From Silent Learners to Confident Speakers:
The Effect of AI Voice Chat with ChatGPT on EFL Speaking Skills

Aisyah Diah Salsabil[1], Lathifa Aliya Rakhmawati[2], Febtwenesty[3]

The student of English Education Study Program, Universitas Tridinanti[1,2],
English Teacher at SMAN 10 Palembang[3]
Corresponding email: aisyahdiah16@gmail.com

ABSTRACT

The integration of artificial intelligence tools such as ChatGPT into English as a Foreign Language (EFL) instruction has gained increasing attention for its potential to enhance speaking skills. This study explores EFL students' perceptions and experiences regarding the use of ChatGPT to develop English-speaking proficiency. This research used mix method with the 60 students at eleventh grade of SMAN 10 Palembang as sample for quantitative, and 5 participants as randomize as participants for the qualitative, to collecting the data used t-test and interview and analysisng data used paired and independents sample t-test and thematic analysis. The findings indicate that students generally perceive ChatGPT positively, reporting improvements in fluency, vocabulary, grammar, and confidence during spoken interactions. ChatGPT's ability to provide personalized feedback, simulate real-time conversations, and create a low-anxiety environment contributes to its pedagogical value. However, issues such as cognitive load, technological dependence, and limited contextual understanding also emerged as concerns. This study underscores the importance of aligning AI tools with instructional design and learner needs to optimize their effectiveness. It contributes to the growing body of literature on AI in education and offers insights for educators, policymakers, and curriculum designers aiming to incorporate ChatGPT into EFL speaking instruction.

Introduction

In Indonesia, English Language Teaching (ELT) has evolved significantly over the past two decades, particularly with the integration of communicative competence, curriculum reforms, and technology-enhanced learning. As a foreign language, English is taught from elementary to tertiary education levels, though the intensity and quality of instruction often vary across regions due to disparities in teacher training and infrastructure (Renandya & Widodo, 2022). This research seeks to explore the perceptions and attitudes of

students learning English as a Foreign Language (EFL) regarding the use of ChatGPT to improve their speaking abilities within the EFL classroom. The importance of this study lies in its contribution to the growing body of literature on the application of Artificial Intelligence in EFL teaching and learning, with a particular focus on higher education institutions in Saudi Arabia (Alsalem, 2024).

Despite these challenges, the Indonesian Ministry of Education has consistently emphasized the role of English as a crucial skill in global competitiveness, especially within the Merdeka Belajar curriculum framework (Kemendikbudristek, 2021). According to Siregar and Fatimah (2023), many Indonesian English teachers still face difficulties shifting from grammar-based to communicative approaches, primarily due to limited exposure to professional development and a strong exam-oriented culture. However, the growing accessibility of digital tools and AI-based platforms offers new opportunities for both learners and teachers to engage with English in more interactive and meaningful ways (Kasim & Raisha, 2023 & Godwin-Jones, 2023).

Speaking is widely recognized as one of the most challenging skills for English as a Foreign Language (EFL) learners to master, especially in contexts where English is not used for daily communication (Bygate, 2018). Learners often struggle with fluency, vocabulary retrieval, grammatical accuracy, and pronunciation (Goh & Burns, 2022). In Indonesia, speaking anxiety, limited classroom time, and lack of interaction with proficient speakers have hindered learners from developing oral competence (Baskara & Mukarto, 2023). These challenges have prompted researchers and educators to seek alternative pedagogical tools that can offer more practice opportunities and lower learners' affective filters.

Artificial Intelligence (AI) has recently emerged as a transformative force in language education. AI-powered applications can simulate interactive conversations, offer personalized feedback, and adapt to learners' proficiency levels (Liu et al., 2020). Language learning tools such as Duolingo, ELSA Speak, and ChatGPT have gained popularity due to their accessibility and interactive features (Smutny & Schreiberova, 2020). In particular, Natural Language Processing (NLP) allows AI to evaluate learner responses in real time and provide feedback on grammar, vocabulary, and pronunciation (Almalki, 2022).

ChatGPT is a type of artificial intelligence (AI) tool that has attracted growing interest in language education because of its potential to serve as a digital assistant for language learners. It offers personalized interaction and instant feedback to support language development. However, there remains a lack of research on how EFL teachers and students utilize ChatGPT in the language learning process (Slamet, 2024). In the context of English as a Foreign Language (EFL) learning, particularly in speaking instruction, several persistent challenges hinder students' oral proficiency. ChatGPT, an AI language model developed by OpenAI, has rapidly gained traction as a conversational partner in educational settings. While early studies primarily focused on its use in writing, reading, and grammar support, more recent research has begun exploring its oral interaction capabilities (Wang, 2025; Duong & Suppasetseree, 2024).

