

AI IN ASSESSING ENGLISH ACADEMIC WRITING: OPPORTUNITIES, CHALLENGES, AND FUTURE DIRECTIONS

Amanullayeva Kamola Muminovna

PhD, Acting Professor of the Samarkand

Branch of ISFT Institute.

miskamolka@yahoo.com,

Saloxiddinova Gulmira Sirojiddin qizi

Master's student in Foreign Language and Literature

at the Samarkand Branch of ISFT Institute.

gulmirasalohiddinova00@gmail.com

Abstract: *The rapid integration of artificial intelligence into educational settings has opened new avenues for writing assessment, particularly in the context of English academic writing. This article examines the role of AI-based tools in evaluating EFL and ESL learners' academic writing, considering dimensions such as automated feedback, metadiscourse analysis, criteria-based assessment, and the pedagogical implications for writing instruction. Drawing on a synthesis of current scholarship in applied linguistics, language pedagogy, and educational technology, the study argues that AI assessment tools, when thoughtfully implemented, can complement teacher feedback, support learner autonomy, and raise the quality of academic writing outcomes. The article also identifies key challenges, including issues of transparency, cultural bias, and the need for teacher mediation, and concludes with recommendations for integrating AI assessment into writing curricula.*

Keywords: *AI assessment, English academic writing, EFL writing, automated feedback, criteria-based assessment, metadiscourse.*

1. Introduction

Academic writing in English has long been recognized as one of the most demanding competencies for students in higher education, particularly for those whose first language is not English. With the growing internationalization of universities and the rising demand for research published in English, the ability to produce well-structured, coherent, and rhetorically appropriate academic texts has become indispensable. Assessment of such writing, however, poses considerable challenges for educators, especially in contexts where class sizes are large, feedback cycles are slow, and assessment criteria are inconsistently applied (Khodjieva, Ardiana, & Rum, 2025). Against this backdrop, artificial intelligence (AI) has emerged as a promising solution for scaling and standardizing writing assessment.

The spread of English as a medium of instruction across Central Asia and beyond has intensified interest in evaluating academic writing quality (Tursunov, Nematov, & Mamayoqubova, 2021). Researchers have noted that standardized testing instruments, such as the IELTS, are widely used to benchmark academic language proficiency, yet their relationship to actual academic writing performance remains complex (Fayziev *et al.*, 2020). AI-driven assessment tools offer an alternative or complementary approach that can provide immediate, detailed, and scalable feedback on writing quality.

This article explores how AI is currently being used to assess English academic writing, reviewing the theoretical foundations of writing assessment, the affordances and limitations of AI tools, and the pedagogical implications for EFL and ESL contexts. The discussion draws on a range of recent scholarship in applied linguistics, language teaching methodology, and educational technology, with particular attention to research from Uzbekistan and broader Central Asian contexts where English academic writing instruction is undergoing significant transformation.

2. Theoretical Background

2.1 The Nature of Academic Writing Assessment

Writing assessment in academic contexts involves multiple, interrelated dimensions. These include textual features such as coherence, lexical complexity, grammatical accuracy, and argumentation, as well as rhetorical dimensions that reflect the writer's awareness of audience, purpose, and disciplinary conventions. The assessment of academic writing has traditionally been characterized by subjectivity, and considerable research effort has gone into developing robust, transparent criteria for evaluating academic texts (Khodjieva, 2021).

Criteria-based assessment has gained significant traction in Central Asian higher education. Khodjieva (2022) discusses the difficulties encountered in implementing criteria-based assessment in Uzbekistan, noting that while such approaches offer greater transparency and objectivity, teachers and students alike face challenges in interpreting and applying rubrics consistently. Building on this, Zumrad (2023) explores the principles of assessing critical thinking through criteria-based frameworks in Uzbekistan, arguing that well-designed rubrics need to capture not only surface-level linguistic features but also higher-order cognitive processes such as analysis, synthesis, and evaluation.

