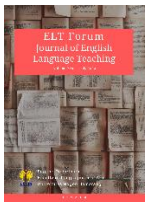
ELT FORUM 15 (Special Issue) (2026)

Journal of English Language Teaching

http://journal.unnes.ac.id/sju/index.php/elt



Mapping thematic evolution and pedagogical strategies of Generative AI in EFL writing

Khofifah Eva Septyana[✉1], **Kushardiyanti Novinda**[1]

[1]Universitas Negeri Semarang, Indonesia

Article Info

Article History:
Received on 1 May 2026
Approved on 5 August 2026
Published on 17 August 2026

Keywords: Generative AI; EFL academic writing; AI literacy; thematic evolution

Abstract

The mass adoption of Generative AI has brought a massive landscape shift in academic writing for EFL students. This technology acts as a powerful cognitive scaffold for students on finishing their academic writing tasks such as brainstorming ideas, checking grammar accuracy, paraphrasing a sentence and as sparring partner. Nevertheless, the tools also act as double-edged sword because its comprehensive aid that triggers the scaffolding paradox, cognitive offloading, ethical issue or academic integrity. However, these perspectives remain distributed across several research domains, making it harder to understand how the field is developing as an integrated research area. The study employs a hybrid study design using bibliometric informed scoping review to examine the thematic structure and the development of research on generative AI in EFL academic writing. The research aims to quantitatively map the thematic evolution of AI in EFL writing globally from 430 Scopus-indexed articles from January 2023 to June 2026, and also qualitatively synthesize the condition when AI act as substitute or act as cognitive behavior mechanism of overreliance on using Generative AI, and aims to offer a strategy or framework about practical AI literacy in EFL academic writing. Rather than treating Generative AI as inherently beneficial or harmful, it examines the emerging tension between AI as cognitive assistance and AI as a potential substitute for EFL learners. The study mainly proposes an evidence-informed conceptual framework that conceptualizes effective Generative AI assisted writing as a process in which this tool remains subordinate to learner critical thinking and judgment.

✉Correspondence Address:
B3 Building FBS Unnes
Sekaran, Gunungpati, Semarang, 50229
E-mail: khofifah.eva77@students.unnes.ac.id

p-ISSN 2252-6706 | e-ISSN 2721-4532

INTRODUCTION

The introduction of generative AI has caused a paradigm change in the field of education, especially in terms of its capacity to create learning content, personalize learning experiences, and provide automated methods for assessment and feedback (Sung et al., 2024). Within EFL academic writing, the potential benefits of GenAI are particularly increasing especially during their academic writing process,including brainstorming, outlining, drafting, paraphrasing, linguistic revision and feedback (Lee & Jeong, 2026).

According to (Pandey & Solanki, 2026) Academic integrity is one major issue in education. A related problem is the decline of core competencies. This over-reliance risks triggering 'skills atrophy,' where student writers' independent reasoning and composition capabilities severely deteriorate due to the bypassing of necessary learning struggles (Arifani et al., 2026). Furthermore, the pervasive copying of automated outputs homogenizes academic prose, resulting in the erasure of students authorial voices (Pham, 2026).

The lack of critical evaluation use of AI for assignments process significant ethical concerns within academic system. Furthermore, GenAI platforms provide infinite, non-fading, and effortless scaffolding, which inadvertently traps students in a state of permanent cognitive reliance (Hamamra et al., 2026). To dismantle this paradox and restore student agency, a structured pedagogical response is urgently required. Rather than reverting to retrogressive bans, higher education institutions must cultivate 'Critical AI Literacy' , equipping student writers with the meta-cognitive capacities to negotiate meaning with machines while actively preserving their epistemic ownership (Stanko et al., 2026)

Previous studies have examined many aspects of GenAI in education, including students' perceptions, academic integrity, critical thinking, self regulated learning. However, these strands are frequently examined separately. Research focusing on writing assistant may emphasize productivity and language support, while research focusing on academic writing may emphasize misuse and authorship, and also research which focus on students' perception may not addressing how GenAI changes the distribution of cognitive responsibility during the writing process. The tention matters in academic writing because writing is not merely a product-oriented activity, rather a process-oriented one which learner develop argument, metacognition, judgement, and disciplinary thinking. If GenAI used only to produce a finished one, it might not help the student on fostering their critical thinking, so AI literacy is really crucial in order to bridge the gaps from substitution to assistant role.

The rapid growth of literature in this field for the past few years has become largely and heterogeneous that each study only provides partial representation of its development. Bibliometric mapping can reveal the structural shift of research while a scoping review can deep dive on the breadth based on evidence from the themes identified through the mapping process.

The study addresses two research questions:

1. What is the thematic evolution and trends of research on Generative AI in EFL academic writing context between 2023-2026 based on bibliometric mapping using VOSviewer?
2. What is the pedagogical strategy or framework to foster AI literacy and learner critical thinking and learning-related tensions are reported in the core literature?

The goal of this hybrid study is to map the scholarly landscape and establish an actionable framework for integrating Critical AI Literacy in educational settings. Specifically, the study aims to: (1) spatially and quantitatively map keyword co-occurrence clusters to prove the marginalization of cognitive autonomy issues in current landscape; (2) qualitatively synthesize the core literature and formulate a standardized, process-oriented writing workflow that utilizes GenAI as an assistant to foster their critical thinking and literacy on their EFL academic writing class.

The contribution of the study is therefore not to determine whether GenAI should be universally adopted or rejected. Instead, it seeks to identify the conditions and conceptual relationship that distinguish AI-supported learning from AI-substitute performance. Furthermore, it suggests a strategy and framework on how to use GenAI in EFL academic writing context.

The theoretical intersection of second language acquisition and education technology has long relied on Lev Vygotsky's sociocultural theory learning takes place in the Zone of Proximal Development (ZPD) and is fundamentally a socially influenced process. According to this paradigm, "scaffolding" is a short-term supportive structure that helps the learner finish tasks that they couldn't finish on their own. As the learner internalizes the information and achieves autonomous mastery, fading necessitates the methodical, gradual withdrawal of support, converting procedural guidance into declarative, decontextualized cognitive skills (Arifani et al., 2026; Hamamra et al., 2026).

The integration of Generative AI, introduces a critical disruption known as the 'Scaffolding Paradox'. Unlike a human teacher who actively reduces assistance to push the student toward independent thinking, an AI chatbot will continue to rewrite, refine, and generate text for the student indefinitely. This lack of a fading mechanism subverts the ZPD (Waziana et al., 2025). Instead of internalizing language structures and moving from assisted performance to independent mastery, students are trapped in a state of perpetual assistance (Bouzar et al., 2024). When writing difficulties arise, students default to a passive state, allowing the machine to perform the complex cognitive work of drafting, synthesizing, and revising. The result can create the illusion of competence, increase risk of false sense of efficacy (Bouzar et al., 2024; Waziana et al., 2025).

