Journal of English Language Teaching and Applied Linguistics
Volume 01, Issue 01, 43-53.
e_ISSN: 3110-8822
https://e-journal.gomit.id/index.php/jaiela/oai
DOI: 10.58723/jaiela.v1i1.56

Harnessing Generative AI to Support EFL Academic Writing: A Systematic Review of English Education Students

Desti Angraini
UIN Sulthan Thaha Saifuddin,
INDONESIA

*** Corresponding author:**
Desti Angraini, UIN Sulthan Thaha Saifuddin, INDONESIA. ✉ desti.angraini@uinjambi.ac.id

Article Info

Article history:

Received: July 4, 2025
Revised: Agustus 10, 2025
Accepted: Agustus 11, 2025

Keywords:

Academic Writing
Artificial Intelligence in Language Learning
English Education Students
Generative AI

Abstract

Background of Study: Generative Artificial Intelligence (AI) has increasingly influenced higher education, offering innovative possibilities for English language teaching, particularly in academic writing. While its adoption is expanding, systematic evidence on its roles, benefits, and limitations for English education students remains limited.
Aims and Scope of Paper: This paper aims to systematically review empirical studies on the integration of generative AI into English education at the tertiary level. The focus is on identifying its pedagogical applications, impacts on learning, and challenges in the context of academic writing development.
Methods: Following the PRISMA 2020 guidelines, relevant peer-reviewed studies published between 2020 and 2025 were collected from major academic databases. Inclusion criteria targeted studies involving English education students in higher education settings that implemented generative AI tools as part of teaching and learning activities.
Results: The review highlights key benefits of generative AI, including enhanced learner autonomy, improved writing quality through AI-assisted feedback, and increased engagement. However, challenges were identified, such as ethical concerns, the risk of over-reliance, and the need to develop students' critical literacy skills in using AI-generated content.
Conclusion: Integrating generative AI into English education has the potential to significantly improve academic writing skills and student engagement. Nevertheless, effective implementation requires clear pedagogical frameworks, ethical guidelines, and targeted training to ensure responsible and balanced use. The findings provide practical insights for educators, curriculum designers, and policymakers in optimizing AI integration into English language teaching and learning.

To cite this article: Angraini, D. (2025). Harnessing Generative AI to Support EFL Academic Writing: A Systematic Review of English Education Students. *Journal of English Language Teaching and Applied Linguistics, 1*(1), 43-53. https://doi.org/10.58723/jaiela.v1i1.56

INTRODUCTION

Background of the Study:
In today's highly interconnected and globalized society, English language proficiency has evolved into an indispensable skill for academic success, professional advancement, and intercultural communication. The pursuit of effective language acquisition and communicative competence has thus become a priority in higher education worldwide. Parallel to this demand, rapid technological advancements particularly in Artificial Intelligence (AI) have begun to reshape the educational landscape, offering unprecedented opportunities to enhance language learning outcomes. AI has been shown to augment a range of language-specific competencies. For example, empirical studies have demonstrated its capacity to improve reading comprehension (Xu et al., 2019), facilitate repetitive practice through automated systems (Kim, 1989), and support pronunciation accuracy (Noviyanti, 2020). Beyond these domain-specific applications, AI also enables broader pedagogical

functions such as automated grading, instant feedback provision, adaptive learning pathways, and intelligent tutoring systems (Pokrivčáková, 2019). Categorizes AI applications in education into three primary domains: learner-facing tools, which directly engage students in learning activities; teacher-facing tools, which assist educators in tasks such as assessment and lesson planning; and system-facing tools, which aid administrators in managing and analyzing institutional data. In the field of English Language Teaching and Learning (ELT/L), AI systems can process vast datasets, utilize natural language processing for speech, writing, and listening, and apply linguistic patterns to deliver personalized and contextually relevant support.

