Exploring the Landscape of Generative AI (ChatGPT)-Powered Writing Instruction in English as a Foreign Language Education: A Scoping Review

Feiwen Xiao (肖斐文) [ID]
The Pennsylvania State University

Siyu Zhu (朱思宇)
The Hong Kong Polytechnic University

Wen Xin (辛闻)
The University of Kansas

Abstract

Purpose: Recent advancements in artificial intelligence (AI)-powered technologies, particularly ChatGPT, have sparked significant interest in English as a Foreign Language (EFL) education. This review aims to explore the landscape of ChatGPT's application in EFL writing.

Design/Approach/Method: Given the nascence of this field, the study conducts a scoping review by analyzing 16 empirical studies published before December 2023 to investigate the role of ChatGPT in EFL writing.

Findings: The review explores the current and potential uses of ChatGPT in EFL writing, highlighting its dual role as both a writing assistant and an assessment tool. On one hand, ChatGPT is widely acknowledged for providing real-time feedback that enhances writing quality and efficiency. On the other hand, challenges and concerns remain prevalent.

Corresponding author:
Feiwen Xiao, Department of Learning and Performance Systems, The Pennsylvania State University, 199 Fischer Rd, University Park, PA 16802, USA.
Email: ffx5014@psu.edu

Originality/Value: The findings reveal key gaps in the literature, such as the need for more inter-disciplinary research, the adaptation of AI models to meet the linguistic and cultural needs of EFL learners, and the integration of multimodal AI tools. The review emphasizes the importance of critical thinking and information literacy training for educators and students while addressing ethical considerations. These insights offer a roadmap for future research and the practical implementation of AI in language education, providing valuable guidance for different stakeholders.

Introduction

Recent developments and advancements in artificial intelligence (AI) and natural language processing have given rise to many AI-powered language technologies. ChatGPT, as one of the most forefront generative artificial intelligence (GenAI) tools, has received widespread attention due to its capability of generating human-like conversations in real time. While AI-driven conversational agents (CAs), such as Siri, Alexa, and Google Assistant have been widely used across a variety of domains for decades, they are limited by their rule-based approaches and often struggle with complex queries (Kusal et al., 2022; Wellnhammer et al., 2020; Xiao et al., 2023). In contrast, ChatGPT (OpenAI, 2023) distinguishes itself as a more advanced conversational system and can generate more contextually appropriate responses. Such a feature has unsurprisingly been embraced by English as a Foreign Language (EFL) learners because the application can not only offer them a personalized and interactive learning experience (Ray, 2023), but also has the potential to address some common challenges EFL learners face in the learning of writing, including but not limited to a lack of constructive feedback for revisions (Zhang & Hyland, 2023), insufficient metacognitive support for higher-level writing skills, such as planning and organization (Aldabbus & Almansouri, 2022), and inadequate vocabulary and grammatical knowledge (Gayed et al., 2022). For example, research has indicated the usefulness and effectiveness of ChatGPT in brainstorming (Barrett & Pack, 2023), enhancing lexical diversity (e.g., Kartal, 2023), proofreading (e.g., Özçeliˇk, 2023), and improving writing efficiency (e.g., Yan, 2023). Moreover, the application can also help students improve their writing skills by providing examples of well-structured essays and guiding them through the revision process (Hong, 2023; Kim et al., 2023). Despite its usefulness, studies have also raised concerns about students' overreliance on the tool, potentially impeding the development of language skills and self-regulation (Barrot, 2023). Research also suggests that while ChatGPT provides rich information and assists in writing, it also presents risks of plagiarism, misuse, and

ethical concerns related to generating misinformation and biased content (AlAfnan et al., 2023; Zimmerman, 2023). In other words, current research perspectives on Generative Pre-training Transformer (GPT) in writing education are diverse, but the results are mixed. These necessitate a scoping review of the use of ChatGPT in EFL writing practices. Conducting a scoping review can yield a more holistic understanding of ChatGPT's applications in EFL writing contexts and contribute to a refined view of its impact on language education. This knowledge will be valuable to language educators, policymakers, and researchers in this evolving field.

Rationale

Although published studies have explored the applications and challenges of ChatGPT in EFL contexts from multiple perspectives, few studies have synthesized and analyzed the existing body of literature to offer an overview of the current state of affairs of ChatGPT in foreign language education. This scoping review addresses this gap by reviewing empirical studies on the applications, benefits, and challenges of ChatGPT in foreign language education. By providing a comprehensive understanding of the use of the tool in EFL writing with a focus on how such a tool can be leveraged to enhance EFL learning outcomes, this review contributes to the growing literature on AI-powered writing tools in foreign language education, offering pedagogical insights, recommendations, and directions for future research.

Review questions

This review specifically addresses the following four questions:

1. What are the general characteristics of the existing literature on ChatGPT in EFL writing?
2. How is ChatGPT currently being used in EFL writing?
3. What are the prevailing attitudes toward the use of ChatGPT in EFL writing?
4. What are the key benefits and challenges associated with incorporating ChatGPT in EFL writing?

Methodology

A scoping review was carried out to explore the landscape of ChatGPT in EFL writing. Such a methodology was chosen due to its suitability for mapping out the breadth and depth of research in emerging fields (Munn et al., 2018). This approach allows the inclusion of a wide range of published studies, providing a holistic view of the research landscape (Peters et al., 2021).

To conduct the scoping review, a systematic search of academic databases, including Google Scholar, Web of Science, Scopus, and ProQuest was performed using the key terms in Table 1.

Table 1. Search terms.