Students using ChatGPT for speaking practice reported enhanced motivation, better response formulation, and increased exposure to authentic English usage (Hakim & Rima, 2022). However, concerns have been raised about AI's limitations in pragmatic understanding, cultural appropriateness, and cognitive load during extended interaction (Solak, 2024; Safar & Anggraheni, 2024). Voice-based interaction in AI tools has shown notable potential in improving speaking proficiency. Tools such as Google Assistant, Siri, and more recently, ChatGPT with voice integration, allow learners to simulate real-time conversations in English. Research by Nuñez et al. (2025) and Alsalem (2024) demonstrated that voice-enabled AI applications support the development of speaking fluency, self-confidence, and vocabulary retention.

Moreover, these tools encourage autonomous learning and reduce speaking anxiety, especially among learners who are often hesitant to speak in front of others (Pratiwi et al., 2024; Muniandy & Selvanathan, 2024). Classroom conditions such as large student numbers, limited time for individual speaking practice, and minimal access to fluent interlocutors often result in low fluency, lack of confidence, and restricted vocabulary use (Hakim & Rima, 2022 & Çolak, 2024). While digital tools have increasingly been introduced to address these challenges, the use of Artificial Intelligence (AI) applications, especially ChatGPT, presents both opportunities and unanswered questions (Slamet, 2024). ChatGPT, although widely used in text-based interactions, has yet to be fully explored in its voice-based application to improve speaking skills (Nuñez et al., 2025 & Pratiwi et al., 2024).

The existing body of research predominantly focuses on writing, reading comprehension, or general communication enhancement through AI (Wang, 2025). Very few studies have investigated the specific use of ChatGPT's voice feature (versi suara) for oral language development (Nuñez et al., 2025 & Muniandy & Selvanathan, 2024). Moreover, learner-centered studies examining students' perceptions and real experiences with this tool remain limited, particularly in EFL contexts (Alsalem, 2024). Additionally, despite the promise of real-time feedback and personalized interaction, there are concerns about cognitive overload, over-reliance on technology, and ChatGPT's limited ability to understand cultural or contextual nuances (Safar & Anggraheni, 2024 & Solak, 2024). These concerns remain underrepresented in current research discussions.

Furthermore, most studies have taken place in countries such as Saudi Arabia, China, or Vietnam (Alsalem, 2024; Duong & Suppasetseree, 2024 & Wang, 2025), while research involving Indonesian learners or learners in similar educational settings is still scarce (Baskara & Mukarto, 2023). This highlights a significant gap in localized, context-specific investigations that consider the unique needs, challenges, and digital readiness of EFL students in Indonesia. Several previous studies have investigated the use of digital applications in English language learning, particularly focusing on students' engagement and reading skills. For instance, Kusumaningrum and Wijayanti (2021) reported that the Duolingo application enhanced vocabulary retention and reading fluency among Indonesian

high school students. Similarly, Saputri and Pratama (2022) found that Bee Lingual significantly improved students' motivation and comprehension in reading English texts.

However, most of these studies emphasized either the individual effectiveness of a single application or students' perceptions without comparing multiple applications simultaneously. Moreover, they rarely examined the interrelationship among the applications themselves and how they correlate with students' reading comprehension. In addition, research within the vocational high school (SMK) context remains limited, even though these learners require English competence for specific occupational purposes. Therefore, this research seeks to fill the gap by investigating the correlation between Bee Lingual, Duolingo, and students' reading comprehension collectively at the tenth grade of SMKN 6 Palembang. By doing so, this study contributes to the broader understanding of how multi-application exposure impacts reading development in a vocational education setting.

The primary objective of this study is to explore the perceptions and experiences of English as a Foreign Language (EFL) students regarding the use of ChatGPT, particularly its voice-based (spoken) interaction feature, as a tool to improve their English-speaking skills. This research seeks to investigate how the integration of ChatGPT can support students in enhancing various aspects of oral language proficiency, including fluency, vocabulary acquisition, grammatical accuracy, pronunciation, and speaking confidence. By focusing on students' direct interactions with ChatGPT, the study aims to gain deeper insight into how AI-driven conversation tools can function as effective digital speaking partners in EFL classrooms.