Literature Review:
A growing body of research has explored the pedagogical potential and practical challenges of AI integration in ELT/L (Crompton et al., 2024). Kovalenko and Baranivska (2024) emphasize that AI's capacity to deliver personalized learning experiences can substantially enhance learner engagement, provided the underlying systems are carefully designed and capable of iterative improvement. However, they caution against excessive dependence on AI, noting that authentic interaction with educators remains essential for cultivating communicative competence and contextual understanding. Similarly, Koraishi (2023) highlights the transformative role of generative AI tools most notably ChatGPT in streamlining curriculum development, accelerating material design, and enabling more learner-centered instruction. Such tools, by accommodating individual learning needs, have the potential to make language instruction more adaptive and inclusive. Complementing these insights, an experimental study by Wale and Kassahun (2024) demonstrated the positive impact of AI-assisted writing tools, specifically Writerly and Google Docs AI, on EFL writing instruction. Students in the experimental group significantly outperformed their counterparts in task achievement, coherence and cohesion, lexical resource, and grammatical range and accuracy. Furthermore, these students reported highly favorable perceptions of the tools, describing them as engaging, effective, goal-oriented, and supportive attributes that collectively enhanced their motivation and writing outcomes.

Gap Analysis:
While the scholarly discourse on AI in ELT/L has expanded considerably, much of the existing literature concentrates on its general applications or teacher-oriented functionalities. Far less attention has been directed toward generative AI a subset of AI capable of producing original, human-like text in the specific context of English education students, i.e., pre-service teachers and prospective language professionals. In particular, there is a scarcity of systematic syntheses examining how generative AI supports academic writing development and fosters learner autonomy at the tertiary level. This gap is significant given the accelerating adoption of AI-powered writing assistants, collaborative platforms, and virtual learning environments over the past five years.

Rationale of the Study: The recent proliferation of generative AI technologies has disrupted traditional paradigms of teaching and learning, especially in academic writing instruction. Tools such as ChatGPT, Writerly, and Google Docs AI now offer instant feedback, facilitate idea generation, and support revision processes capabilities that could transform how English education students acquire and refine writing skills (A. M. A. Ausat et al., 2023). However, without an evidence-based understanding of their pedagogical implications, the risk of unbalanced adoption, ethical misuse, or over-reliance remains high. Therefore, a systematic review is warranted to consolidate current evidence, critically evaluate its strengths and limitations, and guide best practices for sustainable and ethical integration in higher education.

Purpose of the Study: Guided by the PRISMA 2020 framework (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*), this study systematically reviews empirical literature published between 2020 and 2025 on the integration of generative AI in English education at the university level (Page et al., 2020). Specifically, it aims to:

1. Examine how generative AI technologies are being utilized by English education students to enhance writing skills and promote learner autonomy.

2. Identify the pedagogical benefits, limitations, and challenges of current implementations to inform future research, instructional design, and policy development.

By addressing these objectives, the review seeks to contribute to a nuanced understanding of generative AI's role in language education, offering actionable insights for educators, curriculum designers, and policymakers striving to optimize AI integration in academic writing instruction.

METHOD

Research Design:

This study employed a systematic review design guided by the PRISMA 2020 (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) framework (Haddaway et al., 2022). The PRISMA protocol was selected to ensure methodological transparency, replicability, and comprehensive coverage of relevant literature. According to Tong et al. (2012), PRISMA is applicable for systematic reviews that include synthesis (e.g., meta-analysis) and those without statistical synthesis, making it ideal for this review, which synthesizes heterogeneous empirical studies. The focus was on peer-reviewed empirical research published between January 2020 and April 2025 that investigated the integration of generative AI tools into English education at the tertiary level, particularly in relation to academic writing skills and learner autonomy.

Participants, Population, and Sampling: As this review synthesizes secondary data, "participants" refer to the study samples within the included primary research. Across the selected studies, participants predominantly consisted of undergraduate English education students often pre-service teachers studying at higher education institutions in various regions, including Asia, Africa, Europe, and the Middle East. Sampling techniques reported in the original studies included purposive sampling, convenience sampling, and random assignment depending on research aims.

Instrumentation in Primary Studies: While no primary instruments were used in the review itself, the included studies utilized various data collection tools, such as:

1. Analytic Writing Rubrics (e.g., IELTS Writing Descriptors, adapted academic writing scales)
2. Perception and Attitude Questionnaires (Likert-scale based)
3. Semi-structured Interview Guides for qualitative insights
4. Performance-based Tasks (essay writing, report writing, or project submissions)

Scoring methods in experimental studies generally assessed task achievement, coherence and cohesion, lexical resource, and grammatical range and accuracy. Psychometric validation of instruments in the original studies often involved expert review and inter-rater reliability checks (e.g., Cohen's kappa).

