ENHANCING EFL STUDENTS' WRITING PROFICIENCY THROUGH AI-ASSISTED WRITING TOOLS

Qosimjonov Muhammadg'ofir Hoshimjon o'g'li
Master's Graduate, Namangan State Institute of Foreign Languages
Namangan State Institute of Foreign Languages
Department of Tourism and Translation
Trainee Teacher
Email: qmuhammadgofir@gmail.com
Phone: +998 99 646 09 46

Abstract: The rapid development of artificial intelligence (AI) has introduced new possibilities for English as a Foreign Language (EFL) writing instruction. AI-assisted writing tools, including automated writing evaluation systems, Grammarly, ChatGPT and other generative artificial intelligence applications, can provide learners with immediate feedback on grammar, vocabulary, sentence structure, organization and aspects of academic style. This article examines the potential of AI-assisted writing tools to enhance EFL students' writing proficiency through a synthesis and analysis of recent empirical and review-based research. Particular attention is given to automated written corrective feedback, generative AI-supported revision, writing quality, learner engagement, motivation and academic integrity. Evidence from empirical studies indicates that AI-supported writing can contribute to improvements in several dimensions of writing, although the effects are not uniform across all linguistic and higher-order writing dimensions. For example, a 2023 quasi-experimental study involving 67 EFL university students found no statistically significant advantage of Grammarly-assisted feedback over teacher feedback in syntactic and lexical complexity, accuracy and fluency [1]. In contrast, other studies involving ChatGPT have reported significant improvements in selected writing outcomes, motivation and self-efficacy [2–5]. A 2026 meta-analysis synthesizing 19 empirical studies and 57 effect sizes reported a large overall effect of ChatGPT on EFL writing, with Hedges' g = 0.906, although the authors also conducted a more conservative analysis because multiple outcomes from the same studies may violate statistical independence [6]. The evidence therefore suggests that AI-assisted writing tools are most educationally valuable when they complement, rather than replace, teacher instruction, student reflection and independent writing. The article concludes that effective AI integration requires pedagogical guidance, critical evaluation of AI-generated feedback, transparent academic-integrity policies and a process-oriented approach to writing.

Keywords: Artificial intelligence; AI-assisted writing; EFL; ChatGPT; Grammarly; automated writing evaluation; academic writing; writing proficiency; automated feedback; second language writing

Introduction

Writing is one of the most cognitively and linguistically demanding skills for learners of English as a Foreign Language. Unlike receptive skills, academic writing requires learners to generate and organize ideas, select appropriate vocabulary, construct grammatically accurate sentences, establish logical relationships between propositions and adapt language to a particular communicative purpose and audience. Consequently, EFL students frequently require repeated opportunities for drafting, feedback, revision and reflection.

The development of artificial intelligence has considerably expanded the range of technological tools available to support this process. Automated Writing Evaluation (AWE)

systems have been used in writing education for many years, while recent generative AI systems have introduced more flexible forms of interaction. Traditional AWE systems generally focus on evaluating and providing feedback on features such as grammar, vocabulary, mechanics, organization and writing quality. Generative AI systems such as ChatGPT can additionally engage in dialogue, explain suggested corrections, propose alternative formulations, assist with brainstorming and provide feedback on drafts.

A systematic review by Shi and Aryadoust examined 105 published studies on automated writing evaluation and found considerable heterogeneity in research settings, language environments and educational levels. Their review also identified an imbalance in validation research, with evaluation-related evidence receiving considerably more attention than evidence concerning the correspondence between AWE systems and authentic writing tasks [7, pp. 771–795]. These findings are important because improvement in automated scores does not automatically demonstrate development of a learner's independent writing competence.

Research specifically focused on AI-assisted EFL writing has expanded rapidly since the introduction of generative AI. Barrot argued that ChatGPT has potential as a tutor and source of language input for second-language writing, while also identifying limitations related to accuracy, intelligibility, writing pedagogy and academic integrity [8, Article 100745]. Such limitations mean that AI-generated suggestions cannot automatically be considered equivalent to expert teacher feedback.