Technology terms	Language learning terms
ChatGPT	Writing
ChatGPT (3.5)	Automated Essay Evaluation (AEE)
Generative AI	Assessing Writing
Generative Artificial Intelligence	Writing Assessment
Artificial Intelligent	Feedback
Conversational Agent	English as a Foreign Language
Dialogue System	EFL
Chatbot	ESL
Conversational AI	Bilingual
Natural Language Processing	Non-native Speaker
Large Language Model	Learn English

Specifically, the search identified potential literature using the following Boolean conjunction: ("ChatGPT" OR "ChatGPT (3.5)" OR "Generative AI" OR "Generative Artificial Intelligence" OR "Artificial Intelligence" OR "Conversational Agent" OR "Chatbot" OR "Conversational AI" OR "Large Language Model") AND ("English as a Foreign Language" OR "EFL" OR "English as a Second Language" OR "ESL" OR "Bilingual" OR "Non-native Speaker" OR "Learn English") AND ("Writing" OR "Automated Essay Evaluation" OR "AEE" OR "Assessing Writing" OR "Writing Assessment" OR "Feedback"). Drawing upon the snowball technique (Ridley, 2012), manual searches of key journals in the fields of artificial intelligence, education, and language learning, were also conducted to ensure a comprehensive coverage.

A total of 2,310 candidates were identified through both automated and manual searches. The criteria in Table 2 were then followed to exclude studies that were irrelevant to the scope of the review. For example, conceptual studies and research that focused on Generative AI tools other than ChatGPT were excluded. The employment of the criteria screened out 108 articles. The full texts of all these candidates were then reviewed to determine their eligibility for inclusion in the review.

A total of 16 pieces of scholarship were identified as suitable for the scoping review. Key information from each article, including research design, research site, participants, sample size, and the status of English (ESL vs. EFL), was extracted. Table 3 provides a detailed overview of the articles.

Results

General characteristics of the literature reviewed (Review question 1)

As shown in Table 3, the reviewed literature can be characterized by a broad geographic distribution, diverse research methodologies, and an exclusive focus on ChatGPT. All the studies were

Table 2. Inclusion and exclusion criteria for literature search.

Criterion	Inclusion criteria	Exclusion criteria
Empirical focus	Empirical studies investigating applications, challenges, and potentials of ChatGPT in English as a Foreign Language (EFL)/ English as a Second Language (ESL) writing	Non-empirical studies such as theoretical papers and opinion pieces
Scope	Studies focusing on ChatGPT in ESL/EFL writing	Studies not focusing on AI-powered writing assistant tools or their applications, challenges, and potentials in ESL/EFL writing
Participants	Participants who are either EFL or ESL learners	
Language of publication	Studies published in English	Studies published in languages other than English

Note. This review was carried out in December 2023; as a result, studies published after the date were not included.

published in 2023, highlighting the recent surge of interest in ChatGPT in language education as well as the nascence of the technology. The research predominantly employed mixed methods (47%) and explanatory designs (33%). A noticeable trend across these studies is the involvement of both teachers and students as participants, though a significant portion of the studies focuses on university students (60%). Geographically, the research spanned various countries, including the United States, Vietnam, China, Japan, Korea, and Arabic-speaking regions, with a concentration in Asia (approximately 67% of the studies). Sample sizes vary widely, ranging from small-scale studies with fewer than 10 participants (20%) to large-scale research, such as Mizumoto and Eguchi (2023) involving over 12,000 student essays. While the majority of studies on this topic are journal articles, a notable exception is Schlosser's (2023) master's thesis. All the research was conducted in the setting of universities, many of which explored participants' perceptions of and experiences with the efficacy of ChatGPT in supporting language learning, particularly in improving writing skills.

Uses of ChatGPT in EFL writing (Review question 2)

Two themes emerged from the review regarding the use of ChatGPT in EFL writing: ChatGPT has been used as a writing assistance tool and a writing assessment tool.

ChatGPT as a writing assistance tool. Though not designed for the purpose of writing development, the review found that ChatGPT has emerged as a versatile writing assistance tool in facilitating EFL

Table 3. Studies included in this scoping review.

Citation	Country	Study design	Setting	Participants	ESL/ EFL	Sample size
Barrett and Pack (2023)	United States	Cross-sectional questionnaire design	University	Teachers ($n = 68$) and students ($n = 158$)	ESL	246
Nguyen (2023)	Vietnam	Mixed method	University	Teachers	EFL	20
Schlosser (2023)	Spain	Mixed method	University	University students	EFL	42
Xiao and Zhi (2023)	China	Explanatory study	University	University students	EFL	5
Yan (2023)	China	Multi-method qualitative approach	University	University students	EFL	35
Mohamed (2024)	Japan	Mixed method	University	University teachers	EFL	10
Nurseha (2023)	Japan	Case study	University	University students	EFL	20
Marzuki et al. (2023)	United Kingdom	Case study	University	University teachers	EFL	10
Guo and Wang (2023)	China	Explanatory study	University	University teachers ($n = 5$) and students ($n = 50$)	EFL	55
Geçkin et al. (2023)	Arab	Explanatory study	University	University teachers ($n = 5$) and students ($n = 43$)	EFL	48
Mizumoto and Eguchi (2023)	International	Qualitative study	University	Students' essay ($n = 12,100$)	EFL	12,100
Song and Song (2023)	China	Randomized controlled trial	University	University students ($n = 50$)	EFL	50
Hawanti and Zubaydulloevna (2023)	Indonesia	Quasi-experiment	University	University students ($n = 73$)	EFL	73

Table 3. (continued)