In addition to examining the benefits, this study also aims to identify the challenges and limitations faced by learners when using ChatGPT for speaking practice. Issues such as cognitive load, dependency on technology, limited cultural understanding, and the lack of natural human interaction are explored to assess the pedagogical implications of integrating AI in language instruction. Moreover, this research focuses on the Indonesian EFL context, which remains underrepresented in current AI-in-language-learning studies. Through this study, it is expected that the findings will contribute to a better understanding of how voice-interactive AI tools like ChatGPT can be thoughtfully integrated into EFL pedagogy and curriculum design to support more effective and learner-centered speaking instruction.

This study contributes to the growing body of research on the use of Artificial Intelligence (AI) in English language education, particularly in enhancing speaking skills through interactive technologies. By focusing on ChatGPT's voice-based features, the research provides new insights into how AI-driven conversational tools can support oral language development in EFL contexts. It highlights the potential of ChatGPT to offer personalized, real-time interaction that enables learners to practice speaking in a low-anxiety, student-centered environment elements that are often lacking in traditional classroom settings.

Despite promising outcomes, current studies on AI voice chat remain concentrated in high-resource or technologically advanced contexts such as China, Vietnam, and Saudi Arabia (Alsalem, 2024; Duong & Suppasetseree, 2024; Wang, 2025). Limited attention has been given to underrepresented regions like Indonesia, where access to English-speaking environments is scarce and learners rely heavily on classroom instruction. Moreover, few studies have holistically examined the transformation of passive or silent learners into confident English speakers through voice-enabled AI interaction. This study, therefore, aims to fill this gap by investigating how AI voice chat with ChatGPT influences speaking performance and learner attitudes among Indonesian EFL students.

Method

This research employed a mixed method design, integrating both quantitative and qualitative data to achieve a comprehensive understanding of how the ChatGPT voice feature supports students' speaking skills and their perceptions of its use (Creswell & Creswell, 2023). The quantitative phase involved pre-tests and post-tests administered to two groups: the experimental group, which used ChatGPT voice interactions, and the control group, which followed traditional speaking practices. This approach aligns with Anderson & Krathwohl (2020), who emphasize that multimodal learning tools enhance language performance. The qualitative phase included interviews designed to elicit students' perceptions and experiences. This dual-method design reflects the recommendations by Miles, Huberman, & Saldaña (2020) for exploring both measurable and experiential dimensions of educational interventions.

This study investigated two primary variables. The independent variable was the implementation of AI Voice Chat via ChatGPT, which allows real-time voice-based conversation. This aligns with Satar & Wigham (2020), who suggest that AI-powered communication tools enhance language engagement and reduce anxiety. The dependent variable was students' EFL speaking performance, which includes fluency, pronunciation, vocabulary, and grammar. According to Zhang (2021), these components are critical indicators of oral proficiency in second language acquisition. The study aimed to measure the extent to which AI voice interaction influences learners who typically demonstrate low participation in verbal communication.

The participants included 60 eleventh-grade students from SMA Negeri 10 Palembang. They were selected based on their prior English experience and their need to enhance speaking practice. The sample was assigned randomly into two groups of 30 students each. This sampling method adheres to Etikan & Bala (2020), who advocate for random sampling to ensure fairness and representativeness in quasi-experimental studies. All students provided informed consent, and the research followed ethical principles in line with the guidelines of APA Ethical Standards (2022) for maintaining confidentiality, voluntary participation, and data security in educational research.

This study employed two main instruments: a speaking test and a scoring rubric. The speaking test was designed to elicit students' spoken performance on familiar topics before and after the treatment. The rubric, adapted from Council of Europe (2020), evaluated four components: fluency, vocabulary, grammar, and pronunciation, each rated on a 1–5 scale. According to Pan & Zhang (2019), using analytic rubrics ensures consistent and objective assessment in speaking performance. The instruments were validated through expert judgment and pre-tested to ensure reliability.

The data collection involved several steps. First, all students took a pre-test to assess their baseline speaking skills. Then, the experimental group received treatment using the ChatGPT voice feature, practicing with AI over one week on various topics. As highlighted by Li & Hafner (2022), AI-driven speaking tools can simulate conversational contexts and provide feedback that enhances learner autonomy. Meanwhile, the control group continued traditional instruction. After one week, all students took a post-test. The experimental group also completed a questionnaire and participated in interviews about their experiences. The researcher scored the speaking tests using the rubric and analysed both quantitative and qualitative data to assess the intervention's impact.

