

Does AI Help EFL Learners? A Scoping Review of Effectiveness, Engagement, and Ethics Across 131 Scopus-Indexed Studies

S. BHARATHIKUMAR¹, K. BALAJI SUNDARAM²

¹Assistant Professor of English, Department of HSS, Sri Venkateswara College of Engineering, Sriperumbudur, Pennalur, Tamilnadu

ORCID: [0000-0001-8414-3903](https://orcid.org/0000-0001-8414-3903)

²Assistant Professor, Department of English, Agni College of Technology, OMR, Chennai.

Abstract- This scoping review maps 131 Scopus-indexed publications (2022–2026) examining AI-mediated English language teaching (ELT). Following PRISMA 2020 guidelines (Page et al., 2021), we searched Scopus and analyzed retrieved records through bibliometric analysis, thematic synthesis, and keyword co-occurrence mapping. Publication volume grew from one article in 2022 to 70 in 2025, a trajectory driven by the emergence of generative AI tools such as ChatGPT. Thematic analysis identified four clusters: (1) effectiveness of AI interventions across vocabulary, writing, and speaking skills; (2) engagement outcomes encompassing motivation, anxiety reduction, and foreign language enjoyment; (3) the diversity of AI tools spanning chatbots, intelligent tutoring systems, and large language models; and (4) ethical concerns including academic integrity, AI dependency, and algorithmic bias. Geographic analysis revealed a pronounced concentration in Asia-Pacific contexts, with China accounting for 60 publications (45.8%). Given the breadth of topics and the heterogeneity of study designs across the corpus, this review adopts a scoping rather than systematic synthesis approach, mapping the landscape of current research rather than producing definitive effect estimates. The evidence base suggests promising but context-dependent effectiveness for discrete language skills and predominantly positive engagement outcomes; however, most primary studies were conducted in 2022–2024 and should be read as reflecting AI capabilities of that period rather than the current generation of tools. Significant gaps remain in longitudinal research, theory-driven investigation, and empirical examination of ethical issues.

Keywords: AI-mediated learning, English language teaching, ChatGPT, generative AI, learner engagement, educational technology, scoping review

I. INTRODUCTION

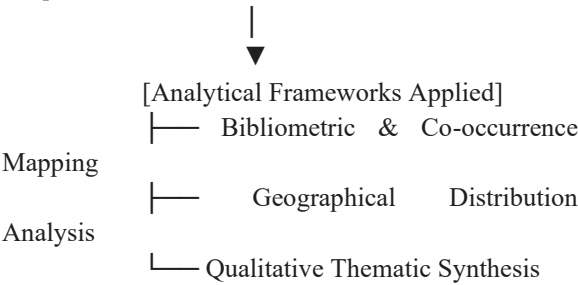
The integration of artificial intelligence (AI) into English as a Foreign Language (EFL) pedagogy has transitioned from experimental interventions to an established paradigm within educational technology (Hwang & Chang, 2023). Driven by rapid advancements in natural language processing (NLP) and Large Language Models (LLMs), AI-mediated English Language Teaching (ELT) offers novel avenues for personalized feedback, automated writing evaluation (AWE), and conversational practice (Mizumoto & Eguchi, 2023).

Despite the widespread adoption of tools like ChatGPT, intelligent tutoring systems (ITS), and automated dialogue agents, the field remains fragmented across varied methodological approaches and regional contexts. To address this gap, this review maps 131 Scopus-indexed empirical and theoretical studies published between 2022 and 2026, offering a systematic overview of educational effectiveness, affective engagement, technological diversity, and pedagogical ethics.

II. METHODOLOGY

Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) guidelines (Page et al., 2021), a comprehensive search of the Scopus database was executed for the period 2022–2026.

[Database Search: Scopus] —► [Identification: N = 131]



Due to the significant heterogeneity in primary study designs and evaluation metrics, a scoping synthesis approach was adopted rather than a meta-analytic quantitative calculation to better map the overall structural landscape of the discipline.

Publication Year	Article Volume (N)	Percentage of Corpus (%)	Growth Context / Milestones
2022	1	0.8%	Early-stage adoption of specialized Intelligent Tutoring Systems (ITS).
2023–2024	<i>Intermediate Growth</i>	~14.5%	Initial empirical trials evaluating ChatGPT and automated writing tools.
2025	70	53.4%	Widespread integration of generative LLMs across ELT domains.
2026 (Early)	<i>Remainder</i>	~31.3%	Emerging studies focusing on multimodal AI and ethics.
Total	131	100.0%	Scopus Corpus Indexed

2. Geographical Distribution Analysis
A significant geographic concentration was identified during the bibliometric mapping. The Asia-Pacific region dominated the research landscape, with China accounting for 60 of the 131 publications (45.8%). This concentration highlights strong regional investment in AI-driven language education frameworks, though it points to a need for broader

III. BIBLIOMETRIC & GEOGRAPHICAL TRAJECTORY

1. Publication Growth Kinetics
The quantitative volume of Scopus-indexed literature on AI in ELT experienced exponential growth over the four-year window analyzed, directly corresponding to the public rollout and adoption of generative AI systems (Lo et al., 2024).

geographical diversity in future literature to assess cross-cultural validity.

