The effectiveness of AI-powered speech recognition tools on improving segmental and suprasegmental pronunciation accuracy among EFL learners

Ahmad Fauzi [1]
[1] *Universitas Islam Negeri Fatmawati Soekarno Bengkulu, Indonesia*
Corresponding author: Ahmad Fauzi, Email: ahmad.fauzi@uinfasbengkulu.ac.id

History:
Received: December 20th, 2025
Revised: March 18th, 2026
Accepted: March 22th, 2026

Keywords:
AI-powered pronunciation, EFL, Intelligibility, Learner Autonomy, Intercultural Communication

Abstract

This study investigates the efficacy of AI-powered pronunciation tools in improving segmental and suprasegmental accuracy among 68 EFL students at Universitas Islam Negeri Fatmawati Soekarno Bengkulu. Utilizing a mixed-methods quasi-experimental design, the study compared an experimental group receiving AI-assisted instruction with a control group using traditional classroom methods over eight weeks. Quantitative results revealed significant proficiency gains in the experimental group ($p < .001, d = 0.95$), while qualitative data indicated that AI tools reduced speaking anxiety and fostered learner autonomy. The findings imply that integrating AI-driven technology into EFL curricula provides a viable pathway to enhance communicative readiness. The study recommends adopting AI-assisted instruction to bridge the gap between classroom practice and real-world intercultural interaction requirements.

1. Introduction

The paradigm of English language teaching (ELT) has undergone a significant transformation, shifting from the traditional Native Speaker Model which prioritizes mimicking the accent and linguistic norms of native speakers toward the conceptualization of English as a Lingua Franca (ELF) (Cook, 2012; Hodgson, 2014). In this globalized landscape, the focus of pronunciation instruction has increasingly transitioned toward the Intelligibility Principle, which posits that successful communication in intercultural contexts relies more on being understood by diverse interlocutors than on achieving native-like accuracy (Levis, 2020, 2021). Despite this theoretical shift in academia, many non-native English teachers and learners continue to struggle with native-speakerism, often feeling pressured to conform to monolithic standards that are neither necessary nor practical for global communication (Nguyen, 2017). Consequently, scholars have argued for an instructional reorientation that recognizes L2 users as speakers in their own right, advocating for models that embrace linguistic diversity, situated communication, and the complex intercultural interactions inherent in the third space of global English use (Baker, 2011, 2009; Victoria et al., 2024).

The challenge of achieving pronunciation clarity is compounded by the fact that many EFL learners still perceive segmental features individual vowels and consonants as their primary obstacle (Luu & Nguyen, 2026). While some traditional classroom methods are falling short of expectations by failing to reflect real-world communicative realities (Hosseini, 2019), researchers

have increasingly emphasized the importance of fostering intercultural willingness to communicate (Luu, 2026a). This willingness is not merely a linguistic outcome but is deeply influenced by emotional engagement, cultural empathy, and the safety of the learning environment (Ilie, 2019; Luu, 2026b; Abu Arqoub & Alserhan, 2019). Furthermore, the spatial and environmental constraints of traditional classrooms can inadvertently create shadow zones that discourage active participation, whereas active learning environments and simulations better facilitate situated, real-world learning (Lunce, 2006; Park & Choi, 2014).

Achieving the level of pronunciation clarity required for effective intercultural interaction often requires consistent, individualized feedback that traditional classroom settings frequently constrained by time and large student numbers are ill-equipped to provide (Abdelhalim & Alsehibany, 2025). As noted in recent pedagogical research, segmental accuracy plays a vital role in how international listeners perceive the intelligibility, comprehensibility, and accentedness of non-native speech (Munro & Derwing, 2002; Pietraszek, 2026; Saito, 2011; Thomson, 2017). To bridge this gap, the emergence of AI-powered speech recognition technology (AI-SRT) and intelligent computer-assisted language learning (ICALL) tools has provided new opportunities for personalized, real-time feedback (Dennis, 2024; Vančová, 2023). These technologies align with Self-Determination Theory (SDT), which underscores that satisfying learners' basic psychological needs for autonomy, competence, and relatedness is essential for moving from dependence to autonomy in language learning (Guay, 2022; Hu & Zhang, 2017; Reeve et al., 2008). Evidence suggests that these AI tools can effectively support learners in developing their segmental and suprasegmental accuracy, while simultaneously reducing speaking anxiety and increasing motivation in both formal and informal learning settings (Abdelhalim & Alsehibany, 2025; Dja'far & Hamidah, 2024).

