect sizes (r).

Table 4 Wilcoxon Signed-Rank Results for Pronunciation, Fluency, Vocabulary, and Grammar

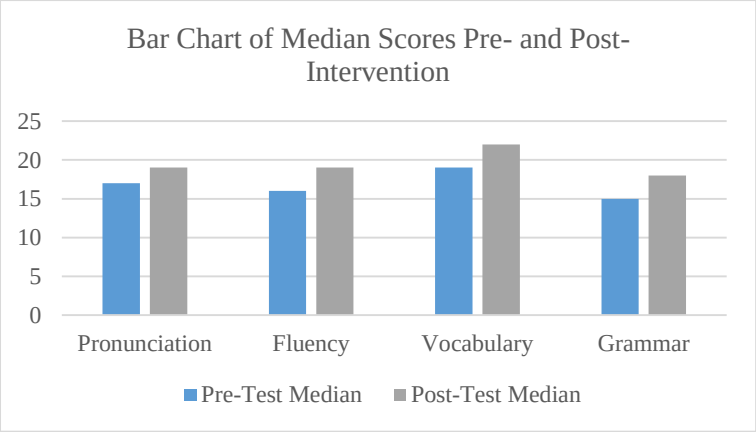
Parameter	Pre-Test Median	Post-Test Median	Z-score	Effect r	p-exact (Two-tail)
Pronunciation	17	19	6.2683	0.895	3.55×10^{-15}
Fluency	16	19	6.0941	0.888	1.42×10^{-14}
Vocabulary	19	22	6.1512	0.887	7.11×10^{-15}
Grammar	15	18	5.7532	0.839	6.13×10^{-11}

Across all four parameters, the Wilcoxon Signed-Rank test yielded statistically significant improvements ($p<0.05$), confirming that post-test scores were consistently higher than pre-test scores. The Z-scores ranged from 5.7532 to 6.2683, reinforcing the strength of the observed differences. Furthermore, effect size estimates (r) ranged from 0.84 to 0.90, indicating that these improvements were statistically significant and highly meaningful in practical terms. According to commonly accepted

benchmarks, effect sizes above 0.80 are generally considered large, which underscores the robust impact of the intervention on learners' speaking performance.

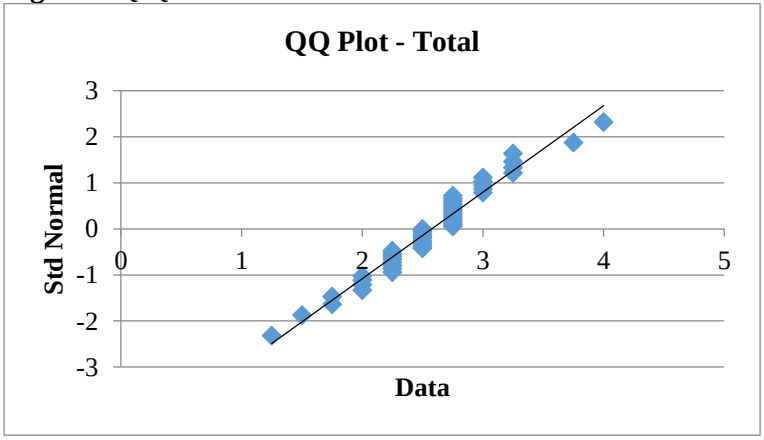
Figure 3 visually represents these improvements, showing the median scores for pre-test and post-test across the four parameters. The clear upward shift in scores highlights the effectiveness of the intervention in enhancing students' speaking performance.

Figure 3 Median Scores for Speaking Parameters Before and After Intervention



Before conducting the paired samples t-test, the normality of the total difference score was assessed using the Shapiro-Wilk test and a Q-Q plot. The Shapiro-Wilk test yielded a p-value of 0.177 ($p > 0.05$), indicating that the normality assumption was not violated. Additionally, the Q-Q plot (Figure 4) shows that the data points closely follow the diagonal reference line, further confirming the suitability of a parametric approach.

Figure 4 Q-Q Plot for Total Difference Score



To assess the impact of the intervention on students' overall speaking proficiency, a paired samples t-test was performed. The analysis confirmed a significant improvement, with post-test scores exceeding

pre-test scores by an average of 3 points. The 95% confidence interval (-2.73 to -2.43) indicates that this increase is highly unlikely to be due to random variation.

