

A METHODOLOGICAL FRAMEWORK FOR ENHANCING EFL LEARNERS' SPEAKING AND WRITING COMPETENCE THROUGH ARTIFICIAL INTELLIGENCE TECHNOLOGIES

Kholboeva Durdona

Master's student, Webster University

Email address: xolboyevadurdona02@gmail.com

Khamzaev Jamshid

Associate Professor, Tashkent University of Information Technologies named after Muhammad al-Khwarizmi

Email address: hamzayevjf@gmail.com

Abstract: *Developing speaking and writing competence remains one of the most persistent challenges in English as a Foreign Language (EFL) instruction, since traditional classrooms rarely provide the individualized practice and immediate feedback these productive skills require. Building on the authors' earlier conference thesis, which surveyed AI-based tools such as chatbots, natural language processing (NLP) applications, and speech-recognition systems, this article develops that work into a structured methodological framework intended for classroom implementation. Using a qualitative literature-synthesis design, the study analyzed sixteen empirical and meta-analytic sources published between 2021 and 2026 to determine which categories of artificial intelligence tools are most consistently associated with measurable gains in speaking fluency, pronunciation accuracy, writing accuracy, writing organization, and learner motivation. The analysis consolidated the instructional sequences reported across these studies into a five-stage cycle covering AI-mediated input, interactive production, automated feedback, learner self-revision, and adaptive re-practice, designed to structure classroom use of AI tools rather than leaving such use incidental. The findings indicate that chatbot- and speech-recognition-based tools are most strongly associated with oral fluency and pronunciation gains, while natural-language-processing checkers and generative AI tools are most strongly associated with writing accuracy and organizational quality; generative AI shows the broadest reported influence on learner motivation and autonomy. The discussion situates these findings against the reviewed literature, notes implementation constraints such as unequal technological access and learner over-reliance on automated corrections, and proposes recommendations for lesson design, teacher preparation, and future empirical validation of the proposed cycle in university EFL classrooms in Uzbekistan.*

Keywords: *artificial intelligence, EFL methodology, speaking skills, writing skills, chatbots, natural language processing, speech recognition, generative AI, adaptive learning, instructional design.*

INTRODUCTION

The rapid diffusion of artificial intelligence (AI) technologies into education has reshaped expectations about what individualized language instruction can look like at scale. Across educational contexts worldwide, AI-mediated tools are increasingly used to personalize instruction, automate feedback, and extend practice opportunities

beyond the physical classroom [8], [9]. In English as a Foreign Language (EFL) settings specifically, speaking and writing remain the two productive skills most resistant to large-class, teacher-centred instruction, since both require repeated practice, timely correction, and sustained motivation that a single teacher cannot always provide to every learner [13], [8].

The authors' earlier conference thesis, presented at the Republican Scientific-Practical Conference held in Navoiy in March 2026, reviewed the general landscape of AI-based tools - chatbots, NLP applications, and speech-recognition systems and argued that these technologies could substantially improve fluency, accuracy, and motivation in EFL classes [7]. That thesis, however, remained at the level of a general survey: it did not specify which categories of tools serve which productive sub-skills, nor how such tools should be sequenced within a lesson or course. The present article addresses this scientific problem directly.

The aim of this study is to develop a literature-grounded methodological framework that specifies which AI tool categories most reliably support EFL speaking and writing, and how these tools can be organized into a coherent instructional cycle rather than used as isolated add-ons. In line with this aim, the study pursues three objectives: (1) to classify the AI tools discussed in recent literature (2021-2026) according to their reported effect on specific speaking and writing sub-skills; (2) to consolidate the recurring instructional sequences identified across this literature into a single, replicable classroom cycle; and (3) to identify the practical and ethical constraints that must be addressed for such a methodology to be implemented sustainably in ordinary teaching conditions, including those typical of higher-education classrooms in Uzbekistan.

