

The Use of Text-to-Speech (TTS) Feature on ChatGPT in Developing Students' Speaking Skills in Terms of Accuracy, Fluency, and Comprehensibility

Shiva Putri Maulida¹, Ssamtidar², Hasriani G³

English Education Study Program, Faculty of Languages and Literature, Universitas Negeri Makassar, Makassar, South Sulawesi, Indonesia

E-mail: samsidar@unm.ac.id

Article History:

Received: 11 Juli 2026

Revised: 28 September 2026

Accepted: 08 Oktober 2026

Keywords: *Text-to-Speech (TTS), ChatGPT, EFL Learning, Fluency, Accuracy, Comprehensibility*

Abstract: *This study aimed to investigate the use of the Text-to-Speech (TTS) feature in ChatGPT in improving the speaking skills of seventh-grade students at SMPN 8 Makassar. Employing a pre-experimental one-group pretest-posttest design, the study involved 36 students selected through purposive sampling. The treatment was conducted over four sessions utilizing TTS-based activities, including pronunciation, shadowing practice, dialogue simulation, and interactive speaking tasks. Students' speaking performance was assessed across three dimensions, namely fluency, accuracy, and comprehensibility, using Heaton's (1988) scoring rubric. The findings revealed measurable improvements in all three aspects: fluency increased from 2.28 to 2.92 (28%), accuracy from 2.14 to 2.61 (22%), and comprehensibility from 2.67 to 3.60 (35%). The proportion of students classified as "very poor" declined significantly from 69.4% in the pre-test to 19.4% in the post-test. In comparison, 8.3% of students reached the "fair" category, and 2.8% achieved the "good" category. The paired sample t-test yielded a value of 10.937, exceeding the t-table value of 2.030, confirming a statistically significant improvement, further supported by a large effect size (Cohen's $d = 1.82$). Although the N-Gain score of 0.18 indicated a low magnitude of gain relative to the maximum possible improvement, these results suggest that ChatGPT's TTS feature is a useful and accessible tool for enhancing EFL students' speaking skills.*

INTRODUCTION

Speaking is one of the most essential productive skills in English language learning because it enables learners to communicate ideas, express opinions, and participate effectively in social and academic interactions. However, for learners of English as a Foreign Language (EFL), speaking remains one of the most challenging skills because it requires linguistic competence,

confidence, fluency, and spontaneous communication (Al Hosni, 2022). Many EFL students still experience difficulties in oral communication due to fear of making mistakes, limited vocabulary, low participation, lack of confidence, dependence on their first language (Ventik, 2017), as well as limited speaking opportunities, overcrowded classrooms, and inadequate learning facilities (Zrekat & Sohbani, 2022).

To address these challenges, instructional innovations are needed to provide students with greater opportunities for independent speaking practice. Conventional English instruction often emphasizes grammatical accuracy rather than communicative competence, limiting opportunities for authentic oral interaction (Kut et al., 2022). Consequently, integrating technology into language learning has become increasingly important to support speaking development through more engaging and interactive learning experiences.

Recent advances in artificial intelligence (AI) have transformed language education by offering intelligent, interactive, and personalized learning experiences. AI-powered applications, such as chatbots, speech recognition systems, and adaptive language learning platforms, enable students to practice speaking with immediate feedback while reducing communication anxiety (Li & Ni, 2021). Among these innovations, text-to-speech (TTS) technology provides accurate pronunciation models with natural intonation, helping learners improve pronunciation, rhythm, and stress patterns (Jufriadi, 2024).

One AI application that incorporates TTS technology is ChatGPT. Its text-to-speech feature enables learners to listen to accurate pronunciation, engage in simulated conversations, and receive assistance with vocabulary and sentence construction (Qorunnada et al., 2023). These features may reduce speaking anxiety, improve learners' confidence, and increase their readiness to communicate. Since speaking performance is also influenced by psychological factors such as speaking anxiety and fear of negative evaluation (Kalantar, 2024), supportive technologies and responsive practice partners can further enhance learners' confidence and oral performance (Nikkels et al., 2023).

Several previous studies have explored the integration of AI and TTS in English speaking instruction. Pratiwi (2024) highlighted the potential of ChatGPT for speaking practice but conducted only a conceptual review. Jufriadi (2024) found that ChatGPT improved students' pronunciation, vocabulary, fluency, and comprehension, although the study did not specifically utilize its text-to-speech feature. Meanwhile, Qorunnada et al. (2023) reported that TTS software improved students' pronunciation but focused only on pronunciation rather than overall speaking performance.

