

AI TECHNOLOGIES FOR ASSESSING STUDENTS' PRONUNCIATION

Olimova Shakhlobonu Khurshedjonovna

Assistant Teacher, Department of Languages, Zarmed University

Abstract:

The advancement of artificial intelligence (AI) has significantly impacted language learning, particularly in the assessment of students' pronunciation. This paper explores various AI-driven technologies used for evaluating pronunciation accuracy, fluency, and intelligibility in second-language learners. It also discusses the advantages and limitations of AI-based pronunciation assessment compared to traditional methods.

Keywords: Artificial intelligence, pronunciation assessment, speech recognition, machine learning, deep learning, language learning, phonetic analysis, neural networks.

Introduction

Pronunciation is a critical component of language learning, yet assessing it objectively remains a challenge. Traditional evaluation methods, such as teacher assessments and peer feedback, are often subjective and time-consuming. AI technologies provide a promising alternative by offering automated, data-driven assessments that ensure consistency and accuracy.

Research on AI technologies for assessing students' pronunciation has been conducted by various scholars globally and within Uzbekistan. Yassine El Kheir, Ahmed Ali, and Shammur Absar Chowdhury conducted a comprehensive review titled «Automatic Pronunciation Assessment -- A Review», [3] discussing recent advancements and challenges in computer-aided pronunciation training. Fu-An Chao, Tien-Hong Lo, Tzu-I Wu, Yao-Ting Sung, and Berlin Chen proposed the «3M» approach in their study «An Effective Multi-view, Multi-granularity, and Multi-aspect Modeling Approach to English Pronunciation Assessment», [2] integrating various features to enhance automatic pronunciation assessment. In Uzbekistan, a study titled «Creation of An Intelligent System for Uzbek Language Teaching» focused on developing an intelligent system for teaching Uzbek as a foreign language, utilizing phonemic recognition technology. The «Uzbek Speech Corpus and Automatic Speech Recognition» project developed the first open-source Uzbek speech corpus, aiming to aid automatic speech recognition and other speech-related tasks. These studies contribute to the advancement of AI-driven pronunciation assessment tools, both globally and within Uzbekistan.

The use of technology to measure deviations from a model pronunciation dates back to the 1990s (Shufang, 2021), when existing systems assessed segmental and sentence-level features without predefined criteria. By the early 2000s, ASR-based systems had begun evaluating pronunciation in ways that differed from human raters over extended speech samples, though assessments of individual phonemes remained comparable. Since native speakers collected the results, an additional challenge emerged in assessing pronunciation in a foreign or second language. According to Lee et al. (2010), exposure to comprehensible input enhances learners' semantic processing, which in turn facilitates the development of syntactic processing when producing speech. During communication, speakers receive feedback from their interlocutors, helping them



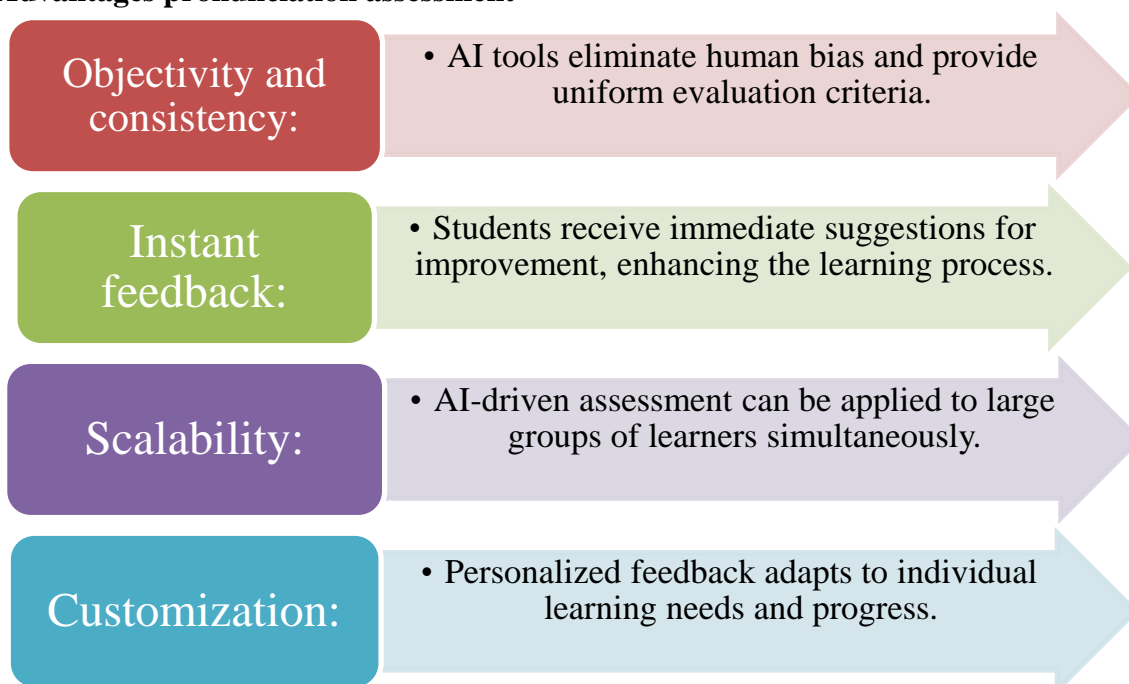
refine their pronunciation. In this context, robots can play a crucial role in motivating young learners by providing corrective feedback on mispronunciations. Despite extensive research on system functionality, comparatively little attention has been given to the pedagogical benefits and challenges of integrating such technology into pronunciation training. Shufang (2021) describes this interaction as human-computer or human-machine communication, incorporating aspects of speech recognition, natural language understanding, and speech synthesis. The latter, in particular, serves as a key link between artificial intelligence and this form of communication [4, 13].