To understand the psychological mechanics of the dependency trap, one must look to John Sweller's Cognitive Load Theory (CLT). In academic writing, GenAI reduces extraneous load by automating tedious tasks such as searching for synonyms or correcting spelling, thereby freeing up working memory to focus on high-level argumentation and conceptual synthesis (germane load). However, empirical evidence indicates that in the absence of Critical AI Literacy, students suffer from a phenomenon known as 'Cognitive Offloading (Waziana et al., 2025). Instead of using the cognitive capacity freed by the AI to engage in deeper analysis, students offload both germane and intrinsic loads to the machine.

In education landscape with massive growth of Gen AI adaptation, the traditional metrics used to evaluate L2 writing competenence need to be improved (Zhang & Liu, 2025). To counteract the dependency trap, educational theory must move beyond technical AI literacy toward a robust framework of 'Critical AI Literacy' (Parker et al., 2025). Technical AI literacy focuses on the functional ability to operate AI tools, such as knowing how to write prompts or generate text. Critical AI Literacy, however, is an epistemic competency. It requires students to view GenAI not as a passive tool or an absolute authority, but as a fallible, biased, and conversational thought partner.

According to (Pandey & Solanki, 2026) ,The Critical AI Literacy Framework that adopted from the OECD and UNESCO framerowk, as comprises in three core dimensions: (1) Cognitive Dimension: Understanding the underlying probabilistic mechanics of LLMs as word-prediction machines, enabling students to anticipate hallucinations and recognize that AI outputs are statistical averages rather than objective truths; (2) Evaluative Dimension: The ability to critically analyze, interrogate, and verify AI-generated texts, cross-referencing claims with verified empirical databases and actively detecting bias; and (3) Ethical Dimension: Developing a strong sense of academic integrity, recognizing the boundary between co-creation and plagiarisme, and maintaining a transparent disclosure policy of how AI was utilized in the writing process. In this framework, writing is redefined not as the production of words, but as the active curation, negotiation of meaning, and critical refinement of ideas.

METHODS

This study adopts a Hybrid Bibliometric informed scoping review design to map the development of research on GenAI in EFL academic writing and to identify emerging themes, conceptual tensions, and areas that require a further integration. This section consists of several steps taken to select the articles used for scoping review. The methodological pocess consisted of four phases: first, literature identification and bibliometric dataset construction; second, bibliometric mapping and thematic cluster identification; third, scooping review and thematic synthesis; and fourth, conceptual integration and framework development.

The primary data was retrieved from Scopus, which provides rigourus coverage of peer reviewed literature. The search strategy was developed using Boolean string operators to cluster three pivotal domain: Technology (gen AI tools), context (EFL academic writing), and Focus (dependency and cognitive outcome). It also applied the restrictive critea on language limit to English and only several subject areas. The search was from January 2023 to June 2026. The finalized advanced search string entered into scopus was follows : (TITLE-ABS-KEY (("Artificial Intelligence" OR "Generative AI" OR "GenAI" OR "ChatGPT" OR "Large Language Models" OR "LLM") AND ("EFL" OR "English as a foreign language" OR "academic writing") AND ("dependency" OR "reliance" OR "critical thinking" OR "AI literacy" OR "academic integrity")) AND PUBYEAR > 2022 AND PUBYEAR < 2027) AND PUBYEAR > 2022 AND PUBYEAR < 2027 AND (EXCLUDE (SUBJAREA,"MEDI") OR EXCLUDE (SUBJAREA,"MULT") OR EXCLUDE (SUBJAREA,"BUSI") OR EXCLUDE (SUBJAREA,"NEUR") OR EXCLUDE (SUBJAREA,"HEAL") OR EXCLUDE (SUBJAREA,"BIOC") OR EXCLUDE (

SUBJAREA,"DECI") OR EXCLUDE (SUBJAREA,"MATE") OR EXCLUDE (SUBJAREA,"IMMU") OR EXCLUDE (SUBJAREA,"ENVI") OR EXCLUDE (SUBJAREA,"ECON") OR EXCLUDE (SUBJAREA,"CENG") OR EXCLUDE (SUBJAREA,"PHYS") OR EXCLUDE (SUBJAREA,"MATH") OR EXCLUDE (SUBJAREA,"NURS") OR EXCLUDE (SUBJAREA,"AGRI") OR EXCLUDE (SUBJAREA,"PHAR") OR EXCLUDE (SUBJAREA,"ENER") OR EXCLUDE (SUBJAREA,"DENT") OR EXCLUDE (SUBJAREA,"CHEM") OR EXCLUDE (SUBJAREA,"ENGI")) AND (LIMIT-TO (DOCTYPE,"ar")) AND (LIMIT-TO (LANGUAGE,"English"))

The PRISMA-ScR framework was used to guide this review. This hybrid design was chosen because unlike a systematic review which focus on evaluate the quality of intervension, this approach employed to identify the gaps and to map the trend of existing research. A hybrid Bibliometric informed scoping review was used because the massive growth of the adoption of Generative AI in EFL context especially in academic writing and the raising concern of the double sword effect followed on the usage. A total of 430 scopus indexed journal article was selected and exported into CSV to be identified with open refine to remove duplication and screened using Rayyan following the PRISMA-ScR standard. The title and abstract were selected based on the PCC with the inclusion and exclusion as follows:

Table 1. Eligibility criteria based on PCC framework

Criterion (PCC & Method)	**Inclusion**	**Exclusion**	**Reason**
Population	University/Higher Education/College students) English as a Foreign/Second Language (EFL/ESL).	K-12 students, teachers,lecture, faculty members and researcher.	*Wrong population*
Concept	Generative AI (ChatGPT, Gemini, Claude) reated to critical thinking, SLR, and overreliance	Traditional AI (Grammarly , Google Translate).	*Wrong concept / Non-GenAI*
Context	(Academic writing, essay, thesis writing).	EFL skills such as speaking, listening, atau reading), and research outside EFL context.	*Wrong context*
Publication type	peer-reviewed journal article written in english.	conference proceedings),thesis, bookchapter, and non English article.	*Wrong publication type / Wrong language*
Text availability	Available Full text article	Not available full rext article/only title and abstract.	*Full-text not available*

