Generative AI in EFL academic writing: a critical review and implications for higher education in Uzbekistan

Yusupova Nurjahon Shuxratovna
Student, Samarkand Oriental University
yusupovanurjahon08@icloud.com

Yokubjonova Nodirabegim Alisherovna
Teacher, Samarkand Oriental University
nodirabegimyokubjonova@gmail.com

FOR CITATION

Yusupova N.S., Yokubjonova N.A. Generative AI in EFL academic writing: a critical review and implications for higher education in Uzbekistan // Modern Science, Technology and Innovation: Proceedings of the International Scientific and Practical Conference (ZFTI–2026, 5 November 2026). – 2026. – P. 28–37. – https://doi.org/10.66073/ReFocus-Conf/051106

QR code — link to the DOI page

ABSTRACT

Generative artificial intelligence (GenAI) is increasingly used to support academic writing in English as a Foreign Language (EFL) education. This focused narrative review critically examines selected empirical studies, research syntheses, and institutional guidance published between 2023 and 2026. It distinguishes improvements in AI-assisted texts from evidence of lasting, independent writing competence and considers feedback, critical thinking, learner autonomy, and academic integrity. The discussion incorporates emerging research from Uzbekistan while acknowledging the limited generalizability of individual institutional studies. The synthesis suggests that AI feedback can support revision when students critically evaluate suggestions, verify sources, and retain responsibility for their writing. As a conceptual contribution, the article proposes a five-stage instructional framework comprising independent planning, targeted AI assistance, critical evaluation, student-led revision, and reflection. The framework has not been empirically tested. The article concludes by identifying priorities for responsible classroom implementation and future research in Uzbek higher education.

Keywords: generative artificial intelligence, EFL academic writing, AI-assisted feedback, learner autonomy, critical thinking, academic integrity, Uzbekistan.

АННОТАЦИЯ

В статье представлен целенаправленный нарративный обзор избранных эмпирических исследований, научных обзоров и рекомендаций организаций, опубликованных в 2023–2026 годах и посвящённых использованию генеративного искусственного интеллекта (GenAI) в академическом письме студентов, изучающих английский язык как иностранный (EFL). Разграничиваются улучшение текста при поддержке ИИ и развитие устойчивых навыков самостоятельного письма. Рассматриваются обратная связь, критическое мышление, автономия обучающихся и академическая добросовестность с учётом исследований, проводимых в Узбекистане. Анализ показывает, что помощь ИИ может быть полезной при условии критической оценки рекомендаций, проверки источников и сохранения ответственности студента за текст. Предлагается концептуальная пятиэтапная модель обучения: самостоятельное планирование, адресная помощь ИИ, критическая оценка, самостоятельное редактирование и рефлексия. Модель пока не прошла эмпирическую проверку. Определены направления её практического применения и дальнейших исследований в высшем образовании Узбекистана.

Ключевые слова: Генеративный искусственный интеллект, академическое письмо, студенты EFL, ChatGPT, преподавание английского языка, академическая добросовестность, высшее образование Узбекистана.

ANNOTATSIYA

Ushbu maqolada 2023–2026-yillarda chop etilgan tanlangan empirik tadqiqotlar, ilmiy sharhlar va institutsional tavsiyalar asosida generativ sun'iy intellektning (GenAI) ingliz tilini xorijiy til sifatida o'rganuvchi (EFL) talabalarning akademik yozuviga ta'siri narrativ sharh usulida tahlil qilinadi. Sun'iy intellekt yordamida tayyorlangan matn sifatining yaxshilanishi bilan mustaqil yozish kompetensiyasining barqaror rivojlanishi farqlanadi. Teskari aloqa, tanqidiy fikrlash, o'quvchi mustaqilligi va akademik halollik masalalari, jumladan, O'zbekistonda olib borilgan tadqiqotlar muhokama qilinadi. Tahlil sun'iy intellekt tavsiyalarini tanqidiy baholash, manbalarni tekshirish va matn uchun mas'uliyatni talabaning o'zida saqlash muhimligini ko'rsatadi. Maqolada mustaqil rejalashtirish, maqsadli AI yordami, tanqidiy baholash, talaba boshchiligidagi tahrir va refleksiyadan iborat besh bosqichli konseptual model taklif etiladi. Mazkur model hozircha empirik sinovdan o'tkazilmagan. O'zbekiston oliy ta'limi uchun amaliy tavsiyalar va keyingi tadqiqot yo'nalishlari bayon etiladi.