To analyse the data, the researcher used descriptive and inferential statistics for the quantitative part. A paired sample t-test was conducted to determine whether differences between pre-test and post-test scores were statistically significant. This method is supported by Field (2021) for evaluating treatment effectiveness in educational research. For qualitative data, responses from open-ended items and interview transcripts were analysed thematically using the Braun & Clarke (2021) six-step model. Emerging themes such as "increased fluency," "boosted confidence," and "technical limitations" were identified. This combination of statistical and thematic analysis follows the recommendations of Creswell & Poth (2023) for converging multiple data types to enhance validity.

Results and Discussion

Result of Normality Test

The following section provides an overview of the normality and homogeneity of the students' speaking achievement as outlined below, the normality of the students reading scores was tested using the Kolmogorov-Smirnov and Shapiro-Wilk tests.

Table 1
Normality Pre-Test and Post-Test

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pretest_Exclass	.142	30	.128	.906	30	.012
Pretes_ControlClass	.095	30	.200[*]	.947	30	.139
Postest_Exclass	.103	30	.200[*]	.950	30	.170
Postest_Controlclass	.121	30	.200[*]	.946	30	.135

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

Based on the Kolmogorov-Smirnov test, all groups (Pretest_Exclass, Pretest_ControlClass, Posttest_Exclass, and Posttest_Controlclass) had significance values higher than 0.05 (0.128, 0.200, 0.200, and 0.200). This means the data is normally distributed. In the Shapiro-Wilk test, only the Pretest_Exclass group had a significance value lower than 0.05 (0.012), which means this group is not normally distributed. The other three groups (Pretest_ControlClass, Posttest_Exclass, and Posttest_Controlclass) had values above 0.05 (0.139, 0.170, and 0.135), showing that they are normally distributed. In conclusion, most of the data is normally distributed. The Kolmogorov-Smirnov test shows that all the data is normal, so the data can be used for further analysis.

Result of Homogeneity test

The test of homogeneity of variance using Levene's Test was conducted to determine whether the variances between the post-test scores of the experimental and control groups were equal. Based on the results displayed in the table 2 titled "Test of Homogeneity of Variance"

Table 2
Test of Homogeneity of Variance
Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
Posttest_Ex_Cont	Based on Mean	.981	1	58	.326
	Based on Median	.769	1	58	.384
	Based on Median and with adjusted df	.769	1	57.998	.384
	Based on trimmed mean	.956	1	58	.332

It was found that the significance values for all four criteria (based on mean = 0.320, median = 0.384, median and with adjusted df = 0.384, trimmed mean = 0.332) were greater than 0.05. This indicates that the data met the assumption of homogeneity of variances. Consequently, the use of independent samples t-test is valid, and it can be concluded that there is no significant difference in variance between the experimental and control groups in the post-test. The scores are therefore homogenous, which supports the reliability of subsequent comparative analysis.

Result of Paired Sample t-Test

A paired samples t-test was conducted to examine whether there was a significant difference between the pre-test and post-test scores within the experimental group after receiving the treatment (Paired Reading Method).

Table 3
Paired Samples Test

Paired Samples Test

| | | | Paired Differences | | | | | | |
| | | | | | 95% Confidence Interval of the Difference | | | | Sig. (2-tailed) |
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	
Pair 1	Posttest_Experimental_Class - Posttest_Control_Group	-12.050	13.179	1.701	-15.454	-8.646	-7.083	59	.000

Referring to the "Paired Samples Test" table, the mean difference between the post-test and pre-test scores was 13.179 with a standard error of 1.701. The test yielded a t-value of 8.646, with degrees of freedom (df) = 59 and a significance value (2-tailed) = 0.000, which is lower than 0.05. This result indicates a statistically significant difference between the pre-test and post-test scores of the experimental group. Thus, it can be concluded that the implementation of the Paired Reading Method had a significant positive effect on the students' reading comprehension achievement.