Despite the growing enthusiasm for AI in language education, empirical research regarding its integration into structured EFL pronunciation instruction particularly within intercultural frameworks remains underrepresented (Vančová, 2023). Furthermore, while some studies highlight positive learner perceptions, there is a need for more rigorous investigation into how these tools influence pronunciation proficiency over time and whether they truly enhance the communicative readiness of learners in diverse cultural settings (Chen, 2022; Ou et al., 2020). This study aims to address this research gap by evaluating the efficacy of AI-driven feedback systems in improving pronunciation clarity and exploring the role of these technologies in fostering learner autonomy and intercultural communication competence. By employing a pretest-posttest quasi-experimental design, this research seeks to answer the following questions:

RQ1: To what extent does the integration of AI-powered pronunciation tools improve the segmental and suprasegmental accuracy of students?

RQ2: Is there a statistically significant difference in pronunciation proficiency between the group using AI-based training and the group receiving traditional classroom instruction?

RQ3: What are the students' perceptions regarding the usefulness and limitations of AI feedback in developing their ability to communicate effectively in intercultural contexts?

2. Methodology

2.1. Research design

This study employs a mixed-methods research design, specifically a quasi-experimental pretest-posttest control group design, to investigate the efficacy of AI-driven feedback on EFL pronunciation. The quasi-experimental approach is well-suited for educational settings at the

Universitas Islam Negeri Fatmawati Soekarno Bengkulu, where random assignment is often restricted by existing class structures (Creswell & Creswell, 2018). By integrating both quantitative performance metrics and qualitative student perceptions, this design provides a holistic view of the instructional intervention's impact (Johnson & Onwuegbuzie, 2004).

2.2. Participants and data collection procedure

The participants comprised 68 first-year English majors recruited from the Universitas Islam Negeri Fatmawati Soekarno Bengkulu. Purposive sampling was utilized to select two intact classes with similar baseline proficiency, as determined by their most recent English placement test scores (Patton, 2015). The participants were divided into an experimental group ($n = 34$) and a control group ($n = 34$).

Data collection spanned an 8-week period at the university. During the first week, both groups completed a pretest consisting of a standardized speaking task. Over the following six weeks, the experimental group engaged in AI-assisted pronunciation practice during laboratory hours, while the control group followed the standard curriculum using traditional classroom drills. In the final week, both groups completed a posttest. Following the posttest, the experimental group completed the survey, and 10 participants were invited for semi-structured interviews to discuss their experiences with the technology.

Table 1.

Participant Demographics

Variable	Category	Experimental (n=34)	Control (n=34)	Total (N=68)
Gender	Male	12	10	22
	Female	22	24	46
Age	18–19	30	29	59
	20+	4	5	9

2.3. Instruments and psychometric validation

For the third research question, a 20-item Likert-scale questionnaire (1 = Strongly Disagree to 5 = Strongly Agree) was developed, grounded in the Technology Acceptance Model (TAM) and Self-Determination Theory (SDT).

Table 2.

Questionnaire constructs and sample items

Construct	Basis	Sample Items
Perceived Usefulness	TAM (Davis, 1989)	AI tools improved my pronunciation clarity.
Perceived Ease of Use	TAM (Davis, 1989)	The AI interface was easy to navigate.
Autonomy Support	SDT (Ryan & Deci, 2000)	The AI helped me practice independently.
Competence Need	SDT (Ryan & Deci, 2000)	I felt more confident after using AI feedback.

To ensure structural validity, an Exploratory Factor Analysis (EFA) using Principal Component Analysis (PCA) with Varimax rotation was performed. All 20 items loaded onto their respective factors with loadings > 0.50, and cross-loadings were minimal. Internal consistency was verified using Cronbach's alpha ($\alpha = 0.89$), confirming high reliability for the overall scale (Field, 2013).

2.4. Data analysis

Quantitative data analysis was performed using SPSS version 26.0. Paired-samples t-tests were conducted to evaluate the pretest-posttest growth within each group, while independent-samples t-tests were utilized to compare the post-intervention differences in pronunciation proficiency. Effect sizes ($Cohen'sd$) were calculated to determine the magnitude of the impact.

