Journal of English Language and Education
ISSN 2597- 6850 (Online), 2502-4132 (Print)
Journal Homepage: https://jele.or.id/index.php/jele/index

OPEN ACCESS
Article

ChatGPT Voice Mode to Reduce Speaking Anxiety and Improve Students' English Speaking Ability

https://doi.org/10.31004/jele.v11i6.3396

Roobby Hamdanur, Mutia Elviani, Ghustiva Liani[abc]
[123]Politeknik Negeri Medan, Indonesia
Corresponding Author: roobbyhamdanur@polmed.ac.id

A B S T R A C T

This study aimed to examine changes in students' speaking anxiety and English speaking ability following the use of ChatGPT Voice Mode as a supplementary speaking practice tool. Specifically, it investigated whether students demonstrated changes in speaking anxiety and speaking performance after participating in voice-based practice activities. The study employed a one-group pretest-posttest design involving 26 Business Administration students at State Polytechnic of Medan, Indonesia. Speaking anxiety was measured using an instrument adapted from the Foreign Language Classroom Anxiety Scale (FLCAS), while speaking ability was assessed across six aspects based on Brown and Abeywickrama and the Common European Framework of Reference for Languages (CEFR). During the intervention, students engaged in speaking practice using ChatGPT Voice Mode. [The duration, frequency, and instructional procedures of the intervention should be specified here.] The findings showed that the mean speaking score increased from 48.93 in the pretest to 59.83 in the posttest, indicating a gain of 10.90 points. Meanwhile, the mean speaking anxiety score decreased from 68.88 to 60.69, representing a reduction of 8.19 points. A paired-samples t-test revealed a statistically significant reduction in speaking anxiety, $t(25) = -4.551$, $p < .001$, with a large effect size (d_z = 0.893). For speaking ability, [report the paired-samples t-test result, p-value, confidence interval, and effect size]. Overall, the findings indicate that the use of ChatGPT Voice Mode was associated with improvements in students' speaking performance and reductions in speaking anxiety. These preliminary findings suggest that ChatGPT Voice Mode may serve as a supplementary pedagogical tool for supporting both the affective and linguistic dimensions of English speaking development.

Keywords: *Chatgpt Voice Mode, Speaking Anxiety, English Speaking Ability*

Article History:
Received 12th October 2026
Accepted 31st October 2026
Published 07th November 2026

INTRODUCTION

English speaking ability is an essential component of foreign language learning because it enables learners to communicate ideas, exchange information, express opinions, and participate in social and academic interactions. Unlike receptive skills, speaking requires learners to produce language in real time while simultaneously managing various linguistic and communicative demands. Therefore, developing speaking proficiency requires sufficient opportunities for learners to use language actively and meaningfully (Thornbury, 2005). Furthermore, speaking competence involves multiple dimensions of oral language performance rather than a single linguistic component, as emphasized in speaking assessment frameworks (H. Douglas Brown, 2019). However, having adequate knowledge of English does not necessarily mean that students are willing or confident to communicate orally. Many English as a Foreign Language (EFL) learners experience anxiety when speaking English, particularly in situations that require them to express ideas spontaneously or perform in front of others. Horwitz et al. (1986) conceptualized foreign language anxiety as a situation-specific form of anxiety associated with foreign language learning, which encompasses communication apprehension, fear of negative evaluation, and test anxiety. Such anxiety may influence learners' willingness to communicate and their participation in classroom speaking activities.

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Roobby Hamdanur, et al.