The object of the research is the process of AI-supported development of EFL learners' speaking and writing competence; the subject is the set of AI tool categories (chatbots, speech-recognition/automatic speech evaluation systems, natural language processing writing checkers, generative AI, and adaptive learning platforms) and their reported pedagogical functions. The working hypothesis is that AI tools differ systematically in which productive sub-skill they support most effectively, and that grouping them into a skill-specific, sequenced instructional cycle yields a more coherent methodology than treating AI as a single, undifferentiated resource. The scientific novelty of the study lies in synthesizing dispersed empirical and meta-analytic findings into one consolidated tool-to-outcome mapping and a concrete five-stage instructional cycle that has not, to the authors' knowledge, been presented in this combined form in the reviewed literature.

LITERATURE REVIEW

Recent research consistently reports positive, though tool-specific, effects of AI on productive language skills, and can be organized thematically into studies of oral/speaking-focused tools, writing-focused tools, and cross-cutting meta-analyses.

Wei [13] conducted a controlled study with 60 university EFL students and found that AI-mediated instruction produced significantly higher achievement across grammar, vocabulary, reading, and writing than traditional instruction, alongside

greater L2 motivation and self-regulation; however, the study measured achievement as a composite construct and did not isolate the specific contribution of speaking practice from other skill domains. Zou et al. [16] examined automatic speech-evaluation software with 40 Chinese EFL learners and reported that the majority of participants perceived measurable gains in speaking skills as a direct result of the automated feedback received, although the study relied primarily on learner self-report rather than independently rated speech samples. Fathi, Rahimi and Derakhshan [3] extended this line of research by linking AI-mediated conversational interaction not only to speaking performance but to willingness to communicate, a motivational construct that earlier speaking-focused studies had largely overlooked.

On the writing side, Chang et al. [1] used a quasi-experimental design as early as 2021 to show that EFL students who received automated writing feedback from Grammarly outperformed a control group receiving no such feedback, establishing an early empirical basis for NLP-based writing tools. Huang, Hew and Fryer [5] subsequently conducted a systematic review of chatbot-supported language learning and concluded that conversational agents are effective for interactive practice but show more limited direct impact on writing accuracy compared with dedicated NLP tools – a distinction that later studies of generative AI complicate. Marzuki et al. [9] interviewed EFL teachers using AI writing tools, including generative AI systems, and found unanimous agreement that these tools improved the content organization and structural coherence of student writing, moving beyond the narrower grammar-correction focus of earlier NLP tools. Song and Song [11] and Teng [12] both examined ChatGPT specifically and reported gains in writing quality and motivation, with Teng [12] adding the important qualification that learners themselves prefer to treat generative AI feedback as a supplementary “companion” rather than a substitute for teacher feedback.

At the level of synthesis, Guan, Li and Gu [4] performed a meta-analysis of generative AI in informal, self-directed English learning and found significant positive effects on both proficiency and self-regulation across 15 studies. Wu et al. [14] meta-analysed AI-enhanced blended learning and confirmed a positive contribution to overall learning achievement beyond blended learning alone. Ren, Stephens and Lee [10] meta-analysed 23 studies on AI and learner self-efficacy and reported a significant positive effect (effect size = 0.758), while Zhang, Jantakoon and Laoha [15] confirmed comparable positive effects of AI tools across educational settings more broadly, though their sample was not limited to language education specifically. Liang et al. [8] mapped the research foci of AI in language education and noted that speaking- and writing-related applications constitute the two most heavily researched productive-skill domains, while Kamalov, Santandreu Calonge and Gurrib [6] and Chaudhry and Kazim [2] situate these developments within the broader, ethically complex transformation of education by AI, cautioning against uncritical adoption.