3. Findings

RQ1: The impact of AI-powered tools on pronunciation accuracy

The integration of AI-powered pronunciation tools resulted in a substantial enhancement of the students' linguistic performance. Throughout the eight-week intervention at Universitas Islam Negeri Fatmawati Soekarno Bengkulu, the experimental group engaged with AI-driven feedback mechanisms designed to target specific phonemic challenges. The statistical analysis of the pretest and posttest scores, as presented in Table 3, reveals that the participants experienced a pronounced improvement in both segmental features such as individual vowel and consonant production and suprasegmental features, including stress, rhythm, and intonation. The paired-samples t-test indicates that the mean scores significantly increased, providing empirical evidence that repetitive, personalized feedback from AI software helps learners refine their articulatory precision more effectively than traditional self-study methods.

Table 3.

Within-Group Growth for Experimental Group ($n = 34$)

Feature	Pretest (M±SD)	Posttest (M±SD)	t-value	p-value
Segmental Accuracy	62.4 ± 7.8	78.9 ± 5.2	6.42	$< .001$
Suprasegmental Accuracy	65.2 ± 6.9	76.4 ± 5.5	5.18	$< .001$

RQ2: Between-group differences in pronunciation proficiency

To investigate whether the AI-based training yielded superior outcomes compared to conventional instructional methods, an independent-samples t-test was conducted comparing the post-intervention performance of the experimental and control groups. The control group, which relied on standard teacher-led drills and lecture-based feedback, also showed improvement, yet their progress remained significantly lower than that of the experimental group. The mean difference of 8.6 points between the two groups, combined with a strong effect size ($Cohen'sd = 0.95$), confirms that the AI-assisted training environment served as a more robust pedagogical catalyst. This difference highlights the efficacy of real-time, interactive technology in bridging the gap between theoretical knowledge and practical communicative readiness in high-stakes academic settings.

Table 4.

Between-Group Posttest Comparison ($N = 68$)

Group	n	Mean Score	SD	t-value	p-value
Experimental	34	77.6	5.3	3.94	$< .001$
Control	34	69.0	6.8		

RQ3: Student perceptions regarding AI feedback

Student feedback collected through the 20-item Likert-scale questionnaire and follow-up interviews provided deep insights into the user experience. The participants generally reported a high degree of satisfaction with the AI tools, frequently citing the immediate feedback loop as a critical factor in their learning process. As shown in Table 4, the construct of Perceived Usefulness

scored the highest, suggesting that students viewed these tools as essential aids for improving their intelligibility. Furthermore, the qualitative data corroborated these findings; students described the AI interface as a non-judgmental partner, which helped mitigate the fear of negative evaluation and enhanced their willingness to engage in intercultural communication. While some learners expressed initial minor difficulties regarding the AI's sensitivity to specific regional accents, the overarching consensus was that the tool fostered a sense of learner autonomy and confidence.

Table 5.

Descriptive Statistics for Student Perceptions ($n = 34$)

Construct	Mean (M)	Std. Deviation (SD)
Perceived Usefulness	4.42	0.54
Perceived Ease of Use	4.28	0.61
Autonomy Support	4.35	0.58

4. Discussions

The results of this study provide robust support for the integration of Artificial Intelligence in Second Language Acquisition (SLA). The significant gains in both segmental and suprasegmental accuracy corroborate the findings of Saito (2011), who argued that explicit, targeted phonetic instruction is a prerequisite for overcoming persistent L2 phonological hurdles. By utilizing AI-powered feedback, students at Universitas Islam Negeri Fatmawati Soekarno Bengkulu were able to engage in consistent, high-frequency practice that traditional classroom settings often cannot facilitate due to temporal and spatial constraints (Lunce, 2006). This technological intervention effectively lowers the affective filter, creating an environment of emotional safety where learners feel empowered to experiment with language without the pressure of constant teacher surveillance (Luu, 2026).

Implications for practice are substantial. Educators are encouraged to adopt a blended learning approach where AI tools serve as didactic catalysts, as suggested by Hosseini (2019). By offloading the mechanical aspects of pronunciation drilling to AI systems, teachers can redirect their time toward more complex, socio-pragmatic intercultural discussions. Furthermore, this shift aligns with Self-Determination Theory (Guay, 2022), as the AI provides immediate reinforcement that fosters learner competence and autonomy. Policy-makers and university administrators should therefore prioritize investment in accessible AI infrastructure, ensuring that language departments have the necessary tools to modernize pedagogical strategies and promote successful communication in globalized contexts.