Data Sources and Search Strategy: Three major academic databases Google Scholar, Wiley Online Library, and Research Gate were systematically searched. The search strategy combined Boolean operators with precise keywords to maximize coverage while minimizing irrelevant results. The search syntax was:

("Generative AI" OR "AI in English Language Education" OR "Artificial Intelligence") AND ("English Education Students" OR "English Language Learners")

Table 1. Inclusion and Exclusion Criteria for the Systematic Review

Criteria Type	Inclusion Criteria	Exclusion Criteria
Language	Articles written in English	Non-English articles
Publication Date	Published between January 2020 – April 2025	Studies published before 2020
Publication Type	Empirical, peer-reviewed journal articles	Conference abstracts, editorials, or non-empirical reports
Focus Area	Focused on integration of generative AI in English education at tertiary level	Research unrelated to English education or without generative AI integration
Scope of Application	Provided measurable learning outcomes related to academic writing, learner autonomy, or skill acquisition	Studies not addressing learning outcomes
Duplicates	Unique records retained	Duplicate records removed

The search and selection process for this systematic review was designed to ensure methodological rigor and alignment with the PRISMA 2020 guidelines. The overarching aim was to identify high-quality, empirical research that provides robust evidence on the integration of generative artificial intelligence (AI) within English education at the tertiary level. To achieve this, a series of clearly defined inclusion and exclusion criteria were established before initiating the literature search, thereby minimizing selection bias and enhancing replicability.

Inclusion Criteria were set to ensure the studies were both relevant to the research objectives and methodologically sound. First, only peer-reviewed journal articles were considered, as this filter ensures that the included studies have undergone scholarly scrutiny and meet academic standards for validity and reliability. Second, the publication date range was limited to January 2020 to April 2025. This timeframe was intentionally chosen to capture the most recent and relevant developments in AI-assisted English education, particularly given the rapid evolution of generative AI technologies such as ChatGPT, Writerly, and Wordtune in recent years. Older studies were excluded to maintain the review's focus on contemporary pedagogical practices and technological affordances.

Third, only articles written in English were included. This decision was based on two primary considerations: (1) ensuring the feasibility of accurate analysis without translation, and (2) capturing the predominant body of AI-related English education research, which is often published in English-language journals. Fourth, the focus area was restricted to the integration of generative AI within English education at the tertiary level. This was a crucial boundary-setting decision, as the tertiary context encompassing universities, colleges, and other higher education institutions—presents distinct pedagogical challenges and opportunities compared to primary or secondary levels. Studies outside this context were excluded to maintain contextual relevance.

Fifth, to ensure that the review was able to draw meaningful and comparable conclusions, only studies that reported measurable learning outcomes were retained. This included outcomes related to academic writing performance, learner autonomy, and language skill acquisition. Studies offering only theoretical perspectives or conceptual discussions without empirical data were excluded, as they did not provide evidence-based findings that could be synthesized for this review.

The Exclusion Criteria were equally important in refining the dataset and preventing dilution of the review's analytical focus. Studies published before 2020 were excluded, as they may not reflect the latest advancements in AI tools and applications. Research unrelated to English education or not integrating generative AI was also excluded, even if it addressed language learning more generally or used non-generative AI technologies (e.g., traditional automated essay scoring systems). Similarly, conference abstracts, editorials, opinion pieces, and non-empirical reports were omitted because they lack the depth and methodological transparency required for systematic synthesis. Additionally, duplicate records arising from searches across multiple databases were removed at an early stage using automated filtering tools. This was critical to prevent redundancy and ensure that each study contributed unique insights to the review.

