

PRONUNCIATION DEVELOPMENT IN EFL LEARNERS USING AI SPEECH RECOGNITION SYSTEMS

Milimo Mundia

English Lecture, Kimyo International University in Tashkent, Namangan Branch

ORCID: <https://orcid.org/0009-0003-7785-8582>

m.milimo@kiut.uz

Abstract: *The ability for students to pronounce words correctly is one of the most important elements of successful communication in an English as a Foreign Language (EFL) environment. Pronunciation has many facets that are difficult for EFL students including phonetic accuracy, word stress and intonation. These aspects of pronunciation may be influenced by a lack of exposure to these sounds in addition to a lack of adequate feedback. With recent advances in Artificial Intelligence (AI), speech recognition technology has opened new avenues in language acquisition. Through real time feedback, AI provides personalized instruction and allows students to practice speaking at will. Therefore, this study examined how effectively AI powered speech recognition can improve the pronunciation of students studying English as a foreign language. The focus was placed on major phonological features including phonetics, stress and intonation. Using a quasi-experimental methodology, eighty university students were randomly assigned to either an experimental or control group. Students completed three different forms of data collection, including; pronunciation assessments, acoustic analyses and a survey asking about their perceptions related to the use of AI technology. Statistically, it was demonstrated that there was significantly better pronunciation from students who utilized AI technology. These results demonstrate the great potential of AI in enhancing EFL student's pronunciation abilities while allowing for autonomous learning. It is suggested that AI based pronunciation tools should be incorporated into EFL curriculum so as to promote accurate and confident communication.*

Keywords: *Artificial Intelligence, Speech Recognition, EFL Learners, Phonetics, Stress, Intonation, Pronunciation Development.*

INTRODUCTION. As globalization continues to expand, English is becoming increasingly prevalent as the primary language of global business, academic communication and technological exchange. As a result, there is a growing need for non-native speakers of English (English as a Second/Foreign Language or ESL/EFL) to achieve proficient pronunciation so that they can be fully understood by others when communicating in English. In addition to being a critical component of developing successful career paths for students interested in pursuing careers in management, accurate pronunciation will facilitate effective verbal communication. Despite decades of formal instruction in English, many EFL learners still experience difficulties pronouncing words correctly in English. Some common reasons why EFL

learners have difficulty with pronunciation include interference from their first language(s), limited access to authentic materials and/or inadequate corrective feedback during instruction. Many traditional instructional strategies do not allow for the same degree of one-on-one interaction necessary to provide individualized corrective feedback and therefore cannot provide adequate identification of learners' phonological errors.

Recently, advancements in the field of Artificial Intelligence (AI) have led to new forms of computer-assisted language learning (CALL) that utilize speech recognition technologies. Examples of speech recognition technologies include Google's Speech-to-Text, Microsoft's Azure Speech Services, ELSA Speak and Duolingo. All of these technologies provide learners with immediate feedback on their pronunciation accuracy based on comparisons of their spoken utterances against those produced by native speaker models. Therefore, learners can continue to refine their phonetic productions and prosody through continued practice using AI-assisted technology. The purpose of this paper is to investigate how AI-assisted speech recognition systems improve pronunciation among EFL learners, focusing specifically on phonetics, word stress, and intonation.

LITERATURE REVIEW. Pronunciation is a key element of obtaining communicative competence in English as a Foreign Language (EFL) as it allows learners to convey their thoughts effectively and engage confidently in meaningful communications. While grammar and vocabulary comprise the structure of the language, pronunciation facilitates that spoken message is perceived clearly. If a learner does not pronounce words clearly, even if their sentence structure is grammatically correct, the listener may interpret their intended message incorrectly, limiting effective communication and potentially decreasing the learner's confidence.

Many researchers agree that for pronunciation to be understandable, learners must master both segmental and suprasegmental features. Segments are defined as individual vowel/consonant sounds. Suprasegments refer to pitch/stress patterns found in syllables and longer sequences of sounds. According to Celce-Murcia et al. (2010), mastering all aspects of these segments/supra-segments allows learners to create a natural flow of speech that is easily understood by listeners [1]. Nevertheless, many EFL learners encounter problems with pronunciation due to interference from their native language(s) and consequently avoid participating in class discussions and communicating verbally. Understanding phonetics allows learners to understand the mechanics behind producing speech and hearing it; thus allowing learners to develop both their articulatory (production) skills and receptive listening skills. Stress on certain words and pitch patterns help learners convey emotions and meaning behind what they say.

