
www.ijte.net

# Generative AI in Language Education: Bridging Divide and Fostering Inclusivity

**Zhenjie Weng**
Duke Kunshan University, China

**Yao Fu**
University of Wisconsin-Whitewater, U.S.A.

.

**To cite this article:**

Weng, Z. & Fu, Y. (2025). Generative AI in language education: Bridging divide and fostering inclusivity. *International Journal of Technology in Education (IJTE)*, *8*(2), 395-420. https://doi.org/10.46328/ijte.1056

International Journal of Technology in Education
2025, Vol. 8, No. 2, 395-420 https://doi.org/10.46328/ijte.1056

# Generative AI in Language Education: Bridging Divide and Fostering Inclusivity

**Zhenjie Weng, Yao Fu**

**Article Info**

***Article History***

Received:
03 November 2024
Accepted:
10 April 2025

***Keywords***

Generative artificial intelligence
Language education
Educational divide
Inclusivity
Systematic review

**Abstract**

This systematic review analyses empirical studies on the implementation of Generative Artificial Intelligence (GenAI) in language education to highlight its impact on promoting inclusivity and bridging educational gaps. The review encompasses quantitative, qualitative, and mixed-methods research, with a focus on inclusivity and educational equity accelerated by GenAI across multiple language education contexts, ranging from K-12 to higher education. Analysis shows that GenAI has the potential to substantially increase learner motivation and confidence by providing real-time feedback and generating engaging learning materials. Such tools can adapt to individual learner needs, support students with special education requirements, and empower economically disadvantaged learners. Furthermore, GenAI facilitates intercultural competence by incorporating diverse cultural contents into language learning. By synthesizing current research, this review delineates a comprehensive understanding of GenAI's role in creating equitable and inclusive language education landscapes, thereby guiding future research, educational practices, teacher training, and policymaking.

## Introduction

In an increasingly interconnected and digital world, the role of language education is evolving. Traditional methods, while still prevalent and valuable, are being augmented and transformed by technological advancements (Crompton et al., 2024). Among these advancements, Generative Artificial Intelligence (GenAI) stands out as a particularly promising tool (Wang & Xue, 2024). GenAI, described as "a recent large language model (LLM) system that can receive inputs such as text and images and use these inputs to generate new content in a range of modalities, including text, images, sound, and video" (Creely, 2024, p. 1), is revolutionizing various sectors, including language education. Examples of GenAI tools include OpenAI's ChatGPT, GPT-4, Playground, and DALL · E3; Anthropic's Claude; Google's Gemini (previously Bard); Stability AI's Stable Diffusion 3; and Runaway's Gen-2 (Law, 2024). These models are capable of producing human-like text, translating languages, and generating personalized learning materials (Heaven, 2020; Lim et al., 2023). By leveraging these capabilities, educators can create more engaging and effective language learning experiences (Jeon et al., 2023). Additionally, GenAI tools can help learners overcome specific challenges and accelerate their language acquisition (Koc & Savas, 2024; Tai & Chen, 2022).

Traditional educational systems often adopt a one-size-fits-all approach, which can marginalize students who do not fit the conventional mold (Creely, 2024; Zhou & Niu, 2015). AI-driven language learning tools, however, can adapt to different learning paces, preferences, and abilities, ensuring that all students have the opportunity to succeed (Grassini, 2023). For example, AI can provide real-time guidance and corrections for non-native speakers (Woo & Choi, 2021) or deliver culturally relevant educational experiences to second language learners (Creely, 2024). Beyond personalizing and diversifying language education, GenAI can also enhance collaborative learning (Toboula & Martinien, 2023). It can assist teachers by automating administrative tasks and offering insights into students' progress, allowing teachers to focus more on fostering meaningful connections and guiding students' learning journeys (Koraishi, 2023). Nevertheless, the integration of GenAI in language education comes with challenges. Concerns about data privacy, algorithmic biases, and the digital divide (Creely, 2024; Grassini, 2023) must be addressed to ensure the equitable distribution of AI's benefits. Ethical considerations are paramount, as deploying AI technologies in education must prioritize the well-being and rights of learners (Huang et al., 2024; Mohamed, 2023; Munoz et al., 2023; Zhu et al., 2023).

With the growing attention on the application of GenAI in education, several systematic reviews have explored AI in education more broadly (e.g., Faisal, 2024; Fu & Weng, 2024; Fu et al., 2024; Ogunleye et al., 2024) or in language education specifically (e.g., Creely, 2024; Crompton et al., 2024; Jeon & Lee, 2024; Jeon et al., 2023; Law, 2024; Li et al., 2024; Liang et al., 2023; Zhai & Wibowo, 2022). Reviews targeting language education typically focus on the potential benefits, affordances, challenges, future directions for AI applications, attitudes toward AI, and an overview of the current state of research. However, they do not concentrate exclusively on how GenAI can be leveraged to narrow educational gaps and promote inclusivity in language education—an area that warrants further exploration. This systematic review stands out in its innovative approach by exclusively synthesizing existing empirical studies on the potential of GenAI to create a more equitable and inclusive language education landscape. The research question guiding this review is: "How can GenAI be used to narrow educational gaps and promote inclusivity in language education?" By systematically analyzing the existing body of research, this review aims to identify areas requiring further exploration. Additionally, its findings can guide the development of AI-driven educational tools that are both effective and ethical, ensuring technological advancements contribute positively to educational equity and inclusivity. Furthermore, this review will provide a solid foundation for future research, pedagogy, teacher professional development, and policy-making, shaping the ongoing conversation about the role of AI in transforming language education.

## Literature Review

Up to this date, diverse review studies (e.g., Crompton, 2024; Jeon et al., 2023; Koc & Savas, 2024; Law, 2024; Liang et al., 2024; Ma et al., 2024; Wang et al., 2024; Yang & Li, 2024; Zhai & Wilbowo, 2022) on GenAI have been conducted, each with a different focus. For example, Crompton (2024) presents a detailed systematic review that investigates the use of AI in English language teaching and learning (ELT/L). The review found that most studies took place in Asia, indicating a regional push towards AI in ELT/L, and most research is conducted within the context of higher education, with a notable gap in studies on adult learners and K-12 education. Additionally, AI has been found to support various aspects of language education, including skill-specific learning, pedagogical

enhancement, and the facilitation of self-regulation in learners. However, challenges such as technical malfunctions, limited AI capabilities, user apprehension, and the standardization of language by AI technologies have been identified. These gaps in the review suggest a need for further exploration of various demographics, educational levels, language skills and subskills, as well as a deeper understanding of AI's limitations and challenges in ELT/L.

Koc and Savas's (2024) meta-synthesis scrutinized studies from various countries and institutional levels and found that common theoretical frameworks, such as the interaction hypothesis and self-determination theory, underpin chatbot-assisted language learning research. There is an increasing amount of published research on AI chatbots, indicating a growing interest and adoption in English language learners; however, general-audience chatbots are used more often than specific-purpose ones, despite the latter's potential for targeted language learning. The methodologies and contexts used are diverse, with a focus on Asia in chatbot research and an underrepresentation in primary schooling. The study recommends the use of more varied methodologies, addressing technical challenges, developing specific-purpose chatbots, and ensuring theoretical grounding for future research. It also suggests curricular advancements for better integration of AI chatbots, potentially utilizing large language models for broader dialogue capabilities. The meta-synthesis underscores that voice-based chatbots can significantly enhance language learning but require thoughtful implementation and further investigation.

Law's (2024) scoping review provides an overview of research in GenAI's application in language education and identifies gaps for future investigation. The findings indicate that GenAI is described with varying terms in the literature, emphasizing the need for a unified terminology for clearer exposure and understanding. Most studies focus on English as a Foreign Language (EFL) and predominantly in higher education settings. The primary areas of research include language teaching and learning (T&L), T&L policy, writing, and assessments, reflecting the impact of ChatGPT. While GenAI is viewed positively for language education, ethical concerns about plagiarism, academic integrity, data privacy, and security have been raised. The review suggests that more empirical studies are needed to assess the long-term impact of GenAI. There is an urgent need for "GenAI literacy" to ensure that teachers understand the privacy and security implications. The recommendation is for continuous professional development for teachers to assist in their due diligence when incorporating GenAI into their teaching practices.

The study conducted by Liang et al. (2024) provides a comprehensive analysis of AI in Language Education (AILEd) research from 1990 to 2020. Using the Technology-Based Learning Review model, the research examines various dimensions such as research methods, participant groups, AI technology, and algorithms, language skills, the role of AI in language education, and learning outcomes. Key findings from the study include: (1) main research areas in AILEd encompass writing, reading, and vocabulary; (2) common AI technologies employed involve Intelligent Tutoring Systems (ITS) and Natural Language Processing (NLP); (3) frequently used AI algorithms in language education studies comprise statistical learning, data mining, machine learning, and natural language parsing; (4) research focused on learning outcomes related to anxiety, willingness to communicate, knowledge acquisition, and classroom interaction, but there is less emphasis on higher-order thinking, complex problem-solving, critical thinking, and collaborative learning; (5) future research is directed toward improving the application of AI in language education. Additionally, the study found that the majority of

AILEd research is quantitative, mainly conducted at the higher education level with large sample sizes and suggests a need for more systematic research on AI's application both within and outside the classroom, focusing on assessment, learning skills, higher-order thinking, and learning anxieties.