The application of these inclusion and exclusion criteria was carried out in multiple screening phases. The first phase involved automated filtering using search parameters such as keywords, date range, and language to exclude non-relevant records. The second phase consisted of title and abstract screening to determine preliminary relevance. Studies passing this phase were subjected to a full-text review, during which the criteria were applied more rigorously. This multi-phase process allowed for both breadth (capturing a wide range of potential studies) and depth (ensuring alignment with the review's focus). In total, the search process initially identified 200 records from leading academic databases, including Google Scholar, Wiley Online Library, and ResearchGate. After removing 75 duplicates, 125 unique records remained. Of these, 60 were excluded during automated screening due to language, date, or publication type filters, leaving 65 articles for manual title and abstract review. The manual screening resulted in the exclusion of 27 articles that did not meet the

thematic or methodological requirements. This left 38 full-text articles for retrieval, of which 28 were inaccessible due to restricted access or missing full-text versions. The final dataset comprised 10 articles that fully met all inclusion criteria and were subsequently analyzed in depth for the synthesis phase.

This deliberate and systematic filtering process ensured that the studies included in the review represent the most relevant, methodologically robust, and contextually appropriate research available. By adhering to these criteria, the review achieves both internal validity ensuring that findings are derived from high-quality sources and external validity, in that the conclusions can reasonably inform pedagogical practices, curriculum design, and policy-making in higher education settings where generative AI integration is being considered. The inclusion and exclusion framework also aligns with best practices in systematic review methodology by promoting transparency (criteria were predefined and consistently applied), reproducibility (other researchers can replicate the search with the same parameters), and focus (only studies directly relevant to the research questions were included). This strengthens the credibility of the review's findings and provides a clear foundation for future research to build upon.

Study Selection Process:
The initial database search identified 200 records. After removing 75 duplicates, 125 records remained. An automated relevance filter excluded 60 articles, leaving 65 records for manual title and abstract screening. Following this screening, 27 articles were excluded due to non-alignment with the inclusion criteria. From the 38 full-text articles retrieved, 28 were inaccessible due to limited availability, resulting in 10 eligible studies for final synthesis. The selection process is illustrated in Figure 1 (PRISMA 2020 flow diagram).

Data Extraction and Analysis Plan:
A structured data extraction template was developed to systematically capture the following information from each study:

1. Author(s) and Year
2. Country/Region of the study
3. Research Design and Methodology
4. Participants and Sample Size
5. Generative AI Tool(s) Used
6. Pedagogical Context and Purpose
7. Measured Outcomes
8. Key Findings
9. Reported Challenges/Limitations

Given the diversity in study designs and measurement approaches, qualitative thematic synthesis was employed rather than meta-analysis. Themes were identified through iterative coding, focusing on patterns in pedagogical benefits, learner outcomes, and implementation challenges.

Scope and Limitations:
While the review followed rigorous PRISMA 2020 protocols, certain limitations must be acknowledged. The final sample size of 10 studies restricts the breadth of generalization. The English-only language filter may have excluded relevant studies in other languages. Additionally, access restrictions on certain paywalled articles may have limited comprehensiveness. Nevertheless, the transparent methodology and inclusion of studies from multiple geographical contexts strengthen the robustness and credibility of the findings.

RESULT AND DISCUSSION

Results:
The synthesis of ten peer-reviewed studies published between January 2020 and April 2025 provides robust empirical evidence that the integration of generative Artificial Intelligence (AI) into tertiary-level English language education substantially enhances students' academic writing skills. Across the

reviewed literature, quantitative research consistently demonstrated statistically significant improvements in multiple micro-level writing features. These include lexical diversity as reflected in a broader range of vocabulary and synonym use cohesion and coherence, and mechanical accuracy, encompassing grammar, punctuation, and spelling. Such gains were particularly evident in intervention-based studies where students received iterative AI-supported feedback during the drafting process. Complementary qualitative evidence, derived from semi-structured interviews, focus group discussions, and teacher reflective reports, corroborated these quantitative findings. Students reported heightened motivation to engage in writing tasks, a reduction in writing-related anxiety, and an increased sense of autonomy in managing the revision process. The immediacy of AI-generated feedback was frequently cited as a motivating factor, allowing students to iterate quickly without waiting for teacher intervention. Furthermore, several studies highlighted how AI tools, such as ChatGPT, Writerly, Google Docs AI, Wordtune, and InstaText, facilitated more independent learning, particularly for students hesitant to seek help in traditional classroom settings (Zhao, 2023). Despite these benefits, the studies also revealed recurring challenges. Foremost among these was a lack of adequate teacher training for AI integration, which often led to inconsistent instructional practices. Digital access disparities especially in resource-constrained contexts limited equitable adoption. Another common barrier was students' limited ability to formulate effective prompts to maximize the relevance and quality of AI-generated feedback. This "prompt literacy" deficit often reduced the pedagogical value of AI-assisted writing.