Kalit so'zlar: Generativ sun'iy intellekt, akademik yozuv, EFL talabalari, ChatGPT, ingliz tilini o'qitish, akademik halollik, O'zbekiston oliy ta'limi.

INTRODUCTION

Academic writing plays an important role in higher education because it requires students to communicate ideas clearly, construct logical arguments, evaluate evidence, and follow academic conventions. For students learning English as a Foreign Language (EFL), these tasks involve both developing language proficiency and learning how academic knowledge is presented in English.

Generative artificial intelligence (GenAI) has created new possibilities for supporting this process. Tools such as ChatGPT can suggest alternative expressions, explain grammatical problems, help students organize ideas, and provide feedback during revision. These functions may be valuable to learners who need additional opportunities to practise English beyond classroom instruction.

However, the ability of an AI system to generate fluent text does not necessarily mean that its use develops students' independent writing competence. A polished essay may reflect a student's own learning, extensive AI assistance, or a combination of the two. If students accept generated text without evaluating its accuracy or understanding its arguments, the final assignment may overstate their actual ability.

Research on ChatGPT and language education has expanded rapidly. Lo et al. (2024) reviewed 70 empirical studies of ChatGPT in ESL/EFL education and identified writing as a prominent application area. Their review also highlighted concerns about inaccurate information, privacy, and academic dishonesty. Li et al. (2025) examined 144 peer-reviewed empirical studies on generative AI in language learning and teaching, documenting the rapid development of this research field. These reviews provide a broad picture of research trends, but their findings should not be interpreted as proof that all forms of AI assistance improve long-term writing proficiency.

Research in Uzbekistan is also beginning to address this issue. A recent case study at Fergana State University examined AI use in academic writing among Uzbek EFL students, considering writing quality, cognitive independence, and technology acceptance (Qosimov et al., 2026). The emergence of such work means that the local research gap should not be described as a complete absence of evidence. Instead, there is a need to critically compare local and international findings and clarify how the available evidence can inform writing instruction.

A further challenge is that studies may measure different outcomes. Some investigate students' perceptions of AI, while others examine the quality of written texts or changes following an instructional intervention. Positive attitudes toward AI do not necessarily demonstrate improvements in writing, and better performance on an AI-assisted assignment does not necessarily indicate stronger independent writing ability. These outcomes must be distinguished when interpreting the literature.

The purpose of this article is to critically review selected research on GenAI-assisted EFL academic writing and identify implications for higher education in Uzbekistan.

The article addresses four research questions:

1. What benefits and limitations of GenAI-assisted writing are reported in the selected literature?
2. How do writing performance, student perceptions, and independent writing competence differ as research outcomes?
3. What pedagogical and ethical challenges should EFL teachers consider?
4. What recommendations and research priorities emerge for Uzbek higher education?

The article contributes a critical synthesis of selected international and Uzbek research and proposes a five-stage framework for integrating AI into writing instruction. It does not present original experimental data or claim that the proposed framework has already been validated.

2. LITERATURE REVIEW

2.1. AI-Assisted Feedback and Writing Performance

Writing feedback can help learners identify errors, improve clarity, and reconsider how ideas are organized. Traditional teacher feedback can be valuable, but it may not always be available immediately or individually. GenAI offers an additional source of feedback that students can access during different stages of writing.

Mahapatra (2024) investigated ChatGPT as a formative feedback tool in undergraduate ESL writing. The study reported improvements in students' academic writing following the intervention and examined their perceptions of the approach. The findings support the potential value of structured AI feedback in a particular instructional setting.

Similarly, Polakova and Ivenz (2024) investigated ChatGPT feedback among 110 university EFL students. Their mixed-methods study reported improvements in several aspects of writing performance. These findings suggest that AI feedback may assist learners with aspects of language and text development when it is incorporated into an instructional process.

