Ying Luo, Yuwei Xia and Xiaofei Lu*

A Systematic Review of the Role of Artificial Intelligence in Second Language Writing Education

https://doi.org/10.1515/dsll-2025-0001
Received January 1, 2025; accepted October 13, 2025; published online November 10, 2025

Abstract: Second language (L2) writing education is undergoing rapid transformation with the increasing integration of artificial intelligence (AI) tools. While research has explored AI applications across diverse educational domains, there remains a lack of systematic reviews addressing the implementation, benefits, challenges, and emerging trends of AI integration in L2 writing education. This review synthesizes 39 empirical studies published between 2019 and 2024, guided by three central questions: (1) What AI tools have been employed, and how have they been integrated into L2 writing education? (2) What benefits and challenges are associated with their use? (3) What emerging trends are shaping the future of AI integration in L2 writing education? Our findings reveal that three major tool types – automated writing evaluation systems, large language models, and specialized writing assistants – are predominantly used to support writing processes, provide feedback, and enhance learner engagement. Most tools were integrated through teacher-guided designs, highlighting the importance of structured mediation. While learners reported gains in grammar accuracy, motivation, and revision ownership, challenges included over-reliance, limited development of higher-order skills, and ethical concerns. Emerging trends point to a shift from surface-level correction toward deeper, metacognitive uses of AI and the embedding of AI in genre-based and reflective pedagogies. However, these developments remain uneven and contingent on instructional design. This review concludes by advocating for theory-informed, equity-oriented integration of AI tools, positioning them not as substitutes for instruction, but as mediational actors within a broader pedagogical ecology.

***Corresponding author: Xiaofei Lu**, Department of Applied Linguistics, The Pennsylvania State University, University Park, PA, USA, E-mail: xxl13@psu.edu. https://orcid.org/0000-0003-2365-2581
Ying Luo, Department of Applied Linguistics, The Pennsylvania State University, University Park, PA, USA, E-mail: yxl869@psu.edu. https://orcid.org/0009-0002-9751-4376
Yuwei Xia, Department of Curriculum and Instruction, The Pennsylvania State University, University Park, PA, USA, E-mail: yxx5304@psu.edu. https://orcid.org/0009-0007-5960-3948

Open Access. © 2025 the author(s), published by De Gruyter on behalf of Chongqing University, China
This work is licensed under the Creative Commons Attribution 4.0 International License.

Keywords: artificial intelligence; systematic review; second language writing education

1 Introduction

Artificial intelligence (AI) technologies have evolved from being a supportive pedagogical tool to becoming an integral aspect of second language (L2) teaching and learning (Hadizadeh 2024; Warschauer and Xu 2024). Against this backdrop, the integration of AI has also marked a pivotal turning point in L2 writing education, with tools such as automated writing evaluation (AWE) systems, specialized writing assistants, chatbots, and generative AI (GenAI) systems reshaping how L2 writing is taught and learned. Recent studies have reported multifaceted uses of AI tools in L2 writing education, such as providing automated feedback used either on its own or along with peer or teacher feedback (e.g., Guo and Wang 2024), supporting various stages (e.g., brainstorming, drafting, and revision) of the writing process (e.g., Hwang et al. 2023), and offering targeted assistance for specialized tasks or applications (e.g., Kwon et al. 2023; Li and Tarp 2024; Mohammad et al. 2023), among others. Existing research has also identified several benefits of integrating AI into L2 writing education, including reduced writing anxiety, increased motivation, and enhanced writing skills (e.g., Biju et al. 2024; Huang et al. 2024; Ibrahim and Kirkpatrick 2024; Liang et al. 2024; Wang 2024). Reflecting these benefits, learners have reported satisfaction with the use of AI tools in L2 writing (e.g., Al-Alami 2024). At the same time, the integration of AI tools in L2 writing education is not without challenges, with concerns related to equity of access, academic integrity, and over-reliance frequently discussed in existing studies (e.g., Al-Alami 2024; Barrot 2023; Zou and Huang 2024).

Several recent systematic reviews, position papers, and meta-analyses have examined the integration of AI technologies in language education in general or the effects of specific AI applications in L2 writing education (Huang et al. 2024; Li 2025; Lo et al. 2024; Pack and Maloney 2024; Warschauer and Xu 2024). However, there remains a lack of systematic reviews addressing the implementation, benefits, challenges, and emerging trends of AI integration in L2 writing education. To address this gap, this study systematically reviews research on AI integration in L2 writing education published between January 1, 2019 and December 1, 2024. Our review is guided by three central questions: (1) What AI tools have been employed, and how have they been integrated into L2 writing education? (2) What benefits and challenges are associated with their use? (3) What emerging trends are shaping the future of AI integration in L2 writing education? Synthesizing findings from 39 empirical studies that meet rigorous selection criteria, this review provides insights into these

critical issues and highlights several future research directions aimed at optimizing AI integration in L2 writing education to enhance L2 writers' learning experiences and outcomes.

2 Literature Review

The development of L2 writing skills poses significant challenges for both learners and instructors. Successful L2 writing requires mastery of multiple interconnected competencies, such as linguistic accuracy, content development, organizational skills, and genre awareness, among many others (e.g., Chang 2016; Liu et al. 2023). Traditional pedagogical approaches often struggle to deliver sufficient personalized support throughout the various stages of the writing process, especially in contexts constrained by large class sizes and limited instructional time. Along with the lack of adequate individualized, real-time support, studies have also documented high levels of writing anxiety among L2 learners (e.g., Al-Alami 2024; Liu et al. 2023).

Recent advancements in AI technologies present promising solutions to the challenges of L2 writing education through various affordances. Notably, AI tools have shown promise in supporting various dimensions of writing development, such as lexical richness, grammatical accuracy, syntactic complexity, cohesion, coherence, text organization, and self-regulated revision (Boudouaia et al. 2024; Guo and Wang 2024; Strobl et al. 2024; Tran 2025). At the same time, concerns about academic integrity and plagiarism have emerged as main challenges in the integration of AI technologies into L2 writing education, with some scholars calling for more reliable methods for detecting AI-generated content (e.g., Alexander et al. 2023; Yan 2023). Other challenges identified in empirical studies include issues of accessibility, over-reliance on AI, authorial voice, and questions about the robustness of AI tools in supporting higher-order writing skills (e.g., Barrot 2023; Esfandiari and Allaf-Akbary 2024; Zou and Huang 2024).

Importantly, recent studies have moved beyond cataloguing the general benefits and drawbacks of AI tools to examine how they can be meaningfully embedded into L2 writing instruction. AI rewriting tools have been used to promote micro-level refinement of learner writing in both in-class or at-home contexts. For example, Al Mahmud (2023) reported that, as learners revised multiple drafts, Wordtune helped them produce more vivid and concrete vocabulary as well as more syntactically complex constructions. As a step further, AWE systems such as Criterion and custom-built platforms such as the Multi-strategy Computer-assisted EFL Writing Learning System (MsCAEWL), have been integrated into extended interventions to support the learners' revision process with varying degrees of teacher scaffolding and learner autonomy (Chen et al. 2024; Han and Sari 2024). For example, Chen et al. (2024)

embedded MsCAEWL, a deep learning-enhanced system, into a 60-day intervention with high school English as a Foreign Language (EFL) learners. Following instruction and training from the teacher, the learners received analytic, trait-specific feedback from the system on thesis statements, content, linguistic complexity, and accuracy over multiple independent feedback-revision cycles, leading to significant improvement in writing quality. By contrast, Han and Sari (2024) implemented combined teacher–Criterion feedback in a college-level EFL course. Learners receiving such hybrid feedback showed greater gains in grammar and mechanics, though not in content development, than those receiving teacher feedback only.

