Jadila: Journal of Development and Innovation in Language and Literature Education
Publisher: Yayasan Corolla Education Center
E-ISSN: 2723-6900
P-ISSN: 2745-9578
Volume 6 Number 4, 2026
Page 231-245

Evaluating the Effectiveness of ChatGPT Voice for Improving English Speaking Skills

Ajah Saputra
Universitas Muhammadiyah Sampit
ajahsaputra@umsa.ac.id
Nur Irma Yanti
Universitas Muhammadiyah Sampit
nurirmayanti@umsa.ac.id
Corresponding email: ajahsaputra@umsa.ac.id

Abstract

The integration of artificial intelligence (AI) into English language teaching has created new opportunities to enhance students' speaking skills through interactive and personalized learning experiences. However, empirical evidence regarding the effectiveness of ChatGPT Voice in vocational high school settings remains limited. This study aimed to examine the effectiveness of ChatGPT Voice in improving the English speaking skills of tenth-grade students in the Accounting and Finance Program at SMKN 1 Kota Besi, Indonesia. A pre-experimental one-group pretest–posttest design was employed involving 24 students selected through total sampling. Data were collected using speaking performance tests administered before and after three instructional meetings integrating ChatGPT Voice into speaking activities. Students' speaking performances were assessed using Brown's analytical speaking rubric, covering pronunciation, grammar, vocabulary, fluency, and comprehension. The data were analyzed using descriptive statistics, the Shapiro–Wilk normality test, a paired-samples t-test, and Cohen's d effect size. The findings revealed that the mean speaking score increased from 57.71 in the pre-test to 75.21 in the post-test. The paired-samples t-test indicated a statistically significant improvement ($t(23) = 7.840$, $p < .001$), while the effect size analysis demonstrated a large educational effect ($d = 1.60$). These findings suggest that ChatGPT Voice can serve as an effective supplementary instructional tool for enhancing English speaking skills in vocational EFL classrooms. Future studies are recommended to involve larger samples, longer intervention periods, and experimental designs with comparison groups to further validate these findings.

Keywords: ChatGPT Voice; artificial intelligence; English speaking skills; vocational high school; EFL learning

A. Introduction

The rapid advancement of generative artificial intelligence (AI) has significantly transformed educational practices by enabling more personalized, adaptive, and interactive learning experiences across various disciplines (Kasneci et al., 2023; Zawacki-Richter et al., 2019). In English as a Foreign Language (EFL) education, AI-powered technologies have shifted language learning from teacher-centered instruction toward learner-centered

E-ISSN: 2723-6900
P-ISSN: 2745-9578
Volume 6 Number 4, 2026
Page 231-245

environments that provide immediate feedback, authentic interaction, and continuous learning opportunities beyond the classroom (Lo, 2023; Tlili et al., 2023). These developments are particularly relevant to speaking instruction because oral communication requires frequent practice, meaningful interaction, and timely corrective feedback, which are often difficult to achieve within conventional classroom settings due to limited instructional time and large class sizes (Derakhshan et al., 2021; Sun, 2020). Consequently, conversational AI has emerged as a promising pedagogical tool that supports students' communicative competence through authentic and individualized speaking practice (Kasneci et al., 2023; Lo, 2023).

Despite these technological advancements, developing speaking proficiency remains one of the greatest challenges for English as a Foreign Language (EFL) learners, particularly in vocational high schools where English is expected to support students' future professional communication. Speaking requires learners to simultaneously process linguistic knowledge, organize ideas, produce comprehensible utterances, and respond appropriately in real-time interactions, making it one of the most cognitively demanding language skills (Derakhshan et al., 2021). In Indonesia, numerous studies have reported that EFL learners frequently encounter difficulties in pronunciation, fluency, vocabulary use, grammatical accuracy, and speaking confidence, resulting in limited participation during classroom communication (Anam & Stracke, 2020; Cirocki & Widodo, 2019). These challenges are even more evident in vocational education, where students are expected to communicate in occupational contexts but often receive limited opportunities to engage in authentic speaking practice because classroom instruction remains constrained by limited instructional time, large class sizes, and teacher-centered learning approaches (Widodo, 2021; Nugroho & Mutiaraningrum, 2020). Consequently, innovative instructional approaches that provide frequent speaking opportunities, authentic interaction, and immediate feedback are increasingly needed to help vocational students develop communicative competence that is relevant to workplace demands.