Although these studies demonstrate the educational potential of AI and TTS, important research gaps remain. Previous research has mainly discussed ChatGPT conceptually, examined general speaking improvement without integrating its TTS feature, or focused solely on pronunciation enhancement. Consequently, empirical evidence regarding the effectiveness of ChatGPT's text-to-speech feature in improving speaking skills comprehensively is still limited.

Based on these considerations, this study aims to investigate the use of ChatGPT's text-to-speech feature in improving the speaking skills of seventh-grade students at SMPN 8 Makassar. Specifically, it examines students' improvement in accuracy, fluency, and comprehensibility after participating in AI-assisted speaking activities. The findings are expected to contribute to the growing body of knowledge on AI-supported language learning while providing practical insights for English teachers seeking innovative strategies to enhance students' speaking performance

LITERATURE REVIEW

Speaking Skills

Speaking is one of the productive language skills that enables learners to express ideas, opinions, and feelings through oral communication. According to Al-Tamimi (2014), speaking is the ability to communicate verbally with others using language effectively. As one of the most challenging skills for English as a Foreign Language (EFL) learners, speaking requires learners to integrate linguistic knowledge with communicative competence to convey messages accurately and understandably.

Brown (2003) identifies five aspects of speaking performance, namely grammar, vocabulary, pronunciation, comprehension, and fluency. However, this study focuses on accuracy, fluency, and comprehensibility because these dimensions are considered the most observable and measurable for beginner-level learners in short-term speaking assessment (Heaton, 1988). Accuracy refers to the correct use of grammar, vocabulary, and pronunciation during oral communication (Thornbury, 2019; Derwing & Munro, 2021). Fluency is the ability to produce speech smoothly with minimal hesitation (Levis, 2020), while comprehensibility refers to how easily listeners can understand the speaker's intended message (Fulcher, 2020). These three aspects collectively represent the essential components of effective oral communication and therefore serve as the indicators for assessing students' speaking performance in this study.

Text-to-Speech (TTS) Technology

Text-to-Speech (TTS) is a technology that converts written text into synthesized speech. Trivedi et al. (2018) define TTS as a process in which written text is analyzed and transformed into digital speech output. Similarly, Ren et al. (2019) describe TTS as a speech synthesis system designed to produce natural and intelligible spoken language from textual input. Within Computer-Assisted Language Learning (CALL), TTS has been widely used to provide pronunciation models, listening practice, and opportunities for independent language learning (Krasnova & Bulgakova, 2014).

The use of TTS offers several pedagogical advantages in language learning. It provides learners with consistent pronunciation input, allows unlimited repetition, and supports self-paced learning, particularly in environments with limited exposure to authentic English speech (Chapelle, 2020). Through repeated listening and imitation, learners can improve pronunciation accuracy, speech rhythm, and confidence in speaking (Amin, 2022). However, TTS also has limitations because it primarily functions as one-way input and cannot provide immediate corrective feedback or replicate the spontaneity of real-life communication (Bione et al., 2017). Therefore, TTS should be viewed as a supplementary learning tool rather than a replacement for interactive speaking practice.

ChatGPT Text-to-Speech Feature

ChatGPT is an artificial intelligence application capable of generating human-like text responses and supporting language learning through various interactive features. One of its features is Text-to-Speech (TTS), which converts generated text into spoken language, enabling learners to listen to accurate pronunciation and natural speech patterns. This feature provides flexible learning opportunities because students can repeatedly access spoken models anytime and anywhere, supporting independent and student-centered learning (Chapelle, 2020).

The ChatGPT TTS feature has the potential to improve speaking performance in several ways. It exposes learners to grammatically correct sentence structures and contextual vocabulary usage, contributing to speaking accuracy (Nation, 2020). Furthermore, the natural rhythm and pacing of synthesized speech can help learners develop fluency by reducing unnecessary pauses and improving speech flow (Levis, 2020). Clear and intelligible audio models also support comprehensibility by helping learners produce spoken language that is easier for listeners to understand (Fulcher, 2020). Although previous studies have reported positive effects of TTS-assisted learning on speaking development, empirical evidence regarding the effectiveness of ChatGPT's TTS feature in Indonesian junior high school EFL classrooms remains limited. Therefore, this study investigates its contribution to improving students' speaking skills, particularly in terms of accuracy, fluency, and comprehensibility.