While GenAI systems such as ChatGPT have some similar functions to AWE systems, they have been employed in more exploratory and dialogic ways. For example, Strobl et al. (2024) implemented a two-week classroom-based intervention in an advanced L2 German writing module, where students compared their own summaries with ChatGPT-generated texts and revised their summaries accordingly. Through rubric-based evaluation and analysis of revision data, the study revealed students' increasing awareness of ChatGPT's linguistic features and limitations, and their critical engagement with the model during revision. Relatedly, customized AI systems, such as Microsoft Copilot and Hwang et al.'s (2023) customized AI-enhanced writing platform, have also been implemented in dialogic ways to support discourse-level instruction and more personalized writing development. Esfandiari and Allaf-Akbary (2024) used Microsoft Copilot to help learners develop the ability to use interactional metadiscourse markers in a 10-session, instructor-led, data-driven learning program. Specifically, students participated in a series of in-class, prompt-based activities with the system to explore the forms and functions of interactional metadiscourse markers, leading to significant gains in their understanding and use of such markers. Hwang et al. (2023) integrated a multimodal, AI-powered system that offered recognition technologies, vocabulary assistance, and real-time feedback into teacher-led sessions and app-based writing activities in a seven-week online EFL writing program focused on narrative tasks. With the assistance of this customized AI-enhanced writing platform, students revised multiple drafts and showed measurable improvements in vocabulary use, sentence structure, and clarity of expression.

In sum, the integration of AI tools into L2 writing instruction reveals not a monolithic trajectory but a spectrum of pedagogical possibilities shaped by tool design, instructional context, and teacher mediation. While rewriting tools like Wordtune promote micro-level refinement through autonomous practice, AWE systems such as Criterion and MsCAEWL offer structured, trait-specific feedback that supports iterative improvement. GenAI systems like ChatGPT are increasingly positioned as dialogic partners rather than merely evaluative agents, enabling deeper content engagement and reflective revision when coupled with human

feedback. Meanwhile, customized AI systems illustrate the growing potential of AI to support discourse-level awareness and task personalization, though their success remains closely tied to teacher guidance. These developments underscore a shift in AI-assisted L2 writing instruction from tool-centered deployment to context-sensitive integration that is driven by pedagogical purposes and that highlights the importance of human-AI complementarity.

Some reviews and position papers have focused on different aspects of the integration of AI technologies in language education in general. For example, Lo et al. (2024) reviewed 70 empirical studies on the use of ChatGPT in EFL/ESL education within a 1.5-year period following its release. They reported several affordances (e.g., increased learning opportunities) and drawbacks (e.g., incorrect information) of ChatGPT use in EFL/ESL education and identified several gaps in existing research in this area (e.g., lack of longer-term studies). Pack and Maloney (2024) provided an in-depth discussion of the importance to raise teacher awareness of a range of important issues, such as ethical considerations, writing skill development, and reliability of AI-generated output, among others, in using AI in language education. In their introduction to the special issue on AI for language learning in *Language Learning & Technology*, which included eight studies exploring different applications of AI in language learning, Warschauer and Xu (2024) pointed out several changing features of AI that should be considered, such as its amplified power, increased flexibility, and integral role in every-day reading, writing, and knowledge production. Several meta-analyses have systematically investigated the effects of specific applications of AI technologies on different aspects of language learning or L2 writing education. For example, a meta-analysis by Wang et al. (2024) reported that AI-guided individualized language learning effectively promoted L2 learners' language development, and one by Huang et al. (2024) found that AI technologies can significantly reduce writing anxiety and computer anxiety.

Collectively, these comprehensive reviews and meta-analyses have offered valuable insights into the benefits, affordances, and challenges of AI integration in language education as well as the benefits of specific AI applications in L2 writing education. They have also raised several issues that should be attended in future research on AI integration in language education. Meanwhile, they have either only included L2 writing education as a small part of the review or focused on specific AI applications in promoting a specific aspect of L2 writing. They have also placed more focus on the capabilities of AI tools, the benefits of AI integration, and the issues to be considered in AI integration, with relatively less attention to the pedagogical implementation of AI tools. A comprehensive review addressing the implementation, benefits, challenges, and emerging trends of AI integration in L2 writing education remains lacking.

This study addresses this gap by providing a systematic review of empirical studies on AI integration in L2 writing education. By synthesizing findings from empirical studies conducted in diverse theoretical frameworks, the review aims to provide a holistic understanding of the ways in which AI tools have been implemented in L2 writing pedagogy, the benefits they provide to L2 writers, and the limitations and challenges faced by L2 writing teachers and learners in their use. In addition to illuminating the current state of AI integration in L2 writing education, the review further aims to identify emerging trends and productive directions to inform future research in this rapidly evolving field.

3 Methodology

The current study adopts a systematic approach to identify, evaluate, and synthesize peer-reviewed journal articles on the integration of AI tools into L2 writing education. Guided by established systematic review protocols (Kitchenham and Charters 2007; Xiao and Watson 2019) and aligned with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Page et al. 2021), the review was conducted in multiple structured phases.

To ensure broad and interdisciplinary coverage across applied linguistics, language education, and AI, we searched six major databases: Web of Science, Scopus, ProQuest, IEEE Xplore, ERIC, and LLBA. These databases were chosen for their complementary strengths: Scopus and ProQuest offer extensive coverage of peer-reviewed journals in education and technology; Web of Science and IEEE Xplore index a wide range of AI-related publications; and ERIC and LLBA specialize in educational and linguistic research. This combination has also been adopted in recent systematic reviews on AI and language education (e.g., Lo et al. 2024), supporting both the rationale and replicability of our database selection.

Two sets of core terms were used to search all databases. The first consisted of AI-related terms, including “artificial intelligence”, “AI writing tools”, “ChatGPT”, “automated writing evaluation”, and “large language models”, and the second consisted of writing-related terms, including “L2 writing”, “second language writing”, and “academic writing”. These terms were joined with the Boolean operator OR, and the truncation symbols were used where appropriate to expand coverage (e.g., “model*” captures both “model” and “models”). In ERIC, we also incorporated technology- and writing-related terms from the database’s built-in system of subject-specific vocabulary (e.g., “computer assisted instruction” and “writing instruction”). Based on the preliminary search results, we added “L2 composition” to queries in Web of Science and Scopus, and “academic discourse” to those in LLBA, as they appeared as keywords in multiple potentially relevant studies retrieved from the

respective databases using the core terms. In addition to database searches, we conducted backward reference tracing (i.e., snowball searching) from the reference lists of key empirical studies and review articles to identify potentially relevant sources not captured by initial queries. This combined strategy helped ensure comprehensive coverage of the recent surge in research on GenAI tools, particularly ChatGPT, in L2 writing education.

This review covers publications from January 1, 2019, to December 1, 2024, a period marked by rapid advances in AI technologies and their growing impact on language education (Strobl et al. 2024; Zou and Huang 2024). This timeframe encompasses the emergence and widespread adoption of GenAI tools such as ChatGPT, which have significantly transformed L2 writing education (Lo et al. 2024; Zadorozhnyy and Lai 2023).