In response to these challenges, ChatGPT Voice-Based Conversational AI has emerged as a promising pedagogical tool for EFL speaking instruction by enabling authentic, interactive, and personalized oral communication. Unlike conventional language learning applications that rely on predetermined exercises, ChatGPT Voice facilitates spontaneous

E-ISSN: 2723-6900
P-ISSN: 2745-9578
Volume 6 Number 4, 2026
Page 231-245

conversations, allowing learners to practice speaking in meaningful contexts while receiving immediate and adaptive feedback (Kasneci et al., 2023). Such interactive learning experiences promote learner autonomy, increase speaking confidence, and provide continuous opportunities for oral practice beyond the limitations of traditional classroom instruction (Lo, 2023; Tlili et al., 2023).

Recent studies have demonstrated the educational potential of ChatGPT and other generative artificial intelligence applications in English language learning. Existing evidence indicates that AI-assisted learning enhances learners' engagement, autonomy, motivation, and language performance through adaptive feedback and personalized learning experiences (Kasneci et al., 2023; Lo, 2023). Moreover, conversational AI has been shown to foster speaking confidence by providing low-anxiety environments that encourage repeated oral practice (Tlili et al., 2023). However, most previous studies have focused on writing support, learners' perceptions, technology acceptance, or higher education contexts rather than objectively examining the effectiveness of voice-based conversational AI in improving speaking performance. Furthermore, empirical studies involving Indonesian vocational high school students remain limited, highlighting the need for further research in vocational EFL settings.

Although previous studies have provided valuable insights into the educational applications of generative artificial intelligence, several important gaps remain unresolved. From a theoretical perspective, existing studies have predominantly emphasized the technological capabilities of AI and students' learning experiences, while limited attention has been given to how voice-based conversational AI facilitates the development of communicative competence through authentic interaction in EFL contexts. Methodologically, many previous investigations have employed qualitative approaches, perception surveys, or correlational designs, providing limited empirical evidence regarding the causal effects of AI-assisted speaking practice on students' speaking performance. Contextually, research focusing on vocational high school students, particularly those enrolled in the Accounting and Finance Program in Indonesia, remains scarce despite the growing demand for English communication skills in vocational education. Furthermore, few studies have examined the implementation of ChatGPT Voice-Based Conversational AI within regular classroom

E-ISSN: 2723-6900
P-ISSN: 2745-9578
Volume 6 Number 4, 2026
Page 231-245

instruction under actual school conditions. These gaps highlight the need for empirical research that objectively evaluates the effectiveness of conversational AI in improving vocational students' English-speaking skills.

In response to these gaps, this study investigates the effectiveness of ChatGPT Voice in improving the English-speaking skills of tenth-grade students in the Accounting and Finance Program at SMKN 1 Kota Besi during the 2024/2025 academic year. Specifically, it examines whether integrating ChatGPT Voice into classroom speaking activities can significantly improve students' pronunciation, fluency, vocabulary, grammatical accuracy, and comprehensibility after three instructional meetings. The findings are expected to provide empirical evidence regarding the pedagogical effectiveness of conversational AI in vocational EFL classrooms.

This study contributes to the growing body of research on AI-assisted language learning in several ways. First, it extends existing literature by examining the effectiveness of ChatGPT Voice-Based Conversational AI in improving English-speaking performance among Indonesian vocational high school students, a context that has received limited scholarly attention. Second, unlike many previous studies that primarily focused on learners' perceptions, attitudes, or technology acceptance, this study objectively measures changes in speaking performance through a standardized analytical speaking rubric. Third, the findings are expected to provide practical insights for English teachers regarding the integration of conversational artificial intelligence into classroom-based speaking instruction, thereby supporting the development of communicative competence required in vocational education and future workplace communication.