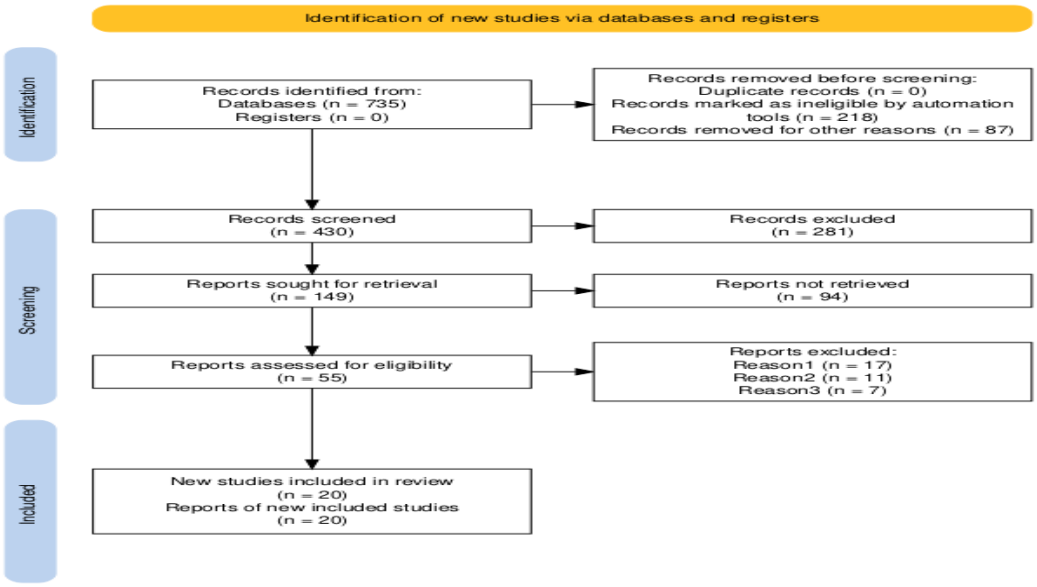


Figure 1. PRISMA diagram workflow

The first screening in scopus just using a keyword above showing 735 articles as result, limited by subject area, language and document type, only 430 articles meet the criteria. The screening mechanism starts by doing a title and abstract screening from 430 after been checked on the duplication using Open Refine and exported it into Rayyan, after the title and abstract screening, there were 281 article that didn't match the PCC framework. This process showed 149 articles left for full text screening. Then, after checked the resources, there were only 55 full text articles that available to download, and only 20 articles really matched the requirements based on the most relevant one to the PCC framework.

FINDINGS AND DISCUSSION

To address RQ1, this study focuses on the result analysis provided in scopus databse to map the the publication trend, the dominance of country, institution and author who did aresearch on Generative AI in EFL context. This research also focuses on finding the gaps and trend from 430 article exported from scopus and to deep dive from the result analysis and mapping using Vosviewer. Co-occurrence analysis of author keyword was performed, applying a threshold of minimum 5 and atleat 47 keyword met the threshold, which shows a massive increase of research about Generative AI on EFL context.

Publication Trend of Gen AI in EFL Academic Writing

Based on the analysis results provided on scopus database from January 2023 to 17 June 2026, the academic literature on Generative AI and its effect on EFL Academic Writing began its progressive development since 2023.The publication peak occurred in 2025 with 188 article published, it is twice higher than in that previous year, and it keep growing just in a half year. Here is the publication growth chart showed in figure 2.

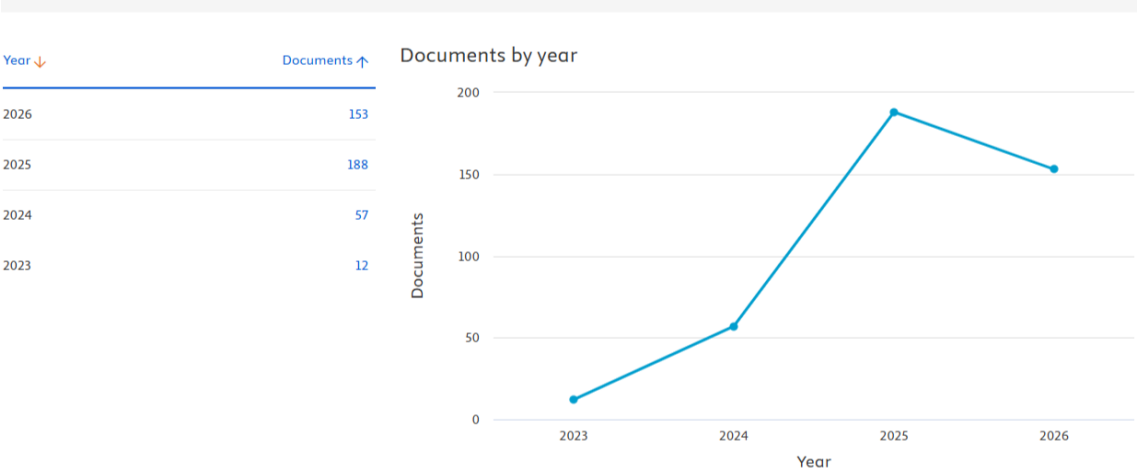


Figure 2. Source: Scopus

Evaluating how Generative AI on EFL academic writing are distributed across 430 article was observed by categorizing these publications by nation, region, affiliation, source, and author, with only the top 10 of each. Firstly, the distribution of academic inquiry focusing on GenAI in EFL academic writing context, that sorted by the country and regional boarder, it is led by China with 91 articles, Saudi Arabia with 35 articles, Indonesia with 31 articles, United States with 27 articles, Viet Nam with 24 articles, Malaysia and UK both with 23 articles, Iran with 21 articles, and Thailand with 16 articles. See figure 3.

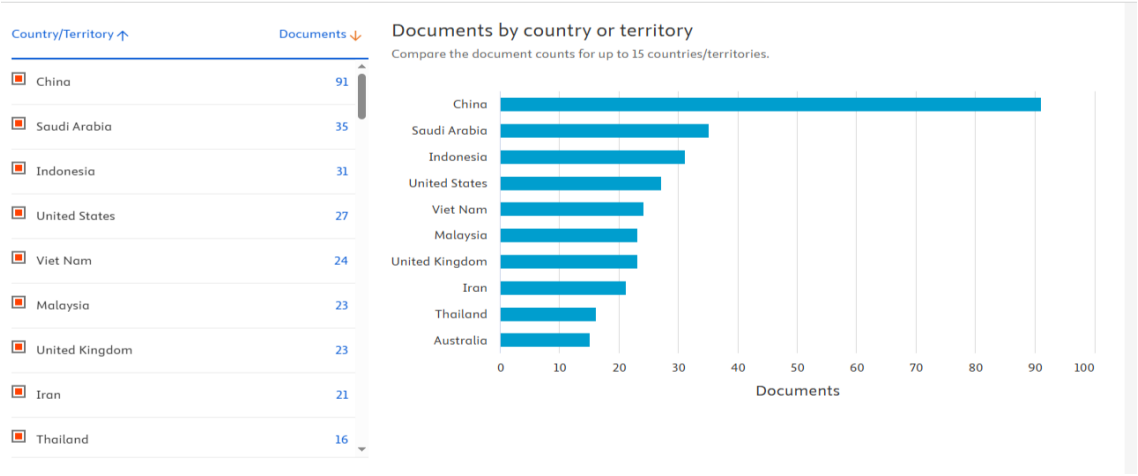


Figure 3. Source; scopus

Secondly, the distribution of Gen AI on EFL academic writing context based on university associations is dominated by Macaa Polytechnic University with 9 articles, Prince Sattam Bin Abdulaziz University with 6 articles, followed with 3 university (Universiti Sains Malaysia, Universiti Kebangsaan Malaysia, University of Dalat) with 5 article each, and them followed with 5 university that published 4 articles that fully shown on the figure 4.