To ensure both relevance and methodological rigor, studies were included only if they met the following criteria: (1) published in peer-reviewed journals between January 2019 and December 2024; (2) explicitly described a research design (e.g., experimental, quasi-experimental, case study, mixed-methods, or survey); (3) provided sufficient methodological detail regarding participants, data collection, and analysis; (4) reported empirical research based on systematically collected and analyzed data; and (5) explicitly integrated AI tools into the pedagogical design of L2 writing instruction. Integration was defined as the use of AI tools for instructional purposes, whether in-class or as structured out-of-class activities aligned with course objectives. Studies in which AI use was entirely self-initiated by students without pedagogical support or curricular alignment were excluded. This intentional focus on pedagogical integration distinguishes our review from others that explore autonomous or incidental student use of AI tools. While such studies can offer insights into learner behavior or attitudes, our goal was to synthesize empirical evidence on how AI tools have been pedagogically enacted to support instructional design, optimize teaching strategies, and enhance structured learning outcomes. This focus ensures the relevance and practical value of the review for educators and curriculum developers seeking to integrate AI tools meaningfully in formal educational settings.

The initial search across the six databases yielded 295 records. After removing 110 duplicates, 185 unique records remained for title and abstract screening. Sixty-five articles were excluded at this stage due to irrelevance, including those that were not empirical studies in peer-reviewed journals (n = 22) and that mentioned AI tools but did not describe their application in writing contexts (n = 15) or did not focus on L2 writing education (n = 28). The remaining 120 articles underwent full-text review. Of these, 78 were excluded for not meeting one or more inclusion criteria: lacking a clear research design (n = 7), providing insufficient methodological detail (n = 14), reporting no empirical data (n = 13), or failing to clearly integrate AI into the

pedagogical design of L2 writing instruction or assessment (n = 47). The 39 remaining studies that met all inclusion criteria were included in the synthesis. The literature selection process is illustrated in a PRISMA-style flow diagram in Figure 1.

While this systematic approach allowed for a refined survey of closely related literature, several limitations should be acknowledged, consistent with challenges commonly reported in educational technology research syntheses (Newman and Gough 2020). Variations in database search functionalities and customized queries may have introduced minor inconsistencies in the search results, despite efforts to standardize procedures across the databases. Additionally, the focus on peer-reviewed journal articles published in English may have limited the diversity of perspectives captured. Furthermore, although explicit inclusion criteria were specified, the decisions to exclude some studies were inevitably based on subjective

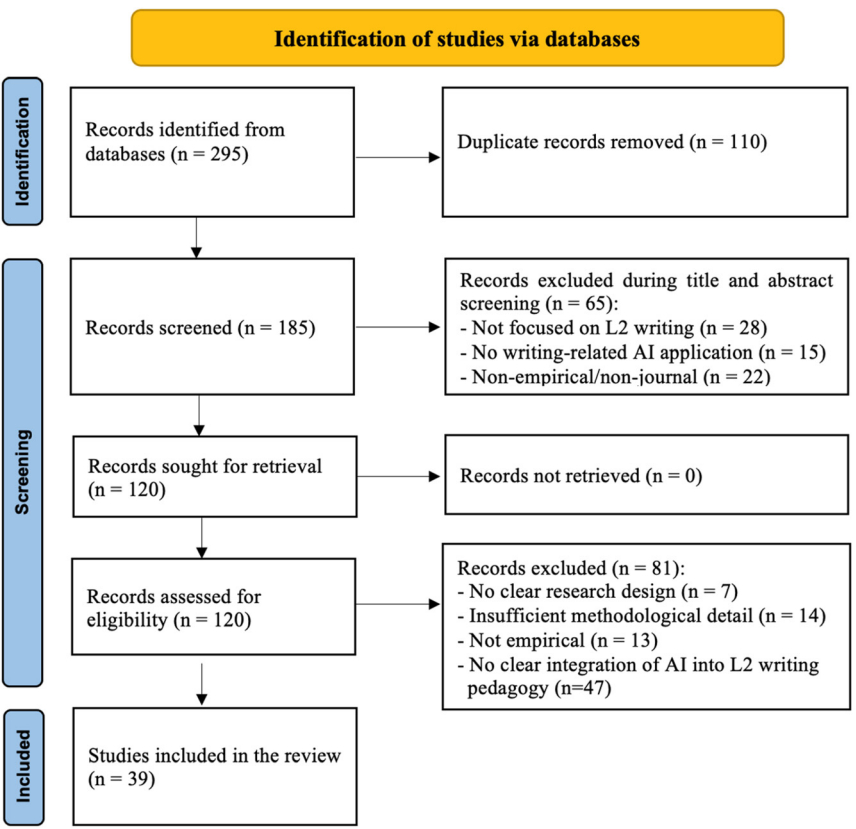

Figure 1: PRISMA flow diagram of the literature selection process.

interpretations of relevance. These limitations notwithstanding, the methodology adopted in the current study is replicable and produced a sizable, representative body of relevant studies for further analysis.

All 39 studies were coded by the first author in line with the study's three research questions. Coding focused on eight dimensions: AI tool type, integration purpose(s), integration procedures, pedagogical outcomes, benefit, perception focus, challenges/concerns, and trend focus. Two rounds of coding were conducted, separated by a two-week interval to allow for reflective distancing and refinement of category definitions. Although inter-coder reliability was not assessed due to the single-coder design, prior research has shown that systematic solo coding with iterative checking and detailed audit trails can yield reliable and valid results, especially when guided by well-defined analytical questions (e.g., Miles et al. 2014; O'Cathain 2010). The complete coding of all 39 studies is provided by the Appendix.

Our analysis of the final set of 39 studies followed principles of qualitative research synthesis (Thomas and Harden 2008). Thematic analysis was applied to identify patterns across the coded dimensions, including study objectives, participant characteristics, research contexts, methodological designs, types of AI tools used, and their instructional functions and outcomes. This analytic process enabled us to synthesize findings across studies and address the three guiding research questions: (1) how AI tools are used in L2 writing education, (2) what pedagogical benefits and challenges are associated with their use, and (3) what emerging trends are shaping AI integration in L2 writing.

4 Findings

4.1 Research Question 1: Integration of AI Tools in L2 Writing Education

Our first research question explored which AI tools have been integrated into L2 writing education and how they have been applied. The systematic review of 39 empirical studies revealed three major categories of AI tools used in L2 writing education (see Figure 2), as well as three main approaches for their integration. Among the 39 reviewed studies, three major categories of AI tools were identified: customized AI systems, automated writing evaluation (AWE) tools, and large language model (LLM)-based generators. Customized AI tools were reported in 9 studies (23.1 %), typically involving tailored systems designed for specific classroom contexts, such as genre-based feedback platforms or AI-supported peer review systems.

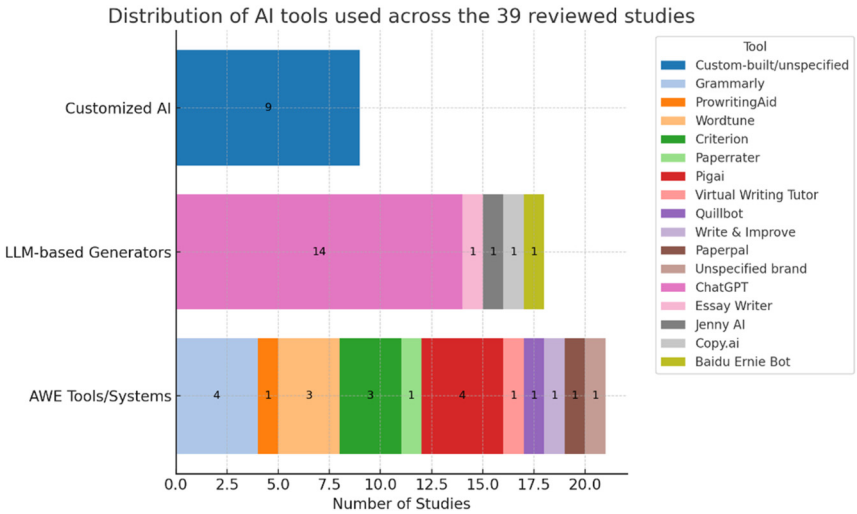

Figure 2: Distribution of AI tools used across the 39 reviewed studies (some studies reported multiple tools).