The test yielded a t-value of -34.62 (df = 48) and a p-value < 0.001, confirming statistical significance. Furthermore, the effect size was exceptionally large (Cohen's d = 4.95, r = 0.98), suggesting that the observed gains were not only statistically significant but also substantially meaningful in practical application.

Table 5 Paired t-Test Results for Total Speaking Score

	Sample Size	**Pre-Test Mean (SD)**	**Post-Test Mean (SD)**	**Mean Difference**	**t-value**	**df**	**Cohen's d**	**Effect r**	**p-value (Two-Tail)**	**95% CI Lower**	**95% CI Upper**
Value	49	17 (±3.44)	19 (±3.16)	-3	-34.62	48	4.95	0.98	1.31×10^{-35}	-2.73	-2.43

Table 5 presents the detailed statistical breakdown, reinforcing the robustness of the findings. These results provide compelling evidence that AI-driven speaking practice led to substantial and measurable improvements in students' speaking proficiency.

The results of this study demonstrate that implementing ChatGPT Voice in English as a foreign language (EFL) oral practice led to significant improvements in key competencies such as pronunciation, fluency, vocabulary, and grammar. The effect sizes (r = 0.84–0.90) indicate statistical significance and a meaningful pedagogical impact, supporting the premise that conversational AI tools can transform language learning (Gunawardena et al., 2024). These findings align with technology-mediated language acquisition theory (Pan & Sun, 2024), which posits that sustained interaction with intelligent systems accelerates productive skill development by combining contextualized practice and adaptive feedback. Furthermore, prior studies on AI in education (Zawacki et al., 2019) emphasize that personalization and immediate feedback are critical factors for optimizing autonomous learning, consistent with the outcomes observed here.

The reduction in language anxiety, attributed to ChatGPT Voice's non-judgmental environment, corroborates research grounded in affective filter theory (Andreou & Baseki, 2012; (Dewaele & Macintyre, 2024). In traditional settings, public correction may inhibit participation (Kauhanen et al., 2024), whereas AI allows learners to experiment without fear of ridicule, as evidenced by notable

fluency gains (+2.8 points). This finding resonates with recent studies highlighting the role of digital tools in democratizing oral practice (Godwin-Jones, 2022), particularly in contexts with limited access to native speakers or trained instructors—a common reality in non-English-speaking countries like Ecuador (Tafazoli, 2024).

A distinctive aspect of this study is its interdisciplinary focus on public health, enabling learners to connect English practice with prior knowledge from their scientific specialization. This holistic approach aligns with Content and Language Integrated Learning (CLIL) principles (Harrop, 2012); (Esteban, 2015), where thematic contextualization enhances motivation and technical vocabulary retention. For instance, the significant vocabulary improvement (+2.7 points) suggests learners practiced linguistic structures and reinforced discipline-specific terminology, a benefit supported by STEM education research (Seitenova et al., 2023).

The practical relevance of these skills beyond the classroom also warrants attention. This experiential format, coupled with instant feedback, explains robust gains in fluency and vocabulary, findings congruent with learner-centered approaches (Ghorbandordinejad & Kenshinbay, 2024). Additionally, studies on gamification and engagement (Yuan & Liu, 2024) suggest that AI's interactivity enhances intrinsic motivation, a key factor in sustaining long-term practice adherence.

However, certain limitations warrant caution. The short intervention duration (three weeks) hinders the assessment of long-term skill retention, a challenge documented in similar studies (Dunlosky et al., 2013); (Jaeggi et al., 2011). Furthermore, the sample was confined to one Ecuadorian public high school, limiting generalizability. Future research could extend the intervention period, diversify student populations, and integrate qualitative analyses to deepen understanding of AI's mechanisms for oral skill development (Essel et al., 2024); (Maphoto et al., 2024). Exploring synergies between AI and blended learning methodologies could further maximize pedagogical benefits (Tan et al., 2025).

CONCLUSION

ChatGPT Voice demonstrated significant effectiveness in enhancing learners' oral proficiency, with measurable gains in pronunciation clarity, speech fluency, lexical diversity, and grammatical accuracy, attributable to repetitive AI-mediated practice and instant corrective feedback.

