Ahmadi Reyhane, MA Student Basarati Ali, Assis. Prof. VIZJA University

THE ROLE OF AI-GENERATED FEEDBACK IN IMPROVING ENGLISH PRONUNCIATION OF EFL LEARNERS

Abstract

This research examines the efficacy of AI-provided feedback, that is, through ChatGPT, in enhancing non-native students' English pronunciation. Twenty-eight Polish EFL learners were assigned to control and experimental groups. Instruction and correct feedback were administered to both groups over the course of eight weeks, but corrective feedback by ChatGPT was administered to the experimental group only. Preand post-test measurements were taken of three most important aspects of pronunciation: accuracy, intonation, and linking. Results showed that the experimental group improved considerably more on all aspects studied compared to the control group. The findings suggest that AI-driven feedback is an appropriate supplement to traditional instruction, which can elicit more autonomous and focused pronunciation improvement.

Keywords: ChatGPT; pronunciation; AI feedback; non-native learners; language learning

1. Introduction

The artificial intelligence (AI) development has been changing the language learning process (Woo & Choi, 2021). In recent years, AI has become an inseparable part of students' education worldwide, helping them improve their language skills. As classrooms become more digitally facilitated, AI is no longer a futuristic and imaginative concept; rather, it is becoming a pivotal part of students' everyday learning experiences. For instance, in the field of language learning, studies indicate that different AI tools, such as ChatGPT, can enrich language inputs and create opportunities for language learners to boost their communicative competence (Hidayatullah, 2024). It is also shown that AI mobile applications have helped improve listening skills (Alrasheedi, 2024). In addition, another study shows that the students who used AI tools had significantly higher reading comprehension scores than the ones who were taught using traditional methods (Hidayat, 2024).

Recently, little attention has been paid to studying how AI can help improve pronunciation. As one of the sub-skills of speaking, pronunciation is a challenging part of language learning for non-native English learners. Thus said, apart from the importance of knowing the right words that match the context, the way they are pronounced is key to being understood in conversations.

Literature Review

Recent advancements in artificial intelligence have significantly influenced language education, particularly in the domain of pronunciation training for non-native English speakers. AI-generated feedback, especially through tools like ChatGPT, offers scalable, personalized, and instant support to learners. Julio (2024) highlighted the effectiveness of integrating ChatGPT into pronunciation practice by generating contextualized drills and responsive corrections, which improved learner autonomy and accuracy in spoken English. Similarly, Koka (2024) demonstrated that AI chatbots can provide real-time feedback that enhances both motivation and phonological

awareness, leading to measurable gains in pronunciation. Takagi (2024) noted that students using generative AI received consistent, unbiased corrections on stress, intonation, and linking, which are crucial for speech fluency. Furthermore, Nuryah (2024) compared AI-generated feedback with teacher-led instruction and found AI to be effective in identifying persistent errors and tailoring exercises to specific learner needs. However, some scholars, such as Behzad (2024), caution that AI tools like ChatGPT may lack emotional nuance and cultural adaptability. Nevertheless, the cumulative evidence suggests that AI-driven feedback can be a valuable supplement to traditional teaching in pronunciation-focused learning environments.

2. Methodology

2.1. Research Design

This study employs an experimental design to investigate the extent to which AI-generated corrective feedback can enhance non-native English learners' pronunciation quality. Both experimental and control groups participated in pre- and post-tests. While all learners in the control group received the same pedagogical materials, only the experimental group was provided with feedback generated by ChatGPT. The study was conducted over eight weeks, with two sessions held each week.

2.2. Participants

In this research, a total of 28 Polish non-native English learners participated. The Oxford Placement Test was used to assess their English proficiency levels. The results indicated that the participants' English skills fell within the B1 range according to the CEFR. The participants were between 14 and 18 years old, and all spoke Polish as their first language.

2.3. Study timeline and procedure

This study was conducted over 16 class sessions (8 weeks, with two sessions per week), during which all four language skills – listening, speaking, reading, and writing - were taught. However, the teaching approach was designed to be learner-centered, encouraging students to produce as much language as possible. Each session lasted 90 minutes, consistent with regular class durations. In both the control and experimental groups, teachers provided students with corrective feedback on their pronunciation. However, in the experimental group, the feedback was generated by AI and delivered by the teacher. In contrast, the control group received traditional correction, where the teacher asked students to repeat words after checking the correct pronunciation in a dictionary.

The study focused on three key speaking features: segmental features (consonants and vowels), intonation, and linking. These features were selected due to their measurability and their compatibility with AI-generated tools, which can effectively support, analyze, and provide targeted feedback on these aspects of pronunciation.

Both the control and experimental groups participated in pre- and post-tests to assess improvements in their speaking skills. In the pre-test, conducted in Week 1 (Session 1), students were recorded while taking a speaking test. Two well-trained teachers assessed these recordings using a standardized rubric that evaluates three areas: accuracy, intonation, and linking, with scores ranging from 0 to 10. The same procedure was repeated in the post-test during Week 8 (Session 16), where students read Text C aloud to measure their final improvement. In the experimental group,

teachers analyzed student errors from the pre-test using the rubric and input them into ChatGPT to generate personalized feedback.