RESEARCH METHOD

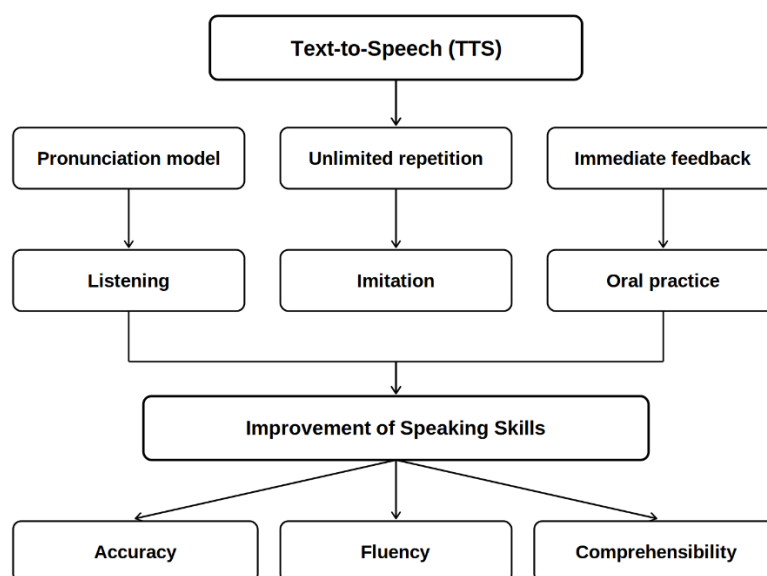
This study employed a pre-experimental research design using a one-group pretest-posttest approach to examine the effectiveness of ChatGPT's Text-to-Speech (TTS) feature in improving students' speaking skills. The design involved administering a pre-test before the treatment and a post-test after the treatment to measure changes in students' speaking performance following the intervention.

The study was conducted at SMPN 8 Makassar during the 2025/2026 academic year. The participants consisted of 36 seventh-grade students selected through purposive sampling. This class was chosen because the students demonstrated relatively low speaking performance and had limited opportunities to practice speaking English during classroom instruction.

The research instrument was an oral speaking test administered as both the pre-test and post-test. Students were asked to perform speaking tasks based on the instructional materials provided during the treatment. Their speaking performance was evaluated using Heaton's (1988) speaking assessment rubric, focusing on three aspects: accuracy, fluency, and comprehensibility. Each student's performance was scored using the established criteria to ensure consistency in the assessment process.

The treatment was conducted over four instructional meetings, integrating ChatGPT's Text-to-Speech feature into speaking activities. Students listened to pronunciation models generated by the TTS feature, practiced through shadowing, participated in dialogue simulations, and completed guided speaking tasks. These activities provided students with repeated exposure to accurate spoken English while encouraging active oral practice throughout the learning process.

The collected data were analyzed using descriptive and inferential statistics. Descriptive statistics were used to determine the mean scores and percentage improvements in students' speaking performance before and after the treatment. Inferential analysis was conducted using a paired-samples *t*-test to determine whether the improvement was statistically significant. In addition, the N-Gain score was calculated to measure the magnitude of students' learning improvement, while Cohen's *d* was employed to determine the practical effect size of the treatment. All statistical analyses were performed using SPSS Statistics.



Gambar 1. Conceptual Framework

FINDINGS AND DISCUSSION

Findings

Students' Improvement in Speaking Skills

Students' speaking performance was assessed based on three aspects: fluency, accuracy, and comprehensibility. The comparison between the pre-test and post-test mean scores indicates that students improved across all aspects of speaking after participating in learning activities that utilized ChatGPT's text-to-speech (TTS) feature.

Table 1. Students' Mean Scores in Speaking Aspects

Speaking Aspect	Pre-test	Post-test	Improvement
Fluency	2.28	2.92	28%
Accuracy	2.14	2.61	22%
Comprehensibility	2.67	3.60	35%

As shown in Table 1, students demonstrated improvement across all assessed aspects of speaking after participating in the TTS-assisted learning activities. Comprehensibility showed the greatest improvement, increasing by 35% from a mean score of 2.67 to 3.60. Fluency also improved from 2.28 to 2.92, a 28% increase, while accuracy increased from 2.14 to 2.61, a 22% increase. Overall, the findings indicate that the implementation of ChatGPT's text-to-speech feature positively contributed to students' speaking performance. Although improvement occurred across all speaking aspects, the highest gain was observed in comprehensibility, followed by fluency and accuracy.

Students' Speaking Performance Classification

To provide a clearer description of students' overall speaking achievement, the percentage distribution of speaking performance before and after the treatment was analyzed.

