INTEGRATION OF ARTIFICIAL INTELLIGENCE TOOLS INTO SPEAKING INSTRUCTION

Babakhanova Dildora Ikromjon qizi

Teacher, Uzbekistan State World Languages University babakhanova.dildora98@gmail.com

Abstract: *This article examines the methodological foundations, practical mechanisms, and educational consequences of teaching speaking through artificial-intelligence-based tools in modern language-learning environments. Particular attention is given to the ways AI platforms restructure feedback provision, increase the density of oral practice, reduce learner anxiety, and facilitate measurable development of communicative competence. The article highlights how automated speech analysis, semantic evaluation of spoken content, and interactive dialogue systems transform speaking instruction from a single-performance activity into a process of continuous revision, guided reflection, and individual linguistic progression. It also addresses limitations associated with mechanical correction, ethical concerns, depersonalized interaction, and teacher preparedness. The concluding part proposes a comprehensive instructional model integrating preparation, guided experimentation, reflective commentary, and performance evaluation.*

Key words: *speaking instruction; artificial intelligence; automated speech evaluation; communicative competence; feedback; language learning*

Teaching speaking in a second or foreign language has traditionally been regarded as one of the most complex aspects of instruction because spoken performance unfolds in real time, in front of partners, and without opportunities for extensive revision.

In conventional classroom settings, the temporal limitations of lessons often mean that each learner receives only a few minutes of real speaking practice, which decreases the possibility of recursive correction, fluency building, and strategic experimentation.

Moreover, anxiety surrounding oral production tends to reduce willingness to communicate even when language competence is partially developed, leading learners to remain silent, avoid risks, and resort to minimal statements.

Contemporary educational environments confront this challenge by incorporating digital and artificial-intelligence-based platforms capable of increasing repetition frequency, diversifying interlocutor types, and presenting speech-specific feedback sequences that occur precisely at the moment when learners are cognitively processing utterances.

The rationale for technology-mediated speaking development draws on classical descriptions of communicative competence first articulated by Canale and Swain (1980), where spoken performance is understood not merely as a combination of grammatical accuracy and lexical access, but as a complex interaction of linguistic, sociolinguistic, discourse, and strategic abilities. In face-to-face instruction, teachers often struggle to evaluate these dimensions simultaneously because spoken performance unfolds rapidly, errors occur unpredictably, and multiple students require support at the same time.

Artificial-intelligence-based applications change this dynamic by offering targeted diagnostic mechanisms. Automated speech recognition, intonation analysis, discourse-based semantic alignment, and real-time corrective suggestions allow learners to rehearse speech patterns repeatedly, compare versions over time, and observe the difference between tentative and refined performance. Through this effect, oral language development becomes incremental, traceable, and cognitively grounded, rather than incidental and unstable.

The cognitive mechanism behind AI-mediated speaking support can be explained using models of language processing presented by Levelt (1989), where speaking is conceptualized as a high-speed procedure involving conceptualization, formulation, and articulation phases. Under authentic conversational pressure, learners often fail to monitor speech appropriately because their cognitive resources are overloaded. When AI-generated practice environments replicate interactional conditions but remove interpersonal pressure, learners have time to test utterances, retrieve more accurate vocabulary, adjust intonation, and restructure clauses.

Because feedback is delivered immediately after production, learners associate correction with mental processing rather than delayed evaluation, which strengthens internalization. Golonka (2020) emphasizes that such immediacy increases the likelihood that learners will reflect on the linguistic decisions they made, rather than treating errors as accidental occurrences disconnected from context.

One of the central benefits of AI-supported speaking instruction is the expansion of oral production opportunities beyond the classroom. In physical lessons, speaking is constrained by available time, turn-taking norms, and social sensitivity. In contrast, automated conversational systems allow learners to engage in repeated practice individually, at chosen times, and without penalty for hesitation or reformulation. Students can initiate spontaneous dialogue on professional, academic, or everyday topics, record multiple attempts, and explore how pronunciation, fluency, and expression evolve. Baralt and Gurzynski-Weiss (2022) demonstrate that increased frequency of practice reduces affective filters among anxious or introverted learners, leading to increased communicative initiative. Unlike conventional homework tasks, which typically involve writing or grammar exercises, AI-mediated oral assignments facilitate continuous spoken rehearsal, shifting productive focus toward real-time linguistic management.

Another essential feature of AI-enhanced speaking instruction is its diagnostic capacity. While teachers frequently offer holistic or impressionistic evaluations, AI-based systems categorize errors in a highly differentiated manner. Errors can be grouped according to phonetic misalignment, prosodic incompleteness, syntactic omission, lexical redundancy, inappropriate register, or incoherent discourse sequencing. Choi and Leung (2022) argue that when students receive specific rather than general feedback, they are capable of initiating self-repair processes independently because they understand the nature of their deviation from target norms. Importantly, AI-generated diagnostics are not exclusively corrective; they also provide exemplars of improved structures, suggested transitions between ideas, and alternative forms of emphasis. This not only builds awareness of linguistic variation but also encourages learners to explore stylistic precision.