Taken together, the reviewed literature converges on two patterns that motivate the present study and that no single reviewed source addresses jointly. First, no single AI tool category addresses all productive sub-skills equally well: chatbots and speech-recognition systems are most strongly associated with oral fluency and pronunciation

gains, whereas NLP checkers and generative AI are most strongly associated with writing accuracy and organization [1], [5], [9], [16]. Second, motivation and learner autonomy rather than accuracy alone emerge repeatedly as outcomes that generative and adaptive AI tools influence particularly strongly [4], [10], [13]. This is the gap the present study addresses: translating a dispersed set of tool-specific findings into one consolidated mapping and a sequenced classroom methodology.

METHODOLOGY

This study employed a qualitative, literature-based synthesis design, extending the descriptive-analytical approach used in the authors’ original conference thesis [7]. This design was selected because the research questions concerned the classification and consolidation of dispersed findings across an already substantial and rapidly growing empirical literature, rather than the collection of new primary data; a synthesis design allowed the systematic comparison needed to answer the study's research questions within the available scope.

Sources were identified through Frontiers in Psychology, ScienceDirect, SAGE Open, Cogent Education, MDPI, and related academic databases, using the search terms “artificial intelligence”, “EFL” , “speaking skills”, “writing skills”, “chatbot”, “natural language processing” and “generative AI”. Sources were included if they were peer-reviewed, published between 2021 and 2026, and reported empirical findings or meta-analytic evidence on the effect of an AI-based tool on EFL/ESL speaking or writing outcomes. Sixteen sources meeting these criteria were retained for detailed analysis; seven of these, comprising the primary empirical and meta-analytic studies, are summarized in Table 1.

Table 1.

Summary of key studies underpinning the proposed framework

Study	AI tool / focus	Sample / scope	Key reported outcome
Wei (2023) [13]	AI-mediated instruction (mixed)	60 university EFL students	Significantly higher achievement, motivation, self-regulation vs. control group
Zou et al. (2023) [16]	Automatic speech-evaluation software	40 Chinese EFL learners	Majority reported improved speaking skills from automated feedback
Fathi et al. (2024) [3]	AI-mediated conversational interaction	EFL learners, experimental design	Improved speaking performance and willingness to communicate
Chang et al. (2021) [1]	Automated writing feedback (Grammarly)	EFL students, quasi-experimental	Experimental group outperformed control group in writing tasks
Marzuki et al. (2023) [9]	AI writing tools (Grammarly, generative AI)	4 EFL teachers, 3 universities	Unanimous report of improved content organization in student writing
Guan, Li & Gu (2024) [4]	Generative AI (meta-analysis)	15 studies	Significant positive effects on proficiency and self-

			regulation
Ren, Stephens & Lee (2026) [10]	AI and self-efficacy (meta-analysis)	23 studies	Significant positive effect on self-efficacy ($g = 0.758$)

Two analytical procedures were applied to the retained sources. First, a comparative content analysis classified each AI tool category discussed in the literature chatbots, speech-recognition/automatic speech evaluation (ASR) systems, NLP grammar and coherence checkers, generative AI, and adaptive learning platforms against six sub-skill outcomes: speaking fluency, pronunciation accuracy, writing accuracy, writing organization, motivation/engagement, and learner autonomy.

Each tool–outcome pairing was rated as Low, Moderate, or High according to the consistency and strength of the effects reported across the reviewed studies; this rating was assigned independently by the two authors and reconciled through discussion where initial ratings diverged. This procedure produced the synthesis presented in Figure 2, which should be read as a qualitative mapping of the literature rather than a pooled statistical estimate. Second, the recurring instructional sequences described across the reviewed empirical studies input exposure, guided production, automated feedback, learner revision, and repeated practice were extracted, compared across sources, and consolidated into the single five-stage instructional cycle presented in Figure 1.