Discussion:

A thematic analysis of the reviewed studies revealed four dominant patterns.

First, generative AI tools demonstrated exceptional capability in improving micro-level aspects of writing, such as grammatical precision, lexical variation, and sentence-level clarity. These tools provided corrective feedback that students could apply immediately, fostering iterative refinement of their drafts. However, the literature consistently indicated that AI's influence on higher-order writing skills including argumentative structuring, critical reasoning, and integration of complex ideas was more limited. This aligns with the view that while AI excels in mechanical correction, it does not inherently cultivate the cognitive and rhetorical competencies essential for advanced academic writing.

Second, AI integration positively influenced learner autonomy and engagement. By enabling real-time feedback and supporting self-regulated learning, AI tools encouraged students to take greater ownership of their work. Studies such as Harunasari (2023) and Songsiengchai (2025) showed that learners who engaged with AI in a structured manner developed stronger reflective practices, became less reliant on teacher intervention for basic corrections, and exhibited increased persistence in tackling complex writing tasks. However, this engagement was conditional dependent on the tool's usability, perceived relevance, and alignment with meaningful learning objectives.

Third, the introduction of AI into writing pedagogy prompted a shift in the teacher's role. Instead of spending substantial time on error correction, educators could focus on higher-level instruction, such as developing critical thinking, fostering originality, and guiding students in rhetorical development. However, this pedagogical shift required substantial teacher readiness including both digital literacy and pedagogical strategies for AI integration (A. Ausat et al., 2023). Teacher readiness is a crucial factor in the successful integration of AI into ELT, as educators' prior experiences with educational technologies influence their willingness to adopt new tools. In addition, ensuring equitable access to AI-supported English learning opportunities requires addressing infrastructural gaps, particularly in remote areas where open and distance education can serve as an effective solution (Riady et al., 2025). Previous studies have shown that teacher engagement with digital platforms such as social media during the COVID-19 pandemic was shaped by factors including perceived usefulness, ease of use, and institutional support (Riady et al., 2022). Without this readiness, the potential for misalignment between AI use and learning objectives increased, risking superficial improvements without deeper skill development.

Fourth, the literature consistently highlighted risks and ethical concerns. Over-reliance on AI emerged as a significant issue, with some students generating entire essays with minimal personal input, raising concerns about academic integrity and authorship authenticity. The risk was compounded by the growing sophistication of AI-generated text, making plagiarism detection increasingly challenging. Bias in AI output, the tendency toward formulaic language, and the erosion of individual writing style were also identified as potential downsides. Moreover, both technical barriers (e.g., unstable internet, platform accessibility) and the cognitive challenge of crafting effective prompts limited the universal benefits of AI.

Implications

These findings hold substantial implications for both pedagogical practice and educational policy. First, generative AI should be positioned as a complementary resource, rather than a replacement, for human feedback. Particularly in advanced writing instruction, nuanced rhetorical skills—such as argumentation, synthesis, and critical evaluation—remain best cultivated through direct teacher-student interaction. Second, teacher professional development should prioritize AI literacy, encompassing not only technical proficiency but also ethical awareness, critical evaluation skills, and strategies for fostering student reflection on AI-generated suggestions. Third, institutions must develop clear policy frameworks governing AI use in academic contexts, balancing innovation with the preservation of academic standards. Finally, addressing the digital divide through infrastructural investment and equitable access policies is essential to prevent AI from exacerbating educational inequalities (Riady et a., 2025).

Research Contribution

This systematic review makes several notable contributions to the emerging scholarship on AI-assisted language learning. To the best of current knowledge, it is the first synthesis based on the PRISMA 2020 framework that focuses exclusively on generative AI in tertiary-level English education. It offers a dual-level analysis: identifying micro-level benefits, such as enhanced grammar and vocabulary, and macro-level limitations in areas like argumentative coherence and critical reasoning. Moreover, it advances a conceptual understanding of the interdependence between technological affordances and pedagogical readiness, illustrating that AI's potential is maximized only when educators are equipped to integrate it purposefully into curriculum design. By doing so, the review provides a foundation for developing balanced and sustainable AI integration frameworks in higher education writing instruction.