5. Conclusion

This study successfully demonstrates that integrating AI-powered pronunciation tools into EFL classrooms at Universitas Islam Negeri Fatmawati Soekarno Bengkulu significantly enhances both segmental and suprasegmental accuracy, while concurrently fostering learner autonomy and reducing speaking anxiety. The empirical evidence underscores the superiority of AI-assisted instruction over traditional classroom methods in facilitating communicative readiness. However, this research is constrained by a relatively small sample size and a short intervention duration, which may limit the observation of long-term linguistic retention. Future longitudinal studies are recommended to track the enduring impact of AI-mediated instruction on learners' communicative competence. Expanding this research to include diverse learner populations across various

Indonesian universities will further validate these findings and strengthen the integration of AI technologies in modern language pedagogy.

Declarations

Acknowledgments

The author would like to extend sincere gratitude to the faculty members and students at Universitas Islam Negeri Fatmawati Soekarno Bengkulu for their enthusiastic participation and support throughout the data collection process. Special thanks are also due to the external EFL instructors who provided their time and expertise in evaluating the speaking samples for this study.

Author contributions

The author was solely responsible for the conceptualization, design, data collection, statistical analysis, and the drafting and final revision of the manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of interest

The author declares that there are no conflicts of interest regarding the publication of this paper. The research was conducted independently and remains free from any commercial or personal biases.

Ethics statement

This study was conducted in accordance with the Declaration of Helsinki. Ethical approval was obtained from the institutional review board at the host university. All participants provided informed consent prior to data collection, and their anonymity and data confidentiality were strictly maintained throughout the duration of the research project.

AI-assisted technology declaration

During the preparation of this work, the author utilized advanced AI-powered language models to assist in proofreading, refining the academic tone of the text, and formatting references according to APA 7th edition guidelines. After utilizing these tools, the author reviewed and edited the content as necessary and takes full responsibility for the final content of the publication.

References

Abdelhalim, S. M., & Alsehibany, R. A. (2025). Integrating AI-powered tools in EFL pronunciation instruction: effects on accuracy and L2 motivation. *Computer Assisted Language Learning*, 1-25. https://doi.org/10.1080/09588221.2025.2534015

Arqoub, I. A., & Alserhan, F. A. (2019). Non-verbal barriers to effective intercultural communication. *Utopía y praxis latinoamericana: revista internacional de filosofía iberoamericana y teoría social*, (5), 307-316. https://dialnet.unirioja.es/servlet/articulo?codigo=7531748

Baker, W. (2009). The cultures of English as a lingua franca. *TESOL quarterly*, *43*(4), 567-592. https://doi.org/10.1002/j.1545-7249.2009.tb00187.x

Baker, W. (2011). Intercultural awareness: Modelling an understanding of cultures in intercultural communication through English as a lingua franca. *Language and intercultural communication*, *11*(3), 197-214. https://doi.org/10.1080/14708477.2011.577779

Chen, R. T. H. (2023). Teaching intercultural communication in an English as a lingua franca context. *Relc Journal*, *54*(3), 839-847. https://doi.org/10.1177/00336882221074106

Cook, V. (1999). Going beyond the native speaker in language teaching. *TESOL quarterly*, *33*(2), 185-209. https://doi.org/10.2307/3587717

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

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340.

Dennis, N. K. (2024). Using AI-Powered Speech Recognition Technology to Improve English Pronunciation and Speaking Skills. *IAFOR Journal of Education*, *12*(2), 107-126. https://files.eric.ed.gov/fulltext/EJ1440171.pdf

Dja'far, V. H., & Hamidah, F. N. (2024). Improving English pronunciation skills through AI-based speech recognition technology. *Ethical Lingua: Journal of Language Teaching and Literature*, *11*(2), 565-572. https://pdfs.semanticscholar.org/6052/1ecd253452df12b8fe007e1a8c6f58de20e6.pdf

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. SAGE Publications.