Empirical evidence nevertheless provides substantial reasons for investigating AI-assisted writing. Song and Song conducted a mixed-methods study with Chinese EFL students and examined writing performance and motivation in AI-assisted language learning [9]. Their study reported improvements following ChatGPT-supported instruction. Similarly, Polakova and Ivenz investigated ChatGPT feedback among 110 university EFL students and reported improvements in several dimensions of writing, including conciseness, grammar, inclusion of key information and use of passive voice [10, Article 2410101]. Mahapatra's mixed-methods intervention study also reported a significant positive impact of ChatGPT used as a formative feedback tool on undergraduate ESL students' academic writing skills [11, Article 9].

At the same time, evidence is not universally positive. Fan's quasi-experimental study of 67 EFL university students found that students who received Grammarly-assisted feedback did not significantly outperform a comparison group receiving teacher feedback alone in syntactic complexity, lexical complexity, accuracy and fluency [1, pp. 1–17]. This finding demonstrates that the mere availability of AI feedback does not guarantee measurable improvement in every dimension of writing.

The purpose of this article is therefore to examine the effectiveness of AI-assisted writing tools from a balanced evidence-based perspective. Rather than assuming that AI necessarily improves writing, the article considers what existing research demonstrates, which aspects of writing appear most responsive to AI support, what limitations remain, and how AI can be integrated into EFL writing instruction without weakening learner autonomy or academic integrity.

Methodology

This article employs a structured literature-based research methodology. The analysis focuses on empirical studies, systematic reviews, research articles and institutional guidance addressing artificial intelligence, automated writing evaluation, automated written corrective feedback, generative AI and EFL/ESL writing.

The literature was selected according to four principal criteria. First, the source had to address writing instruction, writing performance, writing feedback or academic writing.

Second, priority was given to peer-reviewed empirical research and systematic reviews. Third, studies published between 2023 and 2026 were emphasized because this period corresponds to the rapid expansion of generative AI in language education. Earlier AWE research was included where it provided an important empirical or theoretical basis for understanding AI-assisted feedback. Fourth, quantitative findings were included only when they were explicitly reported by the original researchers.

The methodological approach was interpretive rather than a new statistical meta-analysis. Therefore, the numerical results reported below are extracted from previously published studies rather than calculated from newly collected data. This distinction is essential because differences in sample size, intervention duration, assessment instruments, learner proficiency, AI tools and research designs make direct comparison between individual studies inappropriate.

The literature examined includes research involving Grammarly, ChatGPT and broader AWE systems. For example, Fan used a quasi-experimental design with 67 EFL students [1]. Polakova and Ivenz investigated 110 EFL university students through pre-tests, post-tests, questionnaires and focus-group interviews [10]. Mahapatra used three tests and focus-group discussions with tertiary-level ESL students [11]. Teng and colleagues studied 45 EFL learners in Macau using questionnaire and interview data collected after a semester-long writing course [12, Article 100270]. At a broader level, a 2026 meta-analysis synthesized 19 empirical studies and 57 effect sizes concerning ChatGPT and EFL writing [6].

The analysis also incorporates research on ethical and pedagogical issues. UNESCO's guidance on generative AI in education emphasizes human-centred use, data privacy, age-appropriate implementation, ethical validation and institutional policy development [13, pp. 1–44]. Research on academic integrity has similarly emphasized the need for institutions to establish clear policies and provide students and teachers with guidance on responsible AI use [14, pp. 228–239].

Results

The reviewed literature indicates that AI-assisted writing tools can influence several dimensions of EFL writing, but the strength and consistency of these effects vary.

One of the clearest potential benefits concerns immediate feedback. Traditional teacher feedback is constrained by teachers' workload, class size and available instructional time. AI-based systems can provide automated suggestions immediately after students submit a draft. AWE research has repeatedly investigated this possibility because automated systems can provide rapid feedback across multiple writing tasks. Liu's systematic review analyzed 68 articles published between 2008 and 2023 and concluded that AWE feedback has become increasingly prominent in writing classrooms, while also identifying uncertainty concerning the accuracy of some automated error detection and differences between student preferences for teacher and automated feedback [15, pp. 86–105].