Foreign language speaking anxiety has become an important issue in foreign language education because of its potential implications for students' learning experiences and language development. Horwitz (2001) further explained that language anxiety is an important affective factor associated with learners' experiences and achievement in foreign language learning. Similarly, Scovel (1978) discussed the influence of affective factors on second language learning, highlighting the relevance of anxiety in understanding learners' language development. From the perspective of Krashen (1982) Affective Filter Hypothesis, negative emotional conditions may raise learners' affective filter, potentially making language acquisition more difficult. Recent studies have continued to demonstrate the significance of this issue. Ady & Mardiah (2024), in their systematic review of foreign language speaking anxiety research in Southeast Asia, highlighted the continuing relevance of speaking anxiety in foreign language education. Botes et al. (2020), through a meta-analysis of research using the Foreign Language Classroom Anxiety Scale (FLCAS), also demonstrated the importance of foreign language anxiety in understanding language learning outcomes. Furthermore, Xiong et al. (2024), in a systematic review and meta-analysis of 37 intervention studies, found that learning interventions generally had significant effects on reducing foreign language anxiety. These findings indicate the importance of exploring instructional approaches that can provide supportive learning environments and opportunities for students to develop their speaking confidence.

Alongside these challenges, technological developments have created new opportunities for improving foreign language learning. Artificial Intelligence (AI) has increasingly been incorporated into educational contexts, offering alternative ways for learners to access instructional materials, practice language skills, and receive feedback. In particular, AI-powered applications have the potential to facilitate speaking practice through interactive and responsive communication. Unlike conventional learning materials that generally provide predetermined exercises, conversational AI can offer more flexible opportunities for language production and interaction. The development of generative AI has further expanded these possibilities by enabling systems to generate contextually relevant responses and sustain conversational exchanges. Among these technologies, ChatGPT Voice Mode provides a voice-based interaction feature that allows users to communicate with the system through spoken language. This feature enables students to engage in simulated conversations, answer questions, express opinions, practice vocabulary, and respond spontaneously. Consequently, ChatGPT Voice Mode may provide an alternative environment for English speaking practice, particularly for learners who have limited opportunities to communicate in English or feel uncomfortable speaking in front of others.

Previous studies have reported the potential of AI-powered applications to support students' English speaking development. Several applications, including Cake, ELSA Speak, and ChatGPT, have been explored as supplementary tools for speaking instruction. Aqilah (2023), for example, investigated the use of the Cake application in English speaking learning. Khilma & Azkiyah (2026) reported that the application supported improvements in students' speaking skills, particularly in pronunciation, fluency, and vocabulary. Similarly, Syakila et al. (2026) reported the potential effectiveness of Cake in facilitating students' speaking development. Research by Arosy (2025) also examined the use of ELSA Speak and highlighted its potential to facilitate students' pronunciation practice in storytelling activities. In addition, Istikhomahningsih & Setiawan (2025) found differences in students' speaking ability before and after instruction using the Cake learning application. Collectively, these studies suggest that technology-assisted speaking practice can provide additional opportunities for learners to develop various aspects of oral communication. Nevertheless, the findings also highlight the need to explore emerging AI-based tools that offer more interactive and conversational speaking experiences.

In addition to supporting speaking development, AI-powered applications may also contribute to addressing students' speaking anxiety. Conversational AI can provide opportunities for learners to practice speaking in a less socially demanding environment, allowing them to make mistakes, repeat responses, and develop confidence without the

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Roobby Hamdanur, et al.

immediate pressure of communicating in front of classmates. Dewi Lestari et al. (2025), in their study of AI as a speaking partner, highlighted its potential to alleviate students' speaking anxiety. Similarly, Huang & Liu (2025), through a systematic review of empirical studies published between 2004 and 2024, demonstrated that technology can influence foreign language anxiety, although its effects may vary depending on technological characteristics and learning contexts. More specifically, Ding & Yusof (2025) investigated AI-powered conversation bots and reported their potential to enhance L2 speaking skills while reducing speaking anxiety. Their findings suggest that conversational AI may support both the linguistic and affective dimensions of foreign language learning. These studies indicate that the benefits of AI-assisted speaking practice may extend beyond language performance to include students' emotional experiences during oral communication.