Automatic Speech Recognition (ASR) systems, such as Google's Speech-to-Text and IBM Watson, analyze spoken input and compare it against native speaker models. These systems assess phoneme accuracy, stress patterns, and rhythm, providing detailed feedback to learners. AI-based pronunciation assessment tools leverage machine learning algorithms trained on large datasets of native and non-native speech samples. These models can identify mispronunciations, suggest corrections, and adapt to different accents.

Advanced AI tools use phonetic analysis to evaluate pronunciation at a granular level. Systems like Carnegie Speech and ELSA Speak measure acoustic features, including pitch, intensity, and duration, to provide precise feedback.

Recent developments in deep learning have led to neural network models capable of assessing pronunciation with near-human accuracy. These models utilize convolutional and recurrent neural networks (CNNs and RNNs) to analyze speech patterns and detect errors.

Advantages pronunciation assessment



AI models may struggle with regional accents and non-standard pronunciations. Background noise and poor microphone quality can affect assessment accuracy. AI tools may not fully grasp the nuances of speech in conversational settings. Data privacy and bias in AI training datasets remain concerns.



The assessment of students' pronunciation using AI technologies involves a multi-step process integrating speech recognition, machine learning, and linguistic analysis. Initially, automated speech recognition (ASR) systems transcribe spoken language into text, allowing for phonetic and phonological analysis. These systems compare the input speech with native speaker models, identifying deviations in pronunciation accuracy, fluency, and intonation. Machine learning algorithms, trained on extensive datasets of both native and non-native speech, evaluate pronunciation patterns, detect common errors, and refine assessment criteria over time. Deep learning techniques further enhance the precision of pronunciation evaluation by analyzing acoustic features and contextual speech patterns.

AI-driven pronunciation assessment relies on probabilistic scoring algorithms that quantify pronunciation accuracy at the phoneme, word, and sentence levels. Real-time feedback mechanisms provide learners with phonetic transcriptions, waveform analysis, and articulatory visualizations, facilitating targeted pronunciation improvements. Speech synthesis models generate native-like pronunciations, enabling students to compare and adjust their speech accordingly. Additionally, conversational AI and chatbot-based systems simulate real-life interactions, allowing learners to practice pronunciation within meaningful communicative contexts.

The integration of AI technologies in pronunciation assessment presents several challenges. Variability in accents and dialects can impact the accuracy of AI-driven assessments, as speech recognition models may struggle to accommodate diverse phonetic variations. Environmental factors, such as background noise and microphone quality, may introduce recognition errors, affecting the reliability of evaluation. Furthermore, AI lacks the ability to fully interpret the contextual nuances of speech, limiting its ability to assess pronunciation in dynamic conversational settings. Ethical considerations, particularly those related to data privacy and potential biases in AI training datasets, necessitate ongoing refinement of these technologies to ensure fairness and inclusivity in language assessment.