AWE tools, broadly defined to include both traditional and modern rewriting assistants, appeared in 22 studies (56.4 %). This group included Grammarly (4 studies, 10.2 %), Wordtune (3 studies, 7.7 %), Pigai (4 studies, 10.2 %), Criterion (3 studies, 7.7 %), and other tools such as ProwritingAid, Paperpal, Paperrater, Quillbot, Virtual Writing Tutor, Write & Improve, and unspecified AWE systems (each used in one study, 2.5 %). These tools typically provided automated feedback on grammar, vocabulary, punctuation, and sentence structure, with some also offering suggestions on cohesion, clarity, and lexical variety. While more traditional systems such as Pigai and Criterion focused primarily on surface-level accuracy and mechanical correctness, newer rewriting assistants like Wordtune and Quillbot emphasized sentence fluency and stylistic refinement. Overall, the AWE tools reviewed functioned as real-time evaluators and revision aids, helping learners iteratively improve their drafts with minimal teacher intervention.

LLM-based generators, including tools that support full-text generation based on prompts, were used in 18 studies (46.2 %). This category includes ChatGPT (14 studies), Essay Writer, Jenny AI, Copy.ai, and Baidu Ernie Bot (each used in 1 study, 2.5 %). These tools are primarily designed to generate coherent, fluent text in response to user inputs, allowing learners to receive instant drafts, reworded paragraphs, or extended content suggestions. In educational contexts, they were employed to support idea generation, lexical expansion, and structural scaffolding, especially during the early stages of writing or revision. Some tools, like ChatGPT and

Ernie Bot, also offered interactive dialogue-based assistance, enabling students to refine their drafts iteratively through conversational feedback.

Across the reviewed studies, three pedagogical purposes dominated AI tool integration: supporting writing processes, facilitating feedback provision, and enhancing learner engagement and strategy use. The most prevalent purpose was writing process support (35 studies, 89.7 %). AI tools were integrated into different stages of writing in the reviewed studies, including brainstorming, drafting, revising, and editing, to accelerate idea generation, guide text organization, and prompt self-reflection. For instance, Parker et al. (2025) found that doctoral students leveraged ChatGPT and a custom AI tool as co-authors during brainstorming and revision, enhancing metacognitive engagement and reinforcing authorial agency. Rad et al. (2023) similarly reported that EFL students using Wordtune outside class improved their revision engagement and feedback literacy, with structured reflection journals fostering active decision-making and a stronger sense of control during writing.

A substantial portion of the reviewed studies (28 studies, 71.8 %) integrated AI tools to provide feedback that targets both surface-level language features and discourse-level development. AWE systems such as Grammarly, Criterion, and Pigai were widely used to provide timely, consistent, and individualized feedback to facilitate micro-level revision and reduce teacher workload (e.g., Sari and Han 2024; Xu and Zhang 2022; Yildiz and Gonen 2024; Zhang 2020). A few studies also adopted more specialized tools to scaffold higher-order rhetorical competencies. The web-based module for formative assessment in Sawaki et al. (2024), for example, provided feedback and support on content alignment in summary writing. Genre-informed systems such as Link et al.'s (2024) feedback engine delivered move-level guidance aligned with academic discourse conventions, enabling learners to revise with genre-specific awareness.

Additionally, 19 studies (48.7 %) embedded AI use within reflective routines and strategy instruction to enhance learners' critical engagement and use of self-regulatory strategies. For instance, Lee (2024) designed a blended feedback sequence in which students interpreted ChatGPT feedback in relation to peer and teacher suggestions, fostering metacognitive reflection and ownership of revisions. Similarly, Rad et al. (2023) incorporated structured writing logs to encourage learner reflection on AI-generated suggestions, thereby promoting feedback literacy and sustained revision behaviors. These pedagogical approaches signal a paradigm shift from passive feedback consumption to active feedback negotiation.

4.2 Research Question 2: Benefits and Challenges of Integrating AI Tools in L2 Writing Education

The second research question examined the benefits and challenges associated with the integration of AI tools in L2 writing education. Our review identified seven major categories of benefits and four primary types of challenges (see Figure 3), each supported by varying degrees of empirical evidence from the reviewed studies.

4.2.1 Benefits

A prominent benefit across the reviewed studies was the provision of tailored, real-time feedback, reported in 26 studies (66.7 %). AI-mediated responses, particularly those generated by AWE systems such as Grammarly or GenAI systems such as ChatGPT, were frequently described as specific, immediate, and personalized, offering actionable suggestions on important aspects of writing such as language use, coherence, and text organization. Such real-time feedback not only reduced revision time but also encouraged learner autonomy through on-demand correction and

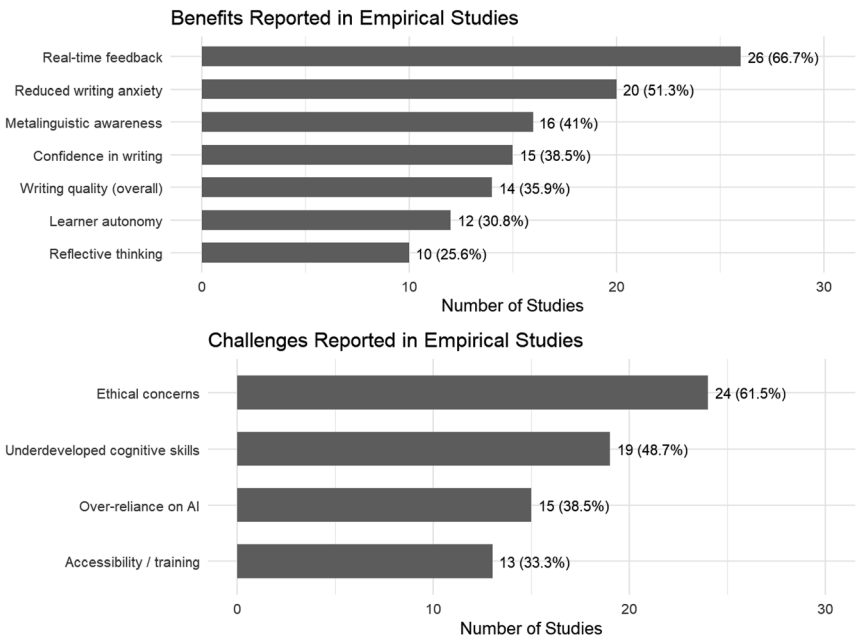

Figure 3: Distribution of reported benefits and challenges of using AI tools in L2 writing education across the 39 reviewed studies (some studies reported multiple benefits or challenges).

reflection. For example, Biju et al. (2024) reported that real-time AI feedback in an assessment context helped improve learners' attitudes about language acquisition and increase their motivation. Similarly, Tsai et al. (2024) found that the instant, varied and personalized feedback provided by ChatGPT removed the wait for instructors' feedback and allowed learners to consider and apply appropriate revisions in a timely manner.

The second most common benefit was reduced writing anxiety (20 studies, 51.3 %). Several studies highlighted how AI tools offered a low-pressure environment for experimentation and revision. For example, Liu et al. (2023) found that students in an AI-assisted writing practicum described feeling more motivated during the drafting process, attributing their reduced stress and fear of making mistakes to the opportunity for repeated practices provided by AWE feedback.