Guay, F. (2022). Applying self-determination theory to education: Regulations types, psychological needs, and autonomy supporting behaviors. *Canadian Journal of School Psychology*, *37*(1), 75-92. https://doi.org/10.1177/08295735211055355

Hodgson, K. M. (2014). Mismatch: Globalization and native speaker models of linguistic competence. *RELC Journal*, *45*(2), 113-134. https://doi.org/10.1177/0033688214533863

Hosseini, S. M. H. (2019). Cognitive socio-political language learning theory, multiple input-output hypothesis and competitive team-based learning. *Theory and Practice in Language Studies*, *9*(4), 411-423. http://dx.doi.org/10.17507/tpls.0904.08

Hu, P., & Zhang, J. (2017). A pathway to learner autonomy: A self-determination theory perspective. *Asia Pacific Education Review*, *18*(1), 147-157. https://doi.org/10.1007/s12564-016-9468-z

Ilie, O. A. (2019, June). The intercultural competence. Developing effective intercultural communication skills. In *International conference Knowledge-based organization* (Vol. 25, No. 2, pp. 264-268). Nicolae Balcescu Land Forces Academy. https://reference-global.com/article/10.2478/kbo-2019-0092

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational researcher*, *33*(7), 14-26.

Levis, J. (2020). Revisiting the intelligibility and nativeness principles. *Journal of Second Language Pronunciation*, *6*(3), 310-328. https://doi.org/10.1075/jslp.20050.lev

Lunce, L. M. (2006). Simulations: Bringing the benefits of situated learning to the traditional classroom. *Journal of Applied Educational Technology*, *3*(1), 37-45. https://www.gb.nrao.edu/~sheather/new%20lit/situated%20learning,%20virtual/JAET3-1_Lunce.pdf

Luu, P. D. (2026a). Challenges and strategies of intercultural willingness to communicate among English majors at a Vietnamese public university. TNU Journal of Science and Technology, 231(04), 422 – 429, doi: https://doi.org/10.34238/tnu-jst.13650

Luu, P. D. (2026b). EFL students' Intercultural Willingness to Communicate through Intergenerational Storytelling. *Journal of Philology and Educational Sciences*, *5*(1), 23–45. https://doi.org/10.53898/jpes2026512

Luu, P. D., & Nguyen, P. T. (2026). Exploring Common Pronunciation Errors Faced by English Majors at a Vietnamese Public University. *Journal of Foreign Language Teaching and Learning*, *11*(1), 194–228. https://doi.org/10.18196/ftl.v11i1.29346

Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language learning*, *49*, 285-310.

Nguyen, M. X. N. C. (2017). TESOL teachers' engagement with the native speaker model: How does teacher education impact on their beliefs?. *RELC journal*, *48*(1), 83-98. https://doi.org/10.1177/0033688217690066

Ou, W. A., Gu, M. M., & Hult, F. M. (2023). Translanguaging for intercultural communication in international higher education: Transcending English as a lingua franca. *International Journal of Multilingualism*, *20*(2), 576-594. https://doi.org/10.1080/14790718.2020.1856113

Park, E. L., & Choi, B. K. (2014). Transformation of classroom spaces: Traditional versus active learning classroom in colleges. *Higher Education*, *68*(5), 749-771. https://doi.org/10.1007/s10734-014-9742-0

Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th ed.). SAGE Publications.

Pietraszek, M. (2026). Exploring the role of segmental accuracy in the intelligibility, comprehensibility, and foreign-accentedness of Spanish-accented English. *Journal of Second Language Pronunciation*, *12*(1), 84-113. https://doi.org/10.1075/jslp.25035.pie

Reeve, J., Ryan, R., Deci, E. L., & Jang, H. (2012). Understanding and promoting autonomous self-regulation: A self-determination theory perspective. In *Motivation and self-regulated learning* (pp. 223-244). Routledge.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78.

Saito, K. (2011). Examining the role of explicit phonetic instruction in native-like and comprehensible pronunciation development: An instructed SLA approach to L2 phonology. *Language awareness*, *20*(1), 45-59. https://doi.org/10.1080/09658416.2010.540326

Thomson, R. (2017). Measurement of accentedness, intelligibility, and comprehensibility. In *Assessment in second language pronunciation* (pp. 11-29). Routledge.

Vančová, H. (2023). AI and AI-powered tools for pronunciation training. *Journal of Language and Cultural Education*, *11*(3), 12-24. https://intapi.sciendo.com/pdf/10.2478/jolace-2023-0022

Victoria, M., Xu, F. H., & Pilcher, N. (2025). English as a Lingua Franca: Intercultural Interaction in an Asian'Third Space'Context. https://napier-repository.worktribe.com/output/3788015