The two intervention reports are useful but cannot be assumed to employ equivalent tasks, measures, or instructional conditions. Without a common measure of later unaided writing, gains in revised assignments should not be interpreted as demonstrated transfer to independent competence. Nevertheless, these studies do not establish that every AI tool, prompt, or classroom approach will produce the same result. Their findings relate to specific participants, writing tasks, and instructional conditions. Differences in proficiency, feedback design, and assessment criteria may affect outcomes.

A further issue is whether improvement in a revised text translates into independent learning. If AI identifies an error and supplies the correction, the student may submit a more accurate sentence without understanding the grammatical principle. Evidence of learning would be

stronger if the student could explain the correction and apply the principle in a new writing task.

Thus, the literature supports a qualified conclusion: AI feedback can assist writing performance, but its contribution to lasting competence needs to be evaluated separately.

2.2. What Systematic Reviews Reveal

Lo et al. (2024) provide a broad overview of empirical research on ChatGPT in ESL/EFL education. Their review included 70 studies published within approximately the first 18 months after ChatGPT became publicly available. Writing was the most frequently studied language skill in their sample, accounting for 29 studies, or 41.4%. The review also identified potential benefits, such as personalized learning support, alongside concerns about incorrect information, privacy, and academic dishonesty.

A key implication of this review is that the number of studies focusing on writing does not automatically indicate that the evidence is conclusive. Lo et al. noted the need for more rigorous and longer-term studies, including objective measures of learning outcomes.

Li et al. (2025) examined a broader set of 144 peer-reviewed empirical studies published in the developing field of generative AI in language learning and teaching. Their review illustrates the expansion of research beyond initial exploratory investigations toward more varied empirical approaches.

These reviews differ in their scope and publication periods. Lo et al. focus specifically on ChatGPT in ESL/EFL education during an early period, whereas Li et al. consider empirical generative AI research in language learning and teaching over a later period. Their findings should therefore be treated as complementary rather than directly interchangeable.

Taken together, they show that writing has received substantial attention, but important questions remain about how students learn from AI feedback, whether gains transfer to unaided writing, and how teachers can assess students' independent competence.

2.3. Critical Thinking and Learner Autonomy

Academic writing involves more than producing grammatically correct sentences. Students must formulate claims, select evidence, consider alternative perspectives, and decide whether their conclusions are justified.

Liu et al. (2025) reviewed 15 studies on the relationship between GenAI and critical thinking in EFL education. Their synthesis identified both potential benefits and risks. AI-supported activities may encourage analysis and problem-solving, but excessive dependence may also reduce opportunities for independent judgement.

This tension is central to learner autonomy. Autonomous writers take responsibility for their ideas, language choices, sources, and revisions. They may use external support, but they do not automatically accept every suggestion.

For example, a student may ask an AI tool to identify weaknesses in an argument. The educational benefit depends on what happens next. If the student evaluates the feedback and decides how to respond, the activity may support critical engagement. If the student simply asks the tool to rewrite the argument and submits the result, the opportunity for independent reasoning may be reduced.

The distinction is not between using AI and never using AI. It is between using AI as a resource for thinking and allowing it to replace the thinking required by the assignment.

https://doi.org/10.66073/ReFocus-Conf/051106

2.4. Evidence from Uzbekistan

The Uzbek higher education context deserves specific consideration because local research conditions, language backgrounds, teaching practices, and institutional expectations may differ from those in the settings represented in international studies.

Qosimov et al. (2026) examined AI-assisted academic writing at Fergana State University. According to the published article, the study involved 351 EFL students and used a mixed-methods design, including questionnaire data, comparative writing tasks, and interviews with a subgroup of participants. The authors reported improvements in some linguistic indicators in AI-assisted writing, while also identifying concerns involving argumentative claims, originality, and students' independent reasoning.

This study is relevant because it examines writing quality and cognitive independence within the same local educational context. Its focus suggests that AI assistance should not be evaluated exclusively through grammar and vocabulary. The quality of argumentation and the learner's role in producing the text also matter.

