QO‘QON DAVLAT PEDAGOGIKA INSTITUTI ILMIY XABARLARI
(2026-yil 10-son)
PEDAGOGIKA
PEDAGOGY
https://doi.org/10.70728/c.series.ped.v08.i10.014

AI-ASSISTED FEEDBACK IN TEACHING ENGLISH WRITING: A REVIEW OF RECENT EMPIRICAL STUDIES

Amonova Adiba Shukrullo kizi
Navoi State University
E-mail address: adibaamonova007@gmail.com

Abstract. This article reviews recent empirical research on the feedback that artificial intelligence generates on students' English writing. Eleven peer-reviewed, open-access sources published between 2020 and 2025 were selected and read in full: two meta-analyses, one systematic review and eight empirical studies, two of them conducted in Uzbekistan. The analysis was thematic and organised around four questions: how automated and AI feedback affects writing performance, how the quality of AI feedback compares with that of teacher feedback, how learners engage with and perceive AI feedback, and what role the teacher plays. Compared with control conditions, automated feedback has a medium overall effect on writing performance, but AI feedback shows no clear advantage over feedback from trained teachers or tutors.

Key words: AI-assisted feedback, automated writing evaluation, ChatGPT, written corrective feedback, EFL writing, learner engagement, teacher feedback.

Аннотация.В статье рассматриваются современные эмпирические исследования обратной связи, которую искусственный интеллект (ИИ) даёт на письменные работы студентов на английском языке. Для обзора отобраны одиннадцать рецензируемых источников открытого доступа, опубликованных в 2020–2025 гг., и каждый из них прочитан полностью. Это два метаанализа, один систематический обзор и восемь эмпирических исследований, два из которых проведены в Узбекистане. Источники проанализированы тематически по четырём вопросам: как автоматизированная и ИИ-обратная связь влияет на качество письма, насколько качество ИИ-обратной связи сопоставимо с обратной связью преподавателя, как обучающиеся с ней работают и как её воспринимают, какую роль играет преподаватель.

Ключевые слова: обратная связь с использованием ИИ, автоматизированная оценка письма, ChatGPT, письменная корректирующая обратная связь, письмо на английском как иностранном, вовлечённость обучающихся, обратная связь преподавателя.

Annotatsiya. Maqolada sun’iy intellekt (SI) vositalarining talabalar ingliz tilida yozgan ishlari yuzasidan fikr-mulohaza berishdagi imkoniyatlari so‘nggi empirik tadqiqotlar asosida

ko‘rib chiqiladi. Buning uchun 2020–2025-yillarda nashr etilgan, ochiq foydalanish mumkin bo‘lgan va ilmiy taqrizdan o‘tgan 11 ta manba tanlandi. Ularning barchasi to‘liq o‘rganib chiqildi. Tanlangan adabiyotlar orasida ikkita meta-tahlil, bitta tizimli sharh va sakkizta empirik tadqiqot mavjud hamda empirik izlanishlarning ikkitasi O‘zbekistonda olib borilgan.

Tadqiqotlar to‘rtta masala doirasida ko‘rib chiqilgan, xususan, avtomatlashtirilgan va SI yordamida berilgan fikr-mulohazalar talabalarning yozish natijalariga qanchalik ta’sir qilishi, bunday tavsiyalarning sifat jihatidan o‘qituvchi bergan fikr-mulohazalardan farqi, talabalarning SI tavsiyalaridan foydalanishi va ularga munosabati, shuningdek, bu jarayonda o‘qituvchining qanday o‘rin tutishi tahlil qilindi.

O‘rganilgan natijalar avtomatlashtirilgan fikr-mulohazalar talabalarning yozish ko‘nikmalariga ijobiy ta’sir ko‘rsatishini bildiradi. Biroq bu ta’sirni ortiqcha baholash to‘g‘ri bo‘lmasligi, chunki SI bergan fikr-mulohazalar malakali o‘qituvchi yoki tyutor tavsiyalaridan yaxshiroq ekanini ko‘rsatadigan aniq dalillar hozircha yetarli emas.