The metadiscourse dimension of academic writing has also attracted growing scholarly attention. Metadiscourse refers to the linguistic resources writers use to organize their texts, engage their readers, and signal their stance toward the content (Ardiana, Hamid, & Khodjieva, 2026). A comparative study of Indonesian and Uzbekistani EFL students' research articles found meaningful differences in the deployment of interactional metadiscourse markers, underscoring the importance of explicit instruction in these conventions for non-native academic writers.

2.2 From Manual to Automated Assessment: The Role of Technology

Technological mediation in writing assessment is not new; computer-assisted language learning (CALL) and automated essay scoring (AES) systems have existed for decades. What is new is the depth and sophistication of current AI tools, which can now analyze rhetorical structure, pragmatic-linguistic features, and even disciplinary discourse conventions. A bibliometric mapping analysis of English as a Foreign Language (EFL) writing assessment research from 2014 to 2024 shows a clear thematic evolution: early work focused on analytic rubrics and rater reliability, whereas more recent studies increasingly foreground feedback automation, artificial intelligence, and learner-centered approaches (Khodjieva, Ardiana, & Rum, 2025). This trajectory reflects a broader shift in the field toward formative, technology-enhanced assessment practices.

Alongside this shift, there has been growing recognition that effective writing instruction requires diverse feedback modalities. Absalamov and Khujakulov (2024) argue that innovative, activity-based approaches to writing development can enhance student engagement and writing quality when coupled with clear feedback mechanisms. The challenge for AI assessment is to replicate and, where possible, extend the depth of feedback that skilled human teachers provide.

3. AI Assessment Tools in English Academic Writing

3.1 Automated Feedback and Its Pedagogical Value

Automated writing evaluation (AWE) systems analyze student texts using natural language processing (NLP) algorithms that score essays on dimensions such as grammar, vocabulary range, organization, and cohesion. From a pedagogical standpoint, the value of such tools lies primarily in their capacity to deliver immediate, consistent, and individualized feedback at scale. For EFL learners, who often struggle with both the linguistic and rhetorical demands of academic writing, timely feedback is especially critical

(Iskandarova, 2026a). Delayed or infrequent feedback remains one of the major obstacles to writing development in large university classes, and AI tools offer a practical solution to this problem.

Research on teaching approaches and methodologies consistently emphasizes the importance of aligning instructional strategies with assessment practices (Hasanova, Abduazizov, & Khujakulov, 2021). When AI feedback is integrated into the writing curriculum as a formative tool rather than a summative judgment, it can support a recursive, process-oriented approach to writing. Students receive feedback on drafts, revise their texts, and receive further feedback a cycle that mirrors best practices in writing pedagogy. Iskandarova (2026b) argues that interactive, technology-enhanced methods in English language teaching can meaningfully improve learner outcomes when instructors take an active mediating role.

3.2 Lexical and Grammatical Dimensions of AI Assessment

One of the most technically mature areas of AI writing assessment concerns lexical and grammatical analysis. Contemporary NLP tools can identify a wide range of lexical features, including lexical diversity, the use of academic vocabulary, and the frequency of domain-specific terminology. Phraseological competence the ability to use formulaic language and fixed expressions appropriately is particularly relevant in academic writing (Nasirova, 2024). Nasirova notes that contextual analysis of phraseostylistic meaning requires sensitivity to register and discourse context, dimensions that are increasingly within the reach of sophisticated AI tools.

The assessment of grammar and syntactic complexity in academic texts is another area where AI tools have demonstrated considerable reliability. Research on relative pronouns, for instance, illustrates the complexity of syntactic choices facing EFL academic writers (Nematov, 2020). Whether a writer deploys relative clauses correctly and effectively is a marker of both grammatical competence and discourse sophistication precisely the kind of

feature that well-trained AI assessors can evaluate. Similarly, research on phraseological units in English teaching underscores the need for assessment instruments that capture both form and function (*Nasirova, 2024*).

Lexical differences between varieties of English also have implications for AI assessment. *Nosirova and Kurbanova (2020)* highlight that American and British English differ not only in vocabulary but also in certain grammatical conventions, suggesting that AI tools need to be calibrated to accommodate the target variety of English and the specific disciplinary norms of the academic community in question.