Despite the growing interest in AI-assisted language learning, research examining speaking anxiety and speaking ability simultaneously remains relatively limited. Although Ding & Yusof (2025) investigated both dimensions, their study employed Mondly as the AI-powered conversation platform, which differs from the technology examined in the present study. Meanwhile, Saptiany et al. (2025) explored the use of ChatGPT Voice Mode among Indonesian English for Specific Purposes (ESP) students through a qualitative research design involving interviews. Their findings indicated that ChatGPT Voice Mode provided a supportive and relatively safe environment for speaking practice, allowing students to practice independently, repeat their attempts, and communicate without the same level of fear of negative judgment experienced in face-to-face interactions. However, their study primarily focused on students' perceptions and experiences rather than quantitatively examining changes in speaking anxiety and speaking performance. Therefore, an empirical investigation using a quantitative approach is needed to complement the existing qualitative evidence and provide further insight into the potential of ChatGPT Voice Mode in supporting both dimensions of English speaking learning.

Considering the research gap, the present study investigates changes in students' speaking anxiety and English speaking ability following the use of ChatGPT Voice Mode as a supplementary speaking practice tool. Specifically, it examines students' speaking anxiety and speaking performance before and after participating in ChatGPT Voice Mode activities. By employing a quantitative one-group pretest-posttest design, this study seeks to provide empirical evidence regarding the potential of voice-based conversational AI to support students' affective and linguistic development in English speaking. Accordingly, this study addresses the following research questions: (1) To what extent does students' speaking anxiety change following the use of ChatGPT Voice Mode? and (2) To what extent does students' English speaking ability change following the use of ChatGPT Voice Mode?

METHOD

Research Design

This study employed a quantitative approach using a one-group pretest-posttest design to examine changes in students' speaking anxiety and English speaking ability following the use of ChatGPT Voice Mode. This design was selected because it allowed the researchers to compare the same participants' speaking performance and anxiety levels before and after the intervention. Given the practical context of the study, a comparison group was not included. Therefore, the design provided preliminary evidence of changes associated with the intervention rather than definitive evidence of its causal effectiveness. The participants were 26 first-semester students of the Business Administration Study Program at State Polytechnic of Medan, Indonesia, consisting of 19 females and 7 males. The participants were selected using [convenience/purposive] sampling based on their availability and suitability for the study. The inclusion criteria were that participants were enrolled in the first semester of the Business Administration Study Program, participated in the ChatGPT Voice Mode speaking activities, and completed both pretest and posttest assessments. All 26 participants were included in the final data analysis.

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

The research design can be represented as follows:

$O_1 \rightarrow X \rightarrow O_2$

where:

O_1 = *Pretest (speaking ability assessment and speaking anxiety questionnaire)*

X = *ChatGPT Voice Mode intervention*

O_2 = *Posttest (speaking ability assessment and speaking anxiety questionnaire)*

Instruments

Two instruments were used to collect the data: a speaking anxiety questionnaire and a speaking assessment rubric. Students' speaking anxiety was measured using a questionnaire adapted from the Foreign Language Classroom Anxiety Scale (FLCAS), originally developed by Horwitz et al. (1986). The questionnaire consisted of 18 items covering three dimensions: communication apprehension, fear of negative evaluation, and test anxiety. The questionnaire was administered before and after the intervention to identify changes in students' speaking anxiety. Meanwhile, students' speaking ability was assessed using a speaking rubric adapted from Brown & Abeywickrama (2019) and the Common European Framework of Reference for Languages (CEFR) (Europe, 2020). The rubric covered six dimensions of speaking performance: pronunciation, grammar, vocabulary, fluency, comprehension, and interactive communication. The same assessment criteria were applied in both the pretest and posttest to maintain consistency in evaluating students' speaking performance.

Procedures

The research was conducted through three main stages: pre-intervention assessment, intervention, and post-intervention assessment. At the beginning of the study, students completed a speaking pretest to assess their initial English speaking ability. They were asked to perform a speaking task based on a predetermined topic, and their performances were evaluated using the speaking assessment rubric. After completing the speaking pretest, students filled out the speaking anxiety questionnaire adapted from the FLCAS to measure their initial levels of anxiety when speaking English. These two assessments provided baseline data for evaluating subsequent changes following the intervention.