Citation	Country	Study design	Setting	Participants	ESL/ EFL	Sample size
Al-Garaady and Mahyoob (2023)	Arab	Mixed method	University	University students ($n = 88$)	EFL	88
Schmidt-Fajlik (2023)	Japan	Explanatory study	University	University students	EFL	69
Bok and Cho (2023)	Korea	Mixed method	University	University students	EFL	71

students' writing, with its generative nature and various pathways to improving their writing (Alharbi, 2023). EFL learners have harnessed ChatGPT to brainstorm, improve coherence, and enhance grammar. A finding across the studies is that learners actively used ChatGPT to generate ideas during the brainstorming stage, though they often critically evaluated and refined the outputs rather than adopting them outright (Xiao & Zhi, 2023). Additionally, students found that providing detailed prompts yielded more useful responses. They emphasized that specific, positive, and structured prompts helped optimize the tool's capabilities. This finding suggests that learners adopt an iterative approach, where ChatGPT serves as an idea generator, but the students take an active role in revising and contextualizing the output (Xiao & Zhi, 2023; Yan, 2023).

Likewise, Yan (2023) identified that students used ChatGPT basically as a text generator, paraphraser, grammar checker, and summarizer. Furthermore, the process of using ChatGPT appears to follow developmental phases, beginning with simple, exploratory tasks and advancing to more complex writing practices. As learners become more familiar with ChatGPT, they integrate additional AI tools to ultimately streamline their writing workflows to avoid plagiarism and enhance text quality. Across the board, the studies suggest that ChatGPT serves as a valuable writing assistant by supporting idea generation, offering immediate feedback, and allowing users to interact critically with the content it produces. However, it requires clear instructions and user interventions to maximize its potential.

ChatGPT as a writing assessment tool. Some studies highlight the potential of ChatGPT to streamline writing feedback and assessment (Al-Garaady & Mahyoob, 2023; Schmidt-Fajlik, 2023). Although not a replacement for human instructors, ChatGPT complements the limitations of human instructors by offering real-time error detection and grammar checking. Bok and Cho (2023) surveyed 73 EFL learners and found that more than half appreciated ChatGPT's usefulness in detecting errors related to vocabulary, accuracy, flow, and grammar.

However, when comparing ChatGPT to human instructors, Al-Garaady and Mahyoob (2023) found that while ChatGPT was effective in identifying superficial errors, it struggled with deep structural and pragmatic issues. For instance, it accurately flagged redundant adverbials and misordered elements, but it was less adept at interpreting the learners' intentions behind certain errors. Moreover, ChatGPT also missed certain grammatical errors, such as article usage and preposition omission, although it provided corrected versions afterward. Despite this, the number of errors identified by ChatGPT exceeded those identified by human instructors, but the effectiveness of these identifications varied based on the quality of prompts used, suggesting that refining prompts could improve accuracy.

Schmidt-Fajlik (2023) compared ChatGPT to other grammar checkers and found that ChatGPT, along with other tools, can effectively detect errors in grammar, punctuation, and spelling. However, ChatGPT's interface was preferred due to its clear explanations and added translation capabilities, making it particularly useful for learners with lower English proficiency. Likewise, ChatGPT was found to provide detailed feedback on grammar, vocabulary, and style. It was particularly effective in suggesting improvements to formality and overall clarity of students' writing. The tool flagged more errors (13 in one sample) compared to other tools like Grammarly (Schlosser, 2023). However, unlike Grammarly, ChatGPT did not always provide automatic, structured feedback, requiring users to ask specific questions for feedback. It also generated irrelevant text, which could confuse students if not reviewed carefully.

Efforts to enhance ChatGPT's ability as a writing assessment tool have also been explored. Mizumoto and Eguchi (2023) finetuned a GPT-3.5 model using prompt engineering and applied it to assess 12,100 essays from the ETS Corpus of Non-Native Written English (TOEFL11). They enriched the model with rubrics derived from IELTS Task 2 Writing band descriptors, integrating these with linguistic features like lexical diversity, syntactic complexity, and textual cohesion. Their research demonstrated that while the base GPT model already provided a reasonable accuracy in grading, the integration of specific linguistic features notably enhanced its performance. Notably, lexical features had the greatest impact on improving the accuracy of assessments. This fine-tuning not only shortened the time taken to rate essays but also increased scoring consistency and enabled the provision of immediate feedback that aligns with established grading standards.

Overall, the evidence points to ChatGPT as a valuable supplementary tool for grammar checking for non-native speakers of English. It offers detailed, accessible feedback, though it may not fully address nuanced issues, such as complex linguistic structures and writers' intent, areas where human instructors excel. Moving forward, writing assessment should incorporate higher-level skills such as organization, style, and argumentation, beyond just grammar checking. Future studies should focus on designing holistic prompts that enable ChatGPT to conduct more comprehensive writing evaluations.

What are the attitudes toward ChatGPT in EFL writing? (Review question 3)

Most studies exploring the integration of ChatGPT into EFL writing focus on the benefits and challenges, largely from the perspective of teachers and students. As an initial probe, a dominant proportion of studies (62.5%) have investigated the perceptions of teachers and students on ChatGPT's role in EFL writing. Both groups recognize its potential, yet express concerns about its broader implications in education.

Teachers, being central to the integration of educational technologies, play a crucial role in determining how ChatGPT is used in the classroom. Nguyen (2023) and Mohamed (2024) found that most teachers agree on the potential benefits of ChatGPT, particularly in terms of facilitating lesson planning, generating teaching materials, and providing additional reading resources. However, many educators remain cautious, especially when it comes to ChatGPT's role in assessment and evaluation. They are concerned about students' potential overreliance on the tool, the accuracy and credibility of its outputs, and its effect on students' critical thinking and academic integrity. These findings are echoed by studies from Nurseha (2023) and Marzuki et al. (2023), which noted that while teachers recognize ChatGPT's usefulness in helping students expand their thinking during the writing process, concerns about originality, academic dishonesty, and plagiarism persist.