3.3 Metadiscourse and Rhetorical Analysis

Beyond grammar and lexis, academic writing assessment must grapple with the rhetorical organization of texts. Metadiscourse analysis provides a systematic framework for examining how writers navigate between their own voice, the voices of other scholars, and the expectations of their readers (*Ardiana, Hamid, & Khodjieva, 2026*). The finding that EFL student writers from different national backgrounds deploy interactional metadiscourse differently has important implications for AI assessment: tools designed for one cultural or academic context may misread or under-evaluate the rhetorical strategies typical of writers from other backgrounds.

Typological awareness is also relevant here. *Saloxiddinov (2023)* notes that typological comparison examining how different languages encode similar meanings through different structural means can illuminate the challenges EFL writers face when transferring rhetorical conventions across languages. An AI assessment tool sensitive to these cross-linguistic influences would be better positioned to distinguish language-transfer errors from principled rhetorical choices.

The use of figurative language in academic writing presents another dimension for AI assessment. *Haydarova and Azzamova (2026)* demonstrate through a comparative and functional analysis of literary works by Mark Twain

and Khudoyberdi Tokhtaboev that figurative devices are integral to the poetic and rhetorical texture of written texts. While academic writing typically employs restrained figurative language, some disciplines and genres permit richer metaphorical expression, and AI assessment tools need to be calibrated accordingly.

4. Challenges in AI-Based Writing Assessment

4.1 Cultural Sensitivity and Linguistic Diversity

A significant challenge for AI-based writing assessment is the risk of cultural and linguistic bias. Most commercial AI tools have been trained primarily on corpora of English academic writing produced by native or near-native speakers in North American or British educational contexts. When such tools are applied to texts written by EFL learners from Central Asia, East Asia, or the Arab world, they may penalize culturally-specific rhetorical conventions that are perfectly acceptable within the student's own academic tradition.

Research on sociolinguistic variation in written language underscores the complexity of this challenge. Hazaymeh (2021) demonstrates that the use of numerals in rhetorical and phraseological expressions varies systematically across languages and cultures, illustrating how features that appear as errors from a monolingual English perspective may in fact reflect systematic cross-linguistic influence. An AI assessment tool that penalizes such features indiscriminately risks being not only inaccurate but also educationally counterproductive. The importance of attending to students with diverse learning profiles in writing classrooms is well-documented. Iskandarova (2026a) notes that teachers regularly face the challenge of supporting students with very different levels of linguistic preparation, background knowledge, and learning styles. AI assessment tools that generate undifferentiated feedback may fail to address the specific needs of individual learners from particular cultural and linguistic backgrounds.

4.2 Transparency, Trust, and Teacher Mediation

A recurring concern in the literature on AI assessment is the perceived lack of transparency in how automated systems arrive at their judgments. Students and teachers alike are often uncertain about the criteria and algorithms underpinning AI feedback, which can undermine trust and reduce the uptake of AI-generated suggestions. Language planning and policy frameworks stress the importance of clarity in assessment goals and procedures as a precondition for pedagogical effectiveness (*Khodjieva, 2021*). AI assessment tools need to be accompanied by clear documentation of their criteria, limitations, and the theoretical frameworks that inform their evaluations.

Teacher mediation emerges as a crucial factor in making AI writing assessment effective. Research consistently shows that students benefit most from assessment feedback when teachers help them interpret and act on the feedback they receive. *Hasanova et al. (2021)* emphasize the importance of distinguishing between different pedagogical roles approaches, methods, procedures, techniques, styles, and strategies noting that effective writing instruction requires teachers to make informed choices among these options. Similarly, integrating AI assessment into writing curricula requires teachers to develop new competencies: understanding what AI tools can and cannot assess, communicating AI feedback to students in pedagogically meaningful ways, and supplementing automated feedback with their own professional judgment.