Table 2. Students' Speaking Score Classification

Criteria	Pre-test	Post-Test
Very Good	0%	0%
Good	0%	2.8%
Fair	0%	8.3%
Poor	30.6%	69.4%
Very Poor	69.4%	19.4%

Table 2 shows a noticeable shift in students' speaking performance after the implementation of the treatment. Before the treatment, 69.4% of the students were categorized as very poor, while the remaining 30.6% were classified as poor. No students achieved the fair, good, or very good categories during the pre-test. Following the treatment, the proportion of students in the very poor category decreased substantially to 19.4%. Meanwhile, the percentage of students in the poor category increased to 69.4%, indicating that many students moved from the lowest performance level to a higher category. Furthermore, 8.3% of the students reached the fair category, while 2.8% achieved the good category. These results demonstrate an overall positive shift in students' speaking performance after participating in the TTS-assisted learning activities.

Paired Sample *t*-test

To determine whether the improvement in students' speaking skills was statistically significant, a paired-sample *t*-test was conducted.

Table 3. Paired Sample *t*-test Result

Analysis	Result
<i>t</i> -value	10.937
<i>t</i> -table	2.030
Significance $p < .05$	

The analysis revealed that the obtained *t*-value (10.937) exceeded the critical *t*-table value (2.030), with a significance level below .05. Therefore, the null hypothesis (H_0) was rejected and the alternative hypothesis (H_1) was accepted. These findings indicate that the use of ChatGPT's text-to-speech feature resulted in a statistically significant improvement in students' speaking skills between the pre-test and post-test.

Effect Size

To examine the practical significance of the treatment, the effect size was calculated using Cohen's *d*.

Table 4. Effect Size

Analysis	Result
Cohen's <i>d</i>	1.82

The analysis yielded a Cohen's *d* value of 1.82, indicating a large effect according to Cohen's (1988) criteria. This result suggests that the implementation of ChatGPT's text-to-speech feature had a substantial practical impact on students' speaking performance beyond statistical significance.

N-Gain Analysis

The effectiveness of the treatment was further examined using the normalized gain (N-Gain) score.

Table 5. N-Gain Result

Analysis	Result
Mean Pre-test	7.06
Mean Post-test	9.03
N-Gain	0.18
Category	Low

The N-Gain analysis produced a score of 0.18, which is categorized as low based on Hake's (1998) classification. Although the normalized gain indicates a relatively low level of improvement, the statistically significant paired-sample *t*-test and the large Cohen's *d* value demonstrate that students consistently improved their speaking performance following the treatment. The increase in the mean score from 7.06 to 9.03 indicates measurable progress after four treatment sessions using ChatGPT's text-to-speech feature.

Discussion

The findings demonstrate that the implementation of ChatGPT's text-to-speech (TTS) feature contributed positively to students' speaking development. Before the treatment, most students were classified as very poor (69.4%), indicating limited oral proficiency. This finding is consistent with previous studies reporting that EFL learners commonly experience speaking difficulties due to limited opportunities for practice, speaking anxiety, fear of making mistakes, and reliance on their first language (Al Hosni, 2022; Ventik, 2017). Throughout the treatment, ChatGPT's TTS feature provided students with consistent pronunciation models and opportunities for repeated speaking practice in a low-pressure learning environment. Such conditions are considered beneficial for reducing speaking anxiety and increasing learners' confidence (Kalantar, 2024).

The classroom observations further supported the quantitative findings. During the pre-test, many students hesitated frequently, produced inaccurate pronunciation, and demonstrated strong first-language interference. After four treatment sessions, students showed noticeable improvements in speech rhythm, pronunciation, and confidence during oral performance. This development suggests that repeated exposure to TTS-based auditory input helped students gradually internalize English pronunciation patterns, supporting the development of more automatic spoken language production (Levis, 2020).

The improvement was evident across all assessed speaking aspects. Fluency increased from 2.28 to 2.92 (28%), indicating that students became more capable of producing speech with fewer pauses and hesitations. This finding supports Nation's (2020) argument that repeated exposure to meaningful language input facilitates the development of automaticity in oral production. The shadowing activities conducted through ChatGPT's TTS feature allowed students to imitate natural speech rhythm and pacing, thereby improving their fluency.

