



A CONCEPTUAL REVIEW OF AI-DRIVEN IPA INSTRUCTION FOR UZBEK EFL LEARNERS

Zulfira Mamatova Fayziqulovna

Teacher, Uzbekistan State World Languages University (UzSWLU)
zulfiramamatova@uzswlu.uz

Murodova Ruxsora

Teacher, Uzbekistan State World Languages University (UzSWLU)
ruxsoramurodova@uzswlu.uz

Mavlonov Elnur

Student, New Uzbekistan University (NewUU)
250077@newuu.uz

Article history:	Abstract:
Received: 30 th October 2025 Accepted: 28 th November 2025	<p>In the contemporary landscape of English as a Foreign Language (EFL) instruction in Uzbekistan, pronunciation remains a significant barrier to communicative competence. Uzbek learners face a profound "phonological gap" due to structural disparities between the Turkic morphology of Uzbek and the Germanic roots of English, resulting in persistent interference errors. This conceptual review explores the integration of Artificial Intelligence (AI) and Computer-Assisted Pronunciation Training (CAPT) to bridge this gap. By leveraging Automatic Speech Recognition (ASR) and visual feedback mechanisms, AI tools offer the individualized drilling necessary to master International Phonetic Alphabet (IPA) features. Furthermore, this paper proposes a pedagogical shift toward a "Flipped Classroom" model, integrating Mamatova's (2023) Project-Based Learning (PBL) framework. This hybrid approach utilizes AI for home-based technical accuracy and reserves classroom time for meaningful, communicative interaction, ultimately reducing the affective filter and fostering learner autonomy.</p>

Keywords: AI in Education, EFL, Uzbek Learners, IPA, Pronunciation, Flipped Classroom, Project-Based Learning.

1 INTRODUCTION

In the contemporary landscape of English as a Foreign Language (EFL) instruction, pronunciation remains a fundamental pillar of communicative competence, directly influencing a learner's intelligibility and self-confidence (Abdelhalim & Alosehibany, 2025). Intelligibility serves as the primary goal for modern phonetic training, as mispronunciations can lead to significant communication breakdowns (Amrate & Tsai, 2025). For example, the phonemic distinction between minimal pairs such as "soup" /su:p/ and "soap" /s@Up/ is critical; an inaccurate production can confuse a listener and frustrate the speaker.

In the context of Uzbekistan, English has gained unprecedented popularity due to its essential role in career development and international communication. However, Uzbek students face a profound "phonological gap" because the Uzbek language—a member of the Turkic family—possesses structural, phonetic, and phonological features that contrast sharply with the 1 Germanic roots of English (Makhmatkulov et al., 2021). These linguistic disparities often manifest as persistent

difficulties in sound production, fluency, and prosodic accuracy.

Traditional classroom methodologies in Uzbekistan have historically relied on teacher-centered, book-oriented, and grammar-translation techniques, which frequently marginalize structured pronunciation instruction (Arshad & Rifky, 2023). Furthermore, the lack of consistent, individualized feedback in overcrowded public school settings limits student progress, as teachers cannot devote sufficient time to the phonetic needs of every learner (Baxramova, 2025). Consequently, traditional classroom drilling is often insufficient to bridge the phonological gap. Artificial Intelligence (AI) tools offer a dynamic solution by providing immediate, individualized feedback and adaptive learning environments that complement human instruction (Egamova, 2025).

2 The Phonological Challenge: Uzbek vs. English

The primary hurdle for Uzbek learners is the phenomenon of L1 interference, or negative transfer, where speakers apply the phonetic patterns of their native tongue to the target language. Because



English and Uzbek are unrelated languages, the number of phonetic interferences is extensive.

2.1 Consonant Interference

Specific "negative transfer" errors are prevalent in the production of English consonants that do not exist in the Uzbek phonemic inventory. A significant stumbling block is the substitution of the interdental fricatives /θ/ and /ð/ (Narmanova, 2025). Because these sounds are absent in Uzbek, students frequently approximate them using the closest available native sounds.

- The voiceless interdental fricative /θ/ is often replaced by the alveolar fricative /s/ or the voiceless stop /t/, causing "think" to be pronounced as /sINk/ or /tINk/ (Temirova, 2021).

- The voiced interdental fricative /ð/ is commonly substituted with /z/ or /d/, leading to "this" sounding like /zIs/ or /dIs/ (Pardayeva, 2025).²

Furthermore, confusion between the bilabial approximant /w/ and the labiodental fricative /v/ is a common issue. Since Uzbek lacks a clear distinction between these two phonemes, learners often replace /w/ with /v/, resulting in the confusion of minimal pairs such as "west" and "vest" (Narmanova, 2025). Additionally, the /ʒ/ sound found in words like "measure" is frequently substituted with the voiceless postalveolar fricative /s/ because the voiced variant does not exist in the Uzbek mother tongue (Jo'rayev & Ergashev, 2025).