Grammatical accuracy is another area in which AI tools may be useful. Grammarly and comparable systems can identify spelling, punctuation, grammatical and some stylistic problems. Fan's study, however, demonstrates that the presence of automated corrective feedback does not necessarily produce statistically significant gains in overall writing quality. The 67 students in the treatment condition received two rounds of Grammarly feedback in combination with teacher feedback, yet they did not significantly outperform the comparison group in syntactic and lexical complexity, accuracy and fluency [1, pp. 1–17]. This result suggests that feedback must be actively processed by learners rather than simply accepted.

Research on generative AI presents additional evidence. In a study of 110 university EFL students, Polakova and Ivenz found improvements in multiple writing categories following ChatGPT feedback, including conciseness, grammar, inclusion of key information,

passive voice, linking verbs and reporting verbs [10, Article 2410101]. The study used a pre-test/post-test design and also incorporated questionnaire and focus-group evidence, thereby combining performance-related and perception-related data.

Mahapatra's intervention study likewise found a significant positive impact of ChatGPT on academic writing skills among undergraduate ESL learners [11, Article 9]. The study treated ChatGPT as a formative feedback tool rather than simply as a text generator. This distinction is pedagogically important because formative feedback is intended to support revision and learning rather than merely produce a finished text.

Motivation and self-efficacy also appear in the literature. Song and Song examined Chinese EFL students using a mixed-methods design and reported improvements in writing skills and writing motivation following AI-assisted instruction [9, Article 1260843]. Teng and colleagues, working with 45 EFL learners in Macau, reported positive effects associated with AI assistance, including writing motivation, self-efficacy, engagement and collaborative writing tendencies [12, Article 100270].

The largest quantitative evidence considered in this article comes from a 2026 meta-analysis of ChatGPT and EFL writing. The researchers synthesized 19 empirical studies and calculated 57 effect sizes. The initial analysis reported Hedges' $g = 0.906$, $SE = 0.088$ and $p < .001$, indicating a large statistically significant overall effect [6]. Importantly, the researchers also recognized that several effect sizes could originate from the same study and therefore performed a more conservative analysis using one independent effect size per study. This methodological caution strengthens the interpretation of the result because non-independent observations can artificially increase the apparent precision of a pooled estimate.

The literature also identifies limitations. A scoping review of 16 empirical studies published before December 2023 found that ChatGPT had already become an important subject of research in EFL writing, but the field remained relatively new and required more systematic empirical investigation [16]. Research involving Saudi university students similarly reported positive effects but identified concerns involving contextual understanding, individuality, creativity and ethical use [17].

Finally, concerns about academic integrity remain significant. Generative AI can produce coherent academic-style text without requiring the learner to formulate every sentence independently. Cotton, Cotton and Shipway therefore emphasize the risks associated with academic dishonesty and the difficulty of detecting AI-assisted misconduct, while recommending proactive institutional policies, training and support [14, pp. 228–239].

Analysis and Discussion

The evidence suggests that AI-assisted writing tools should not be understood as replacements for writing instruction. Their educational value depends primarily on how they are incorporated into the writing process.

A useful distinction can be made between AI as a **substitute** for writing and AI as a **scaffold** for writing. In substitution-oriented use, a student gives a topic to an AI system and submits the generated essay with minimal modification. Such use may reduce the amount of cognitive work performed by the learner and provides limited evidence that the learner has developed writing proficiency. In scaffold-oriented use, the student first produces ideas and a draft, then uses AI to identify weaknesses, compare alternative formulations, ask for explanations and revise the text independently. The second model is more consistent with the pedagogical purpose of feedback.

The evidence on Grammarly illustrates this distinction. Fan's study did not find significant differences between Grammarly-assisted and teacher-feedback groups across syntactic complexity, lexical complexity, accuracy and fluency [1]. One interpretation is that automated correction is not inherently equivalent to learning. A student may accept an

automatically generated correction without understanding why the original construction was problematic. Consequently, the educational effect depends on whether the learner engages cognitively with the feedback.

The same principle applies to generative AI. ChatGPT can explain grammatical concepts, suggest alternative vocabulary, identify unclear sentences and provide organizational recommendations. However, it can also produce inappropriate or inaccurate suggestions. Barrot specifically highlighted limitations involving accuracy and intelligibility in the context of second-language writing [8]. Therefore, EFL students need the ability to evaluate AI output rather than treating it as an unquestionable authority.