Roach (2009) notes that suprasegmental are important for creating fluent and natural speech; hence, learners who successfully apply stress and pitch patterns are

able to go beyond simply mimicking pronunciation to actually communicate authentically[2].

Artificial Intelligence (AI) has revolutionized language learning through its application in Computer Assisted Language Learning (CALL). CALL utilizes AI tools to deliver personalized, interactive and adaptive learning experiences that support learners practicing independently. Through instantaneous and objective feedback, learners can recognize their own errors, develop their pronunciation, build confidence and ultimately learn more efficiently. Studies conducted by Liakin et al. (2015), Cardoso & Liakina (2015) and McCrocklin (2016) indicate that Automatic Speech Recognition (ASR) improves the quality of pronunciation and enhances self-assessment [3]. Furthermore, AI speech recognition tools available today; ELSA Speak, Google Speech-to-Text, Microsoft Azure Speech Services, Duolingo etc., enrich the possibilities for EFL learners. These tools analyze acoustic signals of learners' utterances and compare them to native speaker models; therefore, providing learners with specific feedback and interactive exercises for developing their speaking abilities outside of the traditional classroom environment. Learners can now utilize these accessible and easy-to-use tools independently and develop their ability to speak English beyond the confines of a traditional classroom.

METHODOLOGY. To assess the efficacy of AI-powered speech recognition tools in enhancing pronunciation among EFL learners, this study used a quasi-experimental research methodology employing pre-test/post-test control groups. This design was selected because it is suitable for studying educational environments where randomized assignments are typically unfeasible. The method allowed for a controlled comparison of students receiving AI-assisted pronunciation instruction versus students using traditional instructional methods, providing valid information about the influence of technology on EFL learning outcomes. There were 80 undergraduate students from management departments at a university who participated in the study. The students were sampled based on purposeful selection criteria, dividing them equally into an experimental group and a control group. The experimental group received AI-based pronunciation training, while the control group received traditional instructional techniques. Both groups had similar linguistic backgrounds and proficiency levels, allowing for comparable conditions throughout the study.

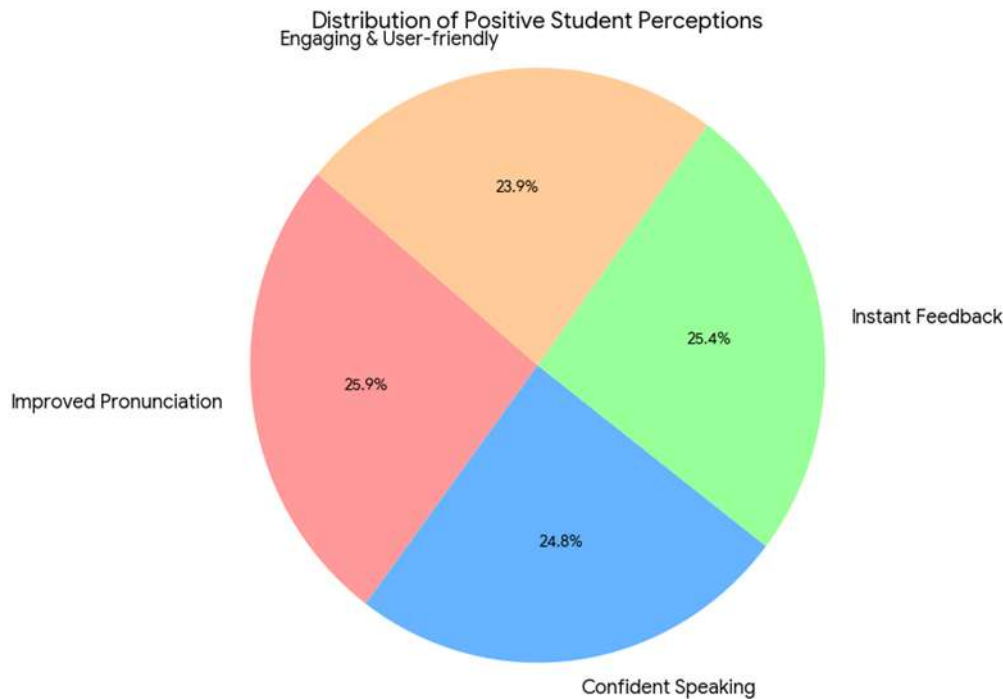
Multiple research instruments were employed to collect data from multiple perspectives. A pronunciation proficiency test was administered prior to the intervention and again following completion of the study to measure growth. Students utilizing AI-powered tools such as ELSA Speak, Google Speech-to-Text, Microsoft Azure Speech Services received personalized feedback and instruction. A pronunciation assessment rubric was developed to evaluate both segmental features (vowels and consonants) and suprasegmental features (stress, rhythm, intonation) of students' pronunciation. To supplement quantitative measures of students' attitudes toward AI-assisted learning, a student perception survey was completed by each participant.