2.2 Vowel Quality and Diphthong Challenges

While consonant errors are noticeable, vowel errors often impact intelligibility more severely.

The Uzbek vowel system is relatively small (consisting of 6 primary vowels) compared to the complex English system (12 monophthongs and 8 diphthongs).

- Tense vs. Lax Vowels: Uzbek learners struggle to distinguish between tense vowels (long) and lax vowels (short). Common errors include conflating /i:/ (sheep) and /ɪ/ (ship), or /u:/ (pool) and /ʊ/ (pull).

- The Schwa /ə/: The concept of vowel reduction in unstressed syllables is foreign to Uzbek speakers. Learners tend to pronounce unstressed vowels with their full quality, disrupting the rhythm of English speech.

2.3 Suprasegmental Interference and Syllable Structure

Beyond individual phonemes, Uzbek learners struggle with the rhythmic flow of English. Uzbek is categorized as a "syllable-timed" language, where each syllable is given roughly equal duration (Baxramova, 2025). In contrast, English is "stress-timed," meaning that stressed syllables occur at regular intervals while unstressed syllables are reduced.

This mismatch results in a "robotic" or flat intonation that lacks the natural rises and falls of pitch necessary for conveying emotional nuance. Additionally, Uzbek syllable structure rarely permits complex consonant clusters (e.g., "strengths"). To cope, learners often insert an intrusive vowel (epenthesis), pronouncing "sport" as "is-port" or "desk" as "des-kuh," which disrupts the flow of speech.³

3 The AI Solution: Mechanisms of Action

AI-powered technologies provide a personalized and scalable means to address these persistent phonological challenges through high-precision feedback systems.

3.1 Automatic Speech Recognition (ASR)

CAPT tools, such as ELSA Speak, SpeechAce, and Talkio AI, utilize Automatic Speech Recognition (ASR) to analyze spoken input and compare it against native speaker models (Abdelhalim & Alsehibany, 2025). These systems provide unbiased, data-driven assessments of phoneme accuracy, word stress, and fluency. For example, ELSA Speak can detect pronunciation errors with an accuracy of more than 95%, offering specific advice on how to correct articulation (Indari, 2023). This immediate corrective feedback is crucial for learners, as it reinforces correct patterns while the attempt is still fresh in their minds.

3.2 Visualization: Spectrograms and Waveforms

AI systems further assist learners by providing visual feedback in the form of spectrograms, waveforms, and articulatory visualizations (Rogerson-Revell, 2021). These tools help students "see" the acoustic features of their speech, such as pitch, intensity, and duration. For Uzbek learners, this is particularly beneficial for distinguishing sounds they cannot easily "hear," such as the subtle difference between the /s/ sound and the /θ/ sound. Visual animations of the oral cavity can demonstrate exactly how to position the tongue and lips for unfamiliar consonants, providing a multi-sensory learning experience (Pardayeva, 2025).

3.3 Anxiety Reduction and the Affective Filter

One of the most significant pedagogical advantages of AI is its ability to create a low-stress, non-judgmental practice environment. Speaking a foreign language often triggers anxiety, particularly when learners fear being judged by peers or teachers for their accent (Egamova, 2025).

AI-powered chatbots and virtual tutors lower this "affective filter," allowing students to rehearse,⁴ make mistakes, and try again privately without embarrassment. This increased sense of safety fosters a higher willingness to communicate (Safi et al., 2025).



4 Implementation: The Flipped Classroom and AI-Integrated PBL

The integration of AI within the Uzbek EFL context necessitates a transition toward the flipped classroom model. In this model, the standard pedagogical sequence is reversed: students engage with theoretical content and repetitive drilling independently at home using AI tools, while classroom time is reserved for collaborative, complex, and communicative tasks.

4.1 Enhancing Learner Autonomy through AI

This model is particularly effective for Uzbek learners, as it offloads the mechanical "listen-and-repeat" drilling required to master IPA phonemes to automated systems. By utilizing ASR at home, students can receive individualized, real-time feedback on their pronunciation accuracy.

This allows them to correct errors in a "safe" environment before entering the social arena of the classroom.

4.2 Integrating Project-Based Learning (PBL)

Central to this pedagogical shift is the enhancement of learner autonomy, a concept extensively explored in the research of Mamatova (2023). Mamatova argues that Project-Based Learning (PBL) acts as a catalyst for developing independent linguistic competence (Kasimova & Mamatova, 2023).

In an AI-driven flipped classroom:

1. Pre-Class (Home): Students use AI tools to proceduralize their knowledge of difficult fricatives (e.g., /T/ and /D/). They track their scores and progress, fostering self-regulation.
2. In-Class (School): The teacher implements Mamatova's PBL framework. Students integrate reading, writing, listening, and speaking into a single "end product," such as a video presentation, a debate, or a poster defense.