However, a single institutional case study cannot represent every university in Uzbekistan. Its results should be interpreted in relation to its sample, measures, research design, and setting. Further studies are needed to determine whether similar patterns occur across universities, proficiency levels, and disciplines.

The contribution of the present article is therefore not to claim that Uzbek students have never been studied. Rather, it is to connect the emerging local evidence with international research and develop a pedagogical interpretation that distinguishes immediate writing performance from independent learning.

2.5. Academic Integrity and Responsible AI Use

AI-assisted writing also raises questions about originality, authorship, source reliability, and transparency. These issues cannot be resolved by assuming that every use of AI is misconduct or that every use is educationally beneficial.

There is an important difference between asking AI to explain a grammatical error and asking it to generate an entire assignment that is supposed to demonstrate independent writing. Whether a particular use is acceptable depends on the assignment's objectives and the rules established by the institution or instructor.

UNESCO's guidance on generative AI in education and research recommends a human-centred approach that considers educational, ethical, and governance issues (Miao & Holmes, 2023). The Committee on Publication Ethics (COPE, 2023) states that AI tools cannot be listed as authors because they cannot take responsibility for the work. Human authors remain accountable for submitted content.

Source verification is equally important. An AI-generated reference should not be included in an academic paper until the original publication has been located and checked. Students and researchers should confirm that the source exists, that the bibliographic information is accurate, and that the source supports the claim for which it is cited.

Responsible AI use therefore requires clear institutional guidance, explicit disclosure rules where applicable, and assessment methods that reward understanding rather than the appearance of a polished text.

https://doi.org/10.66073/ReFocus-Conf/051106

3. METHODOLOGY

3.1. Review Design

This article adopts a focused narrative review of selected literature on GenAI-assisted EFL academic writing. A narrative approach was chosen to compare different types of evidence and develop pedagogical implications rather than calculate a pooled effect size.

The review includes empirical research on AI-assisted writing feedback, systematic reviews of GenAI in language education, research on critical thinking, a local Uzbek case study, and institutional guidance on responsible AI use.

3.2. Source Identification and Selection

The selected publications were located through targeted searches of scholarly publisher pages and research-indexing resources using terms such as “generative AI,” “ChatGPT,” “EFL,” “academic writing,” “writing feedback,” “critical thinking,” “Uzbekistan,” and “higher education.”

Priority was given to publications from 2023 to 2026 because the article focuses on the recent development of generative AI. Earlier or non-empirical institutional guidance was retained where it was directly relevant to academic integrity and responsible implementation.

Sources were included on the basis of their relevance to at least one of the following areas:

1. Empirical outcomes of AI-assisted writing or feedback.
2. Systematic reviews of GenAI in language learning.
3. Critical thinking and learner autonomy.
4. Academic integrity and responsible AI use.
5. Research conducted in Uzbekistan or directly relevant to Uzbek higher education.

The review prioritizes peer-reviewed journal publications and official guidance. Because a complete search log, screening count, and formal risk-of-bias assessment were not documented, the selection process cannot be presented as systematic or exhaustive. Conference papers and other local publications may provide useful context but should not be treated as equivalent in evidential strength to rigorous empirical journal studies.

3.3. Analytical Framework

The selected literature was compared across four analytical dimensions:

- Writing performance: grammar, vocabulary, organization, clarity, and argumentation.
- Learning processes: revision, feedback evaluation, critical thinking, and independent writing.
- Research design: participant groups, instructional context, data sources, and the outcomes measured
- Ethical and contextual issues: source reliability, authorship, transparency, and relevance to Uzbekistan.

The analysis distinguished reported improvements in AI-assisted texts from evidence of learning that can be demonstrated independently. It also considered differences between studies that measured students’ perceptions and those that examined written performance.

https://doi.org/10.66073/ReFocus-Conf/051106

3.4. Scope and Limitations of the Method

This is a selective narrative review, not a systematic review. The precise search dates, database coverage, number of screened records, and exclusion decisions were not documented; these omissions limit reproducibility. It does not claim to have searched every relevant database, followed a complete PRISMA screening protocol, or assessed every eligible publication. No new participant data were collected, and no statistical effect size was calculated.