Several additional benefits were identified among the reviewed studies. Enhanced metalinguistic awareness appeared in 16 studies (41 %), often as a result of students reflecting on lexical and rhetorical choices prompted by AI-generated suggestions. For example, Esfandiari and Allaf-Akbary (2024) reported that targeted feedback from Microsoft Copilot, combined with reflective prompts, enhanced students' awareness of genre conventions and interactional metadiscourse. Increased confidence in writing, cited in 15 studies (38.5 %), was frequently linked to students' perception of AI tools as supportive rather than punitive. For example, Biju et al. (2024) emphasized that the immediacy and clarity of AI assessment fostered more positive attitudes and self-assurance. Furthermore, 14 studies (35.9 %) reported improvements in overall writing quality, often attributed to AI's capacity to support different dimensions of writing. Twelve studies (30.8 %) highlighted improved learner autonomy, suggesting that repeated interaction with AI feedback encouraged learners to take greater ownership over the revision process. For example, Liu et al. (2023) illustrated how students progressed from passively responding to automated feedback to more actively engaging in self-regulated revision cycles. Finally, 10 studies (25.6 %) mentioned the usefulness of AI tools in boosting reflective thinking.

4.2.2 Challenges

Among the various challenges associated with AI integration in L2 writing education, ethical concerns, such as fears about academic dishonesty and blurred lines of authorship, emerged as the most frequently reported issue, appearing in 24 studies (61.5 %). For example, Yan (2023) noted students' concerns about academic integrity and the risk of plagiarism when relying too heavily on AI-generated text, particularly in the absence of clear ethical guidance.

A closely related pedagogical challenge involved the underdevelopment of higher-order cognitive skills, particularly critical thinking, argument construction, and originality of ideas, as documented in 19 studies (48.7 %). For example, Koltovskaia (2020) found that AI-supported revisions often resulted in shallow textual edits, lacking substantive progress in ideation or logical progression.

The issue of over-reliance on AI tools was noted in 15 studies (38.5 %), especially among learners with limited proficiency or writing confidence. Guo, Feng, and Hua (2022) indicated that such students could rely on AI suggestions uncritically, thereby bypassing essential processes of reflection and meaning-making. Moreover, 13 studies (33.3 %) cited accessibility-related barriers such as insufficient digital literacy, limited tool-specific training, or unfamiliarity with interface functionalities that inhibited learners from making informed use of AI resources. Marzuki et al. (2023), for instance, noted that learners without adequate AI training often struggled with interpreting system-generated suggestions and defaulted to surface-level lexical changes.

4.3 Research Question 3: Emerging Trends of AI Integration in L2 Writing Education

Recent research on AI integration in L2 writing education exhibits an increasingly diverse landscape in terms of tool types, pedagogical functions, and instructional design. Among the different types of AI tools, AWE systems (e.g., Grammarly and Criterion) featured in over half of the reviewed studies. However, a notable uptick in the use of GenAI tools such as ChatGPT and specialized AI platforms such as Jenny.ai has emerged in studies published in 2023–2024. This trend signals a transition from form-focused evaluation tools toward generative systems capable of providing meaning-level scaffolding. For example, Kurt and Kurt (2024) illustrated how ChatGPT supported both surface-level corrections and organization-level revision within in-class writing routines. Because of such capabilities, ChatGPT has been employed to supplement teacher feedback in studies that adopt a hybrid instructional model mixing human and AI mediation (Parker et al. 2025). These developments indicate not only a widening range of AI applications but also a conceptual broadening of what constitutes feedback and mediation in L2 writing classrooms.

Beyond tool types, emerging trends point to a strong movement toward pedagogically embedded, reflective uses of AI. Integration is no longer conceived as a technical add-on but increasingly as a curricular component with explicit developmental goals. Several studies (e.g., Lee 2024; Parker et al. 2025; Rad et al. 2023) reported the use of structured prompting routines, reflective writing logs, and

teacher-guided interactions with AI. These practices reflect a pedagogical shift toward repurposing AI to foster metacognitive development by prompting students to re-evaluate higher-order concepts such as argumentation, coherence, and rhetorical stance. Such designs suggest an alignment with learning-as-development perspectives, extending the role of AI beyond output optimization toward deeper forms of learner transformation.

Another notable trend concerns the convergence of AI and assessment, particularly through the use of AI for formative feedback and the cultivation of feedback literacy. Studies such as Rad et al. (2023), Parker et al. (2025), and Biju et al. (2024) proposed AI-assisted writing as a viable alternative to traditional assessment practices. These works advocated repositioning AI from a tool of control to a means of formative dialogue. This conceptual shift from judgment to negotiation resonates with sociocultural paradigms of language learning, which emphasize mediation, reflection, and learner agency (Lantolf and Thorne 2006). In this sense, AI is not merely automating evaluation but reconfiguring the social function of assessment.

Furthermore, adaptive and personalized feedback systems are gaining traction. Chen et al. (2024) and Liu et al. (2023) reported on AI tools tailored to learner proficiency levels and genre-specific conventions, implemented through teacher-designed intervention protocols. These practices reflect a broader trend toward intelligent, context-sensitive mediation, where AI tools function not as generic assistants but as dynamically responsive agents. Crucially, such systems are best embedded in iterative revision cycles, signaling an emerging model of AI use as a dialogic co-constructor rather than a static proofreader.

In sum, the trajectory of AI integration in L2 writing education is evolving from surface-level correction toward deeper pedagogical alignment, reflective engagement, and conceptual transformation. Rather than focusing exclusively on tool effectiveness – already addressed in RQ1 – this section highlights broader developmental logics and instructional rationales. As the field matures, future research should examine how AI mediates identity, authorship, and epistemic stance in writing, and how such mediations are shaped by institutional, cultural, and technological infrastructures. These inquiries will be crucial in determining whether AI serves as a tool of liberation or containment in L2 writing education.

5 Discussion

This review confirms that while AI tools are increasingly embedded in L2 writing classrooms, their pedagogical impact remains uneven. This section reflects on deeper tensions related to tool functionality, learner development, and the evolving role of mediation. One central insight from this review is the increasing conceptualization

of AI not merely as a corrective tool but as a dialogic partner in the writing process. Studies such as Parker et al. (2025) and Guo et al. (2024) showed that when integrated with structured tasks and guided reflection, AI tools can facilitate metacognitive engagement and genre awareness. They argued that the pedagogical value of AI tools in L2 writing education lies less in output polishing and more in their potential to scaffold cognitive processes. This argument aligns with sociocultural perspectives that stress the mediational role of tools in learning. For instance, in van Lier's (2004) concept of "semiotic mediation," tools shape not only behavior but cognition. Similarly, Lantolf and Thorne (2006) argued that tools embedded in goal-directed, socially situated activity become agents of development rather than instruments of delivery.

However, the limitations of AI as a learning scaffold must be critically acknowledged. Studies such as Al Mahmud (2023) and Koltovskaia (2020) cautioned that many AI tools remain insufficiently equipped to support the development of higher-order skills such as argumentation, conceptual elaboration, and audience awareness. Such tools often generate feedback that is either too generic or overly form-focused, which may lead students to conflate surface fluency with rhetorical sophistication. Without human mediation, such interactions risk reinforcing shallow revision habits. This concern echoes Vygotsky's (1978) assertion that effective mediation must move learners through the zone of proximal development (ZPD) to facilitate learning, as well as Swain's (2006) claim that dialogic mediation must create the conditions for internalization – not simply deliver corrections.