When students are personally involved in the creation of projects, they view English not as a "chore" but as a meaningful vehicle for communication. The phonological confidence gained from home-based AI practice directly supports their performance in these collaborative projects, reducing the fear of public speaking.

4.3 The Shifting Role of the Educator

The implementation of this model fundamentally alters the role of the educator in the Uzbek secondary and higher education sectors. The teacher shifts from being a mere "pronunciation checker" to acting as a "facilitator" of meaningful interaction. As AI handles the labor-intensive tasks of error detection, the educator is free to provide high-level guidance, cultural insights, and motivational support. This synergy creates an "optimal

learning environment" that merges technological precision with essential human connection.

5 CONCLUSION

The integration of AI-driven IPA instruction holds significant potential for modernizing English education in Uzbekistan. By addressing the structural phonological gap through real-time feedback, ASR technology, and speech visualization, AI tools empower Uzbek learners to overcome L1 interference and move beyond a "robotic" speaking style. While challenges such as infrastructure gaps and the need for digital literacy persist, the potential for personalized, autonomous, and low-anxiety learning makes AI an indispensable pedagogical ally. Through a strategic implementation of the flipped classroom model—supported by Mamatova's communicative PBL framework—Uzbekistan can foster a new generation of confident, accurate, and independent English speakers who are ready to compete on the global stage.

REFERENCES

1. 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*.
2. Abdurazakova, U. O. (2025). Inclusive project-based learning with differentiated instruction and AI-supported digital technologies for 9th grade English language learners in Uzbekistan. *Journal of Applied Science and Social Science*, 15(9), 48–54.
3. Ahmad, Y. B. (2018). Teaching English pronunciation of suprasegmental features on students of English education. *SHS Web of Conferences*, 42, 00048.
4. Amrate, M., & Tsai, P. (2025). Computer-assisted pronunciation training: A systematic review. *ReCALL*, 37(1), 22–42.
5. Arshad, S. M. B. M., & Rifky, M. (2023). Pronunciation difficulties of Uzbek speaking EFL learners: A study based on IELTS candidates in Uzbekistan. *ResearchGate*.
6. Baxramova, M. M. (2025). The role of speech recognition AI in enhancing listening and pronunciation for Uzbek learners. *Science Bulletin*, 8(2), 197–199.
7. Egamova, M. (2025). Artificial intelligence technologies as a tool for advancing English speaking skills in Uzbek educational settings. *Lingvospectrum*, 12(2), 121–124.



9. Indari, A. (2023). Detection of pronunciation errors in English speaking skills based on Artificial Intelligence (AI). *STKIP Budidaya Binjai*, 15(2), 67–77.
10. Jo'rayev, I. B., & Ergashev, R. S. (2025). Comparing English pronunciation challenges between Uzbek students and teachers: A case study from Uzbekistan. *Universal Journal of Academic and Multidisciplinary Research*, 3(29), 90-105.
11. Kasimova, Z. H., & Mamatova, Z. F. (2023). Project-based learning as an effective teaching approach to promote learner autonomy. *Web of Semantic: Universal Journal on Innovative Education*, 2(3), 255–258.7
12. Education, 2(3), 255–258.7
13. Lao-un, J., & Khampusaen, D. (2025). Developing an AI-powered pronunciation application to improve English pronunciation of Thai ESP learners. *Languages*, 10(11), 273.
14. Makhmatkulov, K., Kushnazarova, Y., Makhmudov, K., & Abdumalikova, S. (2021). Interference in English and Uzbek languages and its elimination. *Academic Research in Educational Sciences*, 2(3), 848–854.
15. Sciences, 2(3), 848–854.
16. Narmanova, S. (2025). Teaching English pronunciation to Uzbek speakers: Challenges and solutions. *Modern Education and Development*, 3(2), 824–845.
17. Olimova, S. K. (2025). AI technologies for assessing students' pronunciation. *Web of Journals*, 3(2), 163–167.
18. Pardayeva, S. A. (2025). Challenges of teaching English pronunciation to Uzbek learners: A case study. *Innovative Academy*.
19. Rogerson-Revell, P. M. (2021). Computer-Assisted Pronunciation Training (CAPT): Current issues and future directions. *RELC Journal*, 52(1), 1–17.
20. Safi, M. S., Ullah, H., & Khan, R. (2025). Integrating TBLT with digital tools: Impacts on learners' oral performance. *Journal of English Pedagogy and Applied Linguistics*.
21. Temirova, H. B. (2025). Phonetic interference in comparative and contrastive phonetics: An analysis of pronunciation difficulties in Uzbek and English language learners. *Zenodo*.
22. Temirova, H. (2021). The problems of English pronunciation of Uzbek students. *Economy and Society*.
23. Uzakova, G. R. (2024). Improving pronunciation in primary English learners through AI-based language apps. *Central Asian Journal of Academic Research*, 5(11), 117–120.8