2.4. Instruments

The instruments used in this study included both pre-test and post-test speaking assessments designed to evaluate participants' pronunciation proficiency. A standardized pronunciation rubric, adapted from existing sources (specific citation to be added), was employed to assess three core areas: accuracy, intonation, and linking. This rubric used a 0–10 scoring scale and was applied consistently by two trained evaluators across both testing phases. For the experimental group, ChatGPT served as a tool for generating personalized feedback, including targeted pronunciation drills (e.g., minimal pairs, tongue twisters) and contextual speaking tasks. In contrast, the control group received traditional teacher-written feedback, which was documented using standardized teacher feedback forms based on the same rubric criteria.

Findings

Results from the current experimental study suggest that AI-generated feedback improved pronunciation performance of non-native English learners in the experimental group relative to the control group. An interrater reliability analysis was done for the pre- and post-tests using a rubric for intonation and linking with scoring out of ten. Initially, all participants had aggregated sub-scores of "Pronunciation," "Intonation," and "Linking" at similar means across levels, indicating matching baseline performance medians.

After a period of eight weeks, the experimental group exhibited gains in all assessed areas relative to their control counterparts. In the pronunciation category for the experimental group, students improved their averages from 4.1 to a post-test average of 7.3, representing an upward shift from "average" to "good" band. In comparison, the control group's average score also improved from 4.3 to 5.5, with average band still being their final placement.

Experimental group members also reported newfound success with intonation, with the group increasing their average score ranging from 3.8 to 6.9. The control group, in turn, improved slightly gaining 4.0 to 5.1. The experimental group also benefitted from improved expressive intonation due to drills focusing on rhythm and pitch at the sentence level to their recording output that were scripted in ChatGPT based on their capture.

With regard to linking, the experimental group increased from 3.5 to 7.0, suggesting a dramatic shift from minimal linked speech towards smoother and more natural prosody. The scores of the control group's linking increased by only a small margin, from 3.6 to 4.9. The computer-generated recommendations for practice at the word-to-word level of smoother transfer—particularly with contextual tongue twisters and phrasal reading drill—proved effective in creating fluency.

Qualitative feedback from teachers also confirmed that experimental group students were more involved with the AI-supported practice and more independent and motivated. Teachers observed more self-correction and higher involvement in speaking practice among this group. The control group students used more memorized corrections and habitual repetition of pronunciation patterns, with less internalization of feedback observable.

3. Discussion and conclusion

The findings suggest that the use of AI-generated feedback in pronunciation instruction has some pedagogical advantages. The significant improvement in all aspects covered by the experimental group, particularly linking and intonation, demonstrates the effectiveness of ChatGPT's specific, data-driven comments to create advanced speaking skills. Compared to traditional practices that primarily rely on repetition and generic correction, the AI can tailor responses to individual learner profiles.

REFERENCES

- 1. Alrasheedi, S. (2024). The Effect Of Using AI Applications To Develop Efl Listening Comprehension Skills Among University Students. *Conhecimento & Diversidade*, *16*(44), 601-637.
- 2. Behzad, S. (2024). Language learning meets generative AI: Utilizing large language models for *metalinguistic explanations* [Doctoral dissertation, Georgetown University]. Georgetown University Institutional Repository.
- 3. Hidayat, M. T. (2024). Effectiveness of AI-Based Personalised Reading Platforms in Enhancing Reading Comprehension. *Journal of Learning for Development*, *11*(1), 115-125.
- 4. Hidayatullah, E. (2024). The impact of TalkPal. AI on English speaking proficiency: an academic inquiry. *Journal of Insan Mulia Education*, 2(1), 19-25.
- 5. Julio, C. (2024). *Evaluation of speech recognition, text-to-speech, and generative text artificial intelligence for English as foreign language learning speaking practices.* Doctoral dissertation, Tokyo Denki University. TDU Institutional Repository.
- 6. Koka, N. A. (2024). An insight into the efficiency of artificial intelligence (AI)-chatbot as digital tutors for enhancing learners' motivation and performance in linguistics courses. *Pakistan Journal of Life and Social Sciences*, 22(2), 19812–19830.
- 7. Nuryah, A. (2024). *The comparison between lecturer and AI-powered platform feedback in assessing English pronunciation* Undergraduate thesis, UIN Syarif Hidayatullah Jakarta. UINJKT Repository.
- 8. Takagi, N. (2024). Generative AI in English language education: An annotated bibliography. *Ibaraki Christian University Journal*, 58(1), 113–127.
- 9. Woo, J. H., & Choi, H. (2021). Systematic review for AI-based language learning tools. *arXiv* preprint arXiv:2111.04455.