The non-judgmental, low-pressure environment provided by ChatGPT Voice substantially decreased learners' speaking anxiety, fostering increased participation and risk-taking in language production, which directly correlated with improved fluency and confidence.

Integrating health-related topics aligned with learners' science curriculum (CLIL methodology) enhanced contextualized language acquisition, enabling students to transfer domain-specific knowledge into meaningful English communication, thereby reinforcing both technical vocabulary retention and interdisciplinary competence.

The interactive nature of ChatGPT Voice, including real-time dialogue simulation and adaptive responses, boosted learner engagement and intrinsic motivation, creating a dynamic practice ecosystem that mirrors authentic conversational scenarios.

While results were promising, the short intervention period (3 weeks) and homogeneous sample (science-track students in rural Ecuador) limit generalizability. Future studies should extend the duration, diversify participant demographics, and incorporate qualitative insights to explore long-term retention and cultural applicability of AI tools in EFL contexts.

BIBLIOGRAPHIC REFERENCES

Ahmed, S. K. (2024). How to choose a sampling technique and determine sample size for research: A simplified guide for researchers. *Oral Oncology Reports*, *12*, 100662. https://doi.org/10.1016/j.oor.2024.100662

Andreou, G., & Baseki, J. (2012). Phonological and Spelling Mistakes among Dyslexic and Non-Dyslexic Children Learning Two Different Languages: Greek vs English. *Psychology*, *3*(8), Article 8. https://doi.org/10.4236/psych.2012.38089

Annamalai, N., Eltahir, M. E., Zyoud, S. H., Soundrarajan, D., Zakarneh, B., & Al Salhi, N. R. (2023). Exploring English language learning via Chabot: A case study from a self determination theory perspective. *Computers and Education: Artificial Intelligence*, *5*, 100148. https://doi.org/10.1016/j.caeai.2023.100148

Cavalcanti, A. P., Barbosa, A., Carvalho, R., Freitas, F., Tsai, Y.-S., Gašević, D., & Mello, R. F. (2021). Automatic feedback in online learning environments: A systematic literature review. *Computers and Education: Artificial Intelligence*, *2*, 100027. https://doi.org/10.1016/j.caeai.2021.100027

Denston, A. (2021). *Review of the Literature on Speaking and Listening Skills*. College of Education, Health, and Human Development University of Caterbury. https://assets.education.govt.nz/public/Documents/Ministry/Changes-in-education/Review-of-the-Literature-on-Speaking-and-Listening-Skills-.pdf

Dewaele, J.-M., & Macintyre, P. (2024). The two faces of Janus? Anxiety and Enjoyment in the Foreign Language Classroom. *ResearchGate*, *4*(2), 100–119. https://doi.org/10.14746/ssllt.2014.4.2.5

Donghwa, L., Hong-hyeon, K., & Seok-Hyun, S. (2023). *Development research on an AI English learning support system to facilitate learner-generated-context-based learning*. *5*(71), 629–666. https://doi.org/10.1007/BF02961473

Du, J., & Ben Kei, D. (2024). *Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice—ScienceDirect*. *6*. https://doi.org/10.1016/j.caeai.2024.100230

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving Students' Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Essel, H. B., Vlachopoulos, D., Essuman, A. B., & Amankwa, J. O. (2024). ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based conversational large language models (LLMs). *Computers and Education: Artificial Intelligence*, *6*, 100198. https://doi.org/10.1016/j.caeai.2023.100198

Esteban, S. G. (2015). Integrating Curricular Contents and Language through Storytelling: Criteria for Effective CLIL Lesson Planning. *Procedia - Social and Behavioral Sciences*, *212*, 47–51. https://doi.org/10.1016/j.sbspro.2015.11.297

Ghorbandordinejad, F., & Kenshinbay, T. (2024). Exploring AI-Driven Adaptive Feedback in the Second Language Writing Skills Prompt: AI Technology in Language Teaching. *EIKI Journal of Effective Teaching Methods*, *2*(3), Article 3. https://doi.org/10.59652/jetm.v2i3.264

Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology*, *26*(2), 5–24.