Limitations

The scope of this review is constrained by several factors. First, the relatively small sample size of ten studies limits the generalizability of its conclusions. Second, the language restriction to English publications may have excluded relevant studies in other languages, particularly those conducted in non-English-speaking contexts with different pedagogical traditions. Third, methodological heterogeneity including differences in study design, participant demographics, and assessment measures precluded meta-analytic synthesis. Fourth, access limitations to certain paywalled articles may have introduced publication bias, as studies with less favorable results are less likely to appear in open-access formats. Finally, the review focuses primarily on short-term interventions, leaving the long-term effects of AI integration on writing development relatively unexplored.

Suggestions for Future Research

Future research should address these limitations through longitudinal designs that capture sustained effects of AI-assisted writing on learner outcomes over multiple semesters or academic years. Comparative studies across a range of generative AI tools could illuminate which features (e.g., grammar correction, idea generation, style refinement) are most effective for specific writing objectives. There is also a need for research into hybrid feedback models that combine AI-generated suggestions with structured human guidance, potentially yielding a more comprehensive support system for learners. Furthermore, cross-cultural investigations could shed light on how AI's pedagogical affordances vary across educational systems and linguistic backgrounds. Finally, policy-oriented studies should focus on developing ethical guidelines and assessment standards for AI use

in higher education, ensuring that technological innovation is integrated in ways that uphold academic integrity and foster genuine skill development.

Table I. Reviewed Articles

Author(s) & Year	Title	Sample	Methodology	Key Findings	Reported Challenges
Bantalem Derseh Wale & Yirgalem Fentie Kassahun (2024)	*The Transformative Power of AI Writing Technologies: Enhancing EFL Writing Instruction through the Integrative Use of Writerly and Google Docs*	92 second-year EFL university students, Injibara University, Ethiopia (randomly divided into experimental & control groups)	Quasi-experimental pretest–posttest two-group design; mixed methods; data from essays, questionnaires, focus groups, and teacher diaries	Students using Writerly & Google Docs significantly improved in task achievement, coherence, vocabulary, and grammar; higher engagement; promoted collaboration and self-editing; reduced teacher workload	AI feedback not always personalized; risk of over-reliance limiting critical thinking; preference for manual methods among some; initial confusion without guidance
Md Kamal Hossain & Md Abdullah Al Younus (2025)	*Teachers' Perspectives on Integrating ChatGPT into EFL Writing Instruction*	22 tertiary-level EFL teachers from public & private universities, Bangladesh	Phenomenological qualitative study; purposive sampling; open-ended questionnaires; thematic analysis	Teachers found ChatGPT useful for idea generation, grammar correction, and organization; supported process-oriented and collaborative learning	Over-reliance by students; limited critical thinking; access issues; concerns over academic integrity; need for teacher training and integration frameworks
Songsiengchai (2025)	*Implementation of AI: ChatGPT for Effective English Language Learning among Thai Students*	120 Thai university students (aged 19–20)	Mixed methods: quasi-experiment, interviews, field notes	Significant improvement in English skills ($p < .001$); increased motivation and engagement; personalized feedback	Digital divide; need for teacher training; lack of evidence on speaking/listening outcomes
Fatemeh Etaat (2024)	*The Effect of AI-Based Applications on EFL Writing Skill Development: An Inquiry into Integration of AI into Language Learning*	34 intermediate-level Iranian EFL learners (16 experimental, 18 control)	Quasi-experimental; pretest, six periodic tests, posttest over 36 sessions; attitudinal data via questionnaires	Experimental group improved mechanics, lexis, and grammar; reduced writing time; positive attitudes from students & teacher	Over-reliance on AI; limited gains in organization/content development; ethical concerns (plagiarism, autonomy loss); teacher guidance needed