Jeon et al. (2023) diverge from general AI use and systematically review the use of speech-recognition chatbots, specifically within language learning contexts. Their analysis of 32 empirical studies investigates how these chatbots are applied in language learning and their implications for future research, especially with the advent of larger language models (LLMs). The review highlights an increasing trend in research on speech-recognition chatbots since 2020, with a significant spike in 2022. The use of speech-recognition chatbots has been mainly within college settings, targeting English language learning. While the review acknowledges the growing recognition of the benefits of using speech-recognition chatbots for language learning, it calls for diverse research to evaluate different aspects such as cultural influence and language skills beyond speaking and listening. Several research areas remain underexplored, including the need for longer-term studies, involvement at different educational levels apart from higher education, and exploration of chatbot use in naturalistic settings outside the classroom. The paper urges future research to consider implementing chatbots with wearable devices for maximum ubiquity and to further examine the educational potential and roles of LLM-powered chatbots in educational contexts.

The reviews have addressed various aspects of AI use in language education and pointed out future directions for research in this area. For example, several reviews (e.g., Jeon et al., 2023; Law, 2024; Liang et al., 2024) mention the limited investigation of AI use beyond higher education contexts and the lack of longitudinal examinations. Recognizing the contributions of these reviews and to advance the field further, a systematic review that specifically addresses how GenAI can narrow the educational gap and promote equity in language education is needed. In line with this aim, this study unpacks how the use of GenAI can narrow educational gap and promote equity in language education.

## Method

To address the research question, we conducted a systematic mixed studies review (Frantzen & Fretters, 2016) to explore how GenAI tools could be leveraged to bridge the educational divide and foster equity in language education across various contexts, from K-12 to higher education. Adhering to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Page et al., 2021), we selected 32 empirical studies that employed quantitative, qualitative, or mixed methods from four databases for inclusion in our review. In the sections that follow, we outline our search strategy, screening process, and inclusion and exclusion criteria, as well as the methods we used for data analysis.

### Search Strategy

To identify suitable studies for our review, we utilized the search strings detailed in Table 1, which align with our research question, to conduct searches in Scopus, Web of Science (WoS), the Association for Computing

Machinery (ACM) Digital Library, and EBSCOhost. These databases were chosen for their extensive coverage of high-quality peer-reviewed journal articles and conference proceedings relevant to the scope of this systematic review.

Table 1. Search Strings

| Operator | Dimension/Topic | Search Terms |
|---|---|---|
| AND | Generative Artificial Intelligence | "generative AI" OR "generative artificial intelligence" OR "GenAI" OR "chatbot" OR "ChatGPT" |
| | Language Education | "language education" OR "language teaching" OR "language learning" OR "language research" |
| | Bridging Divide or Fostering Equity | "inclusivity" OR "Equity" OR "Diversity" OR "Access" OR "Divide" OR "Inequalities" OR "Gap" |

**Screening Process and Criteria**

We adhered to the PRISMA guidelines (Page et al., 2021) and followed the inclusion and exclusion criteria detailed in Table 2 to facilitate our screening process (see Figure 1).

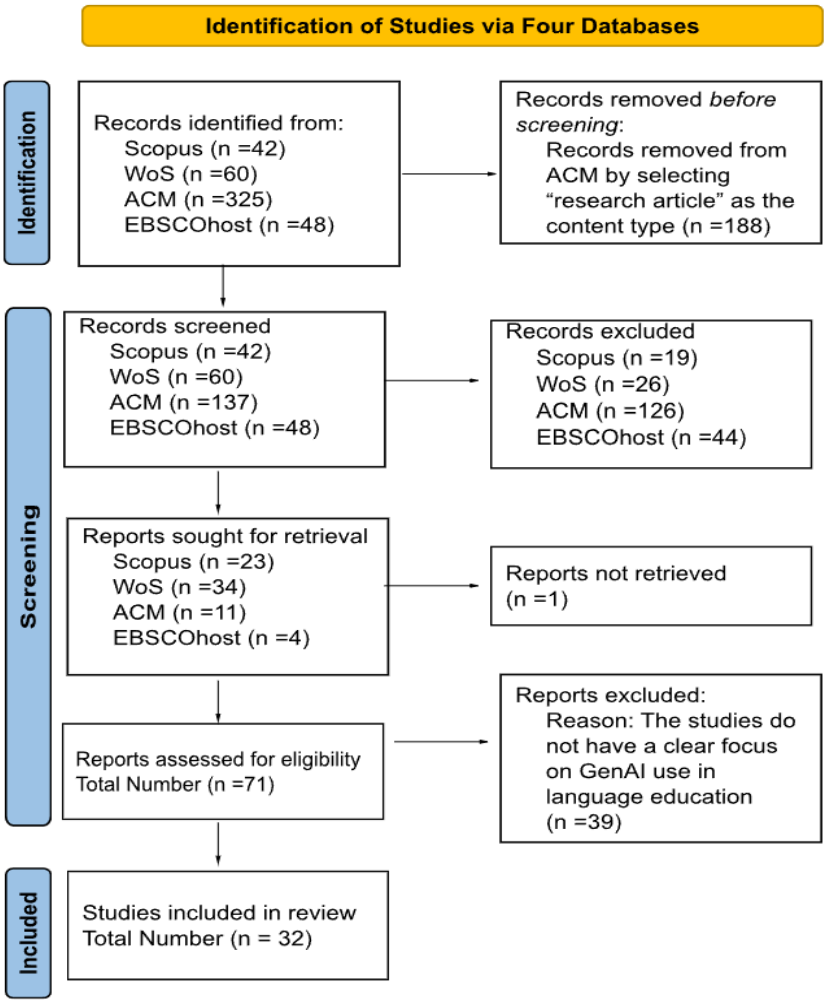


Figure 1. Flowchart of Study Identification

Initially, by applying the search strings, we identified 42 records in Scopus, 60 in Web of Science (WoS), 325 in the ACM Digital Library, and 48 in EBSCOhost. Within the ACM database, we specified "research article" as the content type, which allowed us to filter out 188 records that did not meet our requirements. Subsequently, we screened the titles and abstracts, discarding 19 records from Scopus, 26 from WoS, 126 from ACM, and 44 from EBSCOhost based on our inclusion criteria (refer to Table 2). From this screening, 72 studies were deemed suitable for retrieval: 23 from Scopus, 34 from WoS, 11 from ACM, and 4 from EBSCOhost. Upon further examination of these 72 studies, one was excluded due to the absence of an English full-text version. A thorough read-through of the remaining papers revealed that 39 studies lacked a clear emphasis on the use of GenAI in language education, leading to their exclusion from our review. Consequently, we incorporated 32 empirical studies into our systematic review. To ensure their suitability, we evaluated the quality of these 32 studies using the modified criteria proposed by Dixon-Woods et al. (2006), and the assessment results are presented in Table 3.

Table 2. Inclusion and Exclusion Criteria

| Inclusion Criteria | Exclusion Criteria |
|---|---|
| 1. The study's full text was available in English. | 1.The study's full text was not written in English. |
| 2. The study was published by July 31st, 2024, in peer-reviewed journals or conference proceedings. | 2. The study was not published in a peer-reviewed journal or conference proceedings and/or was published after July 31st, 2024. |
| 3. The study was primary research that collected original data and applied qualitative, quantitative, or mixed methods as the study methodology. | 3.The study did not collect original data via quantitative, qualitative, or mixed methods. |
| 4. The study had a clear focus on leveraging GenAI to bridge educational divides and/or foster equity in second language and/or foreign language education context(s). | 4. The study did not have a clear focus on leveraging GenAI to bridge educational divides and/or foster equity in second language and/or foreign language education context(s). |

Table 3. Quality Assessment of the Included Studies

| Criteria | Totally met | Partially met | Not met |
|---|---|---|---|
| Are the aims and objectives of the research clearly stated? | 32 | | |
| Is the research design clearly specified and appropriate for the aims and objectives of the research? | 32 | | |
| Do the researchers provide a clear account of the process by which their findings were reproduced? | 32 | | |
| Do the researchers display enough data to support their interpretations and conclusions? | 30 | 2 | |
| Is the method of analysis appropriate and adequately explicated? | 32 | | |

**Data Analysis**

To analyze the 32 studies for this review, we applied the advanced convergent QUALITATIVE meta-integration technique (Frantzen & Fetters, 2016) and thematic analysis (Boyatzis, 1998). By employing the advanced meta-integration method, we were able to transform quantitative data and evidence into qualitative ones and synthesize all relevant data and evidence collected from quantitative, qualitative, and mix-methods empirical studies included in this systematic review. Then, we conducted thematic analysis to identify themes related to our research question. To enhance the reliability of our data analysis, we both scrutinized the 32 studies and independently coded each study line-by-line to seek underlying themes. We also compared our codes, discussed the key categories of our codes, and reached consensus on essential themes regarding how to harness GenAI in narrowing the education gap and promoting equity in language education during our meetings.