Student perceptions reflect a similarly nuanced view. According to Xiao and Zhi (2023), students perceive ChatGPT as a powerful tool that supports writing, particularly by enhancing their vocabulary, grammar, and providing immediate feedback. Many students reported using ChatGPT to prepare for standardized tests like IELTS, appreciating the personalized and interactive learning experience it offers. However, they also recognize its limitations, particularly in supporting more advanced writing skills such as structuring arguments and creating coherent, original content. Like the teachers, students are aware of the potential risks of plagiarism and emphasize the need for clear guidance and regulations to ensure the responsible use of ChatGPT in writing.

A cross-sectional study by Barrett and Pack (2023) explored both teacher and student perspectives on the use of ChatGPT at different stages of the writing process, such as brainstorming, outlining, revising, and feedback. Both groups found the tool most helpful in the early stages, such as generating ideas and organizing thoughts, where it serves as a supportive tool rather than replacing the student's creative process. Interestingly, teachers expressed more positive attitudes toward ChatGPT's use in providing feedback and evaluating student work, whereas students were more reserved. Similarly, Schlosser (2023) found that students felt that ChatGPT, while helpful, should be used alongside teacher feedback for more comprehensive error correction. Despite these differences, both groups emphasized the importance of transparent and ethical use of the tool.

Yan (2023) provided further insights into student perceptions by investigating their actual use of ChatGPT in writing tasks. Students frequently praised the writing quality and variety of language styles produced by ChatGPT. However, consistent with other studies, concerns about plagiarism

and the fairness of using AI tools in academic settings were prevalent. Interestingly, due to the foci of L2 writing on composition structure, grammar, and semantic accuracy, students in this study didn't encounter issues indicated by students from the other studies, including lack of originality of texts, and lack of domain expertise (Barrett & Pack, 2023; Xiao & Zhi, 2023). In line with the aforementioned studies, students were concerned about inequity and plagiarism issues brought up by ChatGPT, showing more concern rather than satisfaction.

Despite the powerfulness and versatility of ChatGPT, both teachers and students exhibited mixed attitudes about the use of ChatGPT in EFL writing. Undeniably, they believed in the promise of this technology in language education and expressed curiosity in exploring its proper use. In the meantime, they harbor concerns about issues arising with the advent of ChatGPT, such as potential plagiarism and overreliance. Also, most of the studies sought to understand teachers' perceptions of ChatGPT integration, with fewer efforts examining students' perceptions. Interestingly, there exists a discrepancy in teachers' and students' perceptions of ChatGPT, especially when it comes to writing evaluation and assessment. This asymmetrical attention underscores the importance of comprehensively understanding the perceptions of both stakeholders to ensure effective implementation and address potential concerns.

Benefits and challenges of ChatGPT in EFL writing (Review question 4)

Research has increasingly highlighted the potential benefits of using ChatGPT in EFL writing. First of all, ChatGPT can be accessed at any time and from any place, providing EFL learners with rich language input, and flexible, personalized learning opportunities outside the traditional classroom setting (Mohamed, 2024). ChatGPT's ability to deliver instant feedback is particularly valuable for EFL learners who may not have instant feedback from instructors. Bok and Cho (2023) reported that learners were able to enhance paragraph flow and learn more appropriate vocabulary through ChatGPT's suggestions. The ability to generate well-structured sentences allows learners to produce more readable and coherent texts with fewer errors.

Second, beyond learning gains, ChatGPT is recognized for enhancing student engagement and motivation. For example, Song and Song (2023) found that ChatGPT provided students with real-time feedback on grammar, vocabulary, sentence structure, and organization during writing tasks. The study showed students' significant improvements in all aspects of writing, including organization, coherence, and vocabulary after using ChatGPT. Students also reported higher motivation and engagement, attributing this to the autonomy and personalized feedback offered by the AI. Similarly, Hawanti and Zubaydulloevna (2023) observed a significant reduction in anxiety among students using ChatGPT.

Moreover, ChatGPT can help reduce the burden on teachers in both writing assessment and material design (Mohamed, 2024). It has been praised for its consistency and efficiency in assessing writing, particularly in detecting surface-level errors such as grammar and vocabulary. It also

provides time-efficient, uniform feedback across a large number of essays, reducing teacher work-load and minimizing human biases in scoring (Geçkin et al., 2023; Mizumoto & Eguchi, 2023).

However, several challenges have also been discovered. ChatGPT's capacity to generate human-like text can assist students in brainstorming ideas or even drafting entire essays. While this offers rich language input, it raises concerns about academic integrity, as students might rely on the AI to complete their tasks, thereby undermining their writing development. Issues of plagiarism also arise when students use AI-generated text without proper understanding or credit (Barrett & Pack, 2023; Nguyen, 2023; Song & Song, 2023). Additionally, overreliance on ChatGPT may discourage students from engaging with traditional learning methods or from interacting with human instructors (Marzuki et al., 2023; Mohamed, 2024; Nguyen, 2023; Nurseha, 2023). A critical concern is diminished authorship, where students reported that ChatGPT sometimes alters their original writing significantly, leaving them disconnected from the final product. This concern was noted by the participants in multiple studies, with some expressing fears about losing their voice and style during the revision process (Al-Garaady & Mahyoob, 2023).

