

Generative Artificial Intelligence in Language Learning: A Review of Selected Journal Publications from 2023 to 2024

Xiaoshuang Zhang ¹, Bowen Jing ¹, Jiarong Chen ¹, Lin Luan ^{1*}

¹ School of Humanities, Beijing University of Posts and Telecommunications, Beijing, China

* luanlin@bupt.edu.cn

Abstract: *Generative artificial intelligence (GenAI) based on large language models (LLMs) represented by ChatGPT-3.5 has ushered in promising tools for foreign language learning and provoked great interest in its application in language learning. This study reviews the empirical literature on the application of GenAI in language learning published on five high-impact journals between 2023 and 2024 since the launch of ChatGPT-3.5. Fourteen articles are selected and further analyzed from the perspective of research contexts, research methods and research findings. The review reveals that the current research on GenAI application in language learning mainly targets English and language skills including writing and speaking, focuses on higher education learners and employs mixed methods. Learners generally hold positive attitudes to GenAI-assisted language learning and benefit from GenAI application, while challenges also emerge in this context. Future studies should broaden the research scope by exploring other languages, language skills such as reading and listening, and educational contexts such as elementary education, meanwhile aiming to address these challenges. This study identifies limitations of current research on GenAI-assisted language learning and provides implications for future researchers and practitioners.*

Keywords: generative artificial intelligence (GenAI), language learning, literature review

1. Introduction

The rising pervasiveness of Artificial Intelligence (AI) has led applied linguists to combine it with language teaching and learning processes. As a new milestone of artificial intelligence, Generative Artificial Intelligence (GenAI) based on Large Language Models (LLMs), represented by ChatGPT-3.5 launched by OpenAI in November 2022, has attracted global attention, holding great implications for the field of foreign language teaching (Law, 2024). Through deep learning and massive text data, GenAI powered by LLMs acquires the ability of language understanding, information summary, content creation and language prediction (Jeon & Lee, 2023). The core of language acquisition lies in the learner exposure to and understanding of a large number of language inputs, as well as the construction and acquisition of language through the analysis of situational context (Krashen, 1981). The powerful ability of GenAI to understand and generate natural language based on LLMs has shown great potential in empowering foreign language teaching. As research on language learning assisted by GenAI has seen explosive growth, a systematic review of those empirical studies is needed. Despite that some review studies have been conducted on AI-assisted education, few reviews have focused specifically on the application of the newly-launched GenAI powered by LLMs in the field of second language learning and teaching. Therefore, this study aims to review studies on the application of GenAI powered by LLMs on second language learning published in top-tier journals. The research questions proposed in this study are as follows:

(1) What research contexts have been investigated and what methods have been employed in the selected studies?

(2) What research findings have been discovered in the selected studies?

2. Methodology

Considering that it is ChatGPT-3.5 that makes the education field realize the great potential of GenAI powered by LLMs, the related research starts from November 2022 in a real sense. Therefore, this review focuses on the studies published from 2023 to 2024 and follows the PRISMA guidelines. Following the study by Yang & Li (2024) for reviewing the application of ChatGPT for L2 learning, this paper focuses on five high-impact journals in the field of technology-enhanced language learning, namely, *Computer Assisted Language Learning*, *ReCALL*, *System*, *Language Learning and Technology*, and *Computers & Education*. The key words used in the searching of qualified articles in these targeted journals include: (“generative artificial intelligence” OR “large language model” OR “GenAI” OR “LLM” OR “ChatGPT”) AND (“language learning” OR “language education” OR “language acquisition”). The initial search yielded 109 articles. The retrieved articles were screened on the title and abstract according to the following criteria: a) focused on the application of GenAI powered by LLMs in language learning; b) empirical studies; c) journal articles. The remaining articles were assessed in full text to confirm their eligibility and 14 studies were retained for further analysis. Data extraction focused on research context, participants, target language, target language skills, methods and findings. The extracted data were synthesized using a descriptive and thematic approach.

3. Results

All of the 14 selected studies target English as the foreign language to learn. As to the participants, as shown in Table 1, the majority of selected studies focused on higher education students including undergraduates of different grades. Only two studies (Z.-M. Liu et al., 2024; Wang et al., 2024) pay attention to elementary students. The other three have participants from informal language learning contexts, one investigating IELTS learners in Iran (Fathi & Rahimi, 2024) while the other two recruiting EFL learners from social media (G. L. Liu, et al., 2024; Tram et al., 2024). The first language of participants varies, including Chinese, Arabic, Korean, and Japanese, and most of these participants are Chinese-speakers. And the sample size of participants ranges from small to large according to their research focuses with different research methods. More studies targeting elementary and high school students are also needed, and the target language should be further expanded to other language except English.