B. Research Methodology

Sampling

This study employed a pre-experimental research design using a one-group pretest–posttest design to investigate the effectiveness of ChatGPT Voice in enhancing students' English-speaking skills. This design was considered appropriate because the study aimed to measure changes in students' speaking performance before and after the implementation of the instructional intervention within a single classroom setting. The design enabled the

E-ISSN: 2723-6900
P-ISSN: 2745-9578
Volume 6 Number 4, 2026
Page 231-245

researcher to evaluate the effectiveness of ChatGPT Voice under authentic classroom conditions while maintaining the regular teaching and learning process (Creswell & Creswell, 2018; Cohen, Manion, & Morrison; Fraenkel, Wallen, & Hyun)

The study was conducted at SMKN 1 Kota Besi, Central Kalimantan, Indonesia, during the 2024/2025 academic year. The research context was an English as a Foreign Language (EFL) classroom involving tenth-grade students enrolled in the Accounting and Finance Program (Akuntansi Keuangan dan Lembaga/AKL). The unit of analysis comprised individual students' English-speaking performance as measured through pretest and posttest speaking assessments.

The target population consisted of all tenth-grade students in the Accounting and Finance Program at SMKN 1 Kota Besi. Because only one class was available, total sampling was employed, allowing all students who met the research criteria to participate in the study (Etikan & Bala, 2017). Consequently, the realized sample consisted of 24 students, all of whom completed both the pretest and posttest, resulting in a 100% participation rate. Total sampling was considered appropriate because it minimized sampling bias and ensured that the entire accessible population was represented in the study. Nevertheless, because the research involved only one intact classroom in a single vocational high school, the findings should be interpreted with caution and may not be generalized to other educational settings with different student characteristics or learning environments.

Data Collection

Data were collected using pretest and posttest speaking performance assessments administered before and after the instructional intervention. Speaking performance assessment was selected because it directly measures learners' oral communicative ability in authentic language use rather than relying on self-reported perceptions or questionnaire responses. Performance-based assessment has been widely recommended for evaluating speaking proficiency as it enables students to demonstrate their actual language production across multiple dimensions of speaking competence (Brown & Abeywickrama, 2019).

Before the intervention, all participants completed an individual pretest in the form of an oral presentation to determine their initial speaking proficiency. The speaking test required

E-ISSN: 2723-6900
P-ISSN: 2745-9578
Volume 6 Number 4, 2026
Page 231-245

students to present orally using topics appropriate to their English proficiency level. Their performances were evaluated using Brown's analytical speaking rubric, which assesses five components of speaking, namely pronunciation, grammar, vocabulary, fluency, and comprehension (Brown & Abeywickrama, 2019). The same assessment criteria were employed during the posttest to ensure consistency in measuring students' speaking performance before and after the intervention.

Following the pretest, the instructional intervention was conducted over three classroom meetings, each lasting 4 × 45 minutes, following the school's regular English schedule. During the intervention, students used ChatGPT Voice (Free Version) on their own smartphones to engage in guided speaking activities designed to encourage authentic oral interaction and repeated speaking practice. The instructional topics included Self-Introduction, Greeting and Leave Taking, and Daily Activities, which were selected because they align with the tenth-grade English curriculum for vocational high schools and provide meaningful opportunities for communicative language practice. Throughout each session, students interacted orally with ChatGPT Voice, practiced responding to conversational prompts, revised their spoken responses based on immediate AI feedback, and concluded the lesson by delivering individual oral presentations before the class.

After completing the three instructional meetings, all participants took the posttest, which followed the same procedures, speaking format, assessment rubric, and scoring criteria as the pretest. Using equivalent assessment procedures before and after the intervention enhanced the comparability of students' speaking scores and strengthened the validity of measuring learning gains attributable to the instructional treatment (Brown & Abeywickrama, 2019). To provide a clearer overview of the research procedures and enhance the replicability of the study, the instructional sequence implemented during the pretest, treatment sessions, and posttest is summarized in Table 1.