4.3 Standardized Testing and AI Assessment Alignment

The relationship between AI writing assessment and high-stakes standardized testing is a further area requiring careful attention. IELTS and similar examinations remain powerful gatekeeping instruments in academic contexts across Asia and Central Asia (*Fayziev et al., 2020*). AI tools that are not calibrated to the specific writing task types, assessment criteria, and band descriptors of these examinations may provide feedback that is either irrelevant

to exam preparation or, worse, that leads students to develop writing habits that are penalized in high-stakes contexts.

This issue points to a broader need for alignment between AI assessment instruments and the established standards and frameworks used in EFL/ESL educational contexts. The distinction between ESL and EFL contexts has important implications for writing assessment design, since the communicative needs, available linguistic input, and motivational profiles of learners in these contexts differ substantially (*Tursunov, Nematov, & Mamayoqubova, 2021*). AI tools developed primarily for ESL contexts may need significant adaptation to serve EFL learners effectively.

5. Implications for Writing Instruction and Curriculum Design

The pedagogical implications of AI writing assessment are far-reaching. At the most basic level, the availability of scalable, immediate feedback enables a shift from product-oriented to process-oriented writing instruction a shift that has long been advocated in writing pedagogy but has been difficult to implement at scale due to resource constraints. Kudratova and Ochilova (2022) demonstrate that engaging instructional activities, including games and collaborative tasks, can enhance the learning environment in which writing skills develop; AI feedback can similarly make the feedback process more engaging and iterative.

For writing at the primary and early secondary levels, the integration of technology in language learning has shown promising results. Anorboyeva and Azzamova (2022) document modern approaches to teaching English in primary schools, noting that innovative, activity-based methods can be particularly effective when students are developing foundational literacy skills. While AI writing assessment is currently most relevant at the tertiary level, the principles underlying effective formative feedback immediacy, specificity, and actionability are applicable across educational levels.

The role of listening skills in supporting academic writing development should not be overlooked. Nematov (2020) emphasizes that listening plays a foundational role in language acquisition, and that students who are exposed to extensive spoken academic English develop a more intuitive feel for the register and rhetorical conventions of academic writing. AI assessment tools that analyze only written text may miss this dimension; integrating AI with broader multimodal assessment frameworks could provide a more complete picture of a learner's academic language competence.

Critical thinking skills are also integral to strong academic writing, and AI assessment needs to engage with this dimension if it is to provide genuinely useful feedback. Iskandarova (2025) argues that the development of critical thinking among students is fundamental to academic success, and that teaching approaches that foreground analysis, reflection, and argumentation produce better writers. AI tools that assess only surface-level features grammar, spelling, and basic coherence may inadvertently reinforce a surface-oriented approach to writing, at the expense of deeper rhetorical and intellectual engagement.

6. Discussion

The foregoing analysis suggests that AI holds genuine promise for the assessment of English academic writing, but that its effective deployment requires sustained attention to theoretical, pedagogical, and contextual factors. The strengths of AI assessment scale, consistency, immediacy, and the ability to analyze complex textual features are most fully realized when AI tools are embedded within coherent pedagogical frameworks and supported by knowledgeable teachers.

The concept of terminology and vocabulary in academic discourse is particularly relevant to AI assessment design. Azzamova (2020) argues that terminology has a multi-dimensional character, encompassing both cognitive and communicative functions. AI tools that assess academic writing need to

capture not only whether students use technical vocabulary correctly but also whether they deploy it in contextually appropriate ways a distinction that requires sophisticated semantic and pragmatic analysis.

Proverbs and formulaic expressions, while less central to prototypical academic writing, appear in certain genres and disciplines, and their assessment raises interesting questions about the boundaries of academic register. Ochilova and Nasrullaev (2018) trace the historical development of proverbs and their contemporary usage, while Ochilova (2014) examines the semantic structures of phraseological units in English. This body of work illustrates the depth and complexity of the lexical knowledge required for full academic writing competence knowledge that AI tools are increasingly, though not yet fully, able to assess.