Table 1 Overview of research articles.

Target language skills (number of articles)	Author(s)	Participants		
		Sample size	Education background	First language
Writing (6)	Chen Ziqi et al., 2024	45	Undergraduates	Mixed
	Z.-M. Liu et al., 2024	65	Sixth-grade students	Chinese
	Fathi & Rahimi, 2024	14	Mixed	Persian
	Liu et al., 2023	8	Undergraduates	Chinese
	Yan, 2024	117	Sophomore	Chinese

			undergraduates	
	Wang, 2024	79	Undergraduates	Japanese
Speaking (3)	Wang et al., 2024	99	Undergraduates	Chinese
	Tai & Chen, 2024	85	Sixth-grade students	Chinese
	Guan et al., 2024	48	Undergraduates	Chinese
General skills (5)	Manar Almanea, 2024	135	Undergraduates	Arabic
	G. L. Liu, et al., 2024	867	Mixed	Chinese
	Tram et al., 2024	334	Mixed	Vietnamese
	Liu & Reinders, 2024	24	First-year undergraduates	Chinese
	Lee et al., 2024	22	Undergraduates	Korean

The majority of selected studies employs mixed-method design (79%), which improves the reliability of results. Various methods are employed in these studies for data collection, ranging from questionnaire to audio-recording, and the majority adopts more than one type of them. The most popular methods include questionnaire, followed by interviews and essay/writing journals. For example, several studies investigated EFL learners' AI adoption experience and factors influencing their continuous usage (e.g. G. L. Liu, et al., 2024; Tram et al., 2024) through both questionnaire and interviews. Meanwhile, there is one study collecting data through observation during the writing process (Fathi & Rahimi, 2024) and one study acquiring speaking data through audio-recording (Tai & Chen, 2024).

As illustrated in Table 1, these selected articles focus on different language skills, which can be concluded into three categories, namely, writing, speaking and general skills (referring to the comprehensive language abilities integrating different aspects). Most studies investigate the effects of GenAI on language learners from the aspect of writing skills, such as the writing performance with the assistance of GenAI, the cognitive process in GenAI-assisted multimodal composing, and their attitudes to GenAI-generated corrective feedback. Studies targeting speaking skill mainly delve into whether GenAI influences speaking proficiency, speaking anxiety, communicative competence and willingness to communicate. Moreover, instead of targeting one specific language skill, nearly one-third of the selected articles explore the GenAI effects from the perspective of general language skills, mainly focusing on the overall attitudes to GenAI as a tool for language learning. There is a lack of studies focusing on the effects of GenAI on other language skills, such as reading and listening.

The research findings of the selected studies can be summarized into three themes. In terms of learner attitudes, learners exhibit a relatively high acceptance of GenAI-assisted language learning due to the perceived usefulness (Tram et al., 2024; G. L. Liu, et al., 2024). As to the impact of GenAI in language learning, it is revealed that the utilization of GenAI may benefit language learners in various aspects, including improving learning motivation and confidence (e.g., Z.-M. Liu et al., 2024), developing critical thinking and self-regulated strategies (Liu et al., 2023; Fathi & Rahimi, 2024), and reducing speaking anxiety (Wang et al., 2024). Meanwhile, challenges and concerns are also identified in the selected studies. While GenAI tools generally enhance language learning, the effect may vary due to internal factors. For example, Liu & Reinders (2024) highlighted the individual differences observed in learning motivation. The integration of GenAI in

English learning raises ethical concerns, such as academic dishonesty and over-reliance (Manar Almanea, 2024). GenAI tools occasionally generate unnatural or even inaccurate responses, negatively impacting the learning experience (Lee et al., 2024). The use of AI tools alone is insufficient to sustain long-term language learning outcomes and thus appropriate teacher guidance is urgently required (Guan et al., 2024; Wang, 2024).