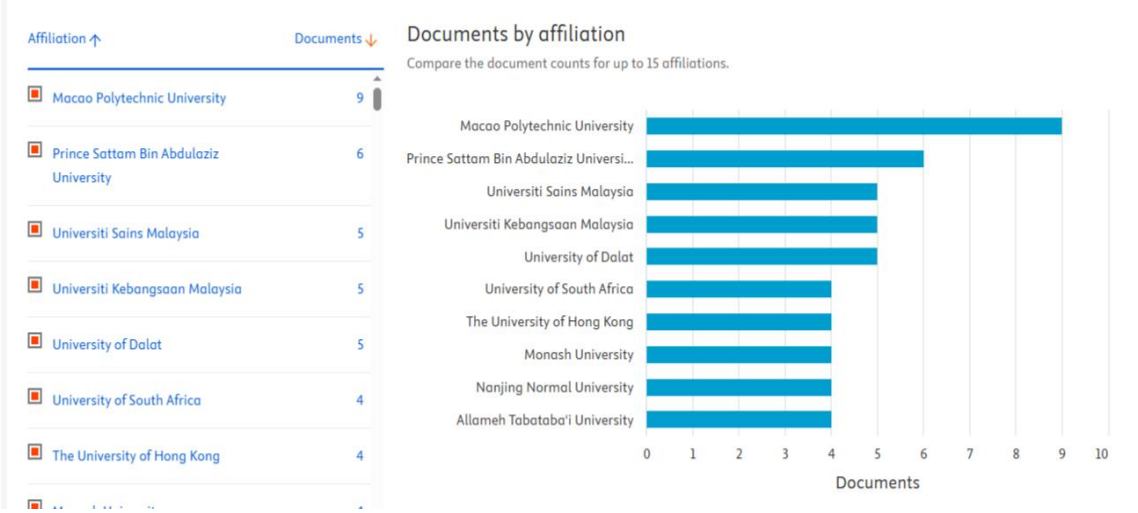


Figure 4. Source: scopus

Third, the distribution of Gen AI in EFL academic writing by authors is led by Cong-Lem, N and Teng, M.F. with 5 articles, followed by Marzuki with 4 articles, and 7 authors with 3 articles each (Derakhsan, A.; Detrick, R.; Hastomo, T, Hwang, G.J; Jamshad, M.; Kim, J.; and Lee, Y.J)

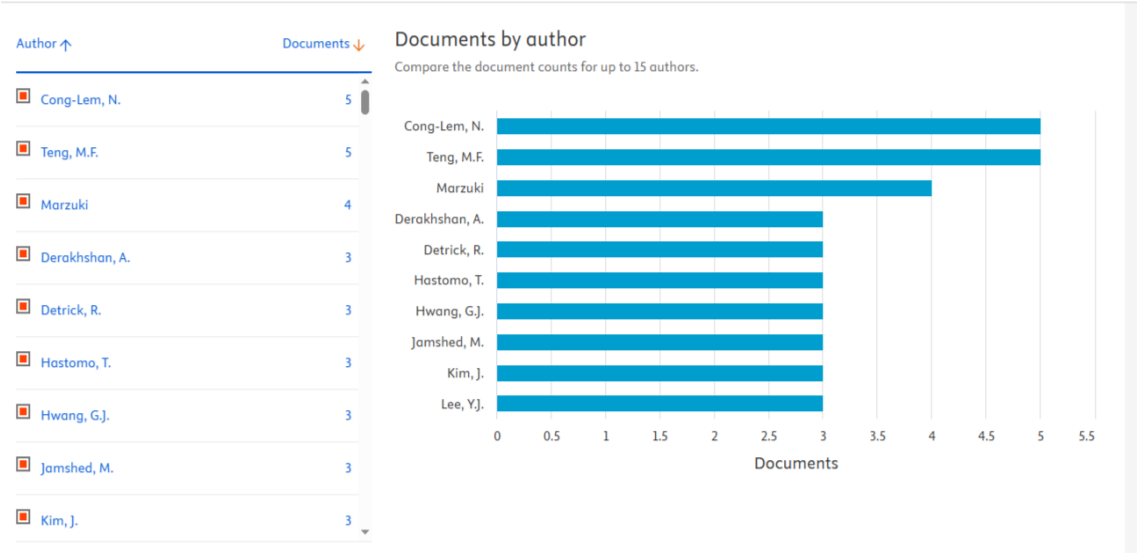


Figure 5. Source; Scopus

The topics that frequently investigated in the field of Generative AI research within EFL context.

In this particular area of study, there are several topics that frequently investigated from VOSviewer mapping. The VOSviewer visualization offers a comprehensive overview of scholarly research within this field, highlighting the dominant trends and the emerging theme. The results of of this metadata showed a central node and highly dense region that dominated by the keyword related to Artificial Intelligence such as Chatgpt, Generative AI. The marginal cluster (emerging themes) appear on the peripheries of the maps with tiny nodes and thin network, we can see the keywords such as critical thinking, AI literacy, ELT, writing skills, learner autonomy, and ethics. From figure below, the occurrences of keyword and cluster can be seen clearly.

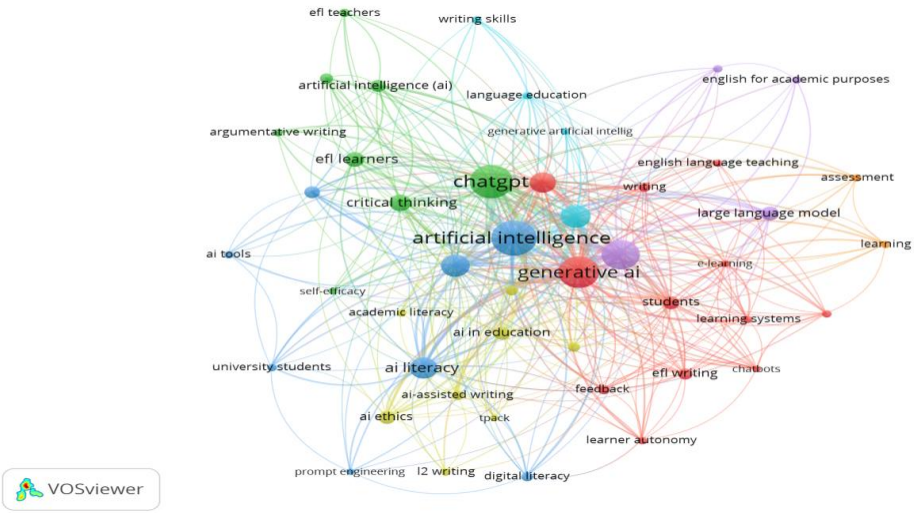


Figure 6. Source: Output VOSviewer Software

Table 2. Keywords by authors

Keyword	cluster	weight<Links>	Total link strength
Chatbots	1	18	26
e-learning	1	18	30
efl writing	1	17	31
english as a foreign language	1	35	123
english language teaching	1	10	17
Feedback	1	16	32
generative ai	1	42	273
learner autonomy	1	15	25
learning systems	1	18	38
self-regulated learning	1	12	23
Students	1	30	89
Writing	1	17	30
argumentative writing	2	8	13
artificial intelligence (ai)	2	16	34
Chatgpt	2	42	257
critical thinking	2	30	76
efl learners	2	20	50
efl teachers	2	6	10
language learning	2	10	19
self-efficacy	2	9	11
ai literacy	3	29	100
ai tools	3	8	12
artificial intelligence	3	43	305
digital literacy	3	12	22
educational technology	3	17	38
higher education	3	32	152
prompt engineering	3	9	14