The meta-integration approach empowered us to analyze data from methodologically diverse research, including 3 quantitative, 14 qualitative, and 15 mixed-methods empirical studies. In addition, questionnaires and interviews were the most popular methods for data collection adopted by the researchers. Out of the 32 studies, 5 are conference papers and 27 are journal articles. It is worth mentioning that 68% of the included studies were published by July 31st in 2024 while only 10 studies were published before 2024 (see Figure 2).

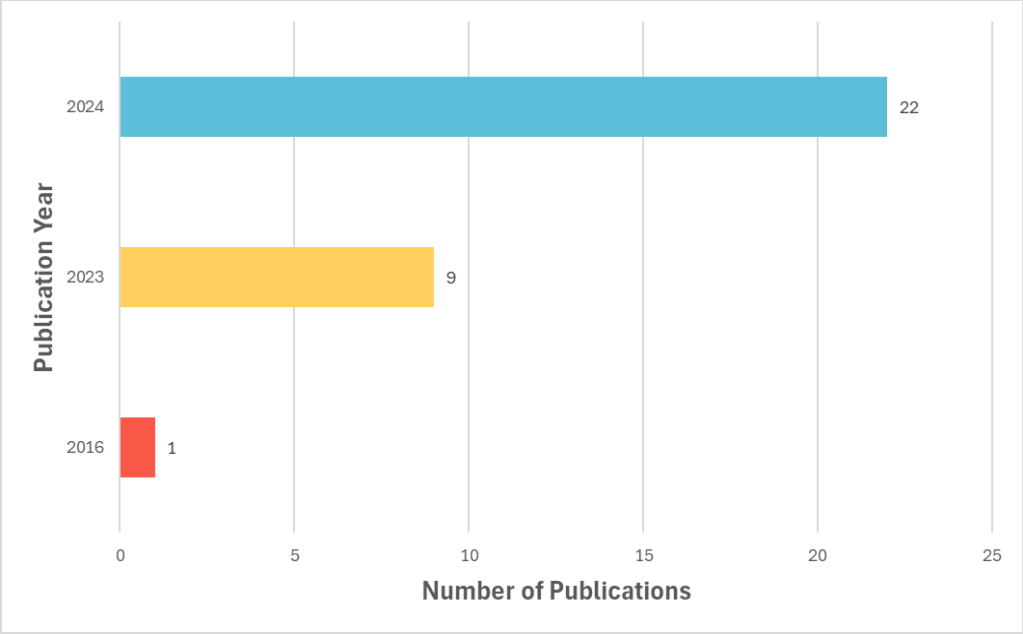


Figure 2. Number of Studies Published by Year

Regarding the geographic distribution of the first author, the majority of the studies were conducted in Asia (N= 23), followed by North America (N= 5), Europe (N= 3), and Occeania (N= 1). China, Turkey, the U.S., and Canada are the only four countries out of a total of 16 that contributed at least two papers, with China as the most productive country (see Figure 3). Besides, 15 studies were conducted via domestic collaborations and 9 studies via international collaborations. Among the 32 included studies, 8 of them were written by a single author, 5 by two authors, 10 by three authors, and 9 by at least four authors.

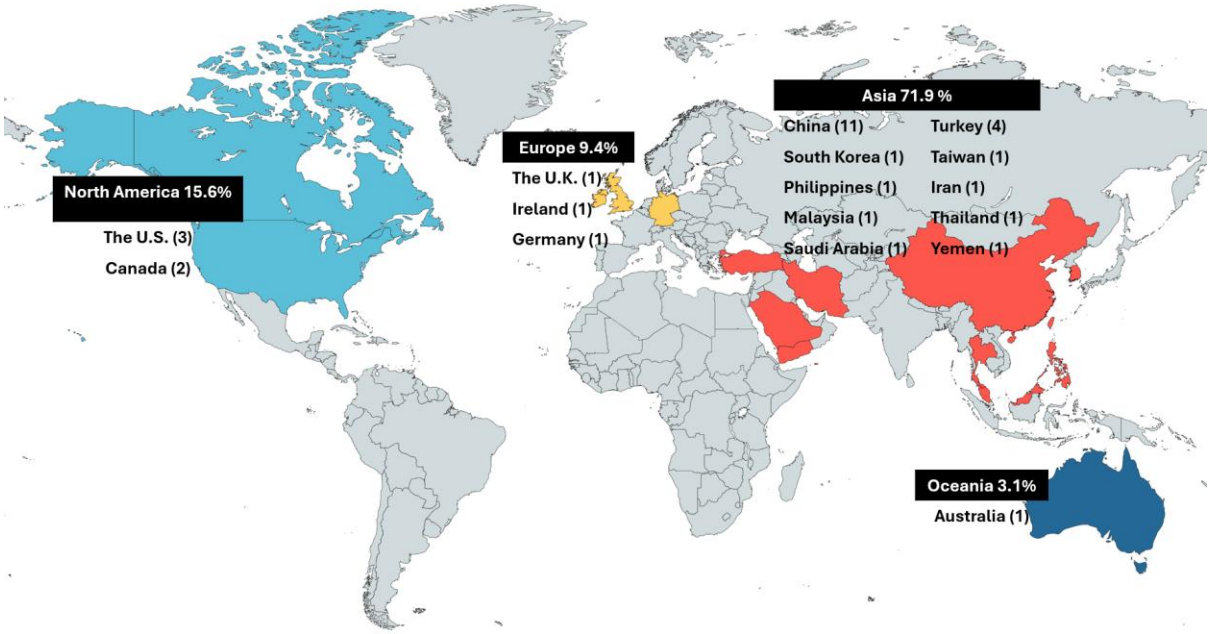


Figure 3. Geographic Distribution of Authorship

In terms of educational levels, 18 studies were conducted in higher education contexts, 9 studies in K-12 education contexts, and 1 study in both K-12 and higher education contexts (see Figure 4). By contrast, 4 studies did not specify the educational levels. We also found that English as a second language or foreign language is the most widely studied language education context that has attracted much research attention. It is interesting to note that ChatGPT is the most popular GenAI tool examined by researchers across various educational levels and language education contexts. Table 4 provides detailed information about different aspects of each included study, consisting of year of publication, first author country/region, author collaboration, methods/data collection, GenAI application, language context, research foci, and education level(s).