A further tension arises from the dual nature of AI's immediacy: while real-time feedback can reduce anxiety and increase learner autonomy (e.g., Biju et al. 2024; Liu et al. 2023), it may also short-circuit the reflective pause essential for deep learning. Ellis (2010) warned that excessively implicit feedback may limit cognitive processing. From a sociocultural lens, this dynamic disrupts the recursive process of meaning-making, potentially flattening the dialogic space needed for development. Thus, educators must balance immediacy with productive delay, and efficiency with intentional disruption, to sustain the generative tension of the writing process.

In light of these dynamics, it is crucial to rethink how AI integration is framed within instructional design. Our findings support the call by Pitychoutis (2024) for reimagining AI not as a replacement for traditional methods but as a means of pedagogical transformation. When embedded within genre-based instruction, iterative revision cycles, and teacher-guided protocols, AI tools can become powerful mediators of learning. Such approaches shift the focus from tool functionality to tool positioning: not what the tool does, but how and when it is used, and with what kind of pedagogical intentionality. This resonates with principles of dynamic assessment (Poehner 2008), which emphasize mediation tailored to learner responsiveness rather than fixed input. It also reflects what Swain (2006) terms the creation of

dialogic spaces, where learners appropriate and reshape mediational means rather than simply receive them.

Finally, this review highlights a critical gap between design and equity. As noted in Khazanchi et al. (2024), access to AI-supported instruction remains uneven, particularly in under-resourced settings. While many studies reported on university-level implementations, there is limited attention to how AI tools can be adapted for learners with limited digital literacy, unstable connectivity, or culturally diverse rhetorical expectations. From a Universal Design for Learning (UDL) perspective (Meyer et al. 2014), equitable AI integration must account for learner variability – not by standardizing tools, but by diversifying support structures. Without such commitments, AI risks deepening stratification rather than expanding opportunity.

Taken together, the findings of this review underscore the importance of moving beyond technocentric discourse. AI in L2 writing education is not a panacea, nor is it a threat to be resisted. Rather, it is a pedagogical actor whose value depends on the intentionality, reflexivity, and equity with which it is integrated. As Swain (2006) reminds us, language learning is not only a matter of input and output, but of conceptual development through dialogic mediation. This lens urges us to see AI not as a shortcut to accuracy, but as a potential co-constructor of meaning – provided its use is situated within ethically grounded, pedagogically sound, and theoretically informed frameworks.

6 Future Directions

Building on the findings and tensions discussed above, this review identifies several promising directions for future research on AI tools in L2 writing education. As AI technology continues to evolve, there is a critical need for longitudinal studies that assess its long-term impact on L2 learners' writing development. While existing studies (e.g., Nawaz et al. 2024; Song and Song 2023; Wei et al. 2023) provided valuable insights into immediate impacts or short-term benefits, the sustained effects of AI-assisted instruction on writing motivation, coherence, argumentation, and disciplinary identity remain underexplored. Future work should investigate how these tools influence not only performance, but developmental trajectories over time.

Another pressing area is the exploration of hybrid feedback models. Studies like Zeevy-Solovey (2024) and Liu et al. (2023) suggested that blending peer, teacher, and AI-generated feedback may yield synergies that surpass any single source. Future research could examine how these feedback types interact, how learners interpret and prioritize them, and how instructional scaffolds can support productive engagement across modes. Such inquiries would clarify how human-AI feedback partnerships can best support deep learning.

The ethical and pedagogical implications of AI use must remain a central concern. As discussed in Barrot (2023) and Pack and Maloney (2024), fears about plagiarism, dependency, and authorship demand not only regulatory responses but pedagogical ones. Future studies should examine how institutions and instructors can design usage policies, curricula, and reflection tasks that cultivate learners' critical AI literacy – ensuring transparency, responsible use, and ethical awareness.

Access and equity also require urgent attention. Several studies (e.g., Kurt and Kurt 2024) documented infrastructural and pedagogical disparities in AI access. Future research should explore how to design cost-effective, adaptable, and context-sensitive AI systems and teacher training models that serve marginalized learners, rather than privileging already-resourced populations.

Finally, this review highlights a critical gap between design and equity. As noted in Marzuki et al. (2023), access to AI-supported writing instruction remains uneven, particularly in under-resourced or digitally marginalized contexts. While many studies reported on university-level implementations, there is limited attention to how AI tools can be adapted for learners with limited digital literacy, unstable connectivity, or culturally diverse rhetorical expectations. From a UDL (Meyer et al. 2014), equitable AI integration must account for learner variability – not by standardizing tools, but by diversifying support structures. Without such commitments, AI risks deepening stratification rather than expanding opportunity.

7 Conclusions

This systematic review has examined the integration of AI tools into L2 writing education, revealing a rapidly evolving field marked by both pedagogical innovation and theoretical ambiguity. Across diverse instructional contexts, AI has shown potential to mediate learning processes, facilitate revision practices, and enhance learner engagement. However, its contributions remain uneven – contingent upon how tools are positioned within pedagogical design, what forms of mediation are enacted, and whether learners are supported to move beyond surface features toward conceptual transformation.

Rather than treating AI as a static technology, this review has argued for its reconceptualization as a pedagogical actor whose value emerges through dialogic interaction, task-embedded use, and intentional mediation. While AI can provide efficient, accessible feedback, it cannot replace the reflective, socially situated dimensions of learning that human guidance sustains. The findings highlight a fundamental tension: AI's capacity to automate does not guarantee its capacity to educate. Only when deployed within responsive, equity-oriented frameworks can AI

support learners' developmental trajectories rather than reinforce existing limitations.

In sum, this review contributes to a growing body of scholarship that resists technocentric narratives and calls for deeper theoretical grounding in the design, use, and evaluation of AI-mediated writing support. As L2 writing instruction continues to adapt to technological shifts, researchers and educators must remain attuned to the affordances and constraints of AI – not as end points, but as starting points for reimagining pedagogy. The future of AI in L2 writing lies not in more automation, but in more thoughtful mediation.

Research ethics: Not applicable.
Informed consent: Not applicable.
Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Use of Large Language Models, AI and Machine Learning Tools: During the preparation of this manuscript, the authors utilized ChatGPT for proofreading and polishing. All suggested revisions were carefully reviewed, and selected changes were incorporated as deemed appropriate. The authors assume full responsibility for the content of the publication.
Conflict of interest: None declared.
Research funding: None declared.
Data availability: Not applicable.

References

Alexander, Katarzyna, Christine Savvidou, and Chris Alexander. 2023. "Who wrote this essay? Detecting AI-generated writing in second language education in higher education." *Teaching English with Technology* 23 (4): 89–107.

Al-Alami, Suhair E. 2024. "EFL learners' attitudes towards utilizing ChatGPT for acquiring writing skills in higher education: a case study of computing students." *Journal of Language Teaching and Research* 15 (4): 1029–38.

Al Mahmud, Fawaz. 2023. "Investigating EFL students' writing skills through artificial intelligence: wordtune application as a tool." *Journal of Language Teaching and Research* 14 (5): 1395–404.

Barrot, Jessie S. 2023. "Using ChatGPT for second language writing: pitfalls and potentials." *Assessing Writing* 57: 100745.

Biju, Neha, Nasser Said Gomaa Abdelrasheed, Khilola Bakiyeva, K. D. V. Prasad, and Biruk Jember. 2024. "Which one? AI-assisted language assessment or paper format: an exploration of the impacts on foreign language anxiety, learning attitudes, motivation, and writing performance." *Language Testing in Asia* 14: 45.

Boudouaia, Azzeddine, Samia Mouas, and Bochra Kouider. 2024. “A study on ChatGPT-4 as an innovative approach to enhancing English as a foreign language writing learning.” *Journal of Educational Computing Research* 62 (6): 1289–317.