Kalit so‘zlar: sun’iy intellekt yordamidagi fikr-mulohaza, yozma ishni avtomatik baholash, ChatGPT, yozma tuzatuvchi fikr-mulohaza, ingliz tilida yozish, o‘quvchi faolligi, o‘qituvchi fikr-mulohazasi.

INTRODUCTION. Feedback is an important part of teaching writing, although its impact can vary depending on the type of feedback students receive. Research suggests that simply praising students or pointing out their mistakes is not always enough. Feedback becomes more useful when it provides clear information and helps students understand how to improve their work [1; Abstract; Discussion]. However, giving detailed comments on every student's written assignment is not easy, particularly for teachers working with large groups. It takes time and effort that teachers often cannot afford.

Automated Writing Evaluation (AWE) systems have offered a possible way to address this problem. More recently, tools based on large language models, such as ChatGPT, have attracted growing attention in English language teaching. Previous research on automated feedback has generally shown positive effects on students' writing performance [2; Abstract]. Studies on generative AI are more recent, but writing has already become one of the main areas of investigation. A systematic review of research on ChatGPT in ESL/EFL education found that a considerable number of studies focused on writing activities [3; Results, section 3a]. However, the authors also noted that much of the existing research relies on students' self-reported experiences, while fewer studies provide quantitative evidence of actual improvements in writing [3; Abstract; Discussion].

This issue is particularly relevant to English language teachers in Uzbekistan. As AI tools become more common in education, teachers need to understand whether the feedback provided by these tools can genuinely help students improve their writing. Recent studies conducted in the Uzbek context have also addressed this question [4; 162], [5; 91–109]. Their findings offer a useful starting point for considering how AI-assisted feedback can be used in local classrooms and how it compares with feedback provided by teachers.

This article reviews recent empirical research on the use of AI-assisted feedback in teaching English writing. It examines both the benefits and limitations of these tools, with particular attention to their practical use in the classroom. The review focuses on four questions:

1. What effect does automated and AI-generated feedback have on students' writing performance?

2. How does the quality of AI-generated feedback compare with feedback provided by teachers or tutors?

3. How do students use AI-generated feedback, and how do they perceive it?

4. What role does the teacher play when AI tools are used to provide feedback?

MATERIALS AND METHODS

This is a narrative (integrative) review, not a systematic review or a meta-analysis. The search was carried out in October 2026 using a general web search engine, the Directory of Open Access Journals and publishers’ open-access websites, with combinations of the terms AI feedback, ChatGPT feedback, automated written corrective feedback, automated writing evaluation, EFL/ESL writing and Uzbekistan. A source was included if it (a) was published in a peer-reviewed journal; (b) dealt with automated or AI-generated feedback on writing or reviewed such research; (c) reported empirical data or synthesised empirical studies; and (d) was openly accessible, so that the full text could be read and checked. Studies available only as an abstract were excluded, even when they looked relevant. One general meta-analysis of feedback research was added as theoretical background.

The empirical studies vary considerably. They differ in design (quasi-experimental, comparative-evaluative, case study and phenomenological), in context (universities in China, India, the Philippines and an unnamed Asia–Pacific country, a secondary-school sample, and two settings in Uzbekistan), in learners (secondary and tertiary, English as a foreign or second language, and in one case English as a new language) and in tools (Grammarly, GPT-3.5, GPT-4, ChatGPT, Gemini and Perplexity). For each source we recorded the design, participants, tool, outcome measures, main findings and stated limitations, and then grouped and compared the findings under the four research questions. Because the designs and measures differed so much, we did not attempt any statistical pooling; all figures reported below are taken directly from the sources. Page numbers are given where a source has printed pagination. For articles published with article numbers and no fixed pagination, the relevant section of the article is indicated instead.

RESULT AND DISCUSSION

Table 1 gives an overview of the empirical studies included in the review.