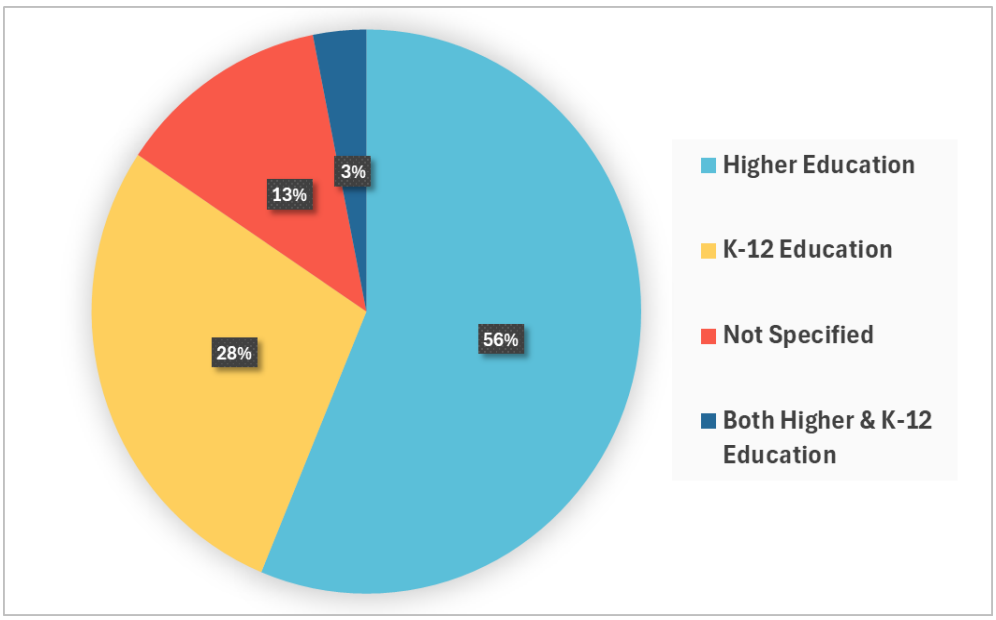


Figure 4. Educational Level(s) of the Included Studies

Table 4. Characteristics of the Included Studies

| Author/ Year | 1st Author Country/ Region & Collaboration | Document Type | Data Collection | GenAI Application | Language Context | Research Foci | Education Level(s) |
|---|---|---|---|---|---|---|---|
| Alenizi et al. (2023) | Saudi Arabia<br><br>Domestic collaboration among three authors | Journal Article | Mixed methods<br><br>Survey questionnaire from 199 English as a foreign language (EFL) teachers | ChatGPT | EFL special education | Examining teachers' attitudes towards using ChatGPT in language instruction | K-12 education |
| Annamalai (2024) | Malaysia | Journal Article | Qualitative study<br><br>Interview data from 26 teachers in secondary schools | ChatGPT | English as a second language (ESL) education | Investigating teachers' perspectives regarding factors that could impact their switching intention from traditional classroom instruction to ChatGPT use in English education | K-12 education |
| Bin-Hady et al. (2023) | Yemen<br><br>International collaboration among four authors in Yemen, Jordan, and Saudi Arabia | Journal Article | Qualitative study<br>online open discussions from 20 Research Gaters | ChatGPT | ESL/EFL learning | Scrutinizing Research Gaters' responses regarding how to use ChatGPT in English language learning for ESL/EFL learners | Higher education |
| Cai et al. (2023) | China<br><br>Domestic collaboration among three authors | Journal Article | Mixed methods<br><br>Questionnaires from 458 college students | ChatGPT | Language learning | Studying college students' attitudes regarding ChatGPT-assisted language learning | Higher education |
| Chen et al. (2024) | China<br><br>International collaboration among seven authors in China and the U.S. | Conference Proceeding | Mixed methods<br><br>Quiz, test and other data collected from 24 university students | RetAssist, an interactive vocabulary learning system | ESL education | Investigating how RetAssist can help improve learners' vocabulary learning | Higher education |
| Chiaráin & Chasaide (2016) | Ireland<br><br>Domestic collaboration | Conference Proceeding | Quantitative study<br><br>Questionnaires from 228 students | Speech-enabled chatbot | Irish language education | Examining students' reactions to the chatbot and their opinions about the | K-12 education |

| Author/ Year | 1st Author Country/ Region & Collaboration | Document Type | Data Collection | GenAI Application | Language Context | Research Foci | Education Level(s) |
|---|---|---|---|---|---|---|---|
| | between two authors | | | | | platform | |
| De Vargas et al. (2024) | Canada<br><br>International collaboration among four authors in Canada and the U.S. | Conference Proceeding | Mixed methods<br><br>Internal testing conducted with four speech-language pathologists (SLP) and two special education teachers along with questionnaires, board creation, open-ended question responses collected from 8 SLPs | QuickPic that creates topic-specific communication boards from photos | Augmentative and Alternative Communication-based language instruction for non-speaking individuals | Studying SLPs' and special education teachers' opinions regarding QuickPic design, interaction, vocabulary quality, and overall use of QuickPic | Not specified |
| Duong & Suppasetseree (2024) | Thailand<br><br>Domestic collaboration between two authors | Journal Article | Mixed methods<br><br>Quasi-experiment (speaking tests, questionnaires, and interviews) with 30 Vietnamese undergraduate students | AI voice chatbot | EFL education | Investigating the effects of an AI voice chatbot on EFL students' English speaking skills | Higher education |
| Fathi & Rahimi (2024) | Iran<br><br>Domestic collaboration between two authors | Journal Article | Qualitative study<br><br>Essays, interviews, observations, reflective journals and other data collected from 14 EFL learners ranging in age from 19 to 25 | ChatGPT | EFL education | Examining how AI-enhanced writing meditation helped EFL students' academic writing | Not specified |
| Foung et al. (2024) | Canada<br><br>International collaboration among three authors in Canada and China | Journal Article | Qualitative study<br>Written reflections of 74 students and focus group interviews with 28 students | ChatGPT | EFL education | Exploring how students used GenAI tools in their writing assessments | Higher education |
| Guo et al. (2024) | China<br><br>Domestic collaboration | Journal Article | Qualitative study<br>Pre- and post-workshop lesson plans and interview | Argumate | EFL education | Studying EFL teachers' chabot-enhanced lesson planning for teaching | K-12 education |

| Author/ Year | 1st Author Country/ Region & Collaboration | Document Type | Data Collection | GenAI Application | Language Context | Research Foci | Education Level(s) |
|---|---|---|---|---|---|---|---|
| | among four authors | | data collected from 10 Chinese EFL teachers | | | argumentative writing | |
| Han et al. (2024) | U.S.<br><br>Domestic collaboration among seven authors | Journal Article | Qualitative study<br><br>Interview data from 16 teachers and 12 students; survey and interview data from 12 parent participants | ChatGPT and Stable Diffusion | Elementary literacy education | Investigating teachers, parents, and students' perspectives on incorporating GenAI in elementary literacy | K-12 education |
| Jeon (2024) | Republic of Korea | Journal Article | Qualitative study<br><br>Interview data and student-chatbot interaction logs from 36 Korean students who were 12 years old | Chatbots created using Google's Dialogflow | EFL education | Examining students' experiences and perspectives about the benefits of using AI chatbots in the primary ESL classroom | K-12 education |
| Karaosmanoglu et al. (2024) | Germany<br><br>Domestic collaboration among seven authors | Conference Proceeding | Qualitative study<br><br>Observations of a group of 6 players aged 14-24 | ChatGPT | French language learning | Exploring how the integration of ChatGPT into Language of Zelda, an educational game, could facilitate French language learning | Not specified |
| Karataş et al. (2024) | Turkey<br><br>International collaboration among five authors in Turkey and Canada | Journal Article | Qualitative study<br><br>Interview data collected from 13 preparatory class students | ChatGPT | Foreign language learning | Studying the students' perspectives on the effect, benefits, and limitations of ChatGPT-assisted language learning | Higher education |
| Kartal (2024) | Turkey | Journal Article | Qualitative study<br><br>Interview and weekly written narratives from 12 English student teachers | ChatGPT | EFL education | Scrutinizing the impact of ChatGPT on English student teachers' thinking skills and creativity during their practicum | Higher education |
| Lee et al. (2024) | China<br><br>International collaboration | Journal Article | Mixed methods<br><br>Pre- and post-test surveys and | AI chatbots | Global Englishes language teaching | Investigating the effects of AI chatbot tasks on preservice teachers' awareness | Higher education |

| Author/ Year | 1st Author Country/ Region & Collaboration | Document Type | Data Collection | GenAI Application | Language Context | Research Foci | Education Level(s) |
|---|---|---|---|---|---|---|---|
| | among three authors in China, the U.S., and South Korea | | interviews from 97 preservice English teachers | | | of Global Englishes | |
| Li, Bonk, & Kou (2023) | U.S.<br>Domestic collaboration among three authors | Journal Article | Qualitative study<br>Content analysis of videos from various language teaching and learning communities on YouTube | ChatGPT | Language learning | Examining YouTubers' perceptions about benefits of ChatGPT use in language learning | Not specified |
| Li, Li, & Cho (2023) | U.S.<br>Domestic collaboration among three authors | Journal Article | Mixed method<br>Writing scores, samples, and reflections from four 9th-grade students | ChatGPT | Chineses language learning | Exploring the effectiveness of ChatGPT use in improving Chinese language learners' Chinese writing | K-12 education |
| Liu, Darvin, & Ma (2024) | China<br>International collaboration among three authors in China and Canada | Journal Article | Mixed methods<br>Questionnaire and interview data from 867 Chinese EFL learners | Large language model platforms, such as ChatGPT and Bing Chat | EFL education | Investigating Chinese EFL learners' acceptance and adoption of GPT chatbots | Higher education |
| Liu, Park, & McMinn (2024) | China<br>Domestic collaboration among three authors | Journal Article | Mixed methods<br>Survey data from 475 university students and interview data from 12 university students | Gen AI tools, such as ChatGPT | Learning academic communication skills, especially in writing, grammar, vocabulary, and reading | Scrutinizing students' perceptions toward GenAI tools for English academic communication | Higher education |
| Mabuan (2024) | Philippines | Journal Article | Mixed methods<br>Focus group discussion and survey data from 115 English language teachers in elementary schools, high schools, and colleges | ChatGPT | English language teaching | Studying teachers' perceptions on the ChatGPT use in English language teaching | K-12 & higher education |
| Özçelik & Ekşi | Turkey | Journal | Qualitative study | ChatGPT | English language | Examining students' | Higher |

| Author/ Year | 1st Author Country/ Region & Collaboration | Document Type | Data Collection | GenAI Application | Language Context | Research Foci | Education Level(s) |
|---|---|---|---|---|---|---|---|
| (2024) | Domestic collaboration between two authors | Article | Observation, fieldnotes, and interview data from 11 undergraduate students | | writing | perceptions about using ChatGPT as a learning assistant in helping them enhance their writing | education |
| Qu & Wu (2024) | Australia<br>International collaboration between two authors in Australia and the U.K. | Journal Article | Quantitative study<br>Survey data collected from 189 Chinese international students enrolled in British universities | ChatGPT | ESL education | Exploring ESL learners' perceptions regarding ChatGPT use for computer-assisted language learning | Higher education |
| Wu et al. (2024) | China<br>Domestic collaboration among four authors | Journal Article | Mixed methods; Pretest and posttest data and dialog logs collected from 44 Mandarin second language learners in the second grade | MSLIPA, an intelligent personal assistant(IPA) | Mandarin second language learning | Investigating the effects of an IPA on Mandarin second language learners' listening and speaking abilities as well as the characteristics of and strategies for their interaction with the IPA | K-12 education |
| Yan (2023) | China | Journal Article | Qualitative study<br>Survey data from eight Chinese undergraduate EFL majors and video-based classroom recordings | ChatGPT | EFL education | Studying students' perceptions about ChatGPT application in L2 writing | Higher education |
| Yeh (2024) | Taiwan | Journal Article | Qualitative study<br>Lesson plans, transcriptions of microteaching sessions, observation notes, and the reflective notes from 13 in-service teachers enrolled in an advanced graduate course | GenAI tools, such as ChatGPT | EFL education | Examining the impact of GenAI tools on inservice EFL teachers' pedagogy | Higher education |
| Yıldız (2023) | Turkey | Journal Article | Quantitative study<br>Pre- and post-test | ChatGPT | ESL education | Researching the effect of ChatGPT-generated dialogues | Higher education |