Consequently, the article's contribution is an interpretive synthesis of selected literature. Its conclusions should not be interpreted as an exhaustive estimate of the effectiveness of GenAI in EFL writing. A future systematic review could build on this work by preregistering a protocol, searching databases such as Scopus and Web of Science, defining explicit eligibility criteria, and formally assessing the quality of included studies.

4. CRITICAL DISCUSSION

4.1. Immediate Performance Is Not the Same as Independent Competence

A recurring issue across the literature is the difference between the quality of an AI-assisted text and a student's ability to write independently. AI feedback may help students correct grammar, improve word choice, and reorganize sentences. Yet these improvements do not automatically demonstrate that learners have acquired the underlying knowledge.

This distinction helps explain why studies should measure multiple outcomes. A final essay can show whether a text meets a set of writing criteria, but it may not reveal how much assistance the student received or whether the student can apply the same skills later.

Assessment can address this limitation by including initial drafts, revision explanations, source-verification tasks, and subsequent writing completed without AI assistance. These measures do not eliminate every difficulty, but they provide a more complete view of learning than the final text alone.

4.2. Why the Research Findings Must Be Interpreted Carefully

The selected studies differ in design and purpose. Empirical intervention studies examine changes under particular instructional conditions, while systematic reviews synthesize broader bodies of research. Local case studies provide context-specific evidence but cannot automatically be generalized to all institutions.

For example, positive results from a guided feedback intervention do not prove that unrestricted AI-generated essay writing will produce the same outcome. Likewise, students' positive attitudes toward a tool cannot establish that their independent writing has improved.

A careful interpretation therefore separates three questions:

1. Does AI help improve a particular piece of writing?
2. Does AI-supported instruction help students learn to write better independently?
3. Under what conditions do benefits occur without unacceptable costs to critical thinking or academic integrity?

These questions are related, but they are not interchangeable. Future studies should specify which outcome they investigate and use methods appropriate to that outcome.

4.3. Implications for Uzbek Higher Education

The emerging Uzbek research makes it possible to move beyond general statements that local evidence is absent. Instead, the priority is to build a stronger and more diverse evidence base.

Universities should examine how AI use varies across English proficiency levels, academic disciplines, assignment types, and teaching approaches. Researchers should also investigate whether students understand the feedback they receive and can transfer improvements to independent tasks.

In classroom practice, AI should complement rather than displace instruction in argumentation, source evaluation, paraphrasing, citation, and revision. These skills remain essential regardless of whether students use AI tools.

Institutions should also distinguish between acceptable assistance and inappropriate delegation. Rules should be communicated before assignments begin, and students should be taught how to comply with them. This is more constructive than relying solely on detection after submission.

4.4. A Five-Stage Framework for Responsible AI-Assisted Writing

Based on the reviewed literature, this article proposes a five-stage pedagogical framework. It is an author-developed conceptual proposal and has not yet been empirically validated.

Stage 1: Independent planning. Students identify the topic, formulate an initial claim, and prepare a basic outline before using AI. An initial outline or short unaided draft provides a record of their own thinking.

Stage 2: Targeted AI assistance. Students use AI to address a specific difficulty, such as grammatical accuracy, paragraph coherence, or the clarity of an argument. The task should define the type of assistance permitted; students may record the question asked and the response received.

Stage 3: Critical evaluation. Students assess whether the AI suggestions are accurate, relevant, and consistent with their intended meaning. Factual claims and references are checked against credible sources; students can briefly justify accepting or rejecting suggestions.

Stage 4: Student-led revision. Students decide which suggestions to accept, adapt, or reject. They remain responsible for the final argument, evidence, wording, and citations, with a revision log where appropriate.

Stage 5: Reflection and transparency. Students briefly describe how AI was used, which changes were made, and what they learned. A brief reflection or AI-use statement can document the process; any disclosure should follow the relevant institutional or journal policy.

The framework connects AI use with the broader aims of writing instruction. It seeks to preserve student agency while allowing learners to benefit from timely feedback. Its effectiveness must be tested through future classroom research.

5. PEDAGOGICAL RECOMMENDATIONS

The critical review leads to six recommendations for EFL writing instruction in higher education.