Table 1. Research Procedure

Meeting	Topic	Learning Activities	Duration
Pretest	Speaking Test	Individual oral presentation	4 × 45 minutes
1	Self-Introduction	Students practiced speaking with ChatGPT Voice and presented the results.	4 × 45 minutes
2	Greeting and	Students conducted guided conversations using	4 × 45

E-ISSN: 2723-6900
P-ISSN: 2745-9578
Volume 6 Number 4, 2026
Page 231-245

	Leave Taking	ChatGPT Voice followed by classroom presentations.	minutes
3	Daily Activities	Students practiced describing their daily routines with ChatGPT Voice and delivered oral presentations.	4 × 45 minutes
Posttest	Speaking Test	Individual oral presentation	4 × 45 minutes

Data Analysis

The collected data were analyzed using both descriptive and inferential statistics to determine the effectiveness of ChatGPT Voice in improving students' English-speaking skills. Descriptive statistics were first employed to summarize students' speaking performance, including the mean, standard deviation, minimum score, and maximum score for both the pretest and posttest. These descriptive measures provided an overview of students' speaking achievement before and after the instructional intervention (Field, 2024).

Prior to hypothesis testing, the data were examined for normality using the Shapiro–Wilk test, as this test is recommended for small sample sizes ($n < 50$) due to its high statistical power in detecting departures from normality (Razali & Wah, 2019). If the data met the assumption of normality ($p > .05$), a paired-samples t-test was performed to determine whether there was a statistically significant difference between students' pretest and posttest speaking scores. Conversely, if the assumption of normality was violated, the Wilcoxon signed-rank test was employed as a non-parametric alternative appropriate for paired observations (Field, 2024).

In addition to statistical significance testing, the magnitude of the instructional effect was determined by calculating Cohen's d effect size. Effect size analysis was included because statistical significance alone does not indicate the practical importance of an educational intervention. Cohen's d provides information regarding the magnitude of improvement resulting from the implementation of ChatGPT Voice, thereby facilitating a more meaningful interpretation of the research findings (Lakens, 2022). The effect size was interpreted as small (0.20), medium (0.50), and large (0.80) following Cohen's classification. All statistical analyses were conducted using IBM SPSS Statistics version 27, with the level of statistical significance set at $\alpha = 0.05$.

Jadila: Journal of Development and Innovation in Language and Literature Education
Publisher: Yayasan Corolla Education Center
E-ISSN: 2723-6900
P-ISSN: 2745-9578
Volume 6 Number 4, 2026
Page 231-245

C. Results and Discussion

1. Results

The purpose of this study was to examine the effectiveness of ChatGPT Voice in improving the English-speaking skills of tenth-grade students in the Accounting and Finance Program at SMKN 1 Kota Besi. To achieve this objective, the students' speaking performance before and after the intervention was analyzed using descriptive and inferential statistics. The results are presented in terms of descriptive statistics, normality testing, hypothesis testing, and effect size analysis.

Table 2. Descriptive Statistics of Students' Speaking Scores

Test	N	Mean	SD	Min	Max
Pre-test	24	57.71	11.13	35	80
Post-test	24	75.21	11.75	55	100

Table 2 summarizes the descriptive statistics of students' speaking scores before and after the implementation of ChatGPT Voice. A total of 24 students completed both the pre-test and post-test. Prior to the instructional intervention, the students obtained a mean score of 57.71, indicating that their initial speaking proficiency was relatively moderate. Following three instructional meetings utilizing ChatGPT Voice, the mean score increased to 75.21, representing an improvement of 17.50 points. In addition to the increase in the mean score, the minimum score improved from 35 to 55, while the maximum score increased from 80 to 100. These findings indicate that improvements occurred not only among high-achieving students but also among students with lower initial speaking proficiency. Furthermore, the relatively similar standard deviations between the pre-test and post-test suggest that score variability remained stable despite the overall improvement in students' speaking performance.