4. Discussion and Conclusion

This study reviews the selected empirical studies on GenAI-assisted language learning from 2023 to 2024 published on five top-tier journals. With the PRISMA approach, detailed analysis was conducted on research contexts, research methods and findings. According to the results, studies on GenAI-assisted language learning are primarily English-targeted and focus on writing, speaking and general skills. Participants are mainly higher education students at campus and several studies recruit online informal English learners as participants. Most studies investigate employed mixed methods with diverse data collection methods. These studies revealed three types of findings, namely, language learners' general acceptance of GenAI-assisted language learning, the positive impact of GenAI tools on language learning, and concerns raised by GenAI application. It can be further proved that the application of GenAI in language learning is a promising research area with the potential to transform language education (Law, 2024). However, the current research also presents several limitations. Firstly, studies targeting non-English are quite limited and language skills such as reading and listening are ignored. Future research should enrich the target language and language skills. Furthermore, the selected studies mainly focus on higher education context and the GenAI application in the context of basic education and other languages needs more attention.

Acknowledgements

This paper was part of the project of “Research on Innovation and Application of Foreign Language Teaching Mode Empowered by Virtual Simulation Experiments in Colleges and Universities in the Context of New Liberal Arts” supported by 2024 Higher Education Research Project of China Association of Higher Education (Grant No. 24WY0303).

References

- Chen Ziqi, Zhu Xinhua, Qi, L., & Wei, W. (2024). L2 students' barriers in engaging with form and content-focused AI-generated feedback in revising their compositions. *Computer Assisted Language Learning*, 1–21.
- Fathi, J., & Rahimi, M. (2024). Utilising artificial intelligence-enhanced writing mediation to develop academic writing skills in EFL learners: a qualitative study. *Computer Assisted Language Learning*, 1–40.
- Guan, L., Zhang, E. Y., & Gu, M. M. (2024). Examining generative AI-mediated informal digital learning of English practices with social cognitive theory: a mixed-methods study. *ReCALL*, 1–17.
- Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, 28.
- Krashen, S. D. (1981). *Second Language Acquisition and Second Language Learning*. Pergamon Press.

- Law, L. (2024). Application of Generative Artificial Intelligence (GenAI) in Language Teaching and Learning: A Scoping Literature Review. *Computers and Education Open*, 6, 100174–100174.
- Lee, J. H., Shin, D., & Hwang, Y. (2024). Investigating the capabilities of large language model-based task-oriented dialogue chatbots and the impact of “suggest replies” function from a learner’s perspective. *System*, 103538–103538.
- Liu, G. L., Darvin, R., & Ma, C. (2024). Exploring AI-mediated informal digital learning of English (AI-IDLE): a mixed-method investigation of Chinese EFL learners’ AI adoption and experiences. *Computer Assisted Language Learning*, 1–29.
- Liu, M. Y., Lawrence Jun Zhang, & Biebricher, C. (2023). Investigating students’ cognitive processes in AI-assisted digital multimodal composing and traditional writing. *Computers & Education*, 104977–104977.
- Liu, M., & Reinders, H. (2024). Do AI chatbots impact motivation? Insights from a preliminary longitudinal study. *System*, 103544.
- Liu, Z.-M., Hwang, G.-J., Chen, C.-Q., Chen, X.-D., & Ye, X.-D. (2024). Integrating large language models into EFL writing instruction: effects on performance, self-regulated learning strategies, and motivation. *Computer Assisted Language Learning*, 1–25.
- Manar Almanea. (2024). Instructors’ and learners’ perspectives on using ChatGPT in English as a foreign language courses and its effect on academic integrity. *Computer Assisted Language Learning*, 1–26.
- Tai, T.-Y., & Howard Hao-Jan Chen. (2024). Improving elementary EFL speaking skills with generative AI chatbots: Exploring individual and paired interactions . *Computers & Education*, 105112–105112.
- Tram, N. H. M., Nguyen, T. T., & Tran, C. D. (2024). ChatGPT as a tool for self-learning English among EFL learners: A multi-methods study. *System*, 103528.
- Wang, C., Zou, B., Du, Y., & Wang, Z. (2024). The Impact of Different Conversational Generative AI Chatbots on EFL Learners: an Analysis of Willingness to Communicate, Foreign Language Speaking Anxiety, and Self-perceived Communicative Competence. *System*, 103533–103533.
- Wang, Y. (2024). Cognitive and sociocultural dynamics of self-regulated use of machine translation and generative AI tools in academic EFL writing. *System*, 126, 103505.
- Yan, D. (2024). Comparing individual vs. collaborative processing of ChatGPT-generated feedback: Effects on L2 writing task improvement and learning. *Language Learning & Technology*, 28(1), 1–19.
- Yang, L., & Li, R. (2024). ChatGPT for L2 learning: Current status and implications. *System*, 124, 103351 – 103351.