



THE EFFECT OF CHATGPT-ASSISTED FEEDBACK ON EFL STUDENTS' ACADEMIC WRITING SKILLS

Moxinur Yusupova

MA in TESOL, Webster University, Tashkent, Uzbekistan

English Teacher, Daholar Maktabi Private School, Khorezm, Uzbekistan

Abstract: The rapid diffusion of large language models into higher education has reopened long-standing debates about the role of feedback in second-language writing development. This article synthesizes the growing body of empirical research on ChatGPT-assisted feedback and its effect on the academic writing skills of English as a Foreign Language (EFL) learners. Drawing on quasi-experimental, mixed-methods, and qualitative studies published between 2023 and 2026, the review examines how ChatGPT-generated feedback influences accuracy, organization, coherence, and learner motivation, and how it compares with teacher-generated and peer feedback. The evidence indicates that structured ChatGPT feedback produces measurable gains in grammatical accuracy, lexical range, and textual organization, and in several studies performs comparably to teacher feedback, while learner engagement depends heavily on feedback literacy and prompt design. The article also considers persistent concerns regarding over-reliance, feedback quality variability, and the erosion of learner autonomy. It concludes with pedagogical recommendations for integrating ChatGPT as a complementary, rather than a substitutive, feedback source within EFL writing instruction, and outlines directions for future longitudinal and cross-proficiency research.

Keywords: ChatGPT; artificial intelligence; EFL writing; automated feedback; academic writing skills; feedback literacy.

Introduction

Academic writing remains one of the most demanding skills for learners of English as a Foreign Language (EFL), requiring simultaneous control of grammar, vocabulary, coherence, and genre-specific conventions. Traditional models of writing instruction rely heavily on teacher-provided feedback, a process that is pedagogically valuable but often constrained by large class sizes, limited instructional time, and the subjective and delayed nature of manual correction. These constraints frequently reduce the frequency and specificity of feedback that individual learners receive, which in turn slows the revision cycles that are central to writing development.

The emergence of generative artificial intelligence, and of ChatGPT in



particular, has introduced a new possibility for addressing these constraints: scalable, immediate, and individualized feedback that is available outside the classroom and independent of instructor workload. Since its public release, ChatGPT has been adopted across a wide range of educational contexts, and a substantial body of applied linguistics research has begun to examine whether AI-generated feedback can meaningfully improve EFL learners' academic writing, and how it compares with conventional feedback sources.

This article reviews and synthesizes this emerging literature. It addresses three guiding questions: (1) What measurable effects does ChatGPT-assisted feedback have on EFL students' academic writing performance? (2) How does ChatGPT-generated feedback compare with teacher-generated and peer feedback in terms of effectiveness and learner perception? (3) What pedagogical and methodological considerations should guide the integration of ChatGPT into EFL writing instruction? The remainder of the article is organized as follows: Section 2 reviews the theoretical background and relevant empirical studies; Section 3 outlines the review methodology; Section 4 synthesizes findings across accuracy, organization, and motivational outcomes; Section 5 discusses pedagogical implications; and Section 6 concludes with limitations and directions for future research.

Literature Review

2.1 Feedback in Second-Language Writing Pedagogy

Feedback has long been regarded as a central mechanism of second-language writing development, functioning to draw learners' attention to gaps between their current interlanguage and target-like production. Written corrective feedback research has traditionally distinguished between direct and indirect correction, focused and unfocused feedback, and has emphasized the importance of feedback literacy — learners' capacity to understand, evaluate, and act on feedback — as a mediating factor in whether feedback translates into actual revision and learning gains (Rahimi, Fathi, & Zou, 2024). Automated feedback tools predating large language models, such as automated writing evaluation (AWE) systems, had already demonstrated that machine-generated feedback could support revision behavior and produce measurable learning gains under formative conditions (Zhu, Liu, & Lee, 2020), setting a precedent for the more flexible, dialogic feedback that generative AI tools now offer.

2.2 ChatGPT as an Emerging Feedback Tool

Unlike rule-based AWE systems, ChatGPT generates feedback through natural-language dialogue, allowing learners to request clarification, ask for explanations, or iteratively refine their drafts. A systematic review of AI-based automated written feedback research situates ChatGPT within a broader



trajectory of increasingly interactive and learner-responsive feedback technologies, while cautioning that the quality and pedagogical soundness of AI feedback remain uneven across tools and contexts (Shi & Aryadoust, 2024). Comparable generative tools, such as Wordtune, have similarly been shown to foster writing feedback literacy and learner engagement when embedded in structured tasks (Rad, Alipour, & Jafarpour, 2024).