Following the initial assessments, students received guidance and instructions from the lecturer on how to use ChatGPT Voice Mode as a supplementary tool for English-speaking practice. The guided instruction was delivered over three meetings, with each meeting lasting 150 minutes. In addition to these instructor-led sessions, students were required to undertake nine independent practice sessions using ChatGPT Voice Mode, with each session lasting 120 minutes. Thus, the intervention consisted of 450 minutes of guided instruction and 1,080 minutes of independent practice, resulting in a total of 1,530 minutes (25.5 hours) of intervention activities. During the practice sessions, students used ChatGPT Voice Mode to engage in spoken interactions and practice expressing ideas in English. The activities involved conversational exchanges, responding to questions, explaining ideas, and practicing spontaneous spoken responses. Students interacted with ChatGPT through voice input and received responses through voice output, allowing them to continue conversations and practice oral communication. The lecturer provided initial guidance on using the application and monitored the implementation of the activities.

After completing the intervention, students undertook a speaking posttest using a topic with a level of difficulty equivalent to that of the pretest to assess changes in their speaking ability. Their performances were evaluated using the same speaking rubric applied during the pretest. Subsequently, students completed the same speaking anxiety questionnaire administered before the intervention to measure changes in their anxiety levels. The pretest and posttest speaking scores, together with the pre- and post-intervention anxiety questionnaire scores, were then compared to identify changes in both outcome variables following the use of ChatGPT Voice Mode.

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Roobby Hamdanur, et al.

Data Analysis

Descriptive statistics were used to calculate the mean and standard deviation of the pretest and posttest scores for both speaking ability and speaking anxiety. Before conducting inferential statistical analyses, the Shapiro-Wilk test was performed to examine the normality of the difference scores between the posttest and pretest measurements. The results indicated that the difference scores for speaking ability were not normally distributed; therefore, the Wilcoxon Signed-Rank Test was used to determine whether there was a statistically significant difference between the pretest and posttest speaking scores. In contrast, the difference scores for speaking anxiety were normally distributed; consequently, a paired-samples t-test was conducted to examine changes in students' anxiety levels. Effect sizes were also calculated to determine the magnitude of the observed changes. For the Wilcoxon Signed-Rank Test, the effect size was expressed as r, whereas Cohen's d_z was used for the paired-samples t-test. The statistical significance level was set at $\alpha = 0.05$. All data were analyzed using the Statistical Package for the Social Sciences (SPSS).

FINDINGS AND DISCUSSION

Findings

The data analysis results from the questionnaire and speaking exams are shown in this section. Data analysis was conducted using descriptive statistics, the Shapiro–Wilk normality test, and inferential statistical tests appropriate for the data distribution characteristic.

Descriptive Statistics Analysis

Descriptive statistical analysis was conducted to determine the change in students' mean scores for speaking anxiety and English speaking performance before and after using ChatGPT Voice Mode. The results of the descriptive analysis are presented in Table 1.

Table 1. Descriptive Statistics of Speaking Ability

Variable	N	Pretest Mean	Pretest SD	Posttest Mean	Posttest SD	Gain
Speaking Ability	26	48.93	9.53	59.83	8.76	**10.90**
Speaking Anxiety	26	68.88	9.90	60.69	9.70	**−8.19**

Based on Table 1, the students' mean speaking performance score increased from 48.93 on the pre-test to 59.83 on the post-test. This change demonstrates an average increase of 10.90 points after students utilized ChatGPT Voice Mode. Meanwhile, the mean speaking anxiety score decreased from 68.88 to 60.69, reflecting a reduction of 8.19 points. Since a lower anxiety score indicates a lower level of anxiety, these descriptive findings suggest improvements in both the students' affective state and speaking ability following the use of ChatGPT Voice Mode.

Normality Test

Normality testing was performed using the Shapiro–Wilk test on the difference scores between the post-test and pre-test. This procedure aims to determine the appropriate inferential statistical method. The decision-making criteria were that a significance value greater than 0.05 indicates insufficient evidence to reject the assumption of normality, whereas a significance value less than 0.05 indicates that the normality assumption is not met. The results of the normality test are presented in Table 2.