university students	3	11	22
academic literacy	4	9	14
ai ethics	4	16	44
ai in education	4	22	48
ai-assisted writing	4	16	22
english as a foreign language (efl)	4	18	30
l2 writing	4	8	12
technology acceptance model	4	16	25
Tpack	4	9	11
academic writing	5	37	226
english for academic purposes	5	6	11
large language model	5	19	50
student perceptions	5	9	15
academic integrity	6	32	147
generative artificial intelligence (genai)	6	11	14
language education	6	15	24
writing skills	6	8	13
Assessment	7	9	17
Learning	7	10	23

The keyword co-occurrence analysis identified several interconnected thematic clusters in the literature on AI and EFL context. It means that the research has developed across several related but distinguishable areas, rather than concentrated on the technology capabilities of gen ai, the network also includes themes related to writing, AI literacy, critical thinking, ethics and assement. By analyzing the keyword of co occurrence from the table before, researcher intrerpret the 7 clusters into 4 broader themes emerge: 1. AI as a learning and writing support, 2. Cognitive engagement and AI literacy, 3. Responsible and ethical AI use, and 4. Academic integrity, learning and assessment.

AI as learning and writing support

The first theme concerns the positioning of AI as a tool for support EFL learning and writing. Cluster 1 was the largest thematic cluster, consisting of 12 keywords, *chatbots, e-learning, efl writing, English as a foreign language, English language teaching, learning systems, and students, and writing.* The presence of both technological and learner-oriented terms suggest that this cluster represents research examining the use of AI and chatbot-based technologies in EFL context and also potential of learning support. The presence of *feedback, learner autonomy* and *self-regulated learning* alsongside *generative ai* suggest that AI can be a form of learning support rather than merely an automated texr-genarted tools. The keyword in cluster 5 also represent the context of writing support, they keyword such as as *academic writing, English for academic purposes, large language model, and students' perception.* This cluster highlight the students' perception of large language model in EFL academic language learning. Generative AI act as interactive partner offer immediate feedback that helps L2 writer correct syntactic errors, paraphrase some words and expand their vocabulary (Barrot, 2026; Lee & Jeong, 2026b). Empirical studies on student perceptions confirm that EFL learners view these tools as liberating devices that dramatically lower writing anxiety and ease the burden of lexical retrieval (Bouzar et al., 2024; Lee & Jeong, 2026b). Furthermore, in English for Academic Purposes (EAP) contexts, under-resourced international students utilize LLMs to translate and organize complex scientific papers, thereby bypassing the traditional linguistic gatekeeping of Western academic publishing and achieving a form of cognitive justice(Pretorius, 2026).

Cognitive Engagement and L2 AI Literacy

The second theme represents a critical transition in the literature from treating AI as an automated editor to framing it as a co-constructive cognitive collaborator. This theme merges the pedagogical

aspects of cluster 2 (*argumentative writing, artificial intelligence, chatgpt, critical thinking, efl learners, efl teachers, language learning, self-efficacy*) with the digital competencies of Cluster 3 (*ai literacy, ai tools, artificial intelligence, digital literacy, educational technology, higher education, prompt engineering, university students*). Rhetorical tasks like argumentative writing demand higher order thinking skilss, including creating framework, integrating evidence, and opposing claims (Stofiana et al., 2025). When students use GenAI uncritically, they risk falling into "cognitive offloading", the mental externalization of logical synthesis to the machine in which can lead to intellectual laziness and "academic deskilling" (Waziana et al., 2025). To counter this cognitive decline, the literature highlights the crucial role of L2 Critical AI Literacy (Yao & Fan, 2026). Highly literate students not to use AI as a substitute for thought; instead, they engage in sophisticated *prompt engineering* and iterative dialogic workflows, treating the AI's synthetic text as a "boundary object" to be analyzed, questioned, and refined (Cong-Lem, 2025; Kim et al., 2026)). This collaborative process, termed Human-AI Distributed Critical Thinking (HADCT), ensures that students maintain active authorial agency while using the machine's processing speed to deepen their own critical analysis (Cong-Lem, 2025; Parker et al., 2025)

Responsible and Ethical AI Use in EFL Classrooms

The third theme addresses the ethical, professional, and institutional challenges of introducing generative tools into L2 writing classrooms. This theme centers on Cluster 4 (*academic literacy, ai ethics, ai in education, ai-assisted writing, English as a foreign language, l2 writing, technology acceptance model*). While traditional technology acceptance models (TAM) explain adoption through perceived ease of use and usefulness, the integration of GenAI in L2 writing is heavily moderated by ethical hesitation and policy ambiguity (Rahayu et al., 2025; Davis & lee., 2026). L2 writers often navigate a state of ethical confusion, struggling to distinguish between healthy, assisted editing and unauthorized writing practices due to a lack of clear institutional policies (Pretorius, 2026). Educators must actively develop the professional competence to design writing tasks that are resilient to AI misuse, establishing clear classroom parameters that encourage transparent, ethical, and responsible tool s.

Academic Integrity, Process Learning, and Assessment.

The final theme addresses the ethical crises arising from unguided AI use and details the process-oriented assessment strategies needed to protect academic standards. This theme synthesizes Cluster 6 (*academic integrity, generative artificial intelligence, language education, writing skills*) with Cluster 7 (*assessment, learning*).

The ease of generating fluent prose has led to a rise in "AI-giarism," where students use LLMs to generate essays and then apply advanced paraphrasing tools (like QuillBot) to hide academic dishonesty and bypass traditional plagiarism checkers (Hadra et al., 2026). Beyond ethical concerns, this practice threatens students' unique voices (Zhou et al., 2026). Because AI outputs are mathematically predictable and formulaic, relying heavily on them causes stylistic homogeneity, systematically erasing the student's unique cultural phrasing, dialectal nuances, and personal "linguistic fingerprint" (Zhou et al., 2026; Pretorius, 2026); Al Hosni, 2025).

To preserve the human element in writing, higher education must move away from evaluating only the final written product and adopt process-oriented assessment models (Singh et al., 2026). These models include tracking drafting histories, requiring reflective dialogic logs of student-AI interactions, and implementing teacher-led oral interviews to establish clear cognitive accountability(Cong-Lem, 2025).