| Author/ Year | 1st Author Country/ Region & Collaboration | Document Type | Data Collection | GenAI Application | Language Context | Research Foci | Education Level(s) |
|---|---|---|---|---|---|---|---|
| | | | data from 60 second-year university students | | | on language learners' motivation and engagement and relevant factors that impact the effectiveness of ChatGPT-generated dialogues in enhancing language learners' motivation | |
| Yu et al. (2024) | China<br><br>International collaboration among three authors in China and Cape Verde | Journal Article | Mixed methods<br><br>Writings and interview data from 46 university students | ChatGPT | Chinese as a second language | Examining how ChatGPT influence African students' Chinese L2 writing proficiency | Higher education |
| Yuan (2023) | U.K. | Journal Article | Mixed methods<br><br>Survey, test and interview data from 74 Chinese elementary school students and interview data collected from two English teachers | Mondly, an AI chatbot developed for language learning | EFL education | Studying how effective AI-chatbots can enhance students' oral English proficiency and communication willingness as well as how teachers can better incorporate AI chatbots into teaching | K-12 education |
| Zhang et al. (2023) | China<br><br>Domestic collaboration among three authors | Journal Article | Mixed methods<br>Interview, pre-post argumentative writings, and pre-post questionnaire data collected from 15 Chinese EFL undergraduate and graduate students who were engaged in argumentative writing training | A chatbot created by the first author using ManyChat | EFL education | Investigating how an training on logical fallacies might help enhance students' EFLL argumentative writing and writing self-efficacy | Higher education |
| Zhao et al. (2024) | China<br><br>Domestic collaboration | Conference Proceeding | Mixed-methods<br><br>Pre-and post survey data collected from | Language Urban Odyssey (LUO) , a serious game design that | Second language acquisition | Examining LUO's effect on improving students' second language acquisition | Higher education |

| Author/ Year | 1st Author Country/ Region & Collaboration | Document Type | Data Collection | GenAI Application | Language Context | Research Foci | Education Level(s) |
|---|---|---|---|---|---|---|---|
| | among eight authors | | six bilingual undergraduate and postgraduate students in their twenties | leverages ChatGPT 3.5's capabilities for simulating real language use | | | |

## Findings

This section discusses how GenAI can contribute to bridging educational gaps and promoting inclusivity in language education based on the selected studies. According to our analysis of the included studies, we found that GenAI can boost language learners' motivation and confidence through multiple ways and individualize instruction to save teachers' time on preparation tasks and provide better learning experiences to students. Also, GenAI can expose language learners to diverse cultures to enhance their cultural understanding and reduce biases in language learning. Further, GenAI can empower students from disadvantaged backgrounds with affordable learning tools although some studies highlighted the divide between free and premium offerings.

### Increasing Learning Motivation and Confidence

GenAI can significantly contribute to increasing language learners' motivation and enhancing their confidence by providing instant feedback on their exercises (e.g., Annamalai, 2024; Jeon, 2024; Karataş et al., 2024; Li, Bonk & Kou, 2023; Fathi & Rahimi, 2024; Li, Li, & Cho, 2023; Liu, Park, & McMinn, 2024; Özçelik & Ekşi, 2024), making the learning process more engaging and less intimidating (e.g., Annamalai, 2024; Bin-Hady et al., 2023; Cai et al., 2023; Chen et al., 2024; Chiarian & Chasaide, 2016; Fathi & Rahimi, 2024; Guo et al., 2024; Jeon, 2024; Karaosmanoglu et al., 2024; Karataş et al., 2024; Kartal, 2024; Qu & Wu, 2024; Lee et al., 2024; Yeh, 2024; Yıldız 2023; Yu et al., 2024; Yuan 2023; Zhang, Zou, & Cheng 2023; Zhao et al. 2024), creating interactive and conversational experiences (e.g., Annamalai, 2024; Karataş et al., 2024; Mabuan, 2024; Wu et al., 2024), offering flexibility that enables learners to learn at anytime and anywhere (e.g., Annamalai, 2024; Cai et al., 2023; Duong & Suppasetseree, 2024), and accessing a vast array of resources they might not otherwise find (e.g., Cai et al., 2023; Mabuan, 2024; Yan, 2023). Regarding providing instant feedback, Karataş et al.'s (2024) qualitative case study provided an in-depth exploration of the nuanced effects of AI on the foreign language learning process within its real-world educational context. Participants in the study reported that with instant feedback on word usage from ChatGPT, they found their motivation bolstered, appreciating the tool's ability to immediately address errors. Also, quite a few studies revealed GenAI-assisted language learning is not only engaging but also reducing learners' anxiety. For instance, Jeon's (2024) study on thirty-six Korean primary school learners regarding their experiences and perceptions of chatbots found that most students reported developing a more positive attitude toward English class when interacting with chatbots, as they provided a non-judgmental environment where they could practice without feeling self-conscious about making mistakes in front of peers or teachers. This reduction in pressure was significant, with students expressing appreciation for the ability to practice independently, at their

own pace, and without the worry of holding back or being corrected by others, and this anxiety-free space encouraged students to engage more willingly in speaking English, contributing to a more constructive learning experience.

GenAI can also understand and generate human-like text, allowing it to engage with language learners in natural and fluid conversations. To give an example, Wu et al. (2024) investigated the effect of an intelligent personal assistant called MSLIPA on Mandarin second language learners. Through comparing the effect between an experimental group (communicated with MSLIPA) and a control group (communicated with their peers), the study found that the students using MSLIPA engaged in more dialogue rounds than a control group, especially following an initial peak indicative of a novelty effect. The experimental group also demonstrated a variety of interaction strategies when faced with communication breakdowns, suggesting a higher level of conversational engagement than the control group. Quantitative data showed the experimental group had significantly more dialogue rounds, with a diverse use of strategies like rephrasing, repeating, and explaining, in contrast to the control group, which mostly abandoned conversations or repeated themselves. This suggests that MSLIPA helps sustain conversations and encourages users to improve their language skills.

Allowing learners to learn at anytime and anywhere is one essential aspect of GenAI that increases learners' motivation and autonomy. GenAI can create virtual environments or simulations that allow learners to practice skills and concepts in a virtual setting. This feature can happen anytime and anywhere, providing practice arenas that would be impossible or impractical in the real world due to constraints of cost, safety, or physical distance. Through an 8-week quasi-experiment conducted with 30 Vietnamese undergraduate students, Duong and Suppasetseree (2024) found that in addition to offering an engaging experience like communicating with a native speaker, the AI voice chatbot examined in their study provides flexibility in language learning by allowing students to practice English speaking at any time and from any location. Finally, GenAI enables learners to access a vast array of resources. Cai et al. (2023), as an example, showed that one of the most frequent language learning strengths in ChatGPT is its diverse resources.

**Individualizing Instruction and Adapting to Different Student Needs**

GenAI technology contributes to inclusive education by providing efficient, user-friendly, and effective means to support communicative interactions in educational settings. It indicates a transformative impact on the way educational resources are created and utilized, promoting a more inclusive and supportive learning environment for students with special needs. Two studies (Alenizi et al., 2023; De Vargas et al., 2024) explored how GenAI can support special education students in language learning. For example, Alenizi et al. (2023) used mixed-methods to explore the attitudes of 199 English as a Foreign Language (EFL) special education teachers towards using ChatGPT for language learning. The insights from the teachers emphasize scaffolding, individualization, collaboration, and support for inclusive practices. They reveal that ChatGPT can be integrated into language instruction for special education students in multiple ways, including (1) breaking down language into smaller, more manageable parts to cater to the individual needs of students, thereby making language learning more accessible; (2) adapting to each student's learning style, pace, and needs with personalized instruction and

feedback, providing an inclusive learning experience; and (3) ensuring that the use of ChatGPT is accessible for all learners, considering diverse needs, providing ongoing training for teachers, and involving continuous evaluation.