Table 2. Normality Test of Difference Scores

Variable	N	Shapiro-Wilk W	p	Interpretation
Speaking Difference	26	0.756	< .001	Non-normal
Anxiety Difference	26	0.979	0.841	Normal

The results of the Shapiro–Wilk test showed that the difference score for speaking performance had a significance value (p-value) of 0.000034 ($p < 0.05$). Consequently, the

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Roobby Hamdanur, et al.

distribution of difference scores for speaking performance did not meet the assumption of normality. Therefore, hypothesis testing for the speaking performance variable was conducted using the non-parametric Wilcoxon signed-rank test. Because the speaking difference scores were not normally distributed, the Wilcoxon Signed-Rank Test was employed. The results of the Wilcoxon Signed-Rank Test for Speaking Ability are presented in Table 3.

Table 3. Wilcoxon Signed-Rank Test for Speaking Ability

Comparison	Z	p	Effect Size (r)	Interpretation
Posttest vs. Pretest	4.482	< .001	0.879	Large

The Wilcoxon analysis produced Z = 4.482, p < .001. Therefore, there was a statistically significant difference between students' speaking performance before and after the intervention. The effect size was r = 0.879, which represents a large effect.

In contrast, the score differences for speaking anxiety demonstrated a significance level of $p = 0.841$, exceeding the 0.05 cutoff. This indicates that the normality assumption was satisfied, as there is no statistical evidence that the difference scores deviate from normality. Therefore, a paired-samples t-test was employed for hypothesis testing on the speaking anxiety variable. The results of the paired-samples t-test are presented in Table 4.

Table 4. Paired-Samples t-Test for Speaking Anxiety

Comparison	Mean Difference	t	df	p	Cohen's dz	Interpretation
Posttest − Pretest	−8.19	−4.551	25	< .001	0.893	Large

The paired-samples t-test results revealed a statistically significant difference between pre-test and post-test scores, with a mean difference of −8.19, $t(25) = -4.551$, and $p < 0.001$. The Cohen's d_z value of 0.893 indicates a large effect size. These findings demonstrate a statistically meaningful change in participants' scores following the intervention. However, since the score difference was calculated as post-test minus pre-test and yielded a negative value, the mean post-test score was actually 8.19 points lower than the pre-test score. Thus, assuming a higher score reflects a better outcome, this finding indicates a post-intervention decrease in scores, despite the large effect size

Discussion

Research Question 1: To what extent does the use of ChatGPT Voice Mode reduce students' speaking anxiety when speaking English?

The results indicate that students' speaking anxiety decreased substantially following the use of ChatGPT Voice Mode. The mean score decreased from 68.88 to 60.69, representing an 8.19-point reduction. More importantly, the paired-samples t-test demonstrated that this difference was statistically significant, t(25) = −4.551, p < .001. The effect size of 0.893 was classified as large. These findings indicate that the reduction was not merely a small descriptive change. The magnitude of the effect suggests that the change in anxiety was meaningful within the studied group.

One possible explanation is that ChatGPT Voice Mode provides students with an additional speaking environment in which they can practice without the same immediate social pressure associated with speaking in front of classmates. Students can make mistakes, repeat responses, reformulate utterances, and continue conversations with an AI partner.

This characteristic may be particularly relevant for students who experience fear of negative evaluation. Horwitz et al. (1986) identified communication apprehension and fear of negative evaluation as important dimensions of foreign language classroom anxiety. When students perceive a speaking situation as less threatening, they may become more willing to participate. The finding is consistent with Dewi Lestari et al. (2025), who investigated AI as a speaking partner and reported its potential to alleviate speaking anxiety. It is also consistent with Ding & Yusof (2025), whose study demonstrated the potential of AI-powered conversation bots to simultaneously support L2 speaking development and reduce speaking anxiety. It is also

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Roobby Hamdanur, et al.

consistent with Saptiany et al. (2025) explored the use of ChatGPT Voice Mode among English for Specific Purposes (ESP) students in Indonesia.