Another challenge is ChatGPT's struggle with providing culturally and contextually sensitive feedback, crucial for grasping pragmatic language use in learning environments. Bok and Cho (2023) noted that while students found ChatGPT's corrections helpful, they expressed frustration about the lack of explanatory feedback, which limited their ability to fully understand their mistakes and develop critical thinking skills in writing. Teachers have also observed that ChatGPT may be unsuitable for lower-intermediate learners, as effective use of the tool might require a B2 level (upper-intermediate) proficiency on the Common European Framework of Reference for Languages (CEFR) scale (Nguyen, 2023). Moreover, the complexity and length of ChatGPT's feedback may be inaccessible to certain learners (Guo & Wang, 2023; Schlosser, 2023). Several other studies have also highlighted the potential for biased language use inherent in the model's training, which could distance learners from authentic English learned through real human interaction (Hawanti & Zubaydulloevna, 2023; Mohamed, 2024).

While ChatGPT offers immediate feedback, it sometimes provides inconsistent textual output, particularly with complex language rules or less common idiomatic expressions (Hawanti & Zubaydulloevna, 2023; Schlosser, 2023). It also struggles with commenting on students' writings in terms of deep meaning, creativity, and coherence, areas where human raters excel. Despite its strengths in identifying formal errors, ChatGPT's feedback lacks the personalization and contextual awareness that human assessments offer. A significant challenge with AI-based assessment tools like ChatGPT is the overemphasis on formal aspects of writing (e.g., grammar and vocabulary). This focus can result in higher scores for grammatically correct essays that lack content depth or creativity, skewing assessments of writing quality (Mizumoto & Eguchi, 2023). Al-Garaady and Mahyoob (2023) found that while ChatGPT could effectively detect grammatical errors, it struggled

with deeper structural and pragmatic issues, which are critical in advanced writing. Human instructors, in contrast, were more effective at recognizing errors related to tone, context, and deeper meaning, highlighting the limitations of AI in fully understanding the complexities of EFL writing.

Discussion

From an initial 2,310 entries gathered from selected databases, 16 studies met the inclusion criteria and were thoroughly examined in this scoping review. The objective was to assess the current landscape of research on the use of GenAI, particularly ChatGPT, in EFL writing instruction and identify gaps for future exploration. Notably, most of the studies were published in 2023, reflecting a surge of interest following the release of OpenAI's ChatGPT (see Section on Review question 3). These areas are of particular relevance due to the recent availability of OpenAI's ChatGPT. As multimodal AI tools like GPT-4 and DALL-E 2 become more widespread, it is expected that research will broaden to cover applications in the near future (Law, 2024).

Besides, most of the studies employed off-the-shelf ChatGPT, which didn't tailor to the specific cultural and linguistic context of learners and teachers. Some teachers have observed that ChatGPT is unsuitable for lower-intermediate learners, suggesting that students must achieve at least B2 on a CEFR proficiency scale to use ChatGPT effectively (Nguyen, 2023). The feedback provided by ChatGPT, often complex and lengthy, may not be accessible to all learners (Guo & Wang, 2023). Plus, successful and effective scaffolding involves the critical pedagogical step of "fading"—gradually removing this support to ensure students can perform independently—and becomes complex (Beed et al., 1991; van de Pol, 2012). As scaffolding is designed to be a temporary support structure, its success not only lies in how well it adapts to and supports the learner's current needs but also in how effectively it can be removed, allowing learners to assume full responsibility for their learning tasks (Sawyer, 2006). There's a risk of overreliance on ChatGPT, where students might depend too heavily on AI for ideas, corrections, and writing strategies, hindering their ability to write independently. Improving ChatGPT's utility in scaffolding involves adapting the large language model to better align with second language learning theories and tailoring prompts to generate more appropriate responses (Al-Obaydi et al., 2023). This could involve training models to recognize and adjust to various proficiency levels, simplifying language as needed, and providing feedback that directly supports learning objectives. Given access to ChatGPT API services, interdisciplinary collaboration is also encouraged to modify off-the-shelf GPT and train GPT-powered CAs for specific learning and teaching purposes based on evidence from second language learning and EFL writing research.

The studies reviewed highlighted the versatility of ChatGPT for EFL writing, with a predominant focus on the university level (see Section on Review question 2). While ChatGPT is being utilized in diverse learning contexts, the majority of studies concentrate on higher education. It is

important to recognize that most studies (67%) investigating the use of ChatGPT in EFL writing primarily focused on learners' and teachers' perceptions in higher education context as distinct areas of inquiry. Also, these studies often explore perceptions separately, without diving deeply into the actual practices involving ChatGPT by either students or teachers. Perceptions can be highly contextualized based on how teachers and learners use ChatGPT. This gap between perceptions and practices highlights an urgent need for more empirical research studies. Future studies should also consider ways of integrating ChatGPT into EFL writing education across different grade levels.

As for attitudes toward ChatGPT, not surprisingly, despite the controversy raised about its possible threats to academic integrity and education equity, both teachers and students share positive attitudes toward using this tool for language learning and teaching. One of the most outstanding benefits for students frequently mentioned in the existing literature is its capability to respond to varied requests in a personalized and efficient way (Mohamed, 2024; Nguyen, 2023). For teachers, ChatGPT can help generate lesson plans, tests, and reading materials referring to different proficiency levels (CEFR, etc.) if given proper interactive prompts (Koraishi, 2023; Nguyen, 2023), and help with automatic essay scoring and giving feedback. Notably, despite teachers' positive perspectives on this application of AI in grading and feedback, it's noteworthy that students were less enthusiastic about the idea of teachers' use of AI to evaluate their work (Barrett & Pack, 2023). This discrepancy highlights the need for a balanced approach that considers both the efficiencies offered by AI and the importance of maintaining human intervention in language education.