Despite these challenges, AI technologies offer significant advantages in pronunciation training by providing immediate, personalized feedback and increasing opportunities for autonomous learning. The continued development of AI-driven pronunciation assessment tools contributes to the advancement of language education, promoting more effective and accessible pronunciation training for learners worldwide.

In Uzbekistan, the assessment of students' pronunciation through AI technologies is an emerging field, with several notable applications and developments:

1. Intelligent systems for Uzbek language teaching (iieta.org):

Researchers have developed intelligent systems specifically designed for teaching the Uzbek language. These systems utilize phonemic recognition technology to analyze learners' speech, identifying pronunciation errors by comparing spoken input to native speaker models. The system's accuracy ranges from 67% to 95%, highlighting its potential in language instruction.

2. Talkio AI (talkio.ai):

Talkio AI is an innovative web application that simplifies the process of learning Uzbek. It employs state-of-the-art voice recognition and text-to-speech technologies, enabling learners to



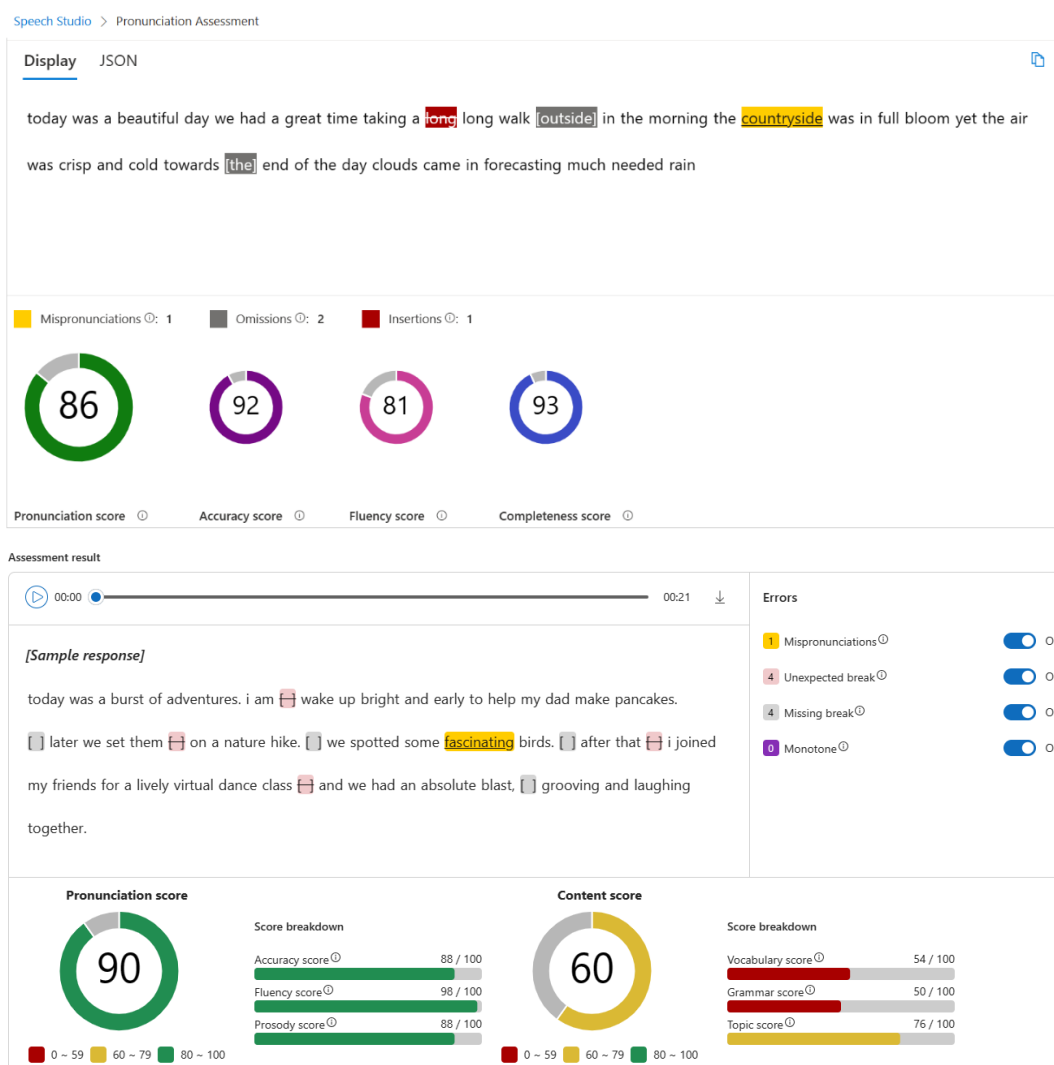
engage in realistic conversations akin to speaking with a native speaker. The app provides instant feedback on pronunciation and sentence structure, facilitating immersive learning experiences.