The finding can also be interpreted in relation to Krashen (1982) affective filter hypothesis. When learners experience lower levels of anxiety and other negative emotional conditions, the affective barrier to language learning may be reduced. In the present context, ChatGPT Voice Mode may provide students with repeated opportunities to use English without the immediate fear of public evaluation. However, the findings should not be interpreted as proof that ChatGPT Voice Mode alone caused the reduction. The absence of a control group means that other factors, such as repeated practice, increased familiarity with the speaking tasks, or general learning experiences, may also have contributed to the observed reduction. Nevertheless, the combination of an 8.19-point reduction, $p < .001$, and large effect size provides substantial evidence of meaningful change within the participant group.

Research Question 2: To what extent does the use of ChatGPT Voice Mode improve students' English-speaking ability?

The descriptive results show that the mean speaking score increased from 48.93 to 59.83, representing a gain of 10.90 points. Because the difference scores were not normally distributed, the Wilcoxon Signed-Rank Test was used. The test demonstrated a statistically significant difference, $Z = 4.482$, $p < .001$, with an effect size of $r = 0.879$, classified as large. The findings therefore demonstrate a substantial improvement in students' speaking performance following the use of ChatGPT Voice Mode.

One possible explanation is the opportunity for repeated oral practice. Speaking is a productive skill that requires learners to formulate and produce language in real time. Thornbury (2005) emphasizes the importance of practice in developing speaking skills. Through repeated interaction with ChatGPT Voice Mode, students can obtain additional opportunities to produce spoken English beyond conventional classroom interaction.

The six assessment dimensions used in this study—pronunciation, grammar, vocabulary, fluency, comprehension, and interactive communication—also reflect the multidimensional nature of speaking competence. Improvement in speaking may occur because conversational AI allows students to practice several dimensions simultaneously. For example, students must pronounce words while producing responses, select appropriate vocabulary, organize grammatical structures, comprehend the AI's spoken responses, maintain fluency, and respond appropriately within the conversation.

This finding is consistent with previous technology-assisted language learning research. Aqilah (2023), Khilma & Azkiyah (2026), and Syakila et al. (2026) reported positive outcomes associated with technology-supported English speaking learning. Arosy (2025) similarly demonstrated the potential of AI-assisted technology to support pronunciation development. The present study extends these findings by examining a generative voice-based conversational system rather than focusing solely on application-based speaking exercises. Ding & Yusof (2025) provide particularly relevant evidence because they investigated AI-powered conversation bots and their relationship with L2 speaking development. The present findings are consistent with the argument that conversational AI can provide learners with additional opportunities to engage in oral language practice.

From a sociocultural perspective, Vygotsky (1978) theory emphasizes the importance of mediated interaction in learning. Although interaction with AI is not equivalent to interaction with a human teacher or peer, AI can function as an additional mediated environment through which learners practice language and receive conversational input.

The Relationship Between Speaking Anxiety and Speaking Improvement

An important finding of this study is the simultaneous occurrence of two changes: speaking ability increased by 10.90 points, while speaking anxiety decreased by 8.19 points. The effect sizes for both outcomes were also large. Speaking: $r = 0.879$. Speaking anxiety: $|dz| = 0.893$. This pattern is theoretically meaningful because linguistic and affective factors can interact in language learning. Students who experience high anxiety may avoid speaking,

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Roobby Hamdanur, et al.

reducing their opportunities for practice. Conversely, students who obtain more opportunities to practice may gradually become more familiar with speaking English and potentially experience greater confidence.

ChatGPT Voice Mode may support this process by providing a relatively accessible environment for repeated speaking practice. Students can engage in conversation without necessarily having to perform immediately in front of an entire class. The results therefore suggest that ChatGPT Voice Mode may have potential as a dual-purpose pedagogical tool. It can provide opportunities for speaking development while also potentially reducing affective barriers to oral communication. This finding is particularly relevant to the research gap identified earlier. Previous research has often emphasized either speaking performance or anxiety. The present study brings the two outcomes together in one intervention and demonstrates that both changed in favorable directions following the use of voice-based AI.