Despite the limited number of intervention studies, the results collectively highlight the potential of ChatGPT as an effective tool in improving language learning from both cognitive and affective dimensions. Current insights are mostly from students' or teachers' self-reported data and qualitative studies. Future studies call for more empirical studies investigating students' actual self-regulated use patterns across different contexts, including both structured classroom and unstructured out-of-class settings. More experimental studies are needed to further explore and validate ChatGPT's effectiveness across diverse language learning contexts and learner populations. A broader and more robust evidence base is necessary to ascertain the generalizability and long-term impact of ChatGPT integration in EFL language education. Moreover, comparative studies examining ChatGPT against other AI-powered language learning tools and instructional approaches would provide valuable insights into its relative efficacy and unique contributions to EFL learning.

The ethical concerns surrounding ChatGPT's use in EFL writing, especially regarding plagiarism, are prominent. Given the absence of global regulations, some institutions, particularly in English-speaking countries like the United Kingdom, have banned its use in assignments. Yan (2023) proposed that any direct or lightly edited use of AI-generated texts constitutes plagiarism. However, banning AI outright may not be a sustainable approach. Instead, policies should

encourage responsible use. Nguyen (2023) suggested protocols like submitting both a rough draft and an AI-assisted final version. Policymakers and educators must collaborate to ensure proper application (Perkins & Roe, 2023). However, issues like data privacy and biases remain underexplored. Future efforts should include addressing underexplored issues like data privacy, AI hallucinations, and biases. Institutions and educators must work together to ensure that AI tools like ChatGPT are used ethically while supporting student learning (Rane, 2023).

A less discussed aspect of AI literacy in the studies reviewed might suggest that it has been largely overlooked in current EFL writing education. Teaching EFL writing in the era of Generative AI goes beyond language instruction; it requires reinforcing information literacy and critical thinking skills. Many studies have noted ChatGPT's inconsistent feedback and generation of false positives (Ray, 2023), but how students and teachers handle these issues in real writing practice remains unclear. A major concern with AI systems is "hallucination," where the AI generates content that doesn't align with reality or logic (Buselic, 2023). This happens because AI models do not truly understand the content, they identify patterns in the data they were trained on. Although ChatGPT offers personalized learning assistance, it requires users to apply critical thinking and information literacy to assess the quality and originality of its outputs (Buselic, 2023). These skills are essential in the 21st century and should be integrated into language education. A learner's ability to critically evaluate AI-generated content will significantly impact the reliability and effectiveness of their language learning. To optimize the use of ChatGPT, students must be trained to critically assess its outputs and understand when and how to apply them appropriately. This dual focus on language proficiency and cognitive skills will prepare students to use AI tools both effectively and ethically.

Last but not the least, many of the articles reviewed emphasize the need for professional development in prompt engineering for educators. The complexity of large language models like ChatGPT often limits their interpretability, making mastery of prompt engineering essential for effective interaction with the AI (Jacob & Weber, 2023; Knoth, 2024). Educators recognize that specialized training in crafting prompts would significantly improve their ability to use ChatGPT, enriching students' learning experiences (Mohamed, 2024; Nguyen, 2023). The variability in ChatGPT's responses based on how queries are worded highlights the importance of this skill (Hawanti & Zubaydulloevna, 2023). Learners also stand to benefit greatly from prompt engineering, as well-designed prompts can enhance the quality and relevance of ChatGPT's outputs (Yan, 2023). However, this requires a foundational understanding of how AI models generate language. Integrating prompt engineering into the curriculum for both teachers and students is crucial. This skill set will not only improve AI interactions but also deepen understanding of AI's capabilities and limitations. Training in prompt engineering can transform AI challenges into opportunities, fostering creativity, critical thinking, and collaboration in educational settings.

Limitations

One key limitation of this review is that all the studies included were published in 2023, coinciding with the recent surge in interest following the release of ChatGPT. As a result, the review may not capture a sufficiently broad time span of literature, unlike other reviews that cover more extended periods. Future reviews will need to extend and expand this scope as more research emerges to provide a more comprehensive understanding of ChatGPT's role in language education. Another limitation is that many of the findings are based primarily on explanatory or preliminary qualitative results. These studies tend to focus on exploring the initial potential and applications of ChatGPT, but need more longitudinal or empirical evidence to draw more definitive conclusions. Further research, including experimental and longitudinal studies, is required to validate and deepen these initial insights.

Conclusion

Future research should focus on expanding the scope of studies across diverse educational contexts and examining the long-term impacts of AI integration in language learning. Interdisciplinary efforts are also needed to tailor AI models to the unique linguistic and cultural needs of EFL learners. Moreover, studies should explore beyond text-based AI and investigate how multimodal Generative AI tools can be integrated into EFL writing to enhance meaning-making. In addition, the review highlights the importance of addressing ethical considerations, including data privacy and language bias, alongside educating learners about these issues. While educators and students recognize the value of ChatGPT, future studies should emphasize the need to strengthen critical thinking and information literacy to maximize its potential. Many studies also stressed the importance of prompt engineering for effective interaction with AI, underscoring the need for specialized training for both teachers and learners. As ChatGPT and other AI-powered writing tools continue to evolve, it is crucial to maintain a balanced approach that integrates human oversight with AI assistance to foster authentic learning. Ultimately, a collaborative effort between educators and policymakers will be essential to ensure the ethical and effective use of AI in EFL writing education. Overall, this scoping review contributes to the body of knowledge by providing a foundational understanding of ChatGPT's role in current EFL writing and paves the way for future research and policy development, aiming to optimize the use of AI tools in language education while addressing their inherent challenges.