The development of AI assessment tools for specific national and regional contexts is an important direction for future research. Akhmedova (2024) demonstrates how algorithmic text-processing techniques specifically the Boyer-Moore-Horspool algorithm can be adapted for Uzbek text, illustrating that the computational challenges of AI assessment are not insuperable even for non-English language contexts. Adapting similar approaches to the specific features of Uzbek EFL learners' academic writing in English could yield tools that are both linguistically sophisticated and culturally appropriate.

7. Conclusion

This article has examined the role of AI in assessing English academic writing from multiple perspectives: theoretical, methodological, pedagogical, and contextual. The evidence suggests that AI assessment tools, thoughtfully designed and carefully implemented, can meaningfully enhance the feedback that EFL/ESL learners receive on their academic writing. However, realizing this potential requires addressing several significant challenges, including cultural and linguistic bias in training data, the need for teacher mediation, and

the alignment of AI tools with established assessment frameworks and standards.

The growing body of research on EFL writing, assessment, and language pedagogy from Central Asian and other non-Anglophone contexts provides a rich empirical and theoretical foundation for the development of more culturally sensitive and pedagogically effective AI assessment tools. Researchers and practitioners in these contexts are well-positioned to contribute to an international conversation about what constitutes quality academic writing in English and how AI can best support its development and assessment.

Future research should focus on longitudinal studies examining the impact of AI writing feedback on student writing development over extended periods; comparative studies of AI and human assessment reliability across different writing genres, disciplines, and cultural contexts; and participatory design studies that involve EFL teachers and learners in the development and refinement of AI assessment tools. Such research will be essential for ensuring that AI assessment contributes positively to writing education and does not simply replicate or amplify existing inequalities.

REFERENCES:

1. Absalamov, K., & Khujakulov, R. (2024). Enhancing writing skills in school: Innovative activities for success. Conference proceedings: Fostering your research spirit, 5, 192–194.
2. Akhmedova, H. (2024). Organization of word search in Uzbek texts based on Boyer-Moore-Horspool algorithm. Al-Farg'oni avlodlari, 1, 196–201.
3. Anorboyeva, G., & Azzamova, N. (2022). Modern approaches and innovations in teaching English in primary schools. Scientific Progress, 3(2), 656–658.

4. Ardiana, Hamid, S., & Khodjieva, Z. (2026). Indonesian and Uzbekistan EFL students' research articles: Interactional metadiscourse. *Twist International Multidisciplinary Journal*, 20(1), 84–93.
5. Azzamova, N. R. (2020). The multi-dimensional character of terminology. *Colloquium-journal*, 47–48.
6. Fayziev, H., Nematov, B. S., Yakubov, F. J., et al. (2020). A critical review of the IELTS impact on academic performance, the case of Asian students at English speaking universities. *JournalNX*, 6(6), 752–756.
7. Hasanova, N., Abduazizov, B., & Khujakulov, R. (2021). The main differences between teaching approaches, methods, procedures, techniques, styles and strategies. *JournalNX: A Multidisciplinary Peer Reviewed Journal*, 7(02), 371–375.
8. Haydarova, N., & Azzamova, N. (2026). Figurative language as a poetic system in the works of Mark Twain and Khudoyberdi Tokhtaboev: A comparative and functional analysis. *Asian Journal of Scientific Research and Innovations*, 1(1), 94–96.
9. Hazaymeh, O. (2021). A sociolinguistic study of the use of cardinal and ordinal numbers in rhetorical, phraseological expressions in Jordanian spoken Arabic. *SKASE Journal of Theoretical Linguistics*, 18(2).
10. Iskandarova, L. B. (2025). Важность формирования критического мышления у студентов [The importance of developing critical thinking in students]. *International Journal of Advanced Research in Education, Technology and Management*, 2.
11. Iskandarova, L. B. (2026a). Challenges teachers face while coping with different types of learners in a classroom. *International Journal of Artificial Intelligence*, 6(02), 250–253.
12. Iskandarova, L. B. (2026b). New methods of teaching English as a second language (ESL). *International Journal of Artificial Intelligence*, 6(02), 250–253