However, it is important to emphasize that the present study does not statistically test whether anxiety reduction caused speaking improvement. No correlation or mediation analysis was conducted, and the one-group pretest-posttest design does not permit a definitive causal interpretation. The appropriate conclusion is therefore that speaking anxiety decreased and speaking ability increased significantly following the intervention, suggesting that the use of ChatGPT Voice Mode has potential to support both dimensions simultaneously.

The findings of the present study are generally consistent with the growing literature on AI-supported language learning. Aqilah (2023), Khilma & Azkiyah (2026), and Syakila et al. (2026) demonstrated that mobile technology can support English speaking development. Arosy (2025) further showed the potential of AI-supported technology for pronunciation development. However, the present study extends this literature in three important ways. First, it focuses on ChatGPT Voice Mode, which allows dynamic voice-based interaction rather than relying primarily on fixed learning activities. Second, it simultaneously evaluates speaking performance and speaking anxiety. Third, speaking performance is evaluated using six dimensions, providing a broader representation of oral communication.

The findings concerning anxiety are also consistent with Lestari et al. (2025), Ding & Yusof (2025), and Saptiany et al. (2025) who identified the potential of AI-based conversational interaction to support learners experiencing speaking anxiety. The results are also compatible with Huang & Liu (2025) broader conclusion that technology can influence foreign language anxiety. Nevertheless, the present study contributes context-specific evidence concerning a particular type of technology voice-based generative AI.

CONCLUSIONS

This study examined changes in university students' speaking anxiety and English-speaking ability following the use of ChatGPT Voice Mode as a supplementary practice tool. The findings showed a significant decrease in speaking anxiety, with mean scores declining from 68.88 to 60.69, and a significant improvement in speaking ability, increasing from 48.93 to 59.83. Both outcomes demonstrated large effect sizes, indicating meaningful within-group changes. These results suggest that voice-based conversational AI may provide a supportive environment for practicing spoken English, improving speaking performance, and reducing anxiety. The simultaneous improvements highlight its potential to support both affective and linguistic aspects of speaking development. However, the one-group pretest-posttest design without a comparison group limits causal conclusions, as other factors may have influenced the results. Further studies involving larger samples, control groups, and longer interventions are needed to establish stronger evidence of ChatGPT Voice Mode's effectiveness in English speaking instruction.

ACKNOWLEDGEMENTS

The authors would like to express our sincere gratitude to Pusat Penelitian dan Pengabdian kepada Masyarakat (P3M), Politeknik Negeri Medan for providing financial support for this research through the Fundamental Research (Penelitian Fundamental) scheme

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

under the DIPA budget as stipulated in contract number B/329/PL5/PT.01.05/2026. We hope that the findings of this study will contribute to the advancement of English language teaching and learning, particularly through the integration of artificial intelligence in language education.

DECLARATION

Author Contributions: Roobby Hamdanur led the study and manuscript preparation. Mutia Elviani and Ghustiva Liani contributed to data analysis, interpretation, and manuscript revision. All authors approved the final version.
Conflict of Interest: The authors declare no conflict of interest.
Funding: No external funding was received for this study.
Ethical Approval: Ethical considerations and relevant institutional permissions were addressed before data collection.
Informed Consent: Informed consent was obtained from the participating students.
Use of Generative AI: ChatGPT Voice Mode was used as the speaking practice intervention. Generative AI was also used to support language editing of the manuscript.

REFERENCES

Ady, R. S. P., & Mardiah, Z. (2024). Foreign language speaking anxiety: A systematic review of a decade in South East Asia. E*thical Lingua: Journal of Language Teaching and Literature, 11*(2). https://doi.org/10.30605/25409190.722

Aqilah, D. S. (2023). The use of Cake application in learning English speaking skills. *Jurnal Pendidikan Bahasa Inggris Undiksha, 11*(3), 252–257. https://doi.org/10.23887/JPBI.V11I3.47323

Arosy, H. N. (2025). The effectiveness of ELSA Speak in facilitating students' pronunciation in storytelling. *IREELL*. https://jurnalfaktarbiyah.iainkediri.ac.id/index.php/ireel/index

Botes, E., Dewaele, J.-M., & Greiff, S. (2020). The Foreign Language Classroom Anxiety Scale and academic achievement: An overview of the prevailing literature and a meta-analysis. J*ournal for the Psychology of Language Learning,* 2(1), 26–56. https://www.jpll.org/index.php/journal/article/view/botesetal

Brown, H. D. (2019). Lang*uage assessment (*3rd ed.). Pearson Education.