To provide a clearer picture of individual learning outcomes, students' score changes were further examined by comparing their pre-test and post-test performances. Of the 24 participants, 20 students (83.33%) demonstrated higher post-test scores, 2 students (8.33%)

E-ISSN: 2723-6900
P-ISSN: 2745-9578
Volume 6 Number 4, 2026
Page 231-245

obtained identical scores, and 2 students (8.33%) experienced slight decreases in their post-test performance. Overall, these descriptive findings indicate a positive trend in students' English-speaking achievement following the implementation of ChatGPT Voice, although the degree of improvement varied across individuals.

Normality Test

Before conducting the hypothesis testing, the normality of the pre-test and post-test scores was examined using the Shapiro–Wilk test, as it is considered appropriate for studies involving a relatively small sample size (Field, 2024). The test was conducted to determine whether the speaking scores were normally distributed and to identify the appropriate inferential statistical test for hypothesis testing.

Table 3. Results of the Shapiro–Wilk Normality Test

Test	Statistic	df	Sig.	Interpretation
Pre-test	0.984	24	0.951	Normal
Pos-test	0.971	24	0.702	Normal

Table 3 shows that the pre-test and post-test scores were normally distributed, as the significance values exceeded the threshold of 0.05. Specifically, the pre-test yielded a significance value of 0.951, while the post-test produced a significance value of 0.702. Since both datasets satisfied the assumption of normality, a paired-samples t-test was considered appropriate for testing the research hypothesis (Field, 2024).

Hypothesis Testing

The paired-samples t-test revealed a statistically significant difference between the students' pre-test and post-test speaking scores, $t(23) = 7.840$, $p < .001$. The mean difference of 17.50 points indicates that students achieved substantially higher speaking scores after participating in three instructional meetings using ChatGPT Voice. Therefore, the null hypothesis (H_0) was rejected, while the alternative hypothesis (H_1) was accepted, indicating that the implementation of ChatGPT Voice had a statistically significant positive effect on students' English-speaking skills. These findings indicate that integrating ChatGPT Voice into speaking instruction significantly improved the English-speaking skills of tenth-grade vocational high school students.

E-ISSN: 2723-6900
P-ISSN: 2745-9578
Volume 6 Number 4, 2026
Page 231-245

2. Discussion

The present study investigated the effectiveness of ChatGPT Voice in improving the English-speaking skills of tenth-grade students in the Accounting and Finance Program at SMKN 1 Kota Besi. This study was conducted in response to the growing interest in integrating generative artificial intelligence into English language learning and the limited empirical evidence regarding its implementation in Indonesian vocational high schools. The findings demonstrated that students' speaking performance improved significantly following the implementation of ChatGPT Voice during three instructional meetings. This improvement was evidenced by the increase in the mean speaking score from 57.71 in the pre-test to 75.21 in the post-test, representing an average gain of 17.50 points. Inferential statistical analysis further confirmed that the observed improvement was statistically significant, indicating that the integration of ChatGPT Voice positively influenced students' English-speaking performance. Furthermore, the large effect size suggested that the intervention produced not only statistically significant results but also meaningful educational benefits. These findings contribute to the growing body of research on AI-assisted language learning by providing empirical evidence that conversational artificial intelligence can effectively support speaking instruction in vocational EFL classrooms.

The findings of the present study are consistent with the growing body of research demonstrating the positive contribution of generative artificial intelligence to English language learning. The significant improvement in students' speaking performance observed after the implementation of ChatGPT Voice supports previous evidence that AI-assisted learning environments provide learners with greater opportunities to practice language, receive immediate feedback, and participate in meaningful interactions, thereby facilitating language development (Kasneci et al., 2023; Lo, 2023). Rather than replacing the role of teachers, conversational AI appears to complement classroom instruction by extending students' opportunities to practice speaking beyond conventional teacher-centered activities.