ANALYSIS AND RESULTS

The comparative content analysis produced the tool-to-outcome mapping shown in Figure 2. Chatbots were rated High for speaking fluency and Low for pronunciation accuracy, writing accuracy, and writing organization; they were rated High for motivation/engagement and Moderate for learner autonomy. Speech-recognition/ASR feedback applications were rated High for both speaking fluency and pronunciation accuracy,

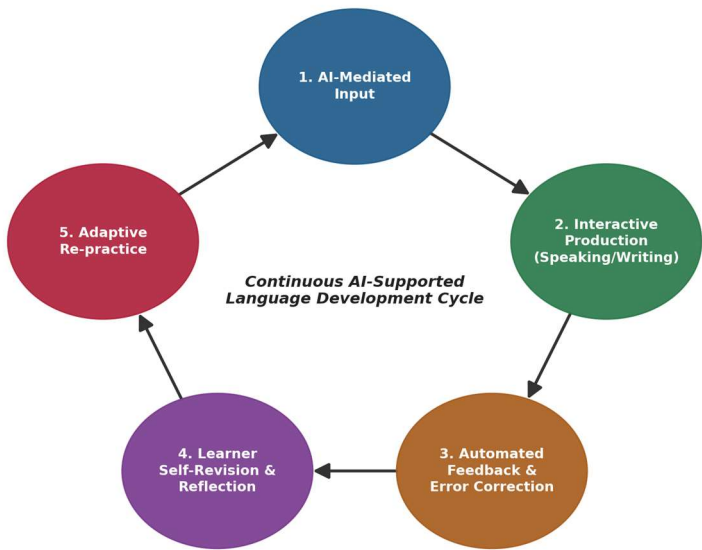


Figure 1. Proposed five-stage AI-integrated instructional cycle for EFL speaking and writing development.

Low for the two writing outcomes, and Moderate for motivation and autonomy. NLP grammar and coherence checkers were rated High for writing accuracy and writing organization, Low for the two speaking outcomes, and Moderate for motivation and autonomy. Generative AI tools were rated High for writing accuracy, writing organization, motivation/engagement, and learner autonomy, and Moderate for speaking fluency; this was the only tool category rated High across four of the six outcome measures simultaneously. Adaptive learning platforms were rated High for motivation/engagement and learner autonomy, and Moderate across the remaining four outcomes.

The second analytical procedure produced the five-stage cycle shown in Figure 1. Stage 1 (AI-mediated input) consists of learner exposure to AI-generated dialogues, readings, or prompts. Stage 2 (interactive production) consists of learners speaking with a chatbot or ASR application, or drafting a written text. Stage 3 (automated feedback) consists of the relevant tool returning itemized feedback: ASR software for speaking, NLP or generative AI checkers for writing. Stage 4 (learner self-revision and reflection) consists of learners reviewing this feedback under teacher guidance and revising their output. Stage 5 (adaptive re-practice) consists of follow-up tasks targeting the specific weaknesses identified, returning to Stage 1 at a recalibrated difficulty level. This five-stage sequence was consolidated directly from the instructional procedures reported in [1], [3], [9], [13], and [16].

Chatbots (e.g., dialogue agents)	High	Low	Low	Low	High	Moderate
Speech-recognition / ASR feedback apps	High	High	Low	Low	Moderate	Moderate
NLP grammar & coherence checkers	Low	Low	High	High	Moderate	Moderate
Generative AI (e.g., ChatGPT)	Moderate	Low	High	High	High	High
Adaptive learning platforms	Moderate	Low	Moderate	Moderate	High	High
	Speaking fluency	Pronunciation accuracy	Writing accuracy	Writing organization	Motivation / engagement	Learner autonomy

Figure 2. Literature-based synthesis of reported AI tool impact on EFL speaking and writing sub-skills (Low / Moderate / High).

Across the meta-analytic sources, motivation and self-regulation were reported at least as consistently as accuracy-related outcomes. Wei [13] reported greater L2 motivation and self-regulated learning strategies among AI-instructed learners relative to the control group. Guan, Li and Gu [4] reported significant effects on self-regulation among learners using generative AI in informal settings. Ren, Stephens and Lee [10] reported a positive effect of AI use on learner self-efficacy of medium-to-large

magnitude ($g = 0.758$) across 23 studies.