Brown, H. D., & Abeywickrama, P. (2019). Language assess*ment: Principles and classroom practices (3rd ed.). Pearson Education*.

Dewi Lestari, H., Hartatik, F., & Amalyasari, M. R. (2025). Availing the amorphous speaking partner of artificial intelligence: An attempt to alleviate speaking anxiety. L*inguapedia,* 9(1). http://ejurnal.uij.ac.id/index.php/LIN

Ding, D., & Yusof, A. M. B. (2025). Investigating the role of AI-powered conversation bots in enhancing L2 speaking skills and reducing speaking anxiety: A mixed methods study. Humanit*ies and Social Sciences Communications, 12(1),* Art*icle* 1223. https://doi.org/10.1057/s41599-025-05550-z

Council of Europe. (2020). *Common European framework of reference for languages: Learning, teaching, assessment – Companion volume. Council* of Europe Publishing.

Horwitz, E. (2001). Language anxiety and achievement. *Annual Review of Applied Linguistics*, *21*, 112–126. https://doi.org/10.1017/S0267190501000071

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

Huang, L., & Liu, M. (2025). The impact of technology on foreign language anxiety: A systematic review of empirical studies from 2004 to 2024. Hu*manities and Social Sciences Communications,* 12(1), 1665. https://doi.org/10.1057/s41599-025-05921-6

Istikhomahningsih, T., & Setiawan, S. (2025). The innovation of cake learning application on student's speaking ability at tenth grade of senior high school. ID*EAS journal on language*

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Roobby Hamdanur, et al.

teaching and learning, linguistics and literature, 13(2), 8269–8276. https://doi.org/10.24256/ideas.v13i2.6105

Khilma, N., & Azkiyah, S. N. (2026). The effectiveness of using the cake app to improve students' speaking ability. Be*stari: Jurnal Pendidikan Dan Kebudayaan,* 7(1), 134–144. https://doi.org/10.46368/bjpd.v7i1.4987

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.

Saptiany, S. G., Hartono, R., Mujiyanto, J., & Rustipa, K. (2025). Exploring the role of chatgpt voice mode in reducing foreign language speaking anxiety among Indonesian ESP students. E*ducational Process International Journal*, *19*(1). https://doi.org/10.22521/EDUPIJ.2025.19.554

Scovel, T. (1978). The effect of affect on foreign language learning: A review of the anxiety research. *Language Learning*, *28*(1), 129–142. https://doi.org/10.1111/J.1467-1770.1978.TB00309.X

Syakila, Muchari, N., Maharani, P. K., Syasya, A. Z., Khatimah, H., Lubis, I. I., & Abcdef, J. (2026). The effectiveness of the CAKE mobile application in improving English speaking skills. J*ournal of English Language and Education*, *11*(3), 607–617. https://doi.org/10.31004/JELE.V11I3.2457

Thornbury, S. (2005). *How to teach speaking*. Pearson Education Limited.

Vygotsky, L. S. (1978). Mind in society the development of higher psychological processes, knox and carol. *Trans, Harvard Business School Press, Cambridge …*.

Xiong, Y., Zhang, Q., Zhao, L., Liu, S., Guan, H., Sui, Y., Feng, J., & Lee, K. M. R. (2024). A meta-analysis and systematic review of foreign language anxiety interventions among students. J*ournal of Language and Social Psychology*, *43*(5–6), 620–650. https://doi.org/10.1177/0261927X241291258

OPEN ACCESS

© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.