Accuracy also improved from 2.14 to 2.61 (22%). Although this increase was lower than the other speaking aspects, it remains theoretically reasonable because accuracy requires learners

to simultaneously master grammatical, lexical, and phonological features of language (Thornbury, 2019; Derwing & Munro, 2021). Through repeated exposure to grammatically accurate and well-pronounced speech generated by ChatGPT's TTS feature, students gradually developed greater awareness of appropriate vocabulary use, pronunciation, and sentence construction (Bione et al., 2017). However, the relatively limited improvement indicates that developing linguistic accuracy requires longer and more intensive practice.

The greatest improvement occurred in comprehensibility, which increased from 2.67 to 3.60 (35%). According to Fulcher (2020), comprehensibility reflects how successfully speakers communicate their intended meaning to listeners. ChatGPT's TTS feature consistently provided clear pronunciation, appropriate stress, and natural intonation, enabling students to imitate intelligible speech models during speaking practice. Consequently, students became more capable of delivering spoken messages clearly and coherently. This finding suggests that TTS is particularly effective in improving communicative clarity among beginner-level EFL learners.

The statistical analysis further supports these findings. The paired-sample *t*-test revealed a statistically significant difference between the pre-test and post-test ($t = 10.937$, $p < .05$), indicating that the observed improvement was unlikely to occur by chance. Furthermore, the Cohen's *d* value of 1.82 indicates a large practical effect of the treatment. Although the N-Gain score (0.18) falls within the low category according to Hake's (1998) classification, this result does not contradict the significant *t*-test because both analyses evaluate different aspects of learning improvement. While the *t*-test demonstrates the reliability of the improvement, the N-Gain score reflects the magnitude of learning gain relative to the maximum possible score.

Several factors may explain the relatively low N-Gain score. The intervention lasted only four treatment sessions, providing limited time for beginner-level students to develop substantial speaking proficiency. In addition, students had relatively little exposure to English beyond classroom activities, while first-language interference continued to influence pronunciation and sentence production. Therefore, the modest N-Gain score should be interpreted as reflecting the short duration of the intervention rather than the ineffectiveness of ChatGPT's TTS feature.

The present findings are consistent with previous studies. Qorunnada et al. (2023) reported that text-to-speech technology effectively improved students' pronunciation through shadowing activities. Likewise, Pratiwi (2024) highlighted the potential of ChatGPT to support speaking practice, while Amin (2022) found that repeated listening and reading activities using TTS applications enhanced learners' fluency and comprehension. Similarly, Bione et al. (2017) concluded that TTS serves as an effective pedagogical tool for improving pronunciation among second-language learners. The present study extends these findings by demonstrating that integrating ChatGPT with its text-to-speech feature contributes not only to pronunciation but also to students' fluency, accuracy, and comprehensibility within an EFL classroom context.

Overall, the findings indicate that ChatGPT's text-to-speech feature effectively improves students' speaking skills across the aspects of fluency, accuracy, and comprehensibility. The combination of statistically significant improvement, a large effect size, and positive classroom observations suggests that TTS-assisted learning provides meaningful support for speaking development. Nevertheless, longer implementation periods may produce greater learning gains and further strengthen students' oral communication skills.

.....

CONCLUSION

This study investigated the effectiveness of the ChatGPT text-to-speech (TTS) feature in improving the speaking skills of seventh-grade students at SMPN 8 Makassar. The findings demonstrate that the implementation of the ChatGPT TTS feature significantly improved students' speaking performance in terms of fluency, accuracy, and comprehensibility. The statistically significant paired-sample *t*-test result and the large Cohen's *d* effect size indicate that the intervention had a meaningful impact on students' speaking development. Although the N-Gain score was categorized as low, this result is likely attributable to the short duration of the treatment, the participants' beginner-level proficiency, and their limited exposure to English outside the classroom. Overall, the findings suggest that the ChatGPT TTS feature can serve as an effective AI-assisted learning tool to support the development of speaking skills in EFL classrooms.

Based on these findings, English teachers are encouraged to integrate ChatGPT's text-to-speech feature into speaking instruction as a supplementary learning tool that provides students with greater opportunities for pronunciation practice, dialogue simulation, and independent speaking activities. Future researchers are recommended to conduct similar studies with longer treatment periods, larger sample sizes, or different educational levels to further examine the long-term effectiveness of AI-assisted speaking instruction and explore its impact on other aspects of English language learning.