2.3 Empirical Evidence on Writing Outcomes

A growing number of quasi-experimental and mixed-methods studies provide direct evidence on the effect of ChatGPT feedback on EFL writing. Polakova and Ivenz (2024), in a quasi-experimental study of 110 university EFL students in the Czech Republic, found statistically significant improvements in conciseness, grammatical accuracy, inclusion of key content, and appropriate use of passive voice following structured ChatGPT feedback, with qualitative focus-group data corroborating learners' positive perceptions of the tool. In a Saudi Arabian tertiary context, Alsofyani and Barzanji (2025) compared ChatGPT-generated feedback with teacher-generated feedback using a pretest–posttest control-group design with 102 participants and found no statistically significant difference between the two feedback sources, suggesting that ChatGPT feedback can be as effective as teacher feedback while producing an overall positive learner attitude toward AI-assisted writing support.

Huang (2026) examined the comparative effectiveness of peer feedback and ChatGPT-generated feedback among 174 Chinese university EFL students enrolled in a mandatory writing course, framing the analysis through engagement theory and feedback literacy; the treatment group receiving ChatGPT feedback, guided by carefully constructed prompts, showed distinct engagement patterns relative to the peer-feedback control group. At the lower proficiency end, studies of A2-level learners have reported that structured, criterion-referenced ChatGPT prompts can produce significant gains in language accuracy and structural organization without regression in any writing subskill, indicating that AI feedback can be adapted to beginner learners when accompanied by clear evaluative scaffolding. Research combining ChatGPT with teacher feedback further suggests a synergistic rather than a purely substitutive relationship: teachers' professional judgment remains important for tailoring AI-generated feedback to individual learners' specific writing challenges, pointing toward blended feedback models rather than full automation.

Beyond short-term gains, longitudinal and cross-genre research has begun to examine whether the benefits of ChatGPT feedback persist over time and across writing tasks, extending earlier single-session designs into multi-week



interventions that track learners' revision behavior across argumentative, descriptive, and report genres. Findings from this emerging strand generally reinforce that consistent, structured exposure to AI feedback sustains gains in accuracy and organization, though improvements in higher-order concerns such as argumentation and critical thinking appear to develop more gradually than surface-level accuracy.

Perception-focused studies add an important qualitative dimension to these outcome measures. Reviews of student perceptions report that learners generally value ChatGPT's immediacy, personalization, and non-judgmental tone, which reduces writing-related anxiety and encourages experimentation with language, while also raising concerns about overuse, the occasional inaccuracy of AI-generated explanations, and the risk of learners accepting feedback uncritically. A systematic literature review applying the PRISMA framework to studies of ChatGPT as an EFL feedback tool concluded that while benefits are consistently reported across contexts, methodological heterogeneity — in sample size, proficiency level, and feedback protocol — limits the generalizability of aggregate effect sizes across the field.

Methodology

This article adopts a narrative literature review methodology, synthesizing peer-reviewed empirical studies, systematic reviews, and mixed-methods investigations published primarily between 2020 and 2026 that examine ChatGPT-assisted or AI-generated feedback in EFL academic writing contexts. Sources were identified through database searches (including Scopus-indexed journals such as Cogent Education, the International Journal of Applied Linguistics, the Journal of Educational Computing Research, and ReCALL) using combinations of the terms "ChatGPT," "AI feedback," "EFL writing," and "academic writing skills." Studies were prioritized for inclusion if they reported quantitative writing-performance outcomes, comparative feedback conditions (AI versus teacher or peer feedback), or learner perception data collected through validated instruments or systematic qualitative procedures. Findings were then thematically organized around three recurring outcome categories identified across the literature: (a) linguistic accuracy and organization, (b) comparative effectiveness relative to human feedback sources, and (c) learner engagement, motivation, and feedback literacy.

Findings and Discussion

4.1 Effects on Linguistic Accuracy and Textual Organization

Across the reviewed studies, the most consistently reported outcome is improvement in surface-level accuracy: grammar, mechanics, conciseness, and appropriate stylistic choices such as voice selection. These gains align with what would be expected from any well-structured corrective feedback



intervention, but the reviewed studies suggest that ChatGPT's capacity to generate individualized, example-rich explanations may accelerate learners' uptake of corrections relative to generic or delayed feedback. Improvements in higher-order textual organization — coherence, paragraph structure, and inclusion of genre-appropriate content — were also reported, though somewhat less uniformly, suggesting that AI feedback may be more reliably effective for form-focused than for discourse-level revision.

4.2 ChatGPT Feedback Compared with Human Feedback Sources

Direct comparisons between ChatGPT feedback and teacher feedback generally report comparable outcomes rather than a clear advantage for either source, with several studies finding no statistically significant difference in post-intervention writing scores between groups. This pattern supports a complementary rather than competitive framing: ChatGPT appears well suited to providing frequent, low-stakes formative feedback that increases the sheer volume of feedback learners receive, while teachers retain an irreplaceable role in diagnosing individual learning needs, contextualizing feedback within course objectives, and modeling academic judgment that purely automated systems cannot replicate.