Contributorship

Feiwen Xiao contributed to the conceptualization of the study, research design, literature search, writing of the manuscript, and data management. Siyu Zhu was involved in the conceptualization of the study and provided critical revisions and edits to the manuscript. Wen Xin contributed to the conceptualization, manuscript revisions and editing to refine the final version.

Declaration of conflicting interests

Funding

ORCID iD

Feiwen Xiao (iD) https://orcid.org/0000-0002-7284-9527

References

Al-Garaady, J., & Mahyoob, M. (2023). ChatGPT's capabilities in spotting and analyzing writing errors experienced by EFL learners. *Arab World English Journals*, *9*(Special Issue on CALL), 3–17. https://doi.org/10.24093/awej/call9.1

wAl-Obaydi, L. H., Pikhart, M., & Klimova, B. (2023). ChatGPT and the general concepts of education: Can artificial intelligence-driven chatbots support the process of language learning? *International Journal of Emerging Technologies in Learning (iJET)*, *18*(21), 39–50. https://doi.org/10.3991/ijet.v18i21.42593

AlAfnan, M. A., Dishari, S., Jovic, M., & Lomidze, K. (2023). ChatGPT as an educational tool: Opportunities, challenges, and recommendations for communication, business writing, and composition courses. *Journal of Artificial Intelligence and Technology*, *3*(2), 60–68. https://doi.org/10.37965/jait.2023.0184

Aldabbus, S., & Almansouri, E. (2022). Academic writing difficulties encountered by university EFL learners. *British Journal of English Linguistics*, *10*(3), 1–11. https://doi.org/10.37745/bjel.2013/vol10n3111

Alharbi, W. (2023). AI In the foreign language classroom: A pedagogical overview of automated writing assistance tools. *Education Research International*, *2023*(1), Article 4253331. https://doi.org/10.1155/2023/4253331

Barrett, A., & Pack, A. (2023). Not quite eye to AI: Student and teacher perspectives on the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education*, *20*(1), Article 59. https://doi.org/10.1186/s41239-023-00427-0

Barrot, J. S. (2023). ChatGPT as a language learning tool: An emerging technology report. *Technology, Knowledge and Learning*, *29*, 1151–1156. https://doi.org/10.1007/s10758-023-09711-4

Beed, P. L., Hawkins, E. M., & Roller, C. M. (1991). Moving learners toward independence: The power of scaffolded instruction. *The Reading Teacher*, *44*(9), 648–655.

Bok, E., & Cho, Y. (2023). Examining Korean EFL college students' experiences and perceptions of using ChatGPT as a writing revision tool. *Journal of English Teaching Through Movies and Media*, *24*(4), 15–27. https://doi.org/10.16875/stem.2023.24.4.15

Buselic, V. (2023, December). Teaching information literacy and critical thinking skills in ChatGPT time. In Proceedings of 2023 International Conference on Computing, Networking, Telecommunications & Engineering Sciences Applications (CoNTESA), Zagreb, Croatia, December 14–15, 2023 (pp. 14–20). IEEE.

Gayed, J. M., Carlon, M. K. J., Oriola, A. M., & Cross, J. S. (2022). Exploring an AI-based writing assistant's impact on English language learners. *Computers and Education: Artificial Intelligence*, *3*, Article 100055. https://doi.org/10.1016/j.caeai.2022.100055

Geçkin, V., Kızıltaş, E., & Çınar, Ç. (2023). Assessing second-language academic writing: AI vs. Human raters. *Journal of Educational Technology and Online Learning*, *6*(4), 1096–1108. https://doi.org/10.31681/jetol.1336599

Guo, K., & Wang, D. (2023). To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing. *Education and Information Technologies*, *29*(7), 8435–8463. https://doi.org/10.1007/s10639-023-12146-0

Hawanti, S., & Zubaydulloevna, K. M. (2023). AI chatbot-based learning: Alleviating students' anxiety in English writing classroom. *Bulletin of Social Informatics Theory and Application*, *7*(2), 182–192. https://doi.org/10.31763/businta.v7i2.659

Hong, W. C. H. (2023). The impact of ChatGPT on foreign language teaching and learning: Opportunities in education and research. *Journal of Educational Technology and Innovation*, *5*(1), 37–45. https://doi.org/10.61414/jeti.v5i1.103

Jacobsen, L. J., & Weber, K. E. (2023). The promises and pitfalls of ChatGPT as a feedback provider in higher education: An exploratory study of prompt engineering and the quality of AI-driven feedback. https://doi.org/10.31219/osf.io/cr257

Kartal, G. (2023). Contemporary language teaching and learning with ChatGPT. *Contemporary Research in Language and Linguistics (ISSN: 2980-2253)*, *1*(1), 59–70. https://doi.org/10.62601/crll.v1i1.10

Kim, S., Shim, J., & Shim, J. (2023). A study on the utilization of OpenAI ChatGPT as a second language learning tool. *Journal of Multimedia Information System*, *10*(1), 79–88. https://doi.org/10.33851/JMIS.2023.10.1.79

Knoth, N., Tolzin, A., Janson, A., & Leimeister, J. M. (2024). AI literacy and its implications for prompt engineering strategies. *Computers and Education: Artificial Intelligence*, *6*, Article 100225. https://doi.org/10.1016/j.caeai.2024.100225

Koraishi, O. (2023). Teaching English in the age of AI: Embracing ChatGPT to optimize EFL materials and assessment. *Language Education and Technology*, *3*(1), 55–72.