Chang, Tzu-Shan. 2016. “The use of genre-based cycle with L1 rhetorical structure in L2 writing class: an exploratory study.” *Studies in English Language Teaching* 4 (4): 518.

Chen, Binbin, Lina Bao, Rui Zhang, Jingyu Zhang, Feng Liu, Shuai Wang, et al. 2024. “A multi-strategy computer-assisted EFL writing learning system with deep learning incorporated and its effects on learning: a writing feedback perspective.” *Journal of Educational Computing Research* 61 (8): 1596–638.

Ellis, Rod. 2010. “EPILOGUE: a framework for investigating oral and written corrective feedback.” *Studies in Second Language Acquisition* 32 (2): 335–49.

Esfandiari, Rajab, and Omid Allaf-Akbary. 2024. “Assessing interactional metadiscourse in EFL writing through intelligent data-driven learning: the Microsoft Copilot in the spotlight.” *Language Testing in Asia* 14 (1): 51.

Guo, Kai, and Deliang Wang. 2024. “To resist it or to embrace it? Examining ChatGPT’s potential to support teacher feedback in EFL writing.” *Education and Information Technologies* 29 (7): 8435–63.

Guo, Qian, Ruiling Feng, and Hua Yuanfang. 2022. “How effectively can EFL students use automated written corrective feedback (AWCF) in research writing?” *Computer Assisted Language Learning* 35 (9): 2312–31.

Guo, Kai, Mengru Pan, Yuanke Li, and Chun Lai. 2024. “Effects of an AI-Supported approach to peer feedback on university EFL students’ feedback quality and writing ability.” *The Internet and Higher Education* 63: 100962.

Hadizadeh, Abbas. 2024. “ChatGPT, the end of L2 academic writing or a blessing in disguise?” *Acuity: Journal of English Language Pedagogy, Literature and Culture* 9 (2): 183–201.

Han, Turgay, and Elif Sari. 2024. “An investigation on the use of automated feedback in Turkish EFL students’ writing classes.” *Computer Assisted Language Learning* 37 (4): 961–85.

Huang, Xiaoli, Wei Xu, Li Fan, and Zhonggen Yu. 2024. “A meta-analysis of effects of automated writing evaluation on anxiety, motivation, and second language writing skills.” *The Asia-Pacific Education Researcher* 33 (4): 957–76.

Hwang, Wu-Yuin, Rio Nurtantyana, Siska Wati Dewi Purba, Uun Hariyanti, Yuniar Indrihapsari, and Herman Dwi Surjono. 2023. “AI and recognition technologies to facilitate English as foreign language writing for supporting personalization and contextualization in authentic contexts.” *Journal of Educational Computing Research* 61 (5): 1008–35.

Ibrahim, Karim, and Robert Kirkpatrick. 2024. “Potentials and implications of ChatGPT for ESL writing instruction.” *IRRODL* 25 (3): 394–409.

Khazanchi, Rashmi, Daniele Di Mitri, and Hendrik Drachsler. 2024. “The effect of AI-based systems on mathematics achievement in rural context: a quantitative study.” *Journal of Computer Assisted Learning* 41 (1): e13098.

Kitchenham, Barbara, and Stuart Charters. 2007. *Guidelines for Performing Systematic Literature Reviews in Software Engineering*. (Technical report EBSE 2007-001). Keele University and Durham University.

Koltovskaia, Svetlana. 2020. “Student engagement with automated written corrective feedback (AWCF) provided by grammarly: a multiple case study.” *Assessing Writing* 44: 100450.

Kurt, Gokce, and Yavuz Kurt. 2024. “Enhancing L2 writing skills: ChatGPT as an automated feedback tool.” *Journal of Information Technology Education: Research* 23 (24): 1–17.

Kwon, S. K., D. Shin, and Y. Lee. 2023. “The application of chatbot as an L2 writing practice tool.” *Language Learning & Technology* 27 (1): 1–19.

Lantolf, James P., and Steven L. Thorne. 2006. *Sociocultural Theory and the Genesis of Second Language Development*. Oxford: Oxford University Press.

Lee, Yoo-Jean. 2024. "Can my writing be polished further? When ChatGPT meets human touch." *ELT Journal* 78 (4): 401–13.

Li, Shaofeng. 2025. "Generative AI and second language writing." *Digital Studies in Language and Literature* 2 (1): 122–52.

Li, Qian, and Sven Tarp. 2024. "Using generative AI to provide high-quality lexicographic assistance to Chinese learners of English." *Lexikos* 34: 397–418.

Liang, Jiaozhi, Fang Huang, and Timothy Teo. 2024. "Understanding Chinese university EFL learners' perceptions of AI in English writing." *International Journal of Computer-Assisted Language Learning and Teaching* 14 (1): 1–16.

Link, Stephanie, Robert Redmon, Yaser Shamsi, and Martin Hagan. 2024. "Generating genre-based automatic feedback on English for research publication purposes." *CALICO Journal* 41 (3): 319–46.

Liu, Chen-Chen, Shi-Jie Liu, Gwo-Jen Hwang, Yun-Fang Tu, Youmei Wang, and Naini Wang. 2023. "Engaging EFL students' critical thinking tendency and in-depth reflection in technology-based writing contexts: a peer assessment-incorporated automatic evaluation approach." *Education and Information Technologies* 28 (10): 13027–52.

Lo, Chung Kwan, Philip Leung Ho Yu, Simin Xu, Davy Tsz Kit Ng, and Morris Siu-Yung Jong. 2024. "Exploring the application of ChatGPT in ESL/EFL education and related research issues: a systematic review of empirical studies." *Smart Learning Environments* 11 (1): 50.

Marzuki, Utami Widiati, Diyenti Rusdin, Inda Indrawati Darwin, Nadaraj Govender, and I. Indrawati. 2023. "The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective." *Cogent Education* 10 (2): 2236469.

Meyer, A., D. H. Rose, and D. Gordon. 2014. *Universal Design for Learning: Theory and Practice*. Wakefield, MA: CAST Professional Publishing.

Miles, Matthew B., A. Michael Huberman, and Johnny Saldaña. 2014. *Qualitative Data Analysis: A Methods Sourcebook*, 3rd ed. Thousand Oaks: Sage.

Mohammad, Taj, A. Alzubi, Mohd Nazim, and Soada Idris Khan. 2023. "EFL paraphrasing skills with QuillBot: unveiling students enthusiasm and insights." *Journal of Pedagogical Research* 7 (5): 359–73.

Nawaz, Samsudeen Sabraz, Mohamed Buhary Fathima Sanjeetha, Ghadah Al Murshidi, Mohamed Ismail Mohamed Riyath, Fadhilah Bt Mat Yamin, Rusith Mohamed, et al. 2024. "Acceptance of ChatGPT by undergraduates in Sri Lanka: a hybrid approach of SEM-ANN." *Interactive Technology and Smart Education* 21 (4): 546–70.

Newman, M., and D. Gough. 2020. "Systematic reviews in educational research: methodology, perspectives and application." In *Systematic Reviews in Educational Research*, edited by Olaf Zawacki-Richter, Michael Kerres, Sandra Bedenlier, Marcus Bond, and Katja Buntins, 3–22. Wiesbaden: Springer VS.

O'Cathain, Alicia. 2010. "Assessing the quality of mixed methods research: towards a comprehensive framework." In *SAGE Handbook of Mixed Methods in Social & Behavioral Research*, edited by Abbas Tashakkori, and Charles Teddlie, 531–55. Thousand Oaks: Sage.

Pack, Austin, and Jeffrey Maloney. 2024. "Using artificial intelligence in TESOL: some ethical and pedagogical considerations." *TESOL Quarterly* 58 (2): 1007–18.

