

ChatGPT for L2 learning: Current status and implications

Lu Yang, Rui Li *

School of Foreign Languages, Hunan University, PR China

ARTICLE INFO

Keywords:

ChatGPT

Evidence-based applied linguistics (EBAL)

Technology-based learning model

Second language (L2) learning

ABSTRACT

Despite a recent proliferation of studies on ChatGPT for second language (L2) learning, there is still a lack of systematic and updated review of its current status. To narrow the gap, this study collected data from 44 selected studies on ChatGPT for L2 learning in terms of six dimensions of the revised technology-based learning model, including ChatGPT, participants, objectives, theories, methodology, and outcomes. The results showed that (1) The most prevalent ChatGPT's roles included content generation, feedback and teaching support. Context control and output customization were the two main prompt patterns. (2) Most studies focused on investigating English-as-a-foreign-language (EFL) college learners with small sample sizes. (3) Learner perceptions (general attitudes, satisfaction, motivation, and engagement) along with writing skills were the major objectives. (4) Social (sociocultural and constructivism), linguistic (input hypothesis, informal digital learning), and cognitive (self-determination theory, autonomy) theories were frequently adopted. (5) Most studies used qualitative, quantitative and mixed methods, with a particular eye on questionnaire surveys, interviews, log data and written texts. (6) Benefits and challenges were summarized from the selected studies. Implications were discussed for future research.

1. Introduction

In the era of digital technology, emerging computer-assisted language learning (CALL) devices or technologies have offered a vast array of opportunities for second/foreign language (L2/FL) acquisition, among which traditional chatbots have received considerable attention, as they could serve as conversational agents/tutors to engage L2 learners in language communication (Li, 2024; Wu & Li, 2024a, 2024b). However, those traditional chatbots, or rule-based chatbots, have a limited capacity to understand L2 learners' inputs and could merely provide pre-defined guidance on certain L2 topics step-by-step, regardless of learners' personalized demands (Kohnke et al., 2023). The advent of OpenAI's ChatGPT in the year 2022 together with the increased use of ChatGPT has dramatically changed traditional L2 learning and teaching scenarios, because ChatGPT could outperform traditional chatbots with clear and specific prompts to harness its full L2 education potentials in a dialogue form (Kohnke et al., 2023), provide L2 learners with enhanced input and human-computer interaction from the interactionist theory perspective (Chapelle, 2005), and "process long input, generate innovative responses to varying prompts, and maintain a continuous chat flow" (Su et al., 2023, p.2). The affordances of ChatGPT for L2 learning have been well documented, including fostering positive learner perceptions (Agustini, 2023; Marzuki et al., 2023), facilitating writing and general language skills (Athanasopoulos et al., 2023; Vera, 2023), and supplementing teacher support in L2 learning (e.g., Barret et al., 2023; Jeon & Lee, 2023; Mohamed, 2024), among others.

* Corresponding author. Lushan South Rd., Yuelu District, Changsha, 410082, Hunan Province, PR China.
E-mail address: liruidianzi@hotmail.com (R. Li).

While a growing body of empirical studies has been steadily conducted, a thorough, systematic review of ChatGPT for L2 learning remains rare (Barrot, 2023), and the accumulation of sufficient empirical evidence also necessitates a panoramic vision of its state-of-the-art research trends. In other words, a comprehensive synthesis, based on the dimensions of revised technology-based learning model (Li, 2022a), could not only contribute to the holistic understanding of ChatGPT for L2 learning that informs pedagogy, but also provide some useful implications for future research. As such, this research seeks to synthesize studies of ChatGPT for L2 learning, with a particular eye on such dimensions as ChatGPT (ChatGPT's roles and prompt patterns), participants (educational levels, sample sizes and target languages), objectives (language skills and learning perceptions), (linguistic, social, cognitive, and others) theories, (general and specific) methods, and outcomes (advantages and challenges).

2. Literature review

2.1. Related studies

While ChatGPT's facilitative effects have been widely obtained, more detailed content with regard to its effectiveness for L2 learning is yet to be discovered.

On the one hand, proponents have claimed the advantages of integrating ChatGPT into L2 learning (e.g., Guo & Wang, 2023; Jeon & Lee, 2023; Vera, 2023). Specifically, Jeon and Lee (2023) contended that ChatGPT could promote English-as-a-foreign-language (EFL) learners' learning efficacy, because it could not only provide learning resources and feedback, but also serve as chat partner and evaluator. Li and colleagues (2023) adopted a mixed-method design to explore the use of ChatGPT on L2 learners' writing skill development, and posited that ChatGPT could afford learners with personalized feedback, content generation, and teaching support. Based on the sociocultural theory, Vera (2023) investigated ChatGPT's pedagogical benefits for EFL learners' language proficiency, and the results demonstrated that learners of the experiment group receiving ChatGPT-based instruction outperformed those of the control group receiving traditional classroom instruction, suggesting that ChatGPT could improve EFL learners' language proficiency more effectively with personalized, interactive, and engaging affordances.