First, establish transparent AI-use policies. Students should know which forms of assistance are permitted for each assignment and when independent writing is required.

Second, teach AI and information literacy. Learners should be able to identify possible inaccuracies, verify references, and distinguish well-written language from reliable evidence.

Third, maintain independent writing practice. Students need opportunities to write without AI so that teachers can evaluate their individual competence and learners can strengthen their skills.

Fourth, assess the writing process. Initial drafts, revision notes, source checks, and short oral or written explanations can complement assessment of final submissions.

Fifth, develop teachers' professional competence. Teacher training should address AI-supported feedback, assessment design, ethical use, and ways to help students evaluate AI suggestions.

Sixth, encourage local research collaboration. Uzbek universities could develop shared research projects using comparable writing tasks, clear assessment rubrics, and transparent research methods. This would make it easier to compare findings across institutions without assuming that a single study represents the entire country.

6. LIMITATIONS AND FUTURE RESEARCH

This article has several limitations. First, it uses a focused narrative review rather than a systematic search and formal quality appraisal. The selected publications are relevant to the research questions but do not represent every study on AI-assisted writing.

Second, the review draws on studies with different research designs and outcome measures. Their results cannot be combined into a single estimate of effectiveness without a more systematic analysis.

Third, the proposed five-stage framework is conceptual. It has not been tested with students or teachers in Uzbekistan, and its educational effectiveness remains an empirical question.

Future research should compare structured AI-assisted feedback with teacher feedback and independent writing practice. Such studies should evaluate both the quality of assisted texts and students' ability to write independently after the intervention.

Research should also examine how students evaluate AI feedback, verify references, and explain their revision decisions. Studies across multiple Uzbek universities would help determine how local context, proficiency, and instructional design influence outcomes.

A systematic review with a predefined search protocol and transparent inclusion criteria would be a useful next step. It could assess the strength of current evidence and identify where new empirical research is most needed.

7. CONCLUSION

GenAI offers useful possibilities for supporting EFL academic writing, particularly through rapid feedback, language assistance, and revision support. Selected empirical studies report improvements in aspects of writing performance, while systematic reviews highlight continuing concerns about reliability, independent learning, and academic integrity.

The key distinction is between improving a particular text and developing the ability to write independently. These outcomes should not be treated as identical. Students need to evaluate AI suggestions, verify information, and remain responsible for their arguments and final texts.

For Uzbek higher education, the emerging local research provides a starting point for more context-sensitive investigation. Universities should develop clear AI-use policies, teach AI liter-

acy, preserve independent writing practice, and assess students' learning processes alongside final submissions.

The five-stage framework proposed in this article offers a practical direction for teaching and future research, but its effectiveness requires empirical testing. Ultimately, AI should be used to support the development of writers who can think critically, communicate accurately, and take responsibility for their academic work.

REFERENCES

1. Committee on Publication Ethics. (2023). *Authorship and AI tools*.
2. Li, B., Tan, Y. L., Wang, C., & Lowell, V. (2025). Two years of innovation: A systematic review of empirical generative AI research in language learning and teaching. *Computers and Education: Artificial Intelligence, 9*, Article 100445.
3. Liu, J., Sihes, A. J. B., & Lu, Y. (2025). How do generative artificial intelligence (AI) tools and large language models (LLMs) influence language learners' critical thinking in EFL education? A systematic review. *Smart Learning Environments, 12*, Article 48.
4. Lo, C. K., Yu, P. L. H., Xu, S., Ng, D. T. K., & Jong, M. S.-Y. (2024). Exploring the application of ChatGPT in ESL/EFL education and related research issues: A systematic review of empirical studies. *Smart Learning Environments, 11*, Article 50.
5. Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments, 11*, Article 9.
6. Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO.
7. Polakova, P., & Ivenz, P. (2024). The impact of ChatGPT feedback on the development of EFL students' writing skills. *Cogent Education, 11*(1), Article 2410101.
8. Qosimov, A., Ne'matov, A., & Shokirova Sharifovna, M. (2026). The integration of artificial intelligence in the academic writing of EFL students in higher education: A case study at Fergana State University. *Cogent Education*.