Page, Matthew J., Joanne E. McKenzie, Patrick M. Bossuyt, Isabelle Boutron, Tammy C. Hoffmann, Cynthia D. Mulrow, et al. 2021. "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews." *BMJ* 372 (8248): 71.

Parker, Jessica L., Veronica M. Richard, Alexandra Acabá, Sierra Escoffier, Stephen Flaherty, Shannon Jablonka, et al. 2025. "Negotiating meaning with machines: AI's role in doctoral writing pedagogy." *International Journal of Artificial Intelligence in Education* 35: 1218–38.

Pitychoutis, Konstantinos M. 2024. "Harnessing AI chatbots for EFL essay writing: a paradigm shift in language pedagogy." *Arab World English Journal Special Issue on ChatGPT*: 197–209. https://doi.org/10.24093/awej/ChatGPT.13.

Poehner, Matthew E. 2008. *Dynamic Assessment: A Vygotskian Approach to Understanding and Promoting L2 Development*. New York: Springer.

Rad, Hanieh Shafiee, Rasoul Alipour, and Aliakbar Jafarpour. 2023. "Using artificial intelligence to foster students' writing feedback literacy, engagement, and outcome: a case of wordtune application." *Interactive Learning Environments* 32 (9): 5020–40.

Sari, Elif, and Turgay Han. 2024. "The impact of automated writing evaluation on English as a foreign language learners' writing self-efficacy, self-regulation, anxiety, and performance." *Journal of Computer Assisted Learning* 40 (5): 2065–80.

Sawaki, Yasuyo, Yutaka Ishii, Hiroaki Yamada, and Takenobu Tokunaga. 2024. "Developing and validating an online module for formative assessment of summary writing with automated content feedback for EFL academic writing instruction." *Language Testing in Asia* 14: 50.

Song, Cuiping, and Yanping Song. 2023. "Enhancing academic writing skills and motivation: assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students." *Frontiers in Psychology* 14: 1260843.

Strobl, Carola, Iryna Menke-Bazhutkina, Niklas Abel, and Marije Michel. 2024. "Adopting ChatGPT as a writing buddy in the advanced L2 writing class." *Technology in Language Teaching & Learning* 6: 1168.

Swain, Merrill. 2006. "Languaging, agency and collaboration in advanced second language proficiency." In *Advanced Language Learning: The Contribution of Halliday and Vygotsky*, edited by Heidi Byrnes, 95–108. London: Continuum.

Thomas, James, and Angela Harden. 2008. "Methods for the thematic synthesis of qualitative research in systematic reviews." *BMC Medical Research Methodology* 8 (1): 45.

Tran, Thi Thanh Thao. 2025. "Enhancing EFL writing revision practices: the impact of AI- and teacher-generated feedback and their sequences." *Education Sciences* 15 (2): 232.

Tsai, Chung-You, Yi-Ti Lin, and Iain Kelsall Brown. 2024. "Impacts of ChatGPT-assisted writing for EFL English majors: feasibility and challenges." *Education and Information Technologies* 29 (17): 22427–45.

van Lier, Leo. 2004. *The Ecology and Semiotics of Language Learning: A Sociocultural Perspective*. Boston: Springer.

Vygotsky, Lev S. 1978. *Mind in Society: The Development of Higher Psychological Processes*. Cambridge: Harvard University Press.

Wang, Yijen. 2024. "Cognitive and sociocultural dynamics of self-regulated use of machine translation and generative AI tools in academic EFL writing." *System* 126: 103505.

Wang, Feifei, Alan Cheung, Amanda Neitzel, and Ching Sing Chai. 2024. "Does chatting with chatbots improve language learning performance? A meta-analysis of Chatbot-assisted language learning." *Review of Educational Research* 95 (4): 623–60.

Warschauer, Mark, and Ying Xu. 2024. "Generative AI for language learning: entering a new era." *Language Learning & Technology* 28 (2): 1–4.

Wei, Ping, Xiaosai Wang, and Hui Dong. 2023. "The impact of automated writing evaluation on second language writing skills of Chinese EFL learners: a randomized controlled trial." *Frontiers in Psychology* 14: 1249991.

Xiao, Yu, and Maria Watson. 2019. "Guidance on conducting a systematic literature review." *Journal of Planning Education and Research* 39 (1): 93–112.

Xu, Jinfen, and Shanshan Zhang. 2022. "Understanding AWE feedback and English writing of learners with different proficiency levels in an EFL classroom: a sociocultural perspective." *The Asia-Pacific Education Researcher* 31 (4): 357–67.
Yan, Dan. 2023. "Impact of ChatGPT on learners in a L2 writing practicum: an exploratory investigation." *Education and Information Technologies* 28 (11): 13943–67.
Yildiz, Hatice, and S. Ipek Kuru Gonen. 2024. "Automated writing evaluation system for feedback in the digital world: an online learning opportunity for English as a foreign language students." *Turkish Online Journal of Distance Education* 25 (3): 183–206.
Zadorozhnyy, Artem, and Wan Yee Winsy Lai. 2023. "ChatGPT and L2 written communication: a game-changer or just another tool?" *Languages* 9 (1): 5.
Zeevy-Solovey, Orly. 2024. "Comparing peer, ChatGPT, and teacher corrective feedback in EFL writing: students' perceptions and preferences." *Technology in Language Teaching & Learning* 6 (3): 1482.
Zhang, Zhe. 2020. "Engaging with automated writing evaluation (AWE) feedback on L2 writing: student perceptions and revisions." *Assessing Writing* 44: 100456.
Zou, Min, and Li Huang. 2024. "The impact of ChatGPT on L2 writing and expected responses: voice from doctoral students." *Education and Information Technologies* 29 (11): 13201–19.

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/dsll-2025-0001).

Bionotes

Ying Luo
Department of Applied Linguistics, The Pennsylvania State University, University Park, PA, USA
yxl869@psu.edu
https://orcid.org/0009-0002-9751-4376

Ying Luo is a Ph.D. student in Applied Linguistics at Penn State University, focusing on advancing AI literacy and competency in language education. Her research explores AI-driven language learning, particularly how AI-human interactions create personalized environments to support ESL/EFL learning. Ying integrates corpus linguistic methods and mixed-methods approaches to examine AI's practical applications and theoretical impacts, aiming to enhance the quality, relevance, and accessibility of language education for diverse learners.

Yuwei Xia
Department of Curriculum and Instruction, The Pennsylvania State University, University Park, PA, USA
yxx5304@psu.edu
https://orcid.org/0009-0007-5960-3948

Yuwei Xia is a Ph.D. candidate at the College of Education at Pennsylvania State University. She holds an M.A. in Applied Linguistics from Columbia University. Specializing in Curriculum and Instruction with an emphasis on second language education, her research focuses on Vygotskian approaches to second and foreign language pedagogy, second language teaching and assessment, theories of second language

acquisition, and technology-integrated learning in second and foreign language classrooms. Her interests extend to vocabulary acquisition and reading-writing development in both Chinese and English.

Xiaofei Lu
Department of Applied Linguistics, The Pennsylvania State University, University Park, PA, USA
xxl13@psu.edu
https://orcid.org/0000-0003-2365-2581

Xiaofei Lu is the George C. and Jane G. Greer Professor of Applied Linguistics and Asian Studies at The Pennsylvania State University. His research interests are primarily in corpus linguistics, computer-assisted language learning, English for Academic Purposes, second language acquisition, and second language writing. He is the author of *Corpus Linguistics and Second Language Acquisition: Perspectives, Issues, and Findings* (2023, Routledge) and *Computational Methods for Corpus Annotation and Analysis* (2014, Springer).