3. Development of language models for continuous Uzbek Speech recognition (pmc.ncbi.nlm.nih.gov):

Efforts have been made to develop language models tailored for continuous Uzbek speech recognition. These models, such as the Uzbek language model (UzLM), are trained on extensive datasets comprising millions of words and sentences. They enhance automatic speech recognition systems by providing probabilistic predictions for word sequences, thereby improving pronunciation assessment accuracy.

4. Uzbek speech recognition corpus (dataoceanai.com):

A comprehensive Uzbek Speech Recognition Corpus has been compiled, featuring recordings from 200 speakers in quiet environments. This dataset serves as a foundational resource for training AI models in speech recognition and pronunciation assessment, contributing to the development of more accurate and effective language learning tools.



The integration of AI technologies in Uzbekistan's language education sector is still in its nascent stages. The development of intelligent systems and language models tailored to the Uzbek language demonstrates a commitment to leveraging AI for educational purposes. However, challenges such as limited data resources, accent variability, and the need for comprehensive linguistic datasets persist. Continued investment in research and the creation of extensive speech corpora are essential to enhance the effectiveness of AI-driven pronunciation assessment tools in Uzbekistan.

Future research should focus on improving AI-based pronunciation assessment by enhancing model adaptability to diverse accents, improving noise robustness, and integrating AI with augmented reality (AR) and virtual reality (VR) for immersive language learning experiences. AI technologies offer a powerful solution for assessing students' pronunciation more accurately and efficiently than traditional methods. While challenges remain, ongoing advancements in AI and deep learning hold the potential to revolutionize pronunciation training and assessment in language education.

References:

1. Abduvalievna RF. Use of educational media as a factor in the integration of foreign language teaching methods. Asian Journal of Multidimensional Research. 2021;10(10):1388-92.
2. Chao FA, Lo TH, Wu TI, Sung YT, Chen B. An effective multi-view, multi-granularity, and multi-aspect modeling approach to English pronunciation assessment [Internet]. 2022 [cited 2025 Feb 11]. Available from: <https://arxiv.org/abs/2208.09110>
3. El Kheir Y, Ali A, Chowdhury SA. Automatic pronunciation assessment – A review [Internet]. 2023 [cited 2025 Feb 11]. Available from: <https://arxiv.org/abs/2310.13974>
4. Vancova, Hana. (2024). AI and AI-powered tools for pronunciation training. Journal of Language and Cultural Education. 11. 12-24. 10.2478/jolace-2023-0022.
5. Хасанова Д. Ш. РОЛЬ АУТЕНТИЧНЫХ МАТЕРИАЛОВ В ОБУЧЕНИИ ЯЗЫКАМ //Proceedings of International Educators Conference. – 2024. – Т. 3. – №. 5. – С. 266-270.
6. Шерматова Б., Ахмеджанова Н., Хасанова Д. Bo'lajak tarjimonlarni tayyorlashga yo'naltirilgan professional tarjima //Зарубежная лингвистика и лингводидактика. – 2024. – Т. 2. – №. 1/5. – С. 96-100.
7. Шухратовна Х. Д. Профессионально-Ориентированное Обучение Английскому Языку Будущих Специалистов В Вузах //Miasto Przyszłości. – 2024. – Т. 44. – С. 339-343.
8. <https://www.iieta.org/journals/ria/paper/10.18280/ria.370617> Creation of an intelligent system for Uzbek language teaching [Internet]. 2023 [cited 2025 Feb 11].
9. <https://issai.nu.edu.kz/uzbek-asr/> Uzbek speech corpus and automatic speech recognition [Internet]. 2023 [cited 2025 Feb 11].