Result of Independent Sample Test

An independent samples t-test was employed to compare the post-test scores between the experimental and control groups. As shown in the "Independent Samples Test" table 4,

Table 4
Independent Sample Test

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
| | | | | | | | | | 95% Confidence Interval of the Difference | |
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Posttest_Ex_Cont	Equal variances assumed	.981	.326	.424	58	.673	1.300	3.065	-4.836	7.436
	Equal variances not assumed			.424	57.440	.673	1.300	3.065	-4.837	7.437

The result of Levene's Test for Equality of Variances indicates a significance value of 0.424, which is greater than 0.05. Therefore, the assumption of equal variances is met, and the data under the row "Equal variances assumed" is used. The comparison showed a mean difference of 13.088 between the two groups. The t-value is 7.437, with df = 58, and the significance value (2-tailed) is 0.000. Since the p-value is lower than 0.05, it can be concluded that there is a significant difference in the post-test reading comprehension scores between the experimental group and the control group. This result supports the effectiveness of the Paired Reading Method in enhancing students' reading comprehension, as students in the experimental group outperformed those in the control group after receiving the treatment.

Interview Results

To support the quantitative findings and gain deeper insights into the students' perceptions and experiences, the researcher conducted semi-structured interviews with five participants from the experimental group. These students were selected through purposive sampling, representing a range of speaking proficiency levels and participation engagement during the treatment phase. The aim of the interviews was to explore how students perceived the use of the ChatGPT voice feature in their speaking practice, including its usability, effectiveness, and any challenges encountered. The interview questions were open-ended and thematically grouped into four areas:

1. Experience and familiarity with ChatGPT Voice
2. Perceived benefits in speaking skills (fluency, vocabulary, pronunciation, confidence)
3. Challenges or limitations during interaction
4. Suggestions for improvement or integration into classroom learning

The qualitative data gathered from these interviews were then analysed using thematic analysis following Braun and Clarke's (2021) six-phase framework. Recurring themes were identified, coded, and interpreted to highlight patterns in student perceptions. These findings offer meaningful insight into how AI-powered speaking tools can influence learner behaviour, motivation, and skill development.

Thematic Analysis Summary (Braun & Clarke, 2021):

Theme	Supporting Quotes/Patterns	Interpretation
Increased Confidence	*"I wasn't scared of making mistakes." (P1); "It didn't judge me." (P2)*	Safe practice space lowered speaking anxiety
Enhanced Fluency & Accuracy	*"I speak faster now." (P3); "It helped me fix grammar." (P4); "Better intonation." (P2)*	Improvements in automaticity, grammar awareness, and pronunciation
Technical/Access Issues	*"Didn't understand me sometimes." (P5); "Bad connection." (P3)*	Highlights voice recognition and internet barriers

Theme	Supporting Quotes/Patterns	Interpretation
Need for Guided Support	*"I wish there were prompts." (P2); "Teachers should guide us first." (P1)*	Indicates importance of scaffolding and hybrid (AI + teacher) models
Motivation & Engagement	*"It was fun!" (P4); "I liked pretending it was real conversation." (P5)*	Authenticity of practice increases learner motivation
Desire for Assessment	*"There should be progress tracking." (P4)*	Learners want measurable feedback to monitor their development

Discussion

The findings of this study confirm that the use of ChatGPT with voice interaction significantly enhances students' English-speaking performance. Participants who engaged in AI-driven conversations showed marked improvements in fluency, vocabulary accuracy, grammatical structure, and pronunciation compared to those learning via conventional classroom methods. These results echo prior studies by Nuñez et al. (2025), which emphasized that AI-mediated speaking practice fosters more autonomous and effective language output.

Student perceptions collected during this study further underscore the practical value of AI tools. Many students reported increased speaking confidence and reduced anxiety due to the immediate and non-judgmental feedback offered by ChatGPT. This aligns with findings from Alsalem (2024), Duong and Suppasetseree (2024), and Pratiwi et al. (2024), who noted that AI tools encourage low-stakes experimentation and diminish the fear of making mistakes in oral communication. Moreover, Muniandy and Selvanathan (2024) observed that regular interaction with AI platforms contributes to a more positive learner identity in EFL contexts.

Despite the promising outcomes, several limitations were noted. Some students experienced cognitive fatigue and confusion after prolonged use of ChatGPT, especially when navigating longer dialogues or unfamiliar topics. These challenges correspond to prior observations by Safar and Anggraheni (2024) and Solak (2024) concerning the cognitive load and mental strain associated with sustained AI use. Furthermore, a subset of participants indicated that ChatGPT occasionally misinterpreted their utterances or failed to respond appropriately in context-sensitive interactions issues that suggest a current limitation in AI's pragmatic competence (Wang, 2025).