DISCUSSION

The results address the three objectives set out in the introduction. Regarding the first objective classifying AI tools by their reported sub-skill effects the mapping in Figure 2 confirms the study's working hypothesis that AI tools differ systematically in which productive sub-skill they support most effectively, rather than functioning as a single, undifferentiated resource. This finding aligns with, and extends, the observation of Huang, Hew and Fryer [5] that chatbots are strong for interactive oral practice but weaker for writing accuracy, and with Marzuki et al. [9], who found that dedicated writing tools rather than conversational agents drove improvements in content organization.

Regarding the second objective consolidating recurring instructional sequences the five-stage cycle proposed here formalizes a sequence that is implicit but not explicitly systematized in the individual studies reviewed. Where Wei [13] and Fathi et al. [3] each describe input–practice–feedback sequences within their own experimental designs, the present synthesis generalizes this pattern into a tool-agnostic cycle applicable across both speaking and writing instruction, addressing the criticism raised by Huang, Hew and Fryer [5] and Teng [12] that AI tools are frequently used as isolated activities rather than integrated into a sustained pedagogical loop.

Regarding the third objective - implementation constraints the results reinforce a caution raised across several of the reviewed studies. Automated feedback, however immediate, is not pedagogically neutral: Huang, Hew and Fryer [5] note that chatbot-based correction can be inconsistent, and Teng [12] reports that learners themselves prefer to treat generative AI feedback as a companion to, rather than a replacement for, teacher judgement. This corroborates the inclusion of Stage 4 (learner self-revision under teacher guidance) in the proposed cycle as a safeguard against the uncritical acceptance of AI-generated corrections, a risk that becomes more significant as generative AI tools are used more heavily for writing [9], [12].

A further constraint concerns equity of access. Effective use of speech-recognition and generative AI tools presupposes stable internet access and, in many cases, paid subscriptions, which are not uniformly available across regional universities. Kamalov, Santandreu Calonge and Gurrib [6] and Chaudhry and Kazim [2] both flag infrastructural and ethical constraints as central to responsible AI adoption in education; the proposed cycle would therefore need to be paired with institutional investment in access and with structured teacher training, since teachers' own AI literacy substantially determines how well Stage 3 and Stage 4 are facilitated in practice.

This study has several limitations. First, because it relies on a literature synthesis rather than original data collection, the Low/Moderate/High ratings in Figure 2 reflect the authors' interpretive classification of reported findings rather than a pooled statistical effect size, and different reviewers applying the same procedure could arrive at somewhat different ratings for borderline cases. Second, the sixteen sources reviewed, while spanning 2021–2026, are concentrated in East and Southeast Asian

and Middle Eastern EFL contexts, and their findings may not transfer directly to the Central Asian higher-education context without empirical verification. Third, the proposed five-stage cycle has not yet been tested empirically; its validation requires a classroom-based study of the kind conducted by Wei [13] and Fathi et al. [3], which is proposed as a direction for future research.

CONCLUSION

This study set out to develop a literature-grounded methodological framework specifying which AI tool categories most reliably support EFL speaking and writing, and how these tools can be sequenced into a coherent classroom methodology. The analysis shows that AI tools differ systematically in the productive sub-skills they best support, and that a structured five-stage cycle AI-mediated input, interactive production, automated feedback, learner reflection, and adaptive re-practice offers a practical way to sequence these tools rather than deploying them incidentally.

The contribution of this work is twofold: a consolidated tool-to-outcome mapping that can guide practitioners' selection of AI tools for specific lesson objectives, and a replicable five-stage instructional cycle that can be embedded into existing EFL curricula. For practitioners, tool selection should be matched deliberately to the target sub-skill, a dedicated reflection stage should be built into every AI-supported lesson to preserve the teacher's formative role, and institutions should invest concurrently in infrastructure and teacher training on AI literacy. Future research should test the proposed cycle empirically in university EFL classrooms in Uzbekistan, using a pre-and post-test design, to determine whether it yields the sub-skill-specific gains this synthesis predicts.