Author(s) & Year	Title	Sample	Methodology	Key Findings	Reported Challenges
Siti Yulidhar Harunasari (2023)	*Examining the Effectiveness of an AI-Integrated Approach in EFL Writing: A Case of ChatGPT*	16 fourth-semester undergraduate EFL students, Jakarta, Indonesia	One-shot case study; quantitative (pre–post test) and qualitative (questionnaires, ChatGPT chat history)	Improved idea generation, grammar, and organization; higher creativity; overcame writer's block; responsible use of AI	Over-reliance; difficulty crafting effective prompts; distractions; technical issues exporting chat history; concerns over reduced critical thinking

CONCLUSION

This systematic review confirms that the expectations stated in the Introduction namely, that generative AI can be effectively integrated into English language education to enhance academic writing skills are strongly supported by the evidence synthesized in the Results and Discussion. The reviewed studies consistently demonstrate that tools such as ChatGPT, Writerly, Google Docs AI, Wordtune, and InstaText significantly improve micro-level writing aspects, foster learner autonomy, and promote engagement when used within sound pedagogical frameworks. These outcomes align with the study's initial aim to explore both the potential and challenges of AI integration at the tertiary level.

However, the findings also highlight that the benefits of generative AI are not inherently guaranteed; their realization depends on thoughtful instructional design, sustained teacher involvement, ethical safeguards, and the development of students' critical and creative capacities. While AI excels in providing mechanical corrections and lexical enhancements, it is less effective in nurturing higher-order skills such as argumentation and rhetorical structuring without human mediation. Therefore, the optimal use of AI lies in adopting a balanced approach in which technology complements, rather than replaces, human feedback.

Looking ahead, the prospects for further research and application are promising. Longitudinal and comparative studies could deepen our understanding of AI's long-term effects on writing proficiency and academic integrity. Furthermore, the development of hybrid feedback models combining AI-generated suggestions with targeted teacher input offers a practical pathway to maximize learning outcomes. At the institutional level, the establishment of clear AI-use policies, professional development programs, and equitable access to digital resources will be critical in ensuring that the integration of AI into English language education remains both ethical and transformative.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to UIN Sulthan Thaha Saifuddin Jambi and Universitas Jambi for their academic and administrative support during the preparation of this study. Special thanks are also extended to the faculty members, students, and researchers whose empirical work formed the basis of this review. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

AUTHOR CONTRIBUTION STATEMENT

All authors contributed significantly to the overall research process. This included developing the initial research concept, designing the study framework, conducting the literature search, and selecting relevant studies in accordance with the PRISMA 2020 guidelines. The authors jointly analyzed and interpreted the data, discussed the implications of the findings, and collaboratively prepared the manuscript. Each author participated in revising and refining the content to ensure clarity, accuracy, and coherence. The final version of the manuscript represents the consensus and shared intellectual input of all authors, who have read and approved it prior to submission.

REFERENCES

Ausat, A. M. A., Azzaakiyyah, H. K., Permana, R. M., Riady, Y., & Suherlan, S. (2023). The Role of ChatGPT in Enabling MSMEs to Compete in the Digital Age. *Innovative: Journal of Social Science Research*, *3*(2), 622–631. https://doi.org/10.31004/innovative.v3i2.346

Ausat, A., Massang, B., Efendi, M., Nofirman, N., & Riady, Y. (2023). Can Chat GPT Replace the Role of the Teacher in the Classroom: A Fundamental Analysis. *Journal on Education*, *5*(4), 16100–16106. https://doi.org/10.31004/joe.v5i4.2745

Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English Language Teaching: Affordances and Challenges. *British Journal of Educational Technology*, *55*(6), 2503–2529. https://doi.org/10.1111/bjet.13399

Etaat, F. (2024). *The Effect of AI-Based Applications on EFL Writing Skill Development: An Inquiry into Integration of AI into Language Learning* [UiT Open Research Data]. https://hdl.handle.net/10037/33935

Haddaway, N. R., Page, M. J., Pritchard, C. C., & McGuinness, L. A. (2022). PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis. *Campbell Systematic Reviews*, *18*(1), 1230. https://doi.org/10.1002/cl2.1230

Harunasari, S. Y. (2023). Examining the effectiveness of AI-integrated approach in EFL writing: A case of ChatGPT. *International Journal of Progressive Sciences and Technology*, *39*(2), 357–368. https://doi.org/10.52155/ijpsat.v39.2.