Study	Context and participants	Design and tool	Main finding
Escalante et al. [6]	University, Asia–Pacific; 48 + 43 ENL students	Quasi-experiment, 6 weeks; GPT-4 vs human tutor	Both groups improved; no significant difference; preferences evenly split
Fan [7]	University,	Quasi-experiment,	No significant advantage

	China; 67 lower-proficiency students	about 12 weeks; Grammarly + teacher vs teacher only	on any complexity, accuracy or fluency measure
Mahapatra [8]	University, India; 35 + 37 students	Quasi-experiment, 6 hours of instruction; ChatGPT	Experimental group outperformed comparison group; concerns about dependence
Steiss et al. [9]	Secondary school; 198 essays, 50 by English learners	Comparison of 16 trained evaluators vs GPT-3.5	Humans rated higher on 4 of 5 criteria; no difference by language status
Lim & Go [10]	University, Philippines; 22 essays, one instructor	Reflexive case study; ChatGPT	Useful but generic; errors; four strategies for blending with teacher feedback
Kasimova & Babakulova [5]	Lyceum, Tashkent; 32 B1–B2 learners, 96 feedback outputs	Qualitative comparison; ChatGPT, Gemini, Perplexity	Tools differ in detail and fit to level; prompt design matters
Yan & Zhang [11]	University, China; 4 learners	Multiple case study, 5 weeks; ChatGPT	Engagement and revision accuracy higher for high-proficiency learners
Abdullaev et al. [4]	Language institute, Tashkent; 30 intermediate learners	Phenomenological study, 8 weeks; AI feedback	Feedback seen as prompt and empowering; doubts about accuracy; over-reliance

Table 1. Empirical studies included in the review

Studies that compare AI feedback with no AI support often report positive results. In Mahapatra’s quasi-experiment with Indian undergraduates, the group that used ChatGPT outperformed the comparison group on both the post-test and the delayed post-test [8; Results]. The systematic review of ChatGPT research likewise reports several quasi-experimental studies in which experimental groups did better than control groups in writing performance or motivation [3; Results, section 3a]. Once AI feedback is set against feedback from a human, though, the advantage disappears. In the six-week study by Escalante, Pack and Barrett, students who received GPT-4 feedback and those who received feedback from a human tutor both improved, and the difference between the groups was not significant. The authors conclude that AI feedback did not produce superior progress but “can likely be incorporated into ENL essay evaluation” [6; 1, 8–9]. Fan’s twelve-week study with lower-proficiency Chinese students led to a similar result: adding Grammarly feedback to teacher feedback did not significantly improve complexity, accuracy or fluency compared with teacher feedback alone [7; Results; Discussion]. Taken together, these results suggest that AI

feedback is better than no feedback, but the reviewed studies do not show it to be better than feedback from a trained teacher or tutor. As the meta-analysis indicates, the answer also depends on what the control group receives.

Classroom studies point the same way. In Lim and Go’s case study of a university writing course in the Philippines, ChatGPT was useful for flagging areas for improvement and consistently identified the strengths of the essays. Its comments, however, were often generic, and it made mistakes: it marked a correct comma as a comma splice, for instance, and wrongly confirmed compliance with APA style [10; Findings]. In Uzbekistan, Kasimova and Babakulova compared feedback from ChatGPT, Gemini and Perplexity on essays written by B1–B2 learners. ChatGPT gave the most elaborated, teacher-like explanations, which helped B1 learners, although it sometimes overgeneralised rules. Gemini was concise and better suited to more autonomous B2 learners. Perplexity was accurate but brief, and this brevity confused weaker learners [5; 100–105]. A structured prompt produced clearer and more level-appropriate feedback than a minimal one, and none of the tools adapted to the learner’s level on its own [5; 99, 106]. These findings link back to general feedback research. If high-information feedback is the most effective, then the value of AI feedback depends on whether the information it provides is accurate, prioritised and suited to the learner, and the tool itself does not guarantee this.