4.3 Motivation, Engagement, and Feedback Literacy

Learner engagement with ChatGPT feedback is not uniform and appears to be moderated by feedback literacy — learners' ability to interpret and act on the feedback they receive — as well as by the specificity of the prompts used to elicit it. Studies grounded in engagement and feedback-literacy frameworks report that learners with higher feedback literacy extract more actionable revisions from AI feedback and report greater perceived usefulness, whereas learners with lower feedback literacy are more likely to accept AI suggestions passively or to disengage. This finding underscores that the pedagogical value of ChatGPT feedback is not solely a property of the tool itself but emerges from the interaction between the tool, the task design, and the learner's existing metacognitive and self-regulatory skills.

Pedagogical Implications

The reviewed evidence suggests several practical recommendations for EFL writing instruction. First, ChatGPT feedback is most effective when guided by structured, criterion-referenced prompts rather than open-ended requests for correction, since specificity in the prompt appears to shape the specificity and usefulness of the resulting feedback. Second, blended models that combine ChatGPT feedback with teacher oversight appear to capture the scalability benefits of AI while preserving the pedagogical judgment that automated systems lack; teachers can use class time to address higher-order concerns while directing learners to ChatGPT for iterative, form-focused



revision. Third, explicit instruction in feedback literacy — teaching learners how to critically evaluate, question, and selectively apply AI-generated suggestions — should accompany the introduction of ChatGPT into the writing classroom, both to maximize learning gains and to guard against uncritical over-reliance on AI output.

6. Limitations and Future Research

The body of research reviewed here, while expanding rapidly, remains constrained by several methodological limitations. Sample sizes in many studies are moderate and drawn from single institutional contexts, limiting generalizability across proficiency levels, first-language backgrounds, and educational systems. Intervention periods are frequently short, and few studies have tracked writing development longitudinally or across multiple genres, leaving open questions about whether initial gains are sustained or transfer to unassisted writing performance. Additionally, most studies rely on aggregate proficiency measures rather than fine-grained analyses of which specific error types or discourse features respond best to AI feedback. Future research would benefit from longitudinal, cross-proficiency, and cross-genre designs, as well as from closer attention to the role of prompt engineering and feedback literacy instruction as moderating variables in intervention outcomes.

Conclusion

The reviewed literature indicates that ChatGPT-assisted feedback can produce meaningful improvements in EFL students' academic writing, particularly in grammatical accuracy, conciseness, and structural organization, and that in several controlled comparisons it performs comparably to teacher-generated feedback. At the same time, the effectiveness of ChatGPT feedback is not automatic: it depends on structured prompting, learners' feedback literacy, and thoughtful integration alongside — rather than in place of — teacher guidance. As generative AI tools continue to develop, EFL writing pedagogy is likely to move toward blended feedback ecosystems in which AI-generated feedback expands the frequency and immediacy of formative support, while teachers retain responsibility for higher-order instructional judgment. Sustained empirical attention to long-term outcomes, equity of access, and the cultivation of critical feedback literacy will be essential to realizing the pedagogical potential of this technology responsibly.

REFERENCES

1. Alsofyani, A. H., & Barzanji, A. M. (2025). The effects of ChatGPT-generated feedback on Saudi EFL learners' writing skills and perception at the tertiary level: A mixed-methods study. *Journal of Educational Computing Research*, 63(2), 431–463.



<https://doi.org/10.1177/07356331241307297>

2. 2.Huang, J. (2026). Assessing ChatGPT feedback for EFL learners' engagement and writing performance. *International Journal of Applied Linguistics*. <https://doi.org/10.1111/ijal.70155>
3. 3.Polakova, P., & Ivenz, P. (2024). The impact of ChatGPT feedback on the development of EFL students' writing skills. *Cogent Education*, 11(1), 2410101. <https://doi.org/10.1080/2331186X.2024.2410101>
4. 4.Rad, H. S., Alipour, R., & Jafarpour, A. (2024). Using artificial intelligence to foster students' writing feedback literacy, engagement, and outcome: A case of Wordtune application. *Interactive Learning Environments*, 32(9), 5020–5040. <https://doi.org/10.1080/10494820.2023.2208170>
5. 5.Rahimi, M., Fathi, J., & Zou, D. (2024). Exploring the impact of automated written corrective feedback on the academic writing skills of EFL learners: An activity theory perspective. *Education and Information Technologies*. Advance online publication. <https://doi.org/10.1007/s10639-024-12896-5>
6. 6.Shi, H., & Aryadoust, V. (2024). A systematic review of AI-based automated written feedback research. *ReCALL*, 36(2), 187–209. <https://doi.org/10.1017/S0958344023000265>