The improvement identified in this study also aligns with previous findings indicating that conversational AI can reduce communication anxiety and encourage learners to

E-ISSN: 2723-6900
P-ISSN: 2745-9578
Volume 6 Number 4, 2026
Page 231-245

participate more actively in speaking activities. One of the main challenges faced by EFL learners is their reluctance to speak because they fear making grammatical or pronunciation mistakes in front of teachers and classmates. ChatGPT Voice offers an alternative learning environment in which students can practice speaking repeatedly without experiencing direct social pressure, allowing them to build confidence while improving their oral communication skills (Tlili et al., 2023). This learning environment may explain why most students in the present study demonstrated higher post-test scores following the intervention.

Furthermore, the present findings reinforce previous studies reporting that personalized learning supported by artificial intelligence promotes learner autonomy and self-directed learning. Through ChatGPT Voice, students were able to engage in spoken interactions at their own pace, repeat conversations when necessary, and receive responses tailored to their utterances. Such learning experiences encouraged students to become more actively involved in constructing their own language knowledge rather than relying solely on teacher explanations. These findings are consistent with previous research suggesting that adaptive AI technologies foster greater learner engagement and facilitate more meaningful language practice than traditional classroom instruction alone (Lo, 2023; Kasneci et al., 2023).

Although the present findings generally support previous studies, this research also extends the existing literature by providing empirical evidence from the context of Indonesian vocational high school students. Most earlier investigations have focused on university students or have explored learners' perceptions of ChatGPT rather than objectively measuring changes in speaking performance through pretest–posttest assessments. By demonstrating significant improvements in students' speaking scores after three instructional meetings, the present study contributes additional empirical evidence regarding the pedagogical value of ChatGPT Voice in vocational English language education.

One possible explanation for the significant improvement in students' English speaking skills is that ChatGPT Voice provided opportunities for continuous and authentic oral interaction throughout the learning process. Unlike conventional classroom instruction, where speaking practice is often constrained by limited instructional time and unequal participation, ChatGPT Voice enabled every student to engage in individualized

E-ISSN: 2723-6900
P-ISSN: 2745-9578
Volume 6 Number 4, 2026
Page 231-245

conversations at their own pace. This flexibility allowed students to practice repeatedly, receive immediate responses, and refine their spoken production without waiting for teacher feedback. Such learning conditions are consistent with previous studies suggesting that conversational artificial intelligence facilitates active language use by creating interactive learning environments that encourage meaningful communication and learner engagement (Kasneci et al., 2023; Lo, 2023).

Another possible explanation relates to the immediate feedback provided during students' interactions with ChatGPT Voice. Immediate responses enabled students to recognize errors, modify their utterances, and improve subsequent speaking attempts within the same learning session. Unlike delayed teacher correction, which often occurs after classroom activities have ended, AI-generated responses offered continuous opportunities for self-monitoring and self-correction. This process may have enhanced students' awareness of pronunciation, vocabulary use, grammatical accuracy, and fluency during communication. The findings align with recent perspectives on **AI-supported self-regulated language learning (SRLL)**, which emphasize that artificial intelligence functions as a metacognitive learning partner by encouraging learners to monitor, evaluate, and refine their language production through repeated practice and reflection (Wei, 2023; Yang et al., 2024). Through iterative conversations with ChatGPT Voice, students were able to compare their spoken responses with AI-generated feedback, recognize areas requiring improvement, and modify their subsequent utterances. Such reflective learning processes are considered fundamental for developing higher levels of speaking proficiency. Previous studies have similarly reported that adaptive feedback provided by artificial intelligence contributes to more effective language learning by supporting learner autonomy and encouraging independent practice (Tlili et al., 2023; Lo, 2023).

Although the overall findings demonstrated significant improvement, the intervention did not produce identical outcomes for all participants. The descriptive analysis showed that 20 students demonstrated higher post-test scores, two students maintained the same scores, and two students experienced slight decreases in their post-test performance. These findings suggest that the effectiveness of ChatGPT Voice may be influenced by individual learner characteristics rather than by the technology alone. Differences in initial English proficiency,

E-ISSN: 2723-6900
P-ISSN: 2745-9578
Volume 6 Number 4, 2026
Page 231-245

speaking anxiety, learning motivation, willingness to communicate, and digital literacy may have affected the extent to which students benefited from AI-assisted speaking practice. Therefore, while ChatGPT Voice functioned as an effective instructional tool for most participants, its educational impact appeared to vary according to learners' individual needs and learning readiness.