The strongest evidence currently available suggests that AI is particularly useful during the feedback and revision stages. The 2026 meta-analysis found a large overall positive effect for ChatGPT-supported EFL writing [6]. A separate 2026 systematic review of 79 empirical articles published between 2023 and 2025 reported that GenAI use in second-language writing is concentrated particularly around feedback, revision and polishing stages, while other stages of the writing process receive less attention [18]. This finding supports a process-oriented pedagogical model in which AI is introduced after students have engaged in independent planning and drafting.

AI tools can also potentially increase feedback frequency. In a conventional classroom, a student may receive teacher feedback only at specific stages because teachers cannot provide unlimited individualized responses. AI can provide an additional feedback layer between teacher consultations. Students can therefore experiment with revisions and receive immediate responses. This may be particularly valuable for learners who are reluctant to ask repeated questions in class.

However, speed should not be confused with quality. Teacher feedback has contextual advantages. Teachers know the course objectives, students' developmental histories and the requirements of specific assignments. Automated systems can identify linguistic patterns rapidly but may not understand the rhetorical purpose of a text or the disciplinary conventions underlying it. The systematic review by Shi and Aryadoust highlights the continuing need to examine the validity of AWE systems and their correspondence with authentic writing tasks [7, pp. 771–795].

Another important issue concerns higher-order writing skills. Grammar correction is relatively easy to automate compared with evaluating the intellectual quality of an argument. Academic writing requires the development of claims, evidence, reasoning, synthesis and critical engagement with sources. AI-generated suggestions may improve surface-level correctness without necessarily developing these underlying intellectual abilities. Recent systematic research similarly indicates that generative AI tends to be stronger in grammar, coherence and revision support than in higher-order dimensions such as argumentation and creativity [19, Article 101041].

AI-assisted writing should therefore be embedded in explicit instructional sequences. One effective sequence can begin with independent brainstorming, followed by student-generated outlining and drafting. The student can then ask an AI system to identify grammatical problems without rewriting the entire text. The learner evaluates each recommendation, accepts or rejects it, and explains the reason for important changes. The teacher subsequently reviews selected revisions and addresses recurrent problems. This approach transforms AI from an answer-generating mechanism into a metacognitive learning resource.

Prompt design is another factor. Students who simply request "correct my essay" may receive extensive rewriting, which reduces opportunities for learning. More pedagogically appropriate prompts can request identification of errors, explanations of grammatical rules,

questions about unclear ideas or suggestions presented as alternatives rather than direct replacements. In this way, AI feedback can encourage learner decision-making.

The issue of academic integrity requires equal attention. UNESCO recommends a human-centred approach to generative AI and emphasizes ethical, safe, equitable and meaningful implementation [13, pp. 1–44]. Institutions should therefore define what forms of AI assistance are permitted. For example, a university might permit grammar checking, brainstorming and feedback but require students to disclose substantial AI assistance or prohibit the generation of complete assessed assignments. Such distinctions are more educationally useful than treating all AI use as either completely acceptable or completely prohibited.

There is also a need to distinguish AI-assisted learning from AI-generated assessment products. If the objective of an assignment is to assess students' independent writing ability, unrestricted AI generation can compromise construct validity. Conversely, if students are explicitly being taught how to use AI critically as part of contemporary academic literacy, controlled AI use may become part of the learning objective. Assessment design must therefore correspond to the intended learning outcomes.

The current evidence also reveals methodological limitations in the research field. Many studies use relatively small samples, short interventions and context-specific groups. For instance, Fan's research involved 67 students [1], while Teng and colleagues studied 45 learners [12]. Such studies can provide valuable classroom evidence but cannot automatically be generalized to all EFL populations. Learners' proficiency, first language, educational culture, writing genre and familiarity with technology may influence outcomes.

The 2026 meta-analysis provides stronger aggregate evidence because it synthesizes multiple empirical studies, but even meta-analysis cannot eliminate heterogeneity among the original studies [6]. Future research should therefore include larger samples, longer intervention periods, standardized writing assessments and delayed post-tests to determine whether improvements persist after AI assistance is removed.