Kusal, S., Patil, S., Choudrie, J., Kotecha, K., Mishra, S., & Abraham, A. (2022). AI-based conversational agents: A scoping review from technologies to future directions. *IEEE Access*, *10*, 92337–92356. https://doi.org/10.1109/ACCESS.2022.3201144

Law, L. (2024). Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review. *Computers and Education Open*, *6*, 100174. https://doi.org/10.1016/j.caeo.2024.100174

Marzuki, W. U., Rusdin, D., & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education*, *10*(2), Article 2236469. https://doi.org/10.1080/2331186X.2023.2236469

Mizumoto, A., & Eguchi, M. (2023). Exploring the potential of using an AI language model for automated essay scoring. *Research Methods in Applied Linguistics*, *2*(2), 100050. https://doi.org/10.1016/j.rmal.2023.100050

Mohamed, A. M. (2024). Exploring the potential of an AI-based chatbot (ChatGPT) in enhancing English as a foreign language (EFL) teaching: Perceptions of EFL faculty members. *Education and Information Technologies*, *29*(3), 3195–3217. https://doi.org/10.1007/s10639-023-11917-z

Munn, Z., Peters, M. D., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, *18*, 1–7. https://doi.org/10.1186/s12874-018-0611-x

Nguyen Thi Thu, H. (2023). EFL teachers' perspectives toward the use of ChatGPT in writing classes: A case study at Van Lang University. *International Journal of Language Instruction*, *2*(3), 1–47. https://doi.org/10.54855/ijli.23231

Nurseha, I. (2023). EFL teachers' perspectives on how AI writing tools affect the structure and substance of students' writing. *JEET, Journal of English Education and Technology*, *4*(3), 321–349. https://jeet.fkdp.or.id/index.php/jeet/article/view/127

OpenAI. (2023). ChatGPT (Mar 14 version) [Large language model]. https://chat.openai.com/chat

Özçeli̇k, N. P. (2023). A comparative analysis of proofreading capabilities: Language experts vs. ChatGPT. In Bülent Pekdağ (Ed.), *International Studies in Educational Sciences* (Vol. 147, Chapter 8 , pp. 147–161). Serüven Publishing.

Perkins, M., & Roe, J. (2023). Decoding academic integrity policies: A corpus linguistics investigation of AI and other technological threats. *Higher Education Policy*, *37*(3), 633–653. https://doi.org/10.1057/s41307-023-00323-2

Peters, M. D., Marnie, C., Colquhoun, H., Garritty, C. M., Hempel, S., Horsley, T., Langlois, E. V., Lillie, E., O'Brien, K. K., Tunçalp, Ö., Wilson, M. G., Zarin, W., & Tricco, A. C. (2021). Scoping reviews: Reinforcing and advancing the methodology and application. *Systematic Reviews*, *10*, 1–6. https://doi.org/10.1186/s13643-021-01821-3

Rane, N. (2023). Enhancing the quality of teaching and learning through ChatGPT and similar large language models: Challenges, future prospects, and ethical considerations in education. Future Prospects, and Ethical Considerations in Education (September 15, 2023). Available at SSRN: https://ssrn.com/abstract=4599104 or http://dx.doi.org/10.2139/ssrn.4599104

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, *3*, 121–154. https://doi.org/10.1016/j.iotcps.2023.04.003

Ridley, D. (2012). *The literature review: A step-by-step guide for students*. Sage Publications.

Sawyer, R. K. (2006). Introduction: The new science of learning. In R. K. Sawyer (Ed.), *The Cambridge handbook of: The learning sciences* (pp. 1–16). Cambridge University Press.

Schlosser, M. (2023). *Raising university students' awareness of Automated Writing Evaluation tools in the English as a Foreign Language writing classroom to promote uptake of feedback and learner autonomy* (Master's thesis, Universitat Oberta de Catalunya). Universitat Oberta de Catalunya Institutional Repository. http://hdl.handle.net/10609/148505

Schmidt-Fajlik, R. (2023). ChatGPT as a grammar checker for Japanese English language learners: A comparison with Grammarly and ProWritingAid. *AsiaCALL Online Journal*, *14*(1), 105–119. https://doi.org/10.54855/acoj.231417

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, *14*, Article 1260843. https://doi.org/10.3389/fpsyg.2023.1260843

van de Pol, J., Volman, M., & Beishuizen, J. (2012). Promoting teacher scaffolding in small-group work: A contingency perspective. *Teaching and Teacher Education*, *28*(2), 193–205. https://doi.org/10.1016/j.tate.2011.09.009

Wellnhammer, N., Dolata, M., Steigler, S., & Schwabe, G. (2020). *Studying with the help of digital tutors: Design aspects of conversational agents that influence the learning process* [Paper presentation]. 53rd Hawaii International Conference on System Sciences, HICSS 2020, Maui, HI, United States, January 7–10, 2020.

Xiao, F., Zhao, P., Sha, H., Yang, D., & Warschauer, M. (2023). Conversational agents in language learning. *Journal of China Computer-Assisted Language Learning*, *4*(2), 300–325. https://doi.org/10.1515/jccall-2022-0032

Xiao, Y., & Zhi, Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions. *Languages*, *8*(3), Article 212. https://doi.org/10.3390/languages8030212

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, *28*(11), 13943–13967. https://doi.org/10.1007/s10639-023-11742-4

Zhang, Z. V., & Hyland, K. (2023). Student engagement with peer feedback in L2 writing: Insights from reflective journaling and revising practices. *Assessing Writing*, *58*, Article 100784. https://doi.org/10.1016/j.asw.2023.100784

Zimmerman, A. (2023). A ghostwriter for the masses: ChatGPT and the future of writing. *Annals of Surgical Oncology*, *30*(6), 3170–3173. https://doi.org/10.1245/s10434-023-13436-0