The findings of this study also provide important pedagogical implications for English language teaching in vocational education. The significant improvement observed after only three instructional meetings suggests that conversational artificial intelligence can be integrated into regular classroom instruction as a complementary learning resource rather than as a replacement for teachers. English teachers may utilize ChatGPT Voice to increase students' opportunities for speaking practice beyond classroom interaction, provide more individualized learning experiences, and encourage learners to communicate more confidently in English (Holmes et al., 2022; UNECO, 2023; Chang & Sun, 2024). Nevertheless, teachers should continue to guide, monitor, and facilitate students' learning to ensure that artificial intelligence is used appropriately and effectively in achieving instructional objectives. Future studies are recommended to involve larger sample sizes, longer intervention periods, and experimental designs with control groups to obtain more comprehensive evidence regarding the effectiveness of conversational artificial intelligence in English language learning.

D. Conclusion and Suggestion

The present study examined the effectiveness of ChatGPT Voice in improving the English speaking skills of tenth-grade students in the Accounting and Finance Program at SMKN 1 Kota Besi. The findings demonstrated that students achieved higher speaking performance after participating in three instructional meetings integrating ChatGPT Voice into classroom learning. Statistical analysis indicated a significant difference between the pre-test and post-test scores, suggesting that ChatGPT Voice effectively supported students' development of pronunciation, grammar, vocabulary, fluency, and comprehension. These findings provide additional empirical evidence that conversational artificial intelligence can

E-ISSN: 2723-6900
P-ISSN: 2745-9578
Volume 6 Number 4, 2026
Page 231-245

be integrated into vocational English language instruction to facilitate meaningful speaking practice and enhance students' opportunities for oral communication.

Despite these encouraging findings, several limitations should be acknowledged. First, the study involved only one class consisting of 24 students from a single vocational high school, which may limit the transferability of the findings to other educational contexts. Second, the intervention was conducted over only three instructional meetings; therefore, the long-term effects of ChatGPT Voice on speaking development remain uncertain. Future research is recommended to involve larger and more diverse samples, employ experimental designs with comparison groups, and implement longer intervention periods to provide more comprehensive evidence regarding the effectiveness of conversational artificial intelligence in English language learning. Additionally, future studies may explore the use of ChatGPT Voice for developing other language skills, such as listening, reading, and writing, or investigate its effectiveness across different educational levels and vocational programs.

References

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

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

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th ed.). Routledge.

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage Publications.

Derakhshan, A., Fathi, J., & Pawlak, M. (2021). The interplay between teacher credibility, student motivation, and willingness to communicate in efl contexts.

Field, A. (2024*). Discovering statistics using IBM SPSS Statistics (6th ed.).* Sage Publications.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2023). *How to design and evaluate research in education* (11th ed.). McGraw-Hill.

Holmes, W., Bialik, M., & Fadel, C. (2022). *Artificial intelligence in education: Promises and implications for teaching and learning* (Updated ed.). Center for Curriculum Redesign.

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeiffer, F., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models

E-ISSN: 2723-6900
P-ISSN: 2745-9578
Volume 6 Number 4, 2026
Page 231-245

for education. Learning and Individual Differences, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), 410. https://doi.org/10.3390/educsci13040410

Richards, J. C. (2022). T*eaching English through interaction: Principles and practices.* Cambridge University Press.

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*. 10(15), 1-25. https://doi.org/10.1186/s40561-023-00237-x

UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO.

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

Yang, Y., Wen, Y., & Song, Y. (2024). Evaluating AI's impact on self-regulated language learning: A systematic review. *System, 126*, 103484. https://doi.org/10.1016/j.system.2024.103484

Zhai, X. (2023). ChatGPT for next generation science learning. *Journal of Educational Technology Systems.* 29(3), 42-46. https://doi.org/10.2139/ssrn.4331313