Finally, audio recordings were collected from students to ensure accurate pronunciation assessments.

Descriptive statistical analysis (means and standard deviations) was employed to describe general trends regarding students' performance. Paired sample t-tests and independent sample t-tests were utilized to determine whether differences existed between students' performance at various points during the study. Effect sizes were computed using Cohen's *d* to quantify the magnitude of differences between students' performance at different stages of the study. Finally, thematic analysis of qualitative comments obtained from students' surveys helped illuminate additional insight into students' experiences and perceptions related to AI-based pronunciation instruction.

RESULTS. The results of this investigation illustrate a clear benefit for students who received AI-based pronunciation instruction relative to those who received traditional instruction. Prior to beginning either treatment condition, both groups exhibited virtually equivalent proficiency levels on the pre-test administration of the Pronunciation Proficiency Test (mean scores = 52.4 vs. 51.9). Thus, the two groups could be considered equivalent at the start of the study, which is consistent with an appropriate experimental design. By the final testing session at the conclusion of the study, however, a marked disparity in proficiency emerged between the two groups. Specifically, the experimental group which received AI-based instruction attained a pronounced increase in proficiency (post-test mean score = 78.6), substantially exceeding that experienced by the control group (post-test mean score = 64.2). This pronounced increase in proficiency illustrated an increased rate of learning for students who received AI-based instruction compared to those who did not receive AI-based instruction.

Statistical analyses supported these interpretations. Specifically, the paired-sample t-test evaluating change in proficiency over time for students in the experimental group yielded a t-score of 14.21 ($p < .001$), illustrating that changes in proficiency were statistically reliable and occurred reliably over time. Although the control group also evidenced some level of increase in proficiency over time (t-score = 6.32; $p < .05$), increases experienced by the control group were considerably less pronounced than those experienced by members of the experimental group. Further supporting interpretation of the results as being due to effects associated with receipt of AI-based instruction was the finding from an independent-samples t-test ($t = 9.47$; $p < .001$) that at the conclusion of the study there existed a statistically reliable difference between performance of members from the two groups. Lastly, quantifying the magnitude of this difference via calculation of an effect-size statistic (Cohen's *d* = 1.30) demonstrated that this difference represented a "large" effect. Collectively, these findings demonstrate convincingly that AI-based instruction accelerates language acquisition processes and improves EFL learner's ability to accurately produce pronunciation in English as well as improves their overall performance in language acquisition (pic 1).



Pic. 1. Distribution of positive student's perceptions

DISCUSSION. The results of this study confirm that there is substantial evidence that demonstrates the use of artificial intelligence (AI) powered speech recognition tools can be used to develop better pronunciation amongst English as a foreign language (EFL) learner [5]. Those learners utilizing the AI assisted instructional model developed improved pronunciation when compared to those using traditional models. The data collected demonstrated that both segmental features of pronunciation, (vowels and consonants) and suprasegmental features of pronunciation (stress and intonation) were improved due to the use of the AI based model. Learners utilized the instant and objective feedback provided by the AI tool to identify errors in pronunciation and modify them in a timely manner. This process fostered a sense of learner autonomy, self-regulation and continued development. This finding supports prior studies done by Liakin et al. (2015) and McCrocklin (2016). Both studies demonstrated the ability of automatic speech recognition (ASR) systems to transform language education.

There is also a great deal of consistency in the findings of this study and well-established theoretical frameworks. For example, the reduction in learner anxiety seen with the use of AI supported instructional models is reflective of Krashen's Affective Filter Hypothesis. Similarly, the increase in learner confidence and self-efficacy was reflective of Bandura's Social Cognitive Theory; specifically, it demonstrated the effect of consistent feedback from the AI system reinforcing a learner's belief in themselves. Finally, the interactive and experiential nature of the AI supported instructional model

is reflective of Constructivist Learning Theory; specifically, it demonstrated how learners constructed knowledge through engagement, reflection, and practice.

RECOMMENDATIONS. Several recommendations for future study have been identified as a result of the findings of this study. Pedagogically, teachers are encouraged to incorporate AI powered speech recognition tools into their EFL curriculum and implement blended learning models. Additionally, educators are encouraged to include emphasis on suprasegmental features of pronunciation, such as stress and intonation, in addition to opportunities for independent practice.