REFERENCES

1. Chang, T. S., Li, Y., Huang, H. W., & Whitfield, B. (2021). Exploring EFL students' writing performance and their acceptance of AI-based automated writing feedback. In *Proceedings of the 2nd International Conference on Education Development and Studies* (pp. 31-35).
2. Chaudhry, M. A., & Kazim, E. (2022). Artificial Intelligence in Education (AIEd): a high-level academic and industry note. *AI and Ethics*, 2, 157-165. <https://doi.org/10.1007/s43681-021-00074-z>
3. Fathi, J., Rahimi, M., & Derakhshan, A. (2024). Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System*, 121, 103254. <https://doi.org/10.1016/j.system.2024.103254>
4. Guan, L., Li, S., & Gu, M. M. (2024). AI in informal digital English learning: A meta-analysis of its effectiveness on proficiency, motivation, and self-regulation. *Computers and Education: Artificial Intelligence*, 7, 100323. <https://doi.org/10.1016/j.caeai.2024.100323>
5. Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning — Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, 38(1), 237-257. <https://doi.org/10.1111/jcal.12610>

6. Kamalov, F., Santandreu Calonge, D., & Gurrib, I. (2023). New era of artificial intelligence in education: towards a sustainable multifaceted revolution. *Sustainability*, 15(16), 12451. <https://doi.org/10.3390/su151612451>
7. Kholboeva, D., & Khamzaev, J. (2026). Artificial intelligence based approaches to develop speaking and writing skills in English classes. *Materials of the Republican Scientific-Practical Conference "Experience in Implementing Artificial Intelligence Technologies in Science and Education: Problems and Prospects"*, Navoiy State University, Uzbekistan, pp.129-132.
8. Liang, J.-C., Hwang, G.-J., Chen, M.-R. A., & Darmawansah, D. (2021). Roles and research foci of artificial intelligence in language education: an integrated bibliographic analysis and systematic review approach. *Interactive Learning Environments*. <https://doi.org/10.1080/10494820.2021.1958348>
9. Marzuki, Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education*, 10(2), 2236469. <https://doi.org/10.1080/2331186X.2023.2236469>
10. Ren, L., Stephens, J. M., & Lee, K. (2026). The impact of AI on learners' self-efficacy: A meta-analysis. *Behavioral Sciences*, 16(1), 158. <https://doi.org/10.3390/bs16010158>
11. Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, 14, 1260843. <https://doi.org/10.3389/fpsyg.2023.1260843>
12. Teng, M. F. (2024). "ChatGPT is the companion, not enemies": EFL learners' perceptions and experiences in using ChatGPT for feedback in writing. *Computers and Education: Artificial Intelligence*, 7, 100270. <https://doi.org/10.1016/j.caeai.2024.100270>
13. Wei, L. (2023). Artificial intelligence in language instruction: impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology*, 14, 1261955. <https://doi.org/10.3389/fpsyg.2023.1261955>
14. Wu, J., Tlili, A., Salha, S., Mizza, D., Saqr, M., López-Pernas, S., & Huang, R. (2025). Unlocking the potential of artificial intelligence in improving learning achievement in blended learning: a meta-analysis. *Frontiers in Psychology*, 16, 1691414. <https://doi.org/10.3389/fpsyg.2025.1691414>
15. Zhang, J., Jantakoon, T., & Laoha, R. (2025). Meta-analysis of artificial intelligence in education. *Higher Education Studies*, 15(2), 189–199. <https://doi.org/10.5539/hes.v15n2p189>
16. Zou, B., Du, Y., Wang, Z., Chen, J., & Zhang, W. (2023). An investigation into artificial intelligence speech evaluation programs with automatic feedback for developing EFL learners' speaking skills. *SAGE Open*, 13(3). <https://doi.org/10.1177/21582440231193818>