On the other hand, critics have raised concerns about the potential challenges of ChatGPT for L2 learning (e.g., Lee et al., 2023; Mohamed, 2024; Yan, 2023). As highlighted in a recent qualitative study (Yan, 2023), EFL learners expressed concerns with ChatGPT's danger to educational equality and academic honesty in L2 writing practicum, because learners' L2 writing process could be simplified from "reading-writing-revision" to "text-generation and post-editing" when using ChatGPT as a shortcut for writing practice. Adhering to contextual language learning, Lee et al. (2023) integrated ChatGPT into augmented reality (AR) glasses to establish a contextual learning setting and explored its impacts on EFL learners' affective perceptions. By adopting a quantitative approach, they obtained that ChatGPT had limited effects on enhancing learning satisfaction and engagement in immersive learning activities and suggested more investigations into EFL learners' perceptions of ChatGPT be conducted in future research. Drawing on input hypothesis, Escalante et al. (2023) compared the effect of ChatGPT- and teacher-generated feedback on L2 learners' writing skill development and claimed that no significant difference existed between two types of feedback, i.e., ChatGPT-vs. teacher-generated feedback.

2.2. Related reviews

Despite those empirical studies, researchers have also conducted several review studies (e.g., Hwang & Chang, 2023; Kuhail et al., 2023; Smutny & Schreiberova, 2020) on traditional chatbots pertinent to this study. For instance, Huang et al. (2022) systematically synthesized 25 selected studies on chatbots' social, technological, and pedagogical affordances along with challenges for L2 learning. Results indicated that chatbot-supported language learning can offer such advantages as interaction, feedback, teaching support, and recommendation, while challenges of failure to launch long conversations, novelty effect, and increased cognitive load should be cautioned as well. Similarly, Ji et al. (2023) reviewed 24 articles on chatbot roles in L2 education, and found chatbots were commonly adopted for content generation, feedback, assessment and grading, and teaching support. Due to the rapid advancement of speech-recognition technology, Jeon, Lee, and Choi (2023) paid renewed attention to the 32 studies on speech-based chatbots for L2 education regarding publication trends, participants, methodology, research foci, chatbot roles, adopted devices, and learning contexts. They recommended that future researchers should investigate roles of chatbots beyond that of conversation partner and explore the effects of those functions on L2 outcomes. In a recent study, Jeon, Lee, and Choe (2023) presented an overview of the features and benefits of chatbots for L2 education by proposing a three-dimension framework, viz. multimodality, embodiment, and goal-orientation. Their findings suggested that future research should explore ChatGPT's characteristics and affordances through the lens of solid theoretical framework.

2.3. Research statement and questions

While the aforementioned quantitative investigations and qualitative reviews are insightful to ChatGPT for L2 learning research, several important gaps remain underexplored. First, despite the importance attached to ChatGPT and the large amount of research that has been carried out, there is no systematic synthesis of ChatGPT for L2 learning research, and the growing interest in ChatGPT calls for a review on the existing studies for the purpose of acting as a one-step pathfinder where the reader has a comprehensive access to knowledge about ChatGPT for L2 learning. Second, considering that sufficient quantitative investigations of ChatGPT for L2 learning have been done, the mixed findings suggest the necessity to revisit the benefits and challenges of utilizing ChatGPT for L2 learning. Third and importantly, existing reviews only explored the limited dimensions of chatbots that lacked a comprehensive theoretical

underpinning of investigation, e.g., participants, methodology, and research foci. Other important dimensions—technologies, objectives, theories, and outcomes—also warrant detailed scrutiny in ChatGPT for L2 learning (Jeon, Lee, & Choi, 2023; Yang et al., 2024; Zou et al., 2022). In this regard, the current research was carried out by taking revised technology-based learning model as the theoretical underpinning, because it has been widely applied in explaining the multi-dimensionality of CALL technologies for L2 learning, such as language massive open online courses (LMOOCs, Fang et al., 2022), blended language learning (Li, 2022a), and mobile-assisted language learning (MALL, Liu & Hwang, 2023), which would serve as an appropriate framework in this study with a comprehensive coverage of technologies-related (viz. ChatGPT), participants-related, and research-related (e.g., objectives, theories, methods, and outcomes) dimensions.

As apparent in Fig. 1, there were six dimensions as ChatGPT (ChatGPT’s roles and prompt patterns), participants (educational levels, sample sizes and target languages), objectives (language skills and learning perceptions), (linguistic, social, cognitive, and others) theories, (general and specific) methods, and outcomes (advantages and challenges). Specifically, ChatGPT’s roles refer to the use of ChatGPT for content generation, feedback, teaching support, assessment and recommendation, whereas prompt patterns refer to the prompt pattern catalogs that precisely outline the desired content of the materials ChatGPT users intend to generate, such as context control, output customization, zero-shot prompting, error identification, etc. (White et al., 2023). Objectives include L2 learner’s learning perceptions and language skills. Theories are the related social, linguistic, cognitive, and other theories. Based on Jeon, Lee, and Choi (2023), participants cover educational levels, sample sizes and target languages. Methods involve general and/or specific methods employed in the reviewed studies. Last, outcomes include ChatGPT’s benefits and disadvantages.

As a consequence, this study attempts to systematically synthesize studies of ChatGPT for L2 learning in terms of ChatGPT, participants, objectives, theories, methods, and outcomes. The scope of this review is based on the research questions that follow.

1. What are ChatGPT’s roles and its prompt patterns reported in the studies of ChatGPT for L2 learning?
2. What are educational levels, sample sizes and target languages reported in the studies of ChatGPT for L2 learning?
3. What are language skills and learner perceptions reported in the studies of ChatGPT for L2 learning?
4. What are related theories adopted in the studies of ChatGPT for L2 learning?
5. What are research methods reported in the studies of ChatGPT for L2 learning?
6. What are advantages and challenges in the studies of ChatGPT for L2 learning?

3. Methodology

3.1. Literature retrieval

The following literature retrieval procedures were diagrammed in Fig. 2 to exhaustively