Qualitative Thematic Synthesize

To address the RQ 2 on pedagogical strategy used to foster critical thinking on EFL academic writing, researcher presents 20 extraction articles detailing the author and year, country, theory, key cognitive findings, and pedagogical mitigation.

Table 3. The articles' extraction

No	Author & year	country	Theory	Key cognitive finding	Pedagogical mitigation
1	(Thi et al.,	Vietnam	Cognitive	Editing overly complex AI	Limit AI edits to single

	2025)		Load	text increases cognitive load and causes stress.	sentences or paragraph-level outlines.
2	(Hao Tran, 2026)	Vietnam	SDT	Students lose confidence in their skills if their original draf is overwritten.	Require students to keep 60% of their original draf unchanged.
3	(Liu, 2026)	China	Positive psych	High emotional enjoyment with AI masks a lack of grammatical learning	Integrate error-correction tests based on their own AI-generated texts.
4	(Ha, 2026)	Vietnam	Cognitive Load	Students accept fluent AI text uncritically, ignoring logical/argument flaws.	Conduct class-based peer editing of AI text to check logic
5	(Lee & Jeong, 2026a)	South korea	TAM	Immediate text generation increases speed but diminishes planning time.	Introduce pen-and-paper outlining before letting students use AI.
6	(Mekheimer, 2025)	Egypt	TAM	GenAI feedback reduces writing anxiety but fosters spelling/grammar laziness	Use formative multi-draft portfolios; reward edit decisions.
7	(Al-Absi & Abdi, 2026)	Yemen	Cognitive Load	Students suffer from cognitive overload when verifying falsified references.	Establish mandatory reference check protocols using primary sources.
8	(Hamamra et al., 2026)	Palestine	ZPD scaffolding	Absence of fading in ChatGPT scaffolds causes severe cognitive offloading.	Implement Socratic prompts; prohibit block rewriting.
9	(Arifani et al., 2026)	Indonesia	ZPD scaffolding	Drafting thesis proposals with AI risks the complete loss of student agency.	Track draft history; require student defense of thesis arguments.
10	(Stanko et al., 2026)	South Africa	Critical literacy	Failure to critically interrogate AI output fosters uncritical copying.	Integrate explicit 'Critical Chatting' exercises into the syllabus.
11	(Puspita et al., 2025)	Indonesia	Pragmatics	While text organization improves, individual student voice is severely muted.	Assess draf evolution by comparing human and AI drafts side-by-side.
12	Shah et al. (2025)	Bangladesh	SDT	Automated peer review reduces student-student collaboration and relatedness.	Combine AI feedback with mandatory peer-group discussions.
13	(Jahan & Akbar, 2026)	Oman	Critical thinking	Over-reliance on AI parafrasa diminishes students' vocabulary recall.	Mandate handwritten brainstorming maps before computer login.
14	(Lee & Jeong, 2026b)	South korea	TAM	Perceived usefulness encourages students to skip early-stage outlining.	Integrate Socratic dialogues where students explain vocabulary choice.
15	(Alshammari & Alhamazany, 2026)	South Arabia	Personalization	Standardized AI coaching causes a decline in students' individual vocabulary.	Require customized prompting to integrate personal stories.
16	Pattier (2025)	Spain	Integrity	L2 writers struggle to distinguish machine tone from their personal voice.	Train students to identify AI vocabulary patterns (e.g., 'profound').
17	(Al Murshidi et al., 2026)	UEA	TAM	Immediate lexical assistance bypasses necessary struggle for L2 acquisition.	Set writing tasks requiring direct handwritten notes in class.
18	(Hadra et	Jordan	Integrity	AI detectors discriminate	Establish trust-based

	al., 2026)			against non-native writers by tagging them as AI.	transparency logs; ignore raw AI scores.
19	Ironsi (2026)	Cyprus	Learner autonomy	AI-mediated reading tasks increase speed but compromise deep comprehension.	Writen examinations; evaluate reading comprehension without devices.
20	Eassa Ali (2025)	Oman	SRL	Low self-regulation leads directly to copying paragraphs from ChatGPT.	Integrate self-reflection logs detailing exactly what was edited.

To understand the mechanism of AI as substitute and AI as assistant, this study devided the 4 theme that already being described previously into 2 domains to construct a pedagogical framework or strategy in order to create a practical step to foster EFL students critical thinking. In addition, researcher dig dive the previous topic to practical SOPs in EFL academic writing class.

Deconstructing The Scaffolding Paradox and cognitive offloading

Using Vygotsky's (1978) Sociocultural Theory, this study dissects the cognitive mechanics of the dependency trap, concentrating on "The Scaffolding Paradox" (Hamamra et al., 2026). Scaffolding defined as A More Knowledgeable Other (MKO) who offer temporary assists and support a student in completing tasks and learn new skill. The gradual withdrawal of this help, referred to as "fading," which compels the learner to assume greater cognitive responsibility and attain autonomous mastery, is an essential component of effective scaffolding. However, the systematic evaluation shows that EFL writers get an endless, effortless, and non-fading scaffold when they engage with ChatGPT or Quillbot.

Students experience systemic cognitive offloading when they avoid "desirable difficulties," which are the constructive cognitive challenges necessary for long-term schema development and language recall (Jahan & Akbar, 2026). Active thinking is separated from the physical process of writing. An "illusion of competence" is produced as a result. Students think their writing skills have increased because the final article appears perfect, but in actuality, their independent drafting abilities have drastically declined, a condition known as "skills atrophy" according to multiple studies (Mekheimer, 2025). Thus, the Scaffolding Paradox demonstrates that frictionless AI scaffolds will unavoidably function as cognitive crutches rather than learning accelerators unless fading mechanisms are actively incorporated into the pedagogy.

Pedagogical Mitigation: Fostering Critical AI Literacy

To break the cycle of dependency and foster student critical thinking and agency, this study synthesizes AI literacy into two. First, Functional AI Literacy**,** it focuses on basic operational competence, including prompt engineering, synonym generation, and grammar-checking. At this phase, students utilize AI as a mechanical drafting tool to optimize speed and productivity. However, they lack the critical awareness to interrogate the output, leaving them highly vulnerable to translation errors and grammatical complacency (Lee & Jeong, 2026b). Rather than viewing AI as an artificial substitute that replaces human writers, students can utilize Gen AI to do a better problem solving and critical thinking by utilizing it as sparring partner to challenge assumptions, explore alternative perspective and generate diverse hypothesis**.**