In addition to special education students, five studies (e.g., Annamalai, 2024; Bin-Hady et al., 2023; Fathi & Rahimi, 2024; Li, Bonk & Kou, 2023; Yeh, 2024) show that GenAI can personalize learning experiences for language learners in general. For example, Li, Bonk and Kou (2023) explored the integration of ChatGPT in self-directed language learning (SDLL) as perceived by YouTube content creators. The study found that ChatGPT is reshaping language education by enabling a personalized, adaptable, and learner-driven approach. It not only supports the individualization of content but also encourages a more strategic, reflective, and self-sufficient language learning journey, cultivating skills beyond the language being learned. As an example, ChatGPT advocates for a shift from rigid teaching methods to "buffet-style" learning that is ultra-personalized, allowing learners to select content based on their interests, akin to choosing from a buffet. GenAI tools, thus, become adaptive instruments tailored to individual learner desires. Similarly, Yeh (2024) investigates how AI can enhance language teaching by personalizing content and supporting the development of communicative skills. The analysis of study reveals that ChatGPT aids in-service teachers in customizing song lyrics and other materials to closely fit lesson objectives, leveling the content to student comprehension abilities. Overall, the collective narratives reflect how GenAI technology empowers teachers to move beyond traditional teaching models, fostering a learning environment that is more engaging, effective, and student-centered. AI's adaptability, combined with teacher creativity, leads to personalized language education that matches the unique needs and learning styles of students

**Enhancing Cultural Understanding and Reducing Biases in Language Learning**

Five studies demonstrate that the use of GenAI can enhance language learners' understanding of diverse cultures (e.g., Han et al., 2024; Karataş et al., 2024; Mabuan, 2024; Liu, Davin, & Ma, 2024) and reduce their biases in language learning, such as native-speakerism (e.g., Lee et al., 2024). The point that GenAI can strengthen language learners' understanding of diverse cultures is evident in Han et al. (2024). Han et al. (2024) examined how teachers, parents, and students perceive and suspect GenAI systems in elementary school settings. The findings pinpoint that GenAI can provide culturally relevant examples and feedback. This helps teachers go beyond their personal experiences and cultural biases. By generating examples from different cultures, for various language constructs such as active and passive voice, teachers can create more inclusive content. Also, GenAI can create culturally salient stories, like fables, which resonate with a learner's heritage. A parent mentioned using GenAI to generate content about traditional Asian dragons to teach her child about their cultural heritage. Further, by incorporating culturally relevant materials like traditional stories and holiday practices into lessons, teachers can create classes that not only teach a language but also promote cross-cultural understanding amongst students from different backgrounds, and GenAI can serve as a bridge between various family cultures, encouraging the sharing of cultural values and languages, thereby fostering a sense of belonging and a stronger family connection.

Lee et al.'s (2024) study shows how GenAI can reduce biases in language learning. Lee et al. (2024) investigates

the potential of English as a lingua franca (ELF) interaction with AI chatbots in raising Global English (GE) awareness. The survey results and interview data point to GenAI's significant role in reducing native-speakerism and enhancing GE awareness among pre-service teachers in an English education program. Statistically significant differences in GE awareness were found between the control group (CG) and experimental groups (EG1 and EG2) who were exposed to GE activities, including AI chatbot interventions. These activities positively influenced participants' perception of English. Initially, many participants held a native-speaker bias, viewing "standard" English as the authentic benchmark. However, after interacting with AI chatbots, participants reported a critical reflection on their own English and other varieties. This indicates a shift towards a more inclusive understanding of English as an international language, with diverse versions being acceptable and intelligible for communication. Participants in EG2, after engaging with AI chatbot activities, demonstrated increased acceptance and understanding of different Englishes, recognizing their linguistic and cultural value. They reported gaining confidence in their own variety of English and showed more willingness to communicate with speakers of other Englishes.

**Empowering Students with Affordable Learning Tools**

Four studies (Foung et al., 2024; Li, Li, & Cho, 2023; Liu et al., 2024; Mabuan, 2024) illustrate how GenAI applications in language education can offer numerous benefits that can particularly support learners from low-income families. For instance, Li, Li, and Cho's (2023) study explores the potential of ChatGPT in supporting and empowering Chinese language learners (CLLs) whose first language is English to enhance their writing skills. The study involving four emergent Chinese language learners from diverse proficiency levels utilized ChatGPT, revealing substantial improvements in writing skills. The intervention showed that all participants improved their Mandarin writing scores during the GenAI-supported phase. There was a clear rise in mean scores, with the initial phase showing lower means and a significant increase during the intervention. Even after the intervention ceased, all students maintained higher scores than their initial baselines, indicating the lasting impact of the GenAI tool. Reflective accounts from the students underscored a great sense of empowerment. ChatGPT provided them with immediate feedback and assisted with homework, akin to always having a Mandarin teacher available. This support seemed crucial as it allowed for continuous learning and improvement, bridging the gap that might exist due to the lack of linguistic support at home or the inability to afford private tutoring. One student appreciated ChatGPT as their private tutor, indicating the differential impact this tool could make for learners without access to additional educational support. These findings highlight that ChatGPT shows promise as a supportive tool for CLLs from low-income families, reducing educational inequality and promoting equitable access to language learning opportunities.

However, two studies (Foung et al., 2024; Yu et al., 2024) caution that GenAI, despite its potential benefits, could exacerbate equity issues due to economic disparities among students. For example, Foung et al.' s (2024) study revealed while all participants in the study could access basic versions of AI tools like WeCheck!, Grammarly, and ChatGPT, their experiences highlighted the divide between free and premium offerings. Students considered cost when evaluating AI tools, with many relying on free versions and viewing the premium versions as either affordable on a "per day" basis or entirely unattainable due to their financial status, with some perceiving

themselves as "poor." The free versions of these tools offer fewer features than paid ones. For instance, the premium Grammarly provides article summaries and more specific writing suggestions, which are not available in the free version. Therefore, students from lower-income backgrounds may not have the same level of support as those who can afford premium versions of AI tools, contributing to an unequal playing field. The study suggests that higher education institutions should adopt policies to address this inequity, ensuring that students from disadvantaged communities have equal support compared to those who can afford the more feature-rich, paid versions of AI tools.

## Discussion

The potential of GenAI to transform language education is evident through its impact on the creation of inclusive and personalized learning experiences, as well as the enablement of intercultural understanding and accessibility. Its capabilities extend to producing human-like text, translating languages, generating personalized materials, and facilitating more engaging learning experiences—all of which offer profound implications for the field of language education.

GenAI's role in enhancing learning motivation and confidence is marked by its ability to provide instant feedback to learners. As demonstrated by Karataş et al. (2024), this immediate response system can address learners' errors, boosting their motivation and reducing language learning anxiety, as in the findings of Jeon (2024). This aspect of GenAI aligns well with Vygotsky's (1978) Zone of Proximal Development, allowing learners to successfully bridge the gap between what they can do independently and what they can achieve with guidance. Furthermore, GenAI enables the creation of conversational experiences that mimic natural human interactions, as seen in the study by Wu et al. (2024). This technology promotes autonomy in language learning, allowing individuals to engage in language practice in a safe, responsive environment. The ubiquitous nature of GenAI platforms, agnostic of time and place, affirms their role in facilitating self-directed learning and wider access to language learning resources, which is a pivotal factor in lifelong learning and continuous professional development.

Individualization in instruction is another critical domain where GenAI exhibits a transformative influence. The technology's adaptability to learners' unique needs ensures differentiated instruction. Alenizi et al.'s (2023) and De Vargas et al.'s (2024) studies demonstrate how GenAI can cater to special education requirements, enhancing inclusivity. The integration of GenAI in language instruction can align with Universal Design for Learning (UDL) principles (CAST, 2024), offering multiple means of engagement, representation, action, and expression. Cultural understanding and the reduction of biases in language learning, highlighted in studies by Han et al. (2024) and Lee et al. (2024), suggest that GenAI can serve as culturally intelligent tools. They help learners navigate through and appreciate the cultural nuances embedded in language, moving beyond the "one-size-fits-all" approach that often pervades traditional language education models. This fosters an environment of cultural pluralism and reduces the prevalence of native-speakerism, thereby aligning language education with global communicative competencies.

GenAI can contribute to the democratization of language learning by providing high-quality educational

experiences that were previously only accessible to those with greater means. Yet, as Foung et al. (2024) highlighted, a distinction between free and premium services can create a new form of disparity; thus, dialogue on equitable resource allocation must be ongoing. Additionally, GenAI allows for high-quality conversational partners and language resources, democratizing access to language practice with native-like experiences. With tools like MSLIPA improving engagement for Mandarin learners (Wu et al., 2024), students can access quality language practice irrespective of their geographical or socio-economic status. Finally, with AI's scalable nature, adopting GenAI tools in public and low-resource education settings can reduce systemic educational inequalities, ensuring that all students, regardless of their school's funding or resources, have access to advanced and supportive educational tools.

## Conclusions and Implications

The systematic analysis of existing research studies on GenAI's application in language education demonstrates its potential to bridge educational gaps by tailoring learning experiences to individual student needs and promoting inclusivity by providing all learners with equal access to educational opportunities and resources. Nonetheless, it is crucial to approach the integration of GenAI with a discerning and ethical perspective. This will help prevent the worsening of current inequalities and ensure that its benefits are accessible to all learners. Furthermore, while GenAI offers transformative possibilities for language education, to fully and responsibly harness these opportunities requires a collective effort in research, pedagogy, teacher professional training, and policy formulation. Each of these domains must address its own set of inherent challenges, including ensuring equitable access, upholding ethical standards, and creating environments that support lifelong and personalized learning. The subsequent sections outlines future directions for research, pedagogy, professional development for teachers, and policy-making.