- 13.Khodjiev, Z. M. (2021). Language planning and policy proposal, goals and objectives. Emergent: Journal of Educational Discoveries and Lifelong Learning (EJEDL), 2.
- 14.Khodjiev, Z. M. (2022). Difficulties in criteria-based assessment. Ijtimoiy fanlarda innovasiya onlayn ilmiy jurnali, 2(6), 28–31.
- 15.Khodjiev, Z., Ardiana, A., & Rum, E. P. (2025). From feedback to artificial intelligence: A bibliometric mapping analysis of the thematic evolution of EFL writing assessment research trends (2014–2024). International Journal of Learning, Teaching and Educational Research, 24(9).
- 16.Kudratova, S. S., & Ochilova, N. F. (2022). The usage of games in teaching. Архив научных исследований, 2(1).
- 17.Nasirova, R. (2024). Linguistic principles of the contextual analysis of the phraseostylistic meaning of fixed expressions based on English literary material. Зарубежная лингвистика и лингводидактика, 4(1), 72–76.
- 18.Nematov, B. M. (2020). Lexical and grammatical features of relative pronouns. Foreign Linguistics and Linguodidactics, 4(2/S), 77–80.
- 19.Nematov, B. M. (2020). The importance of listening to learn English language. PalArch's Journal of Archaeology of Egypt/Egyptology, 17(7), 8510–8520.
- 20.Nosirova, R. A., & Kurbanova, M. K. (2020). Lexical differences between American and British English. UzACADEMIA, 1, 61–65.
- 21.Ochilova, N. F. (2014). Semantic structures of English phraseological units and proverbs with proper names. Samarkand State Institute of Foreign Languages.
- 22.Ochilova, N., & Nasrullaev, J. (2018). The history of ancient proverbs and their modern usage. Bridge to Science: Research Works, 109.
- 23.Saloxiddinov, M. (2023). The history of linguistic typology. Универсальная индексная библиотека Евразийского журнала академических исследований, 3.

24. Tursunov, G. B., Nematov, B. M., & Mamayovqubova, S. O. (2021). The ESL and EFL in the world today. *ACADEMICIA: An International Multidisciplinary Research Journal*, 11(6).
25. Zumrad, K. (2023). Criterion-based assessment and principles of assessing critical thinking in higher education in Uzbekistan. *IJARETM*, 4(2), 357–363.
26. Amanullayeva Kamola Muminovna. The role of translation in science // *Texas Journal of Philology, Culture and History*. – 2023. – P. 60–62.
27. Amanullayeva Kamola Muminovna. The concept of translation and characteristics of features // *Information Horizons: American Journal of Library and Information Science*. – 2024. – P. 63–66.
28. Nasrullayeva Tozagul Suxrobovna, Amanullayeva Kamola Muminovna. Peculiarities of translation of technical terms, concepts and rendering articles in technical translation // *American Journal of Social and Humanitarian Research*. – 2023. – C. 29–33.
29. Kubayeva Nafisa, Amanullayeva Kamola. The problem of teaching students lexical and phraseological features in translation studies of phrasal verbs in English and Uzbek languages // *Eurasian Journal of Academic Research*. – P. 39–42. – <https://in-academy.uz/index.php/ejar/article/view/37844>
30. Алижонов О.А., Амануллаева К.М., Аскарлова М.Б. Использование термина “концепт” в лингвокультурологии // *Вестник магистратуры*. – 2024. – № 12-2 (159). – С. 76–78.
31. Амануллаева К., Алижонов О. Роль лингвокультурологии в лингвистике // *Ijtimoiy-gumanitar va tabiiy fanlardagi an'anaviy va zamonaviy yondashuvlar dialektikasi*. Халқаро илмий-амалий анжуман. – 27-fevral, 2025. – С. 324–325.
32. Амануллаева Камола Муминовна. О художественном концепте и в целом // *Россия и Узбекистан. Международные образовательные, научные и социально-культурные технологии: векторы развития*. – Челябинск–Ташкент–Бухара–Самарканд, 2020. – С. 271–272.