Second**,** Critical AI Literacy in application mean that Students actively evaluate AI outputs, identify grammatical biases, detect reference hallucinations, and verify historical claims against primary academic literature. This dimension mitigates the automation bias that leads to blind copy-pasting (Stanko et al., 2026; Al-Absi & Abdi, 2026). To ensure that students maintain their cognitive agency when working with AI, we propose a standardized Human-in-the-Loop (HITL) Standard Operating Procedure. This SOP dictates that at every stage of the writing process, the human student must remain the primary decision-maker and editor, negotiating meaning with the machine rather than accepting its output passively. The HITL workflow is structured as follows: Step 1: Ideation (Human Dominant): The student must brainstorm, identify the research gap, and write an initial outline entirely by hand, without any AI assistance. This ensures the foundational thesis reflects the student's unique interest and voice; Step 2: AI Interrogation (Collaborative): The student inputs their outline into the LLM, instructing the AI to act as a skeptic and hostile reviewer (using an adversarial

prompt). The AI must identify logical flaws, weak assumptions, and generate hard questions. The student must respond to these challenges, refining their thesis. Step 3: Draft Generation (Scaffolded): The student uses the AI to generate initial drafts of specific sections using evidence-first prompting. The prompt must explicitly state: Only allow a claim if it is immediately followed by a verified citation from my provided evidence table. Flag any unsupported sentences; Step 4: Human Curation and Verification (Human Dominant): The student must read every sentence generated by the AI. They must verify all citations against primary databases, rewrite formulaic prose to inject their personal authorial voice, and actively edit for flow, tone, and register. Any AI suggestion that cannot be verified must be deleted. Step 5: Transparency and Disclosure (Ethical Compliance): The student must submit a detailed appendix disclosing how AI was utilized, including the exact prompts used, the AI's suggestions, and the student's subsequent revisions

AI literacy framework to foster students critical thinking and mitigate overreliance on EFL academic writing process.

To operationalize this Critical AI Literacy model, this systematic review compiles four standard operating procedures (SOPs) for the university writing classroom:

a. Process-Oriented Assessments: Instructors must transition away from grading only the final written product, which is highly vulnerable to AI generation. Instead, grades should evaluate the drafting process, including pre-writing outlines, tracking draft history, and grading student edit justification logs.

b. Socratic Prompting Pedagogy: Teachers must prohibit students from submitting block-rewriting prompts to AI. Students should be trained to use conversational Socratic prompting, asking Generative AI to identify grammatical errors, explain syntactic weaknesses, or outline logical fallacies in the student's original draft, forcing the student to perform the physical editing manually and write it using their own authoritical voice.

c. Hybrid Project-Based Learning (PBL): Classrooms must orchestrate a balanced hibrida workflow. Students may utilize AI during early-stage brainstorming, outline generation, and initial literature scanning. However, the writing of the core argument and the final draft must be completed during proctored, device-free, pen-and-paper writing sessions to secure independent cognitive performance.

d. AI Transparency Disclosures: Students must append an 'AI Use Statement' at the end of their assignments, explicitly documenting (1) the specific AI engines used, (2) the exact prompts submitted, (3) the raw AI output, and (4) a brief reflection justifying what was integrated, what was modified, and what was rejected during the writing process.

CONCLUSION

From 2023 to 2026, this Bibliometric-Informed Scoping Review (B-ScR) has outlined the paths toward critical AI literacy in EFL academic writing as well as the growing dilemma of GenAI dependency. We used VOSviewer to map the intellectual landscape and graphically validate a significant research gap: the cognitive-affective effects of AI are still significantly underrepresented compared to studies that focus only on technological utility. To help educators maintain human agency, we developed a framework of process using AI assisted writing on EFL academic writing class. This scoping review does, however, have some limitations. Our main search was limited to the Scopus database; in order to increase the network's breadth, future research could incorporate cross-comparative bibliometric data from Dimensions and Web of Science (WoS). Additionally, the chosen timeframe from 2023 to 2026 only captures the initial, nascent stages of the Large Language Model (LLM) boom. Because artificial intelligence (AI) models and institutional frameworks are evolving rapidly on a daily basis, the sustained, long-term impacts of autonomous agentic AI platforms remain volatile and require ongoing, longitudinal observation. Recommendations for future research based on the synthesized findings, future research agendas should focus on three primary trajectories. Initially, future investigators are encouraged to conduct long-term experimental trials to evaluate how the Human-in-the-Loop (HITL) Standard Operating Procedures (SOPs) developed in this study influence students' writing anxiety, motivation, and critical thinking skills across diverse geographical regions. Finally, higher education institutions must cooperate to build formal curricula centered on critical AI literacy, shifting institutional paradigms away from punitive, repressive bans and toward the proactive cultivation of academic integrity.

FUNDING STATEMENT
The authors did not receive any financial support to complete this article.

DECLARATION OF AI USE
The authors declare that during the preparation of this manuscript, Generative AI tools (specifically OpenAI's ChatGPT and Gemini) were utilized solely as a cognitive scaffold to enhance grammatical accuracy, refine sentence structures, and polish the academic register of the draft. The conceptual framework, search string design, screening criteria using OpenRefine and Rayyan, qualitative synthesis of the core literature, and the formulation of the HITL SOP were conceived and executed entirely by the human authors. Following the AI-assisted linguistic refinement, the authors meticulously reviewed, cross-referenced, and revised all generated content to ensure factual accuracy and academic integrity. The authors remain fully responsible for the originality, accuracy, and final content of this publication.

REFERENCES
Al-Absi, T., & Abdi, A. (2026). Empowerment and Complication in AI-Assisted Academic Writing: A Mixed-Methods Study of EFL Postgraduates. *Arab World English Journal*, *17*(2), 177–192. https://doi.org/10.24093/awej/vol17no2.12

Al Murshidi, G., Alfahid, M., & Al Zaabi, A. (2026). How do college students weigh the reliability of AI in academic writing? *Cogent Education*, *13*(1). https://doi.org/10.1080/2331186X.2026.2670040

Alshammari, S. H., & Alhamazany, A. (2026). Artificial Intelligence-Supported English as a Foreign Language Learning: A Narrative Review of Personalisation, Feedback, Engagement and Pedagogical Collaboration. *International Journal of Technology in Education and Science*, *10*(2), 219–236. https://doi.org/10.46328/ijtes.7018

Arifani, Y., Anwar, K., Maruf, N., Stryaningrum, R. W., Masduki, & Sahiruddin. (2026). Can ChatGPT Scaffold Postgraduate Students' Thesis Proposals? Voices of EFL Teachers and Students. *International Journal of Learning, Teaching and Educational Research*, *25*(2), 899–920. https://doi.org/10.26803/ijlter.25.2.40

Barrot, J. S. (2026). Harnessing Generative AI for Developing ELT Materials: Striking a Balance Between Innovation and Ethical Use. *Electronic Journal of Foreign Language Teaching*, *23*, 9–20. https://doi.org/10.56040/jsbt2312

Bouzar, A., El Idrissi, K., & Ghourdou, T. (2024). ChatGPT and Academic Writing Self-Efficacy: Unveiling Correlations and Technological Dependency among Postgraduate Students. *Arab World English Journal*, *2024-April* (Special Issue), 225–236. https://doi.org/10.24093/awej/ChatGPT.15