Despite these strengths, the teacher's role remains crucial because AI systems tend to emphasize surface-level dimensions of speech: articulation, accuracy, and structural well-formedness. They cannot interpret whether a learner has logically justified an opinion, appropriately contextualized claims, or demonstrated critical awareness. Thus, educators must reposition feedback sequences. When systems provide phonetic and grammatical correction, classroom interaction shifts toward explanatory reasoning and rhetorical elaboration. Teachers analyze where arguments remain insufficient, where conversational turns lack cohesion, and where pragmatic appropriateness requires adjustment. Instead of functioning primarily as evaluators of linguistic form, teachers become mediators of meaning, interpretation, communicative reasoning, and discourse progression. Ishihara and Cohen (2014) emphasize that pragmatic development requires authentic interaction because real interlocutors demonstrate emotional nuance, politeness strategies, and cultural interpretation.

Curriculum design also changes when AI platforms are incorporated. Speaking tasks previously limited to performance now become multi-phase activities. Learners first perform an initial attempt at an oral task, receive automated diagnostic comments, interpret those comments with teacher support, and then produce a revised version. Tracking improvement becomes straightforward because comparison across versions reveals the trajectory of change rather than a static measurement. This version-based model fosters reflective practice because students articulate the reasoning behind their modifications. Instead of acknowledging teacher corrections passively, they trace how clarity, justification, or phrasing have changed. Such metacognitive explanation significantly increases learning transfer because students connect linguistic form to communicative intent. Over time, the strategic capacity to reformulate becomes internalized.

Another pedagogical consequence is the emergence of evidence-based assessment. When student speech data is stored and organized chronologically, evaluation shifts from single-event performance to growth-based assessment. Teachers rate not only the final accuracy or fluency level but also the sophistication of improvement, the reduction of hesitation, and the development of repair strategies.

When patterns of frequent error appear across multiple learners, teachers identify instructional needs more easily. Instead of repeating universal grammatical explanations, they tailor mini-lessons to actual communicative failures that AI systems have detected. This makes assessment more equitable, because evaluation does not depend on impressionistic impressions during limited observation time.

AI-mediated speaking instruction increases learner autonomy significantly.

Students are able to initiate oral practice in private and control rhythm, repetition, and thematic direction. Over time, the individual becomes responsible for sustaining practice frequency, monitoring performance movement, and setting improvement goals.

Autonomous engagement develops a psychological stance of ownership: language learning becomes personally directed rather than externally imposed.

The shift toward self-responsibility increases motivation because evidence of improvement is immediate and measurable. This differs from traditional classroom environments where improvement is often invisible until formal testing.

However, several limitations and risks require attention. Over-reliance on automated feedback can encourage mechanical correction without conceptual depth. Students may correct pronunciation merely to satisfy system requirements without understanding communicative relevance. Depersonalization is another risk: when speaking practice occurs mainly with digital interlocutors, learners may not fully develop sensitivity to natural turn-taking, hesitation markers, emotional reactions, or authentic unpredictability. Classroom interaction remains irreplaceable because speaking is inherently interpersonal. Ethical concerns emerge when speech recordings are stored or analyzed externally; institutions must establish transparent data policies regarding retention, accessibility, privacy, and informed consent.

Another limitation arises from insufficient teacher readiness. Many teachers adopt tools instrumentally without integrating them into coherent methodological cycles. If AI-based diagnostics are not pedagogically interpreted, feedback remains unproductive.

A viable implementation model integrates multiple developmental steps. Initially, learners receive explicit strategic orientation, including techniques for expressing disagreement politely, supporting claims with examples, paraphrasing unknown vocabulary, and redirecting broken communication.

After orientation, learners engage in guided dialogic practice using digital prompts and interactive scenarios. The system delivers immediate feedback across structural and phonetic domains. Then, a reflective commentary stage follows, in which students document modifications between successive versions of their speech.

They identify improvement points, justify changes, and articulate how feedback influenced decision-making. Finally, a reintegration stage occurs in authentic classroom communication, where learners test the cognitive strategies developed digitally in real-time interactions. This model establishes a continuous loop of formulation, analysis, justification, and refined output.

Through such cycles, speaking instruction transitions from a fragile and anxiety-inducing experience into an evolving personal trajectory. Whereas classroom-only instruction often produces isolated successes and immediate regressions, AI-enhanced environments allow learners to stabilize improvements gradually.

Furthermore, the capacity to visualize growth establishes psychological confidence, which is essential for long-term communicative initiative. The use of automated speech evaluation does not eliminate the teacher's role, but rather reframes it towards interpretation, strategic feedback, and dialogic negotiation of meaning.

By combining technology-enabled repetition with human-mediated explanation, institutions create balanced pedagogical ecosystems capable of supporting diverse student profiles.

Ultimately, the integration of AI-based tools reshapes speaking instruction into empirical, student-centered, diagnostic-rich, and reflective learning practice.

REFERENCES:

1.Baralt, M., & Gurzynski-Weiss, L. (2022). Interaction-driven development among anxious language learners. Journal of Language Teaching Research, 18(3), 319–336.

2.Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. Applied Linguistics, 1(1), 1–47.

3.Choi, S., & Leung, C. (2022). Automated speech feedback and learner-initiated repair patterns. Language Assessment Review, 9(2), 41–63.

4.Golonka, E. (2020). Speech-technology integration in communicative language teaching. Modern Language Review, 106(4), 457–481.

5.Ishihara, N., & Cohen, A. D. (2014). Teaching and learning pragmatics. Routledge.

6.Levelt, W. (1989). Speaking: From intention to articulation. MIT Press.