CONCLUSION. The reviewed research shows that AI-assisted feedback is a useful but limited resource for teaching English writing. Compared with no feedback, automated feedback has a medium positive effect on writing performance; compared with feedback from trained teachers or tutors, it shows no clear advantage in the reviewed studies. AI feedback is fast, detailed and available at any time. It is, however, less able than experienced teachers to identify the most important problems in a text, and it still makes errors. Learners value it, but what they gain depends on their proficiency and on their ability to evaluate the feedback they receive. The evidence therefore supports a blended model in which the teacher designs the prompts, checks and selects the AI comments, and teaches learners how to use them.

The main contribution of the article is a synthesis of international and Uzbek research organised around four practical questions. Further research is needed in Uzbek universities, in particular controlled studies with larger samples and delayed post-tests, analyses of the revisions learners actually make after receiving AI feedback, and comparisons of different prompts and tools across proficiency levels.

REFERENCES:

[1]. Wisniewski B., Zierer K., Hattie J. The power of feedback revisited: A meta-analysis of educational feedback research // Frontiers in Psychology. – 2020. – Vol. 10. – Article 3087. https://doi.org/10.3389/fpsyg.2019.03087

[2]. Fleckenstein J., Liebenow L. W., Meyer J. Automated feedback and writing: A multi-level meta-analysis of effects on students’ performance // Frontiers in Artificial Intelligence. – 2023. – Vol. 6. – Article 1162454. https://doi.org/10.3389/frai.2023.1162454

[3]. Jong M. S. Exploring the application of ChatGPT in ESL/EFL education and related research issues: A systematic review of empirical studies // Smart Learning Environments. – 2024. – Vol. 11. – Article 50. https://doi.org/10.1186/s40561-024-00342-5

[4]. Abdullaev D., Abdazimov A., Khusainova Z., Makhmudov S., Sagdullaeva D. A phenomenological study on the impact of personalized AI feedback on writing quality: Insights from EFL students in Uzbekistan // Computer-Assisted Language Learning Electronic Journal (CALL-EJ). – 2025. – Vol. 26, No. 7. – P. 162–186. URL: https://callej.org/index.php/journal/article/view/816

[5]. Kasimova M. A., Babakulova D. X. A comparative evaluation of ChatGPT, Gemini, and Perplexity feedback for B1–B2 EFL learners // O‘zbekistonda xorijiy tillar [Foreign Languages in Uzbekistan]. – 2025. – Vol. 11, No. 6 (65). – P. 91–109. https://doi.org/10.36078/1767688041

[6]. Escalante J., Pack A., Barrett A. AI-generated feedback on writing: Insights into efficacy and ENL student preference // International Journal of Educational Technology in Higher Education. – 2023. – Vol. 20. – Article 57. – 20 p. https://doi.org/10.1186/s41239-023-00425-2

[7]. Fan N. Exploring the effects of automated written corrective feedback on EFL students’ writing quality: A mixed-methods study // SAGE Open. – 2023. – Vol. 13, No. 2. https://doi.org/10.1177/21582440231181296

[8]. Mahapatra S. Impact of ChatGPT on ESL students’ academic writing skills: A mixed methods intervention study // Smart Learning Environments. – 2024. – Vol. 11. – Article 9. https://doi.org/10.1186/s40561-024-00295-9

[9]. Steiss J., Tate T., Graham S., Cruz J., Hebert M., Wang J., Moon Y., Tseng W., Warschauer M., Olson C. B. Comparing the quality of human and ChatGPT feedback of students’ writing // Learning and Instruction. – 2024. – Article 101894. https://doi.org/10.1016/j.learninstruc.2024.101894 (accepted manuscript: https://par.nsf.gov/servlets/purl/10524947)

[10]. Lim J. M., Go C. Integrating ChatGPT into teacher feedback: Practical insights for L2 writing instruction // TESL-EJ. – 2025. – Vol. 29, No. 3. https://doi.org/10.55593/ej.29115int

[11]. Yan D., Zhang S. L2 writer engagement with automated written corrective feedback provided by ChatGPT: A mixed-method multiple case study // Humanities and Social Sciences Communications. – 2024. – Vol. 11. – Article 1086. https://doi.org/10.1057/s41599-024-03543-y