## Future Research

Future research could explore the efficacy of different types of GenAI tools across diverse learning environments and populations. Longitudinal studies are needed to assess the long-term impact of GenAI on language acquisition and retention. Additionally, comparative studies between traditional and GenAI-integrated pedagogies could offer deeper insights into the effectiveness of such technologies. Also, there is a significant opportunity for research into the analytics provided by GenAI platforms and how these can inform language learning progress. Interdisciplinary research involving education technology, data science, and linguistics could provide a rounded view of how data-driven insights contribute to or hinder language learning. Further, explorations into the ethical implications of AI in education and potential biases in GenAI outputs remain essential. Research should address how GenAI systems can be audited for fairness, accuracy, and cultural sensitivity to ensure they do not perpetuate biases or inequalities.

### *Pedagogy*

Pedagogical frameworks need to evolve to integrate GenAI technology in ways that complement and enhance

language learning. Teachers could focus on developing personalized learning pathways that utilize GenAI for differentiated instruction to cater to varied educational needs. In addition, GenAI should be woven into blended learning models, combining traditional teaching with AI to optimize language education. Teacher training could incorporate strategies to effectively integrate GenAI into lesson planning and delivery.

*Teacher Professional Development*

Professional development programs must equip teachers with not only the technical know-how for using GenAI tools but also the pedagogical skills required to effectively integrate these technologies into language teaching and assessment. Teachers also need to be prepared to navigate the ethical considerations of utilizing AI in education, including privacy concerns, data security, and mitigating potential biases within AI tools.

*Policy-Making*

Policies need to ensure equitable access to GenAI tools for all students, addressing the concerns identified by Foung et al. (2024) regarding potential disparities arising from premium services. Investment in infrastructure that supports the deployment of GenAI tools in underfunded and rural schools should be prioritized. Besides, policy-making should focus on establishing quality standards and regulatory frameworks for GenAI applications in education. This would involve setting benchmarks for content accuracy, pedagogical soundness, and data privacy. In addition, policy-makers could consider sustained funding models for schools to adopt GenAI tools, ensuring that the advent of AI in education does not exacerbate existing digital divides but rather narrows them. Last, fostering partnerships between educational institutions, technology companies, and government agencies can help to facilitate the responsible development and integration of GenAI across learning environments.

## References

Alenizi, M. A. K., Mohamed, A. M., & Shaaban, T. S. (2023). Revolutionizing EFL special education: How ChatGPT is transforming the way teachers approach language learning. *Innoeduca: International Journal of Technology and Educational Innovation*, *9*(2), 5-23. https://dialnet.unirioja.es/servlet/articulo?codigo=9211301

Annamalai, N. (2024). Factors affecting English language high school teachers switching intention to ChatGPT: A Push-Pull-Mooring theory perspective. *Interactive Learning Environments*, 1–18. https://doi.org/10.1080/10494820.2024.2371928

Bin-Hady, W.R.A., Al-Kadi, A., Hazaea, A. and Ali, J.K.M. (2023), Exploring the dimensions of ChatGPT in English language learning: A global perspective, *Library Hi Tech*. https://doi.org/10.1108/LHT-05-2023-0200

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. Sage.

Cai, Q., Lin, Y., & Yu, Z. (2023). Factors influencing learner attitudes towards ChatGPT-assisted language learning in higher education. *International Journal of Human–Computer Interaction*, 1-15. https://doi.org/10.1080/10447318.2023.2261725

CAST (2024). *Universal design for learning guidelines version 3.0.* Retrieved from https://udlguidelines.cast.org

Chen, Q., Liu, S., Huang, K., Wang, X., Ma, X., Zhu, J., & Peng, Z. (2024, July). RetAssist: Facilitating vocabulary learners with generative images in story retelling practices. In Proceedings of the 2024 ACM Designing Interactive Systems Conference (pp. 2019-2036). https://doi.org/10.1145/3643834.3661581

Chiaráin, N. N., & Chasaide, A. N. (2016, May). Chatbot technology with synthetic voices in the acquisition of an endangered language: Motivation, development and evaluation of a platform for Irish. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16) (pp. 3429-3435).

Creely, E. (2024). Exploring the role of Generative AI in enhancing language learning: Opportunities and challenges. *International Journal of Changes in Education, 1*(3), 158-167. https://doi.org/10.47852/bonviewIJCE42022495

Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English language teaching: Affordances and challenges. *British Journal of Educational Technology*. **https://doi.org/10.1111/bjet.13460**

De Vargas, F. M., Yu, C., Shane, H. C., & Moffatt, K. (2024, May). Co-Designing QuickPic: Automated Topic-Specific Communication Boards from Photographs for AAC-Based Language Instruction. In Proceedings of the CHI Conference on Human Factors in Computing Systems (pp. 1-16). https://doi.org/10.1145/3613904.3642080

Dixon-Woods, M., Cavers, D., Agarwal, S., Annandale, E., Arthur, A., Harvey, J., ... & Sutton, A. J. (2006). Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. *BMC Medical Research Methodology*, *6*, 1-13. https://doi.org/10.1186/1471-2288-6-35

Duong, T., & Suppasetseree, S. (2024). The effects of an Artificial Intelligence voice Chatbot on improving Vietnamese undergraduate students' English speaking skills. *International Journal of Learning, Teaching and Educational Research, 23*(3), 293-321. https://doi.org/10.26803/ijlter.23.3.15

Faisal, E. (2024). Unlock the potential for Saudi Arabian higher education: A systematic review of the benefits of ChatGPT. *Frontiers in Education*, *9*, 1325601. https://doi.org/10.3389/feduc.2024.1325601

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. https://doi.org/10.1080/09588221.2024.2374772

Foung, D., Lin, L., & Chen, J. (2024). Reinventing assessments with ChatGPT and other online tools: Opportunities for GenAI-empowered assessment practices. *Computers and Education: Artificial Intelligence*, 100250. https://doi.org/10.1016/j.caeai.2024.100250

Frantzen, K. K., & Fetters, M. D. (2016). Meta-integration for synthesizing data in a systematic mixed studies review: Insights from research on autism spectrum disorder. *Quality & Quantity*, *50*, 2251-2277. https://doi.org/10.1007/s11135-015-0261-6

Fu, Y., & Weng, Z. (2024). Navigating the ethical terrain of ai in education: A systematic review on framing responsible human-centered AI practices. *Computers and Education: Artificial Intelligence*, 100306. https://doi.org/10.1016/j.caeai.2024.100306

Fu, Y., Weng, Z., & Wang, J. (2024). Examining AI Use in Educational Contexts: A Scoping Meta-Review and Bibliometric Analysis. *International Journal of Artificial Intelligence in Education*, 1-57. https://doi.org/10.1007/s40593-024-00442-w

Grassini, S. (2023). Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences*, *13*(7), 692. https://doi.org/10.3390/educsci13070692

Guo, K., Zhong, Y., Qin, J., & Chu, S. K. W. (2024). Investigating EFL teachers' lesson planning for chatbot-assisted learning of argumentative writing: A TPACK perspective. *Interactive Learning Environments*, 1–22. https://doi.org/10.1080/10494820.2024.2372646

Han, A., Zhou, X., Cai, Z., Han, S., Ko, R., Corrigan, S., & Peppler, K. A. (2024, May). Teachers, parents, and students' perspectives on integrating Generative AI into elementary literacy education. In Proceedings of the CHI Conference on Human Factors in Computing Systems (pp. 1-17).https://doi.org/10.1145/3613904.3642438

Heaven, W. D. (2020). OpenAI's new language generator GPT-3 is shockingly good—And completely mindless. *MIT Technology Review*, *29*, 1–6. https://www.technologyreview.com/2020/07/ 20/1005454/openai-machine-learning-language-generator-gpt-3- nlp/

Huang, F., Wang, Y., & Zhang, H. (2024). Modelling Generative AI acceptance, perceived teachers' enthusiasm and self-efficacy to English as a Foreign Language Learners' well-being in the digital era. *European Journal of Education*, *59*(4), e12770. https://doi.org/10.1111/ejed.12770

Jeon, J. (2024). Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives. *Computer Assisted Language Learning*, *37*(1-2), 1-26. https://doi.org/10.1080/09588221.2021.2021241

Jeon, J., & Lee, S. (2024). Can learners benefit from chatbots instead of humans? A systematic review of human-chatbot comparison research in language education. *Education and Information Technologies*, *29*, 23329–23360. https://doi.org/10.1007/s10639-024-12725-9

Jeon, J., Lee, S., & Choi, S. (2023). A systematic review of research on speech-recognition chatbots for language learning: Implications for future directions in the era of large language models. *Interactive Learning Environments*, *32*(8), 4613–4631. https://doi.org/10.1080/10494820.2023.2204343

Karaosmanoglu, S., Fittschen, E. L., Eyicalis, H., Kraus, D., Nickelmann, H., Tomko, A., & Steinicke, F. (2024, May). Language of Zelda: Facilitating language learning practices using ChatGPT. In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (pp. 1-5).