Nevertheless, the overall student response to ChatGPT was positive, with the majority expressing a desire to continue using it in future learning activities. This resonates with Slamet (2024), who emphasized the motivational impact of AI tools in sustaining learners' engagement in language tasks. The effectiveness of ChatGPT also mirrors patterns found in other EFL contexts such as Saudi Arabia, Vietnam, and China (Alsalem, 2024; Duong & Suppasetseree, 2024; Wang, 2025), indicating that the tool's pedagogical potential transcends geographic boundaries.

What makes this study unique, however, is its focus on the Indonesian EFL context—a setting often underrepresented in global AI-assisted language learning research (Baskara & Mukarto, 2023 & Hakim & Rima, 2022). Indonesian learners face specific challenges such as limited classroom exposure to spoken English, lack of native speaker interaction, and insufficient speaking assessments. In this environment, ChatGPT Voice presents itself as a scalable and cost-effective supplement to formal instruction. However, the success of such interventions depends on proper pedagogical integration and teacher mediation (Çolak, 2024).

This study confirms the efficacy of ChatGPT Voice in improving speaking proficiency among Indonesian EFL learners, it also highlights the need for a critical and guided implementation strategy. With appropriate teacher training and institutional support, AI tools can be harnessed to fill persistent gaps in language education and democratize access to authentic spoken interaction.

Conclusion

This study found that using ChatGPT with voice can help improve students' English-speaking skills. Students who practiced speaking with ChatGPT showed better results in terms of fluency, vocabulary, grammar, and pronunciation compared to those who used conventional learning methods. They also felt more confident and enjoyed using this tool. Although some students faced challenges such as confusion, internet issues, or misunderstandings from the AI, most of them had a positive experience. ChatGPT provided them with more opportunities to speak and practice without feeling nervous. This study shows that ChatGPT can be a useful tool for speaking practice in English classes, especially in Indonesia. ChatGPT can support students by providing more speaking time, feedback, and confidence. Teachers and schools may consider using ChatGPT as part of their lessons to make speaking practice more engaging and beneficial for students.

References

Almalki, G. (2022). The use of AI-powered tools in enhancing EFL learners' oral proficiency: A review of literature. *Arab World English Journal (AWEJ), 13*(1), 75–90. https://doi.org/10.24093/awej/vol13no1.6

Alsalem, S. (2024). AI-assisted speaking practice in EFL classrooms: Learners' engagement and performance. *International Journal of Emerging Technologies in Learning, 19*(2), 112–124. https://doi.org/10.3991/ijet.v19i02.45678

Anderson, L. W., & Krathwohl, D. R. (2020). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* Longman.

Baskara, A. M., & Mukarto. (2023). Speaking anxiety among Indonesian EFL learners: A case study at a senior high school in Central Java. *Indonesian Journal of Applied Linguistics, 13*(2), 287–299. https://doi.org/10.17509/ijal.v13i2.43521

Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. SAGE Publications. https://doi.org/10.4135/9781529780494

Bygate, M. (2018). *Learning language through task repetition*. John Benjamins Publishing. https://doi.org/10.1075/tblt.11

Çolak, E. (2024). Pragmatic limitations of AI-based speaking assistants in EFL settings. *Computer Assisted Language Learning Electronic Journal, 25*(2), 34–47.

Creswell, J. W., & Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). SAGE Publications.

Creswell, J. W., & Poth, C. N. (2023). *Qualitative inquiry and research design: Choosing among five approaches* (5th ed.). SAGE Publications.

Duong, T. H., & Suppasetseree, S. (2024). EFL university students' perceptions of ChatGPT as a speaking assistant: A mixed-methods study. *TESOL Journal, 15*(1), e12547. https://doi.org/10.1002/tesj.12547

Etikan, I., & Bala, K. (2020). Sampling and sampling methods. *Biometrics & Biostatistics International Journal*, *8*(1), 001–002. https://doi.org/10.15406/bbij.2020.08.00249

Field, A. (2021). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publications.