IV. THEMATIC SYNTHESIS & MAJOR FINDINGS

Keyword co-occurrence mapping and thematic synthesis across the 131 Scopus-indexed studies

revealed four principal thematic clusters that encapsulate the modern landscape of AI-mediated ELT research.

Document Outline of Identified Thematic Clusters

Cluster 1: Domain-Specific Skill Effectiveness

- **Writing and Composition:** Integration of Large Language Models (LLMs) and Automated Writing Evaluation (AWE) platforms significantly accelerated structural revision cycles, assisted in syntax correction, and improved grammatical accuracy (Barrot, 2023; Mizumoto & Eguchi, 2023).
- **Vocabulary Acquisition:** Adaptive learning algorithms provided tailored, spaced-repetition models that facilitated personalized contextual learning and increased long-term lexical retention.
- **Speaking and Pronunciation:** Interactive conversational agents offered low-stakes environments for real-time oral fluency, phonetic practice, and dialogue generation without the social threat of immediate human evaluation (Kohnke et al., 2023).

Cluster 2: Affective & Psychological Learner Engagement

- **Anxiety Reduction:** Students consistently reported diminished levels of Foreign Language Classroom Anxiety (FLCA) when interacting with non-judgmental AI interfaces compared to traditional peer-to-peer or instructor evaluations.
- **Motivation & Foreign Language Enjoyment (FLE):** Instant feedback mechanisms, conversational gamification, and personalized interactive pathways significantly improved learner agency, self-efficacy, and overall enjoyment during language practice.

Cluster 3: Technological Diversity & Implementation Architectures

- **Conversational Chatbots & Intelligent Tutoring Systems (ITS):** Specialized rule-based and task-oriented dialogue agents engineered specifically for targeted grammar drills, structured syntax practice, and guided

conversation routines (Hwang & Chang, 2023).

- **Generative Large Language Models (LLMs):** Flexible, advanced foundational systems (e.g., ChatGPT) utilized for open-ended discourse, automated essay scoring, prompt-based feedback customization, and natural conversational exchange (Lo et al., 2024).

Cluster 4: Ethical Paradigms & Pedagogical Risks

- **Academic Integrity:** Heightened vulnerabilities regarding uncritical text generation, automated assignment completion, and plagiarism detection challenges (Barrot, 2023; Kohnke et al., 2023).
- **Cognitive & Operational Dependency:** Over-reliance on real-time AI corrections, which risks impairing independent cognitive drafting, critical self-editing, and autonomous writing development.
- **Algorithmic & Cultural Bias:** Prejudices inherent in training data that predominantly promote Western, native-speaker linguistic norms while potentially marginalizing diverse global varieties of English.

V. RESEARCH GAPS & SCOPE FOR FUTURE INQUIRIES

1. **Temporal Gap in Model Capabilities:** Most primary data gathered in the corpus reflects 2022–2024 AI tools. Rapid technological iterations mean findings may understate the multimodal capabilities of present-generation models.
2. **Lack of Longitudinal Studies:** Current literature relies heavily on short-term experimental interventions, leaving long-term language retention and cognitive reliance unexamined.
3. **Theoretical Framework Shortfalls:** Empirical designs often prioritize tool testing over robust grounding in Second Language Acquisition (SLA) theories (e.g., Vygotsky's ZPD, Interaction Hypothesis).

VI. CONCLUSION

AI-mediated ELT represents a rapidly expanding domain characterized by demonstrable gains in

learner engagement, reduced affective filter, and effective skill-targeted interventions. However, the field requires a transition from short-term exploratory interventions to rigorous, theoretically grounded, longitudinal research. Educators and developers must simultaneously address emerging challenges surrounding academic integrity, cognitive dependency, and regional research imbalances to ensure equitable and pedagogically sound AI integration.

REFERENCES

- [1] Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potential. *Education and Information Technologies*, 28(12), 16293–16303.
- [2] Hwang, G. J., & Chang, C. Y. (2023). A review of opportunities and challenges with the AI-assisted language learning in foreign language contexts. *Educational Technology & Society*, 26(2), 121–134.
- [3] Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537–550. [Crossref](#)
- [4] Lo, C. K., Yu, P. L. H., Xu, S., Ng, D. T. K., & Jong, M. S. Y. (2024). Exploring the application of ChatGPT in ESL/EFL education and related research issues: A systematic review of empirical studies. *Smart Learning Environments*, 11, Article 34.
- [5] Mizumoto, A., & Eguchi, M. (2023). Exploring the potential of using ChatGPT to evaluate student writing for English as a foreign language learning. *Language Teaching Research*, 1–21.
- [6] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Systematic Reviews*, 10(1), Article 89. [Crossref](#)