Gunawardena, M., Bishop, P., & Aviruppola, K. (2024). Personalized learning: The simple, the complicated, the complex and the chaotic. *Teaching and Teacher Education*, *139*, 104429. https://doi.org/10.1016/j.tate.2023.104429

Gutierrez-Heras, C., & Villacrés, M. (2022). *Communicative Language Teaching to Enhance Speaking Skills in B1.1 Level*. *7*(4), 1304–1318. https://doi.org/10.23857/pc.v7i4.3891

Haleem, A., Javaid, M., Qadri, M. A., & Suman, R. (2022). Understanding the role of digital technologies in education: A review. *Sustainable Operations and Computers*, *3*, 275–285. https://doi.org/10.1016/j.susoc.2022.05.004

Harrop, E. (2012). Content and Language Integrated Learning (CLIL): Limitations and possibilities. *Encuentro*, *1*(2), 57–70.

Hossain, K. I. (2024). Reviewing the role of culture in English language learning: Challenges and opportunities for educators. *Social Sciences & Humanities Open*, *9*, 100781. https://doi.org/10.1016/j.ssaho.2023.100781

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences*, *108*(25), 10081–10086. https://doi.org/10.1073/pnas.1103228108

Kauhanen, I., Lanas, M., & Kaukko, M. (2024). (Im)possibilities of parity of participation in school settings in the lives of unaccompanied youth. *International Journal of Inclusive Education*, *28*(14), 3422–3436. https://doi.org/10.1080/13603116.2023.2287474

Maphoto, K. B., Sevnarayan, K., Mohale, N. E., Suliman, Z., Ntsopi, T. J., & Mokoena, D. (2024). Advancing Students; Academic Excellence in Distance Education: Exploring the Potential of Generative AI Integration to Improve Academic Writing Skills. *Open Praxis*, *16*(2). https://doi.org/10.55982/openpraxis.16.2.649

Ningsih, Fitria. (2024). *Analyzing students' English-speaking skills using Speechace: insights from an ai-powered assessment tool*. Pedagogic Research-Applied Literacy Journal. 1. 111-124. 10.70574/9w2prx09.

Pan, L.-C., & Sun, J. C.-Y. (2024). An innovative pictographic glosses design for East Asian EFL vocabulary learners: Effects on retention performance and situational interest. *Language Learning & Technology*, *28*(1), 1–23.

Qasserras, L. (2023). Systematic Review of Communicative Language Teaching (CLT) in Language Education: A Balanced Perspective. *European Journal of Education and Pedagogy*, *4*(6), Article 6. https://doi.org/10.24018/ejedu.2023.4.6.763

Seitenova, S., Khassanova, I., Khabiyeva, D., Kazetova, A., Madenova, L., & Yerbolat, B. (2023). The Effect of STEM Practices on Teaching Speaking Skills in Language Lessons. *International Journal of Education in Mathematics, Science and Technology*, *11*(2), 388–406. https://doi.org/10.46328/ijemst.3060

Tafazoli. (2024). Exploring the Potential of Generative AI in Democratizing English Language Education. *ResearchGate*, *1*(2), 1–11. https://doi.org/10.1016/j.caeai.2024.100275

Tan, X., Cheng, G., & Ling. (2025). 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

Wang, K., Lu, Y., & Pan, Z. (2024). Understanding users' effective use of generative conversational AI from a media naturalness perspective: A hybrid SEM-ANN approach. *Data Science and Management*. https://doi.org/10.1016/j.dsm.2024.09.001

Yılmaz, T., & De Jong, E. (2024). The Role of Native Speaker Peers on Language Learners' Fear of Negative Evaluation and Language Anxiety. *Boğaziçi Üniversitesi Eğitim Dergisi*, *41*(1), 1–17. https://doi.org/10.52597/buje.1389669

Yuan, L., & Liu, X. (2024). The effect of artificial intelligence tools on EFL learners' engagement, enjoyment, and motivation. *Computers in Human Behavior*, *162*, 108474. https://doi.org/10.1016/j.chb.2024.108474

Zawacki, O., Marín, V., Bond, M., & Governeur, F. (2019). Systematic review of research on artificial intelligence applications in higher education -where are the educators? *ResearchGate*, *16*(39), 12–30. https://doi.org/10.1186/s41239-019-0171-0