An especially important research direction is the development of AI literacy. EFL students need not only linguistic competence but also the ability to evaluate automated feedback. AI literacy includes understanding that AI-generated information may be inaccurate, recognizing inappropriate suggestions, protecting personal data, identifying potential bias and acknowledging AI assistance when required. These abilities are increasingly relevant to academic and professional writing.

Overall, the evidence supports a balanced conclusion: AI-assisted writing tools have considerable potential, but their educational effectiveness depends on instructional design. The strongest pedagogical model is not "AI writes for the student" but "AI provides feedback and alternatives that the student evaluates, understands and uses to revise their own writing."

Conclusion

AI-assisted writing tools are becoming an increasingly important component of EFL writing education. Research conducted with Grammarly, ChatGPT and automated writing evaluation systems demonstrates that AI can provide immediate feedback, support revision, facilitate language correction and potentially increase students' motivation and confidence. Empirical studies involving 67, 110 and 45 learners respectively provide evidence of both positive outcomes and important limitations [1, 10, 12]. The 2026 meta-analysis of 19 studies and 57 effect sizes further reported a statistically significant large overall effect of ChatGPT on EFL writing, with Hedges' $g = 0.906$ [6].

Nevertheless, the evidence does not justify the conclusion that AI automatically improves writing proficiency. Some controlled studies have found no significant advantage

of automated feedback over teacher feedback in important linguistic dimensions [1]. AI systems can also generate inaccurate, contextually inappropriate or stylistically unsuitable suggestions. Furthermore, excessive reliance on generative AI can reduce opportunities for independent thinking and create academic-integrity problems.

The most defensible educational position is therefore one of guided integration. AI should complement teacher feedback and student effort rather than replace them. Students should first engage in planning and drafting, then use AI for targeted feedback, explanation and revision. Teachers should retain responsibility for higher-order evaluation, rhetorical development and disciplinary guidance.

For OAK-level academic writing instruction, the central implication is that AI literacy should become part of EFL writing pedagogy. Students need to learn not only how to use AI tools but also how to question, verify and critically evaluate their output. Future empirical research should employ larger and more diverse samples, longer interventions and standardized measures of writing proficiency. Particular attention should be given to whether AI-supported improvements remain after the technology is removed.

In conclusion, AI-assisted writing tools should be viewed as pedagogical instruments rather than autonomous authors. When integrated through a structured cycle of independent writing, AI feedback, critical evaluation, human feedback and revision, these technologies can contribute meaningfully to the development of EFL students' writing proficiency while preserving learner autonomy and academic integrity.

References

1. Fan, N. Exploring the Effects of Automated Written Corrective Feedback on EFL Students' Writing Quality: A Mixed-Methods Study. SAGE Open, 2023, Vol. 13, Issue 2, pp. 1–17.
2. Song, C., Song, Y. Enhancing Academic Writing Skills and Motivation: Assessing the Efficacy of ChatGPT in AI-Assisted Language Learning for EFL Students. Frontiers in Psychology, 2023, Vol. 14, Article 1260843. Article-numbered publication; conventional page range not assigned.
3. Alkamel, M. A. A., Alwagieh, N. A. S. Utilizing an Adaptable Artificial Intelligence Writing Tool (ChatGPT) to Enhance Academic Writing Skills among Yemeni University EFL Students. Social Sciences & Humanities Open, 2024, Vol. 10, Article 101095. Article-numbered publication; conventional page range not assigned.
4. Polakova, P., Ivenz, P. The Impact of ChatGPT Feedback on the Development of EFL Students' Writing Skills. Cogent Education, 2024, Vol. 11, Issue 1, Article 2410101. Article-numbered publication; conventional page range not assigned.
5. Mahapatra, S. Impact of ChatGPT on ESL Students' Academic Writing Skills: A Mixed Methods Intervention Study. Smart Learning Environments, 2024, Vol. 11, Article 9. Article-numbered publication; conventional page range not assigned.
6. Zhang, R., Chen, J. The Effects of ChatGPT on EFL Students' Writing Skills: A Meta-Analysis. System, 2026, Vol. 140, Article 104043. Article-numbered publication; conventional page range not assigned. The meta-analysis synthesized 19 empirical studies and 57 effect sizes.
7. Shi, H., Aryadoust, V. A Systematic Review of Automated Writing Evaluation Systems. Education and Information Technologies, 2023, Vol. 28, Issue 1, pp. 771–795.
8. Barrot, J. S. Using ChatGPT for Second Language Writing: Pitfalls and Potentials. Assessing Writing, 2023, Vol. 57, Article 100745. Article-numbered publication; conventional page range not assigned.
9. Song, C., Song, Y. Enhancing Academic Writing Skills and Motivation: Assessing the Efficacy of ChatGPT in AI-Assisted Language Learning for EFL Students. Frontiers in