Cong-Lem, N. (2025). Exploring Human–Ai Distributed Critical Thinking (Hadct): Pilot Validation of a Critical Thinking Scale with Vietnamese Efl Learners. *Teflin Journal*, *36*(2), 279–302. https://doi.org/10.15639/teflinjournal.v36i2/279-302

Ha, N. T. D. (2026). EFL Students' Attitudes Towards the Use of ChatGPT in Learning How to Write Essays: Vietnamese Context. *Theory and Practice in Language Studies*, *16*(2), 464–472. https://doi.org/10.17507/tpls.1602.13

Hadra, M., Cambridge, K., & Mesbah, M. (2026). Evaluating the accuracy and reliability of AI content detectors in academic contexts. *International Journal for Educational Integrity*, *22*(1), 1–18. https://doi.org/10.1007/s40979-026-00213-1

Hamamra, B., N. Khlaif, Z., & Mahamid, N. (2026). AI literacy and mediated learner autonomy in ChatGPT-supported EFL writing: a qualitative study at a Palestinian university. *International Journal of Educational Technology in Higher Education*, *23*(1). https://doi.org/10.1186/s41239-026-00597-7

Jahan, F., & Akbar, H. (2026). Exploring ChatGPT in EFL classrooms: insights into engagement, proficiency and ethics. *Discover Education*, *5*(1). https://doi.org/10.1007/s44217-026-01158-w

Karmi, K., Pham, T. A., Pham, C. H., Tymoshchuk, N., Ivanytska, N., Nechyporenko, V., Tereshchenko, L., Venzhynovych, N., Banyoi, V., Kharkivska, O., Shkurko, H., Ilkiv, A., Wang, R., Yin, X., Ruan, J., Ma, W., Chukwuere, J. E., Van Den Berg, L., Ogunseyi, T. B., … Barjesteh, H. (2026). LLT Journal: A Journal on Language and Language Teaching

LLT Sanata Dharma University, Yogyakarta, Indonesia. *Arab World English Journal*, *15*(1), 1–21. https://doi.org/10.24507/icicelb.17.06.605

Kim, J., Lee, S. S., Detrick, R., Wang, J., & Li, N. (2026). Students-Generative AI interaction patterns and its impact on academic writing. *Journal of Computing in Higher Education*, *38*(1), 504–525. https://doi.org/10.1007/s12528-025-09444-6

Lee, Y. J., & Jeong, S. H. (2026a). Exploring University ELLs' AI Literacy: An Exploratory Descriptive Study. *Journal of Educational and Social Research*, *16*(1), 94–105. https://doi.org/10.36941/jesr-2026-0241

Lee, Y. J., & Jeong, S. H. (2026b). Implementing AI-integrated English writing: university English language learners writing strategies and perceptions. *International Journal of Evaluation and Research in Education*, *15*(1), 805–814. https://doi.org/10.11591/ijere.v15i1.35461

Liu, F. (2026). AI literacy and foreign language boredom in AI-assisted academic English learning: the roles of engagement and task value. *Frontiers in Psychology*, *17*(April). https://doi.org/10.3389/fpsyg.2026.1795301

Mekheimer, M. (2025). Generative AI-assisted feedback and EFL writing: a study on proficiency, revision frequency and writing quality. *Discover Education*, *4*(1). https://doi.org/10.1007/s44217-025-00602-7

Nguyen, H. N., Nguyen, H. M., & Do, T. P. M. (2025). EFL lecturers' experiences and perceptions towards using ChatGPT in teaching writing: A case study in Vietnam. *International Journal of Education and Practice*, *13*(3), 857–869. https://doi.org/10.18488/61.v13i3.4291

Pandey, V. D., & Solanki, S. (2026). *Governing Generative AI in Education*. https://doi.org/10.4018/979-8-2600-1613-8.ch005

Parker, J. L., Richard, V. M., Acabá, A., Escoffier, S., Flaherty, S., Jablonka, S., & Becker, K. P. (2025). Negotiating Meaning with Machines: AI's Role in Doctoral Writing Pedagogy. *International Journal of Artificial Intelligence in Education*, *35*(3), 1218–1238. https://doi.org/10.1007/s40593-024-00425-x

Pham, L. T. T. (2026). Metadiscourse in ChatGPT-generated and human-written research articles in linguistics. *Indonesian Journal of Applied Linguistics*, *15*(3), 561–572. https://doi.org/10.17509/shgzwd41

Pretorius, M. (2026). 'I want the text to still be me.' *English Text Construction*, *January*. https://doi.org/10.1075/etc.25001.pre

Puspita, R. D., Wardani, D. S., Nurhayati, E., & Moid, I. B. (2025). *E NHANCING A CADEMIC W RITING IN HIGHER EDUCATION: T HE I MPACT OF C HAT GPT AND Q UILLBOT IN AI-D RIVEN P ROJECT -B ASED L EARNING*. *24*.

Rahayu, A., Tarihoran, N., Rahmawati, E., Muslihah, E., Ma'mur, I., & Anita. (2025). Navigating the Digital Writing Landscape: EFL Students' Perspectives on ChatGPT Utilization. *International Journal of Learning, Teaching and Educational Research*, *24*(5), 727–752. https://doi.org/10.26803/ijlter.24.5.37

Stanko, D., Kolegaeva, I., & Rohach, L. (2026). Developing Critical AI Literacy in Academic Writing Classrooms. *Arab World English Journal*, *2026-January* (2), 41–54. https://doi.org/10.24093/awej/AI3.3

Sung, S., Ding, X., Jiang, R., Sereiviene, E., & ... (2024). Using artificial intelligence teaching assistants to guide students in solar energy engineering design. *Journal of* …. https://doi.org/10.1080/10899995.2024.2384340

Thi, X. H. N., Thien, H. V. H., Vuong, K. N., & Nguyen, T. T. (2025). Enhancing writing skills through AI-powered tools: perceived benefits and challenges among Vietnamese EFL students. *Discover Education*, *4*(1). https://doi.org/10.1007/s44217-025-00905-9

Tran, H. (2026). How Does AI Work as a "Game Changer"? Self-Determination and Influencing Factors in EFL Students' Use of AI for Academic Writing. *Technology in Language Teaching and Learning*, *8*. https://doi.org/10.29140/tltl.2026.103723

Waziana, W., Andewi, W., Wibisono, D., Hastomo, T., & Muslihudin, M. (2025). Exploring ChatGPT's Impact on Critical, Creative, and Reflective Thinking Skills: A Mixed-Methods Study in an Indonesian EFL Classroom. *Applied Research on English Language*, *14*(4), 77–114. https://doi.org/10.22108/are.2025.145896.2564

Zhang, S., & Liu, X. (2025). Learner Emotions in AI-assisted english as a second/foreign language learning: a systematic review of empirical studies. In *Frontiers in Psychology*. frontiersin.org. https://doi.org/10.3389/fpsyg.2025.1652806