Karataş, F., Abedi, F. Y., Ozek Gunyel, F., Karadeniz, D., & Kuzgun, Y. (2024). Incorporating AI in foreign language education: An investigation into ChatGPT's effect on foreign language learners. *Education and Information Technologies*, *29*, 19343–19366. https://doi.org/10.1007/s10639-024-12574-6

Kartal, G. (2024). The influence of ChatGPT on thinking skills and creativity of EFL student teachers: A narrative inquiry. *Journal of Education for Teaching*, 1-16. https://doi.org/10.1080/02607476.2024.2326502

Koç, F. Ş., & Savaş, P. (2024). The use of artificially intelligent chatbots in English language learning: A systematic meta-synthesis study of articles published between 2010 and 2024. *ReCALL*, 1-18. https://doi.org/10.1017/S0958344024000168

Koraishi, O. (2023). Teaching English in the age of AI: Embracing ChatGPT to optimize EFL materials and assessment. *Language Education and Technology*, *3*(1). https://l angedutech.com/letjournal/index.php/let/article/download/48/37.

Law, L. (2024). Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review. *Computers and Education Open*, 100174.

https://doi.org/10.1016/j.caeo.2024.100174

Lee, S., Jeon, J., & Choe, H. (2024). Enhancing Pre-Service Teachers' Global Englishes Awareness with Technology: A Focus on AI Chatbots in 3D Metaverse Environments. *TESOL Quarterly*. https://doi.org/10.1002/tesq.3300

Li, B., Bonk, C. J., & Kou, X. (2023). Exploring the multilingual applications of ChatGPT: Uncovering Language Learning affordances in YouTuber videos. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT)*, *13*(1), 1-22. https:doi.org/ 10.4018/IJCALLT.326135

Li, X., Li, B., & Cho, S. J. (2023). Empowering Chinese language learners from low-income families to improve their Chinese writing with ChatGPT's assistance afterschool. *Languages*, *8*(4), 238. https://doi.org/10.3390/languages8040238

Li, B., Lowell, V. L., Wang, C., & Li, X. (2024). A systematic review of the first year of publications on ChatGPT and language education: Examining research on ChatGPT's use in language learning and teaching. *Computers and Education: Artificial Intelligence*, 100266. https://doi.org/10.1016/j.caeai.2024.100266

Liang, J. C., Hwang, G. J., Chen, M. R. A., & Darmawansah, D. (2023). Roles and research foci of artificial intelligence in language education: An integrated bibliographic analysis and systematic review approach. *Interactive Learning Environments*, *31*(7), 4270-4296. https://doi.org/10.1080/10494820.2021.1958348

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, *21*(2), 100790. https://doi.org/10.1016/j.ijme.2023.100790

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. https://doi.org/10.1080/09588221.2024.2310288

Liu, Y., Park, J., & McMinn, S. (2024). Using generative artificial intelligence/ChatGPT for academic communication: Students' perspectives. *International Journal of Applied Linguistic*s, *34*(4), 1437-1461. https://doi.org/10.1111/ijal.12574

Ma, H., Ismail, L. & Han, W. (2024). A bibliometric analysis of artificial intelligence in language teaching and learning (1990–2023): Evolution, trends and future directions. *Education and Information Technologies*. https://doi.org/10.1007/s10639-024-12848-z

Mabuan, R. A. (2024). ChatGPT and ELT: Exploring teachers' voices. *International Journal of Technology in Education*, *7*(1), 128-153.

Mohamed A. (2023). 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*. https://doi.org/10.1007/s10639-023- 11917-z.

Mun˜oz S, Gayoso G, Huambo A, Tapia R, Incaluque J, Aguila O, Cajamarca J, Acevedo J, Rivera H, Arias-Gonza´les J (2023). Examining the Impacts of ChatGPT on student motivation and engagement. *Social Space Journal*, *23*(1). https://socialspacejo urnal.eu/menu-script/index.php/ssj/article/download/156/68.

Ogunleye, B., Zakariyyah, K. I., Ajao, O., Olayinka, O., & Sharma, H. (2024). A systematic review of Generative AI for teaching and learning practice. *Education Sciences*, *14*(6), 636. https://doi.org/10.3390/educsci14060636

Özçelik, P. N., & Ekşi, Y. G. (2024). Cultivating writing skills: The role of ChatGPT as a learning assistant—a case study. *Smart Learning Environments*, *11*(1), 10. https://link.springer.com/article/10.1186/s40561-024-00296-8

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *bmj*, 372. https://doi.org/10.1136/bmj.n71

Qu, K., & Wu, X. (2024). ChatGPT as a CALL tool in language education: A study of hedonic motivation adoption models in English learning environments. *Education and Information Technologies*, 1-33. https://doi.org/10.1007/s10639-024-12598-y

Tai, T. Y., & Chen, H. H. J. (2024). The impact of intelligent personal assistants on adolescent EFL learners' listening comprehension. *Computer Assisted Language Learning*, *37*(3), 433-460.https://doi.org/10.1080/09588221.2022.2040536

Toboula, Z., & Martinien, C. (2023). Exploring the impact of AI-Powered collaborative and interactive NLP apps on EFL teaching in the post-Covid-19 era. https://papers. ssrn.com/sol3/papers.cfm?abstract_id=4398817.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wang, F., Cheung, A. C., Neitzel, A. J., & Chai, C. S. (2024). Does chatting with Chatbots improve language learning performance? A meta-analysis of Chatbot-assisted language learning. *Review of Educational Research*. https://doi.org/10.3102/00346543241255621

Wang, Y., & Xue, L. (2024). Using AI-driven chatbots to foster Chinese EFL students' academic engagement: An intervention study. *Computers in Human Behavior*, 108353. https://doi.org/10.1016/j.chb.2024.108353

Woo, J. H., & Choi, H. (2021). Systematic review for AI-based language learning tools. *arXiv: 2111.04455*. https:// doi.org/10.48550/arXiv.2111.04455

Wu, J., Li, Y., Zhou, J., & Chen, S. (2024). The impact of intelligent personal assistants on Mandarin second language learners: Interaction process, acquisition of listening and speaking ability. *Computer Assisted Language Learning*, 1-26. https://doi.org/10.1080/09588221.2024.2317849

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, *28*(11), 13943-13967.

Yang, L., & Li, R. (2024). ChatGPT for L2 learning: Current status and implications. *System*, *124*, 103351. https://doi.org/10.1016/j.system.2024.103351

Yeh, H. C. (2024). The synergy of generative AI and inquiry-based learning: Transforming the landscape of English teaching and learning. *Interactive Learning Environments*, 1-15. https://doi.org/10.1080/10494820.2024.2335491

Yıldız, T. A. (2023). The impact of ChatGPT on language learners' motivation. *Journal of Teacher Education and Lifelong Learning*, *5*(2), 582-597.

Yu, L., Kong, Q., & Hao, H. (2024). Artificial intelligence-assisted Chinese L2 writing: An empirical study on educational sustainability in Africa. *Rupkatha Journal*, *16*(2). https://doi.org/10.21659/rupkatha.v16n2.01

Yuan, Y. (2023). An empirical study of the efficacy of AI chatbots for English as a foreign language learning in primary education. *Interactive Learning Environments*, 1-16. https://doi.org/10.1080/10494820.2023.2282112

Zhai, C., & Wibowo, S. (2022). A systematic review on cross-culture, humor and empathy dimensions in conversational chatbots: The case of second language acquisition. *Heliyon*, *8*(12).

Zhang, R., Zou, D., & Cheng, G. (2023). Chatbot-based learning of logical fallacies in EFL writing: Perceived effectiveness in improving target knowledge and learner motivation. *Interactive Learning Environments*, 1-18. https://doi.org/10.1080/10494820.2023.2220374

Zhao, Y., Pan, J., Dong, Y., Dong, T., Wang, G., Ying, F., ... & Cao, J. (2024, May). Language urban odyssey: A serious game for enhancing second language acquisition through large language models. In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (pp. 1-7).

Zhou, G., & Niu, X. (2015). Approaches to language teaching and learning. *Journal of Language Teaching and Research*, *6*(4), 798–802. https://doi.org/10.17507/jltr.0604.11

Zhu, C., Sun, M., Luo, J., Li, T., & Wang, M. (2023). How to harness the potential of ChatGPT in education? *Knowl Manag E-Learn*, *15*(2), 133–52. https://doi.org/10.34105/j. kmel.2023.15.008. https://www.researchgate.net/profile/Minhong-Wang-2/ publication/370894024_and_Technology/links/646789029533894cac7e6771/an d-Technology.pdf.

**Author Information**

**Zhenjie Weng**
https://orcid.org/0000-0002-2639-7822
Duke Kunshan University
No. 8 Duke Avenue
Kunshan, Jiangsu
China
Contact e-mail: *zw272@duke.edu*

**Yao Fu**
https://orcid.org/0000-0002-4256-335X
University of Wisconsin-Whitewater
800 W Main St.
Whitewater, WI
U.S.A.