Psychology, 2023, Vol. 14, Article 1260843. Article-numbered publication; conventional page range not assigned.
10. Polakova, P., Ivenz, P. The Impact of ChatGPT Feedback on the Development of EFL Students' Writing Skills. Cogent Education, 2024, Vol. 11, Issue 1, Article 2410101. Article-numbered publication; conventional page range not assigned.
11. Mahapatra, S. Impact of ChatGPT on ESL Students' Academic Writing Skills: A Mixed Methods Intervention Study. Smart Learning Environments, 2024, Vol. 11, Article 9. Article-numbered publication; conventional page range not assigned.
12. Teng, M. F. "ChatGPT is the Companion, Not Enemies": EFL Learners' Perceptions and Experiences in Using ChatGPT for Feedback in Writing. Computers and Education: Artificial Intelligence, 2024, Vol. 7, Article 100270. Article-numbered publication; conventional page range not assigned.
13. Miao, F., Holmes, W. Guidance for Generative AI in Education and Research. UNESCO, 2023, 44 pp.; relevant discussion throughout pp. 1–44.
14. Cotton, D. R. E., Cotton, P. A., Shipway, J. R. Chatting and Cheating: Ensuring Academic Integrity in the Era of ChatGPT. Innovations in Education and Teaching International, 2023/2024, pp. 228–239 in the published journal version; repository version records pp. 1–12.
15. Liu, W. A Systematic Review of Automated Writing Evaluation Feedback: Validity, Effects and Students' Engagement. Language Teaching Research Quarterly, 2024, Vol. 45, pp. 86–105.
16. Xiao, F., Zhu, S., Xin, W. Exploring the Landscape of Generative AI (ChatGPT)-Powered Writing Instruction in English as a Foreign Language Education: A Scoping Review. International Journal of Chinese Education, 2025/2026, Vol. 9, Issue 1. The review examined 16 empirical studies published before December 2023. Article-numbered publication; conventional page range not assigned.
17. Investigating the Impact of ChatGPT as an AI Tool on ESL Writing. International Journal of Computer-Assisted Language Learning and Teaching, 2024, Vol. 14, Issue 1. The study analyzed data from 130 students and included a focus group of six students. Article-numbered publication; conventional page range not assigned.
18. The Influence of GenAI on Second Language Writing Instruction: A Systematic Review. System, 2026, Vol. 140, Article 104081. The review synthesized 79 empirical articles published from 2023 to 2025. Article-numbered publication; conventional page range not assigned.
19. Abudalfa, S. I., Barrot, J. S. Generative Artificial Intelligence for Automated Writing Evaluation: A Systematic Review of Trends, Efficacy, and Challenges. Assessing Writing, 2026, Vol. 68, Article 101041. Article-numbered publication; conventional page range not assigned.
20. Nunes, A., Cordeiro, C., Limpo, T., Castro, S. L. Effectiveness of Automated Writing Evaluation Systems in School Settings: A Systematic Review of Studies from 2000 to 2020. Journal of Computer Assisted Learning, 2022, Vol. 38, Issue 2, pp. 599–620.
21. Wilson, J., Czik, A. Automated Essay Evaluation Software in English Language Arts Classrooms: Effects on Teacher Feedback, Student Motivation, and Writing Quality. Computers & Education, 2016, Vol. 100, pp. 94–109.
22. Tan, S., Cho, Y. W., Xu, W. Exploring the Effects of Automated Written Corrective Feedback, Computer-Mediated Peer Feedback and Their Combination Mode on EFL Learner's Writing Performance. Interactive Learning Environments, 2023, Vol. 31, Issue 10, pp. 7276–7286.