At the Institutional Level, Universities are encouraged to invest in AI supported language laboratories and provide educators with targeted training to promote digital literacy. Furthermore, universities are encouraged to create technology driven language policies. From a Policy Perspective, Ministries of Education are encouraged to support the digital transformation of language learning by incorporating AI tools into national curricula and provide funding for research and innovation related to the integration of AI in language learning.

Future Research Directions Future research directions will focus on examining the long term effects of AI on pronunciation, evaluating the effectiveness of AI across different proficiency levels, investigating whether AI can reduce speaking anxiety and improve fluency, and comparing multiple AI platforms to evaluate their comparative effectiveness.

CONCLUSION. Overall, this study has confirmed that AI powered speech recognition tools improve EFL student pronunciation accuracy, fluency, and learner confidence. Through its provision of individualized feedback, promotion of autonomous learning, and creation of a supportive learning environment, AI technologies facilitate both learner linguistic competence and learner motivation. Supported by established theoretical frameworks, the findings of this study illustrate the pedagogical value of integrating AI into language education.

REFERENCES:

- 1) Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). Teaching pronunciation: A course book and reference guide (2nd ed.). Cambridge University Press.
- 2) Mundia, Milimo. "A theoretical framework for an integrated, technology-enhanced curriculum to foster oral competence in english for management." Partner conferences of the International Scientific Journal Research Focus. Vol. 1. No. 1. 2025.
- 3) Usmanovich, Ibragimov Isroil, and Milimo Mundia. "Issues of improving the organizational and economic mechanism of business management." (2021): 10694-10702.
- 4) Mundia, M. Technology for improving the oral skills of management students технологии для улучшения навыков устной речи у студентов,

изучающих менеджмент menejment yo'nalishi talabalarini og'zaki nutqini takomillashtirish texnologiyasi. FarDU. ILMYIY XABARLAR, 82.

5) Khayrullayeva, N. N. "The Influence of Persian Classical Literature on English Literature." *International Journal of Language Learning and Applied Linguistics* 3.4 (2024): 107-111.

6) Gabdulxakova, Rasima Faridovna. "The phases of formation of speech competences among secondary school students while learning the russian language." *Bulletin of Namangan State University: Vol 1.3*: 67.

7) Tokhirova, Shokhsanam. "The Influence OF Physical Culture on Personality Development." *Maktabgacha va Maktab Ta'limi Jurnal*: 674367.

8) Kamalova, Gulnara Gulyamovna, Tulagan Agzamovich Sagatov, and Nizom Jumakulovich Ermatov. "Studying the individual modification of the ovaries of the women during intensive body height of genitals." *European research: innovation in science, education and technology*. 2018.

9) Turdaliyev, V. M., D. Shotmonov, and B. Egamnazarov. "Amortizator saylentbloki uchun sirpanish podshipnigini ta'mirlash komplektini takomillashtirish." *Механика и технология* 2.3 (2021): 21-26.

10) Ganiyev, A. A., Mavlonov, O. N., Shodmonov, D. A., & Maxmudov, J. I. (2024). Kompyuter tarmoqlarida uzatilayotgan axborotni kriptografik himoyalash usuli. *Международный Журнал Теоретических и Прикладных Вопросы Цифровых Технологий*, 7(1), 73-79.

11) Ziyaboyeva, S., Kazakova, M., Lazokat, D., & Ibrokhimova, M. (2021). Teaching english with the use of multimedia technologies. *Asian Journal of Multidimensional Research (AJMR)*, 10(1), 249-252.

12) Nargiza, O., & Diyora, A. (2025). Building a theoretically informed uzbek-english speech recognition system: practical implementation and evaluation. *Great britain-scientific review of the problems and prospects of modern science and education*, 1(8), 36-41.

13) Tozhibayev, Bahromjon, and Saidulla Isokov. "Youth Social Deviation in the Process of Reforms in Uzbekistan." *IX International Scientific and Practical Conference "Current Problems of Social and Labour Relations"(ISPC-CPSLR 2021)*. Atlantis Press, 2022.

14) Israilova, N. "Developing professional competence of teachers through the heuristic method." *Science and innovation* 2.B9 (2023): 204-205.

15) Muxammadjonovna, Mirzaliyeva Dilafruz. "Oilada kitobxonlik madaniyatini shakllantirish: pedagogik-psixologik yondashuv." *Universal journal of academic and multidisciplinary research* 3.25 (2025): 156-159.