Godwin-Jones, R. (2023). Voice-enabled AI tools for language learning: Capabilities, limitations, and classroom potential. *Language Learning & Technology*, 27(1), 1–15. https://doi.org/10125/73448

Goh, C. C. M., & Burns, A. (2022). *Teaching speaking: A holistic approach* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/9781009099023

Hakim, L. N., & Rima, A. M. (2022). Using ChatGPT to improve speaking fluency among Indonesian EFL learners: A pilot study. *Journal of English Language Teaching and Linguistics, 7*(4), 689–703. https://doi.org/10.21462/jeltl.v7i4.1357

Kasim, U., & Raisha, S. (2023). Improving EFL students' speaking skills using ChatGPT: Students' perspectives. *Journal of English Language Teaching and Applied Linguistics*, 5(2), 47–56. https://doi.org/10.32996/jeltal

Kemendikbudristek. (2021). *Kurikulum Merdeka Belajar*. Jakarta: Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi.

Kusumaningrum, D. E., & Wijayanti, R. (2021). The effectiveness of Duolingo application in improving students' vocabulary mastery and reading fluency. *Journal of English Language Education and Literature*, 6(2), 122–133.

Li, M., & Hafner, C. A. (2022). Exploring the use of artificial intelligence (AI) tools in EFL speaking instruction: A case study of Chatbot-assisted learning. *Computer Assisted Language Learning*, *35*(4), 456–478. https://doi.org/10.1080/09588221.2020.1840303

Liu, J., Wang, Y., & Xu, B. (2020). Artificial intelligence in language education: A systematic review. *Computer Assisted Language Learning, 33*(8), 891–921. https://doi.org/10.1080/09588221.2020.1770290

Miles, M. B., Huberman, A. M., & Saldaña, J. (2020). *Qualitative data analysis: A methods sourcebook* (4th ed.). SAGE Publications.

Muniandy, R., & Selvanathan, T. (2024). Overcoming speaking anxiety with AI-based learning tools: A study on Malaysian undergraduates. *Asian EFL Journal, 26*(1), 144–160.

Nuñez, J., Castillo, S., & Peña, L. (2025). Enhancing speaking skills with AI voice chat: A Latin American university experience. *CALICO Journal, 42*(1), 22–40. https://doi.org/10.1558/cj.21949

Pan, Y., & Zhang, D. (2019). A longitudinal study on the effects of AI-assisted pronunciation training on EFL learners' speaking performance. *Language Learning & Technology*, *23*(1), 112–130. https://doi.org/10.125/llt.2019.112130

Pratiwi, S., Hidayati, A., & Anindita, R. (2024). From fear to fluency: Exploring the use of AI in reducing EFL students' speaking anxiety. *Indonesian Journal of English Language Teaching, 19*(1), 55–72. https://doi.org/10.25170/ijelt.v19i1.2345

Renandya, W. A., & Widodo, H. P. (2022). English language teaching in Indonesia: Policy and practice. In C. Garton & F. Copland (Eds.), *International perspectives on English language teaching* (pp. 95–110). Palgrave Macmillan.

Safar, H., & Anggraheni, N. (2024). Evaluating ChatGPT's effectiveness in promoting EFL speaking performance: A classroom action research. *Journal of English for Academic Purposes, 62*, 101256. https://doi.org/10.1016/j.jeap.2024.101256

Saputri, Y. D., & Pratama, R. (2022). Bee Lingual as a digital media to improve students' reading comprehension: A case study at junior high school level. *Indonesian Journal of EFL and Linguistics*, 7(1), 99–115.

Satar, H. M., & Wigham, C. R. (2020). New forms of interaction: Researching multimodal online communication in language learning. In S. Garton & F. Copland (Eds.), *The Routledge handbook of research methods in applied linguistics* (pp. 216–227). Routledge.

Siregar, E., & Fatimah, A. (2023). Challenges in implementing communicative language teaching in Indonesian EFL classrooms. *Indonesian Journal of Applied Linguistics*, 13(1), 88–99.

Smutny, P., & Schreiberova, P. (2020). Chatbots for learning: A review of educational chatbots and their effectiveness. *International Journal of Educational Technology in Higher Education, 17*, Article 13. https://doi.org/10.1186/s41239-020-00134-0

Wang, X. (2025). ChatGPT and oral English training: Potentials and pitfalls in Chinese university contexts. *Asian-Pacific Journal of Second and Foreign Language Education, 10*(1), 1–20. https://doi.org/10.1186/s40862-025-00280-z

Wang, Y., & Lin, C. H. (2023). The use of generative AI in English speaking classrooms: A case study of ChatGPT in EFL settings. *Computer Assisted Language Learning Journal*, Advance online publication. https://doi.org/10.1080/09588221.2023.2234879