REFERENCES

- Al Hosni, S. (2022). Investigating speaking difficulties among EFL learners: Causes and solutions. *Arab World English Journal*, 13(1), 35–48.
- Al-Tamimi, N. O. M., & Attamimi, R. A. (2014). Effectiveness of cooperative learning in enhancing speaking skills and attitudes towards learning English. *International Journal of Linguistics*, 6(4), 27.
- Amin, E. A.-R. (2022). Using repeated reading and listening-while-reading via text-to-speech apps in developing fluency and comprehension. *World Journal of English Language*, 12(1), 211–219. <https://doi.org/10.5430/wjel.v12n1p211>
- Bione, T., Grimshaw, J., & Cardoso, W. (2017). An evaluation of TTS as a pedagogical tool for pronunciation instruction: The foreign language context. In K. Borthwick, L. Bradley, & S. Thoučny (Eds.), *CALL in a Climate of Change* (pp. 56–61). Research-publishing.net. <https://doi.org/10.14705/rpnet.2017.eurocall2017.689>
- Brown, H. D. (2007). *Teaching by Principles: An Interactive Approach to Language Pedagogy* (3rd ed.). Pearson Longman.
- Chapelle, C. A. (2020). *English Language Learning and Technology* (2nd ed.). John Benjamins Publishing Company. <https://doi.org/10.1075/lllt.61>
- Derwing, T. M., & Munro, M. J. (2021). *Pronunciation Fundamentals: Evidence-Based Perspectives for L2 Teaching and Research*. John Benjamins Publishing Company. <https://doi.org/10.1075/lllt.56>
- Dutoit, T. (1997). *An Introduction to Text-to-Speech Synthesis*. Kluwer Academic Publishers.
- Fulcher, G. (2020). *Practical Language Testing* (2nd ed.). Routledge. <https://doi.org/10.4324/9780429444377>
-

- Heaton, J. B. (1988). *Writing English Language Tests*. Longman.
- Jufriadi. (2024). Students' pronunciation problems in speaking English: A case study in Indonesian senior high school. *Journal of English Language Teaching and Linguistics*, 9(1), 45–56.
- Kalantar, F. (2024). An analysis of students' anxiety in speaking English: Causes and coping strategies. *International Journal of Behavioral and Educational Sciences*, 3(2), 110–120. <https://journal.unm.ac.id/index.php/ijobec/article/view/1872>
- Krasnova, E., & Bulgakova, E. (2014). The use of speech technology in computer-assisted language learning systems. In *Speech and Computer* (pp. 459–466). Springer. https://doi.org/10.1007/978-3-319-11581-8_57
- Kut, B. N., Coşkun, E. A., Zeren, E. H., & Tosun, İ. (2022). The role of topic-based virtual speaking club on Turkish adult EFL learners' emotional attitudes towards speaking in English. *The Journal of Turkish Educational Review*, 2(2), 7–27.
- Levis, J. M. (2020). Revisiting the intelligibility principle. *TESOL Quarterly*, 54(4), 1015–1033. <https://doi.org/10.1002/tesq.600>
- Li, Y., & Ni, H. (2021). The application of artificial intelligence in improving EFL learners' speaking competence. *Journal of Language Teaching and Research*, 12(5), 825–833. <https://doi.org/10.17507/jltr.1205.10>
- Nation, I. S. P. (2020). *Learning Vocabulary in Another Language* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/9781108915622>
- Nikkels, A., Berns, P., & Neijenhuis, K. (2023). Communication partner training for SLT students: Changes in communication skills, knowledge and confidence. *Journal of Communication Disorders*, 102, Article 106381. <https://doi.org/10.1016/j.jcomdis.2023.106381>
- Pratiwi, A. D., Rahmawati, S., & Nugraha, D. (2024). Enhancing students' speaking skills through AI-based applications: A case study of ChatGPT in EFL classrooms. *International Journal of Educational Technology*, 14(2), 101–112.
- Qorunnada, R., Fatimah, S., & Aulya, A. (2023). Improving students' pronunciation using text-to-speech (TTS) technology. *Journal of English Education and Teaching*, 7(3), 250–263.
- Simbolon, M. E. (2019). *Tuturan dalam Pembelajaran Berbicara dengan Metode Reciprocal Teaching*. Media Sahabat Cendekia.
- Thornbury, S. (2019). *How to Teach Grammar* (2nd ed.). Pearson Education Limited.
- Ventik, R. (2017). The factors that hinder students from speaking English. *Journal of English Language Teaching*, 6(2), 15–23.
- Zrekat, Y., & Sohbani, Y. A. (2022). Arab EFL university learners' perceptions of the factors hindering them to speak English fluently. *Journal of Language and Linguistic Studies*, 18(1), 775–790. <https://doi.org/10.52462/jlls.219>
-