Hossain, M. K., & Younus, M. A. Al. (2025). Teachers ' Perspectives on Integrating ChatGPT into EFL Writing Instruction. *T E S O L C o m m u n i c a t i o n S*, *4*(1), 41–60. https://doi.org/10.58304/tc.20250103

Kim, I. J. (1989). A contrastive analysis of English and Korean. *The Journal of the Applied Linguistics Association of Korea*, *2*(1), 45–70. https://doi.org/10.17960/ell.2010.16.1.006

Koraishi, O. (2023). Teaching English in the Age of AI: Embracing ChatGPT to Optimize EFL Materials and Assessment. *Language Education and Technology*, *3*(1), 1–15.

Kovalenko, I., & Baranivska, N. (2024). Integrating Artificial Intelligence in English Language Teaching: Exploring the Potential and Challenges of AI Tools in Enhancing Language Learning Outcomes and Personalized Education. *European Socio-Legal & Humanitarian Studies*, *1*, 86–95. https://doi.org/10.61345/2734-8873.2024.1.9

Noviyanti, S. D. (2020). Artificial intelligence (AI)-based pronunciation checker: An alternative for independent learning in pandemic situation. *The Journal of English Language Teaching in Foreign Language Context*, *5*(2), 162–172. https://doi.org/10.24235/eltecho.v5i2.7246

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., & Moher, D. (2020). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, *372*, 71. https://doi.org/10.1136/bmj.n71

Pokrivčáková, S. (2019). Preparing teachers for the application of AI-powered technologies in foreign language education. *Journal of Language and Cultural Education*, *7*(3), 135–153. https://doi.org/10.2478/jolace-2019-0025

Riady, Y., Arisanty, M., & Kuswanti, E. (2025). The Role of Open and Distance Education in Digital Ecosystem Transformation : A Qualitative Study on Enhancing Educational Accessibility. *Indonesian Journal of Elearning and Multimedia*, *4*(1), 81–89. https://doi.org/10.58723/ijoem.v4i2.386

Riady, Y., Arisanty, M., Kuswanti, E., Sukatmi, S., Karim, M. F., & Ajmal, M. (2025). The Implementation of Open and Distance Education to Advance Educational Access in Remote Areas. *FINGER: Jurnal Ilmiah Teknologi Pendidikan*, *4*(2), 141–149. https://doi.org/10.58723/finger.v4i2.387

Riady, Y., Habibi, A., & Sofyan, S. (2022). Factors affecting teachers' social media use during covid-19. *Cogent Social Sciences*, *8*(1), 1–16. https://doi.org/10.1080/23311886.2022.2115658

Songsiengchai, S. (2025). Implementation of Artificial Intelligence (AI): Chat GPT for Effective English Language Learning among Thai Students in Higher Education. *International Journal of Education and Literacy Studies*, *13*(1), 302–312. https://doi.org/10.7575/aiac.ijels.v.13n.1p.302

Tong, A., Flemming, K., Mcinnes, E., Oliver, S., & Craig, J. (2012). Enhancing transparency in reporting the synthesis of qualitative research : ENTREQ. *BMC Medical Research Methodology*, *Figure 1*, 1–

8. https://doi.org/10.1186/1471-2288-12-181
Wale, B. D., & Kassahun, Y. F. (2024). The Transformative Power of AI Writing Technologies: Enhancing EFL Writing Instruction through the Integrative Use of Writerly and Google Docs. *Human Behavior and Emerging Technologies*, *1*(1), 9221377. https://doi.org/10.1155/2024/9221377
Xu, Z., Wijekumar, K., Ramirez, G., Hu, X., & Irey, R. (2019). The effectiveness of intelligent tutoring systems on K–12 pupils' reading comprehension: A meta-analysis. *British Journal of Educational Technology*, *50*(6), 3119–3137. https://doi.org/10.1111/bjet.12758
Zhao, X. (2023). Leveraging Artificial Intelligence (AI) Technology for English Writing: Introducing Wordtune as a Digital Writing Assistant for EFL Writers. *RELC Journal: A Journal of Language Teaching and Research*, *54*(3), 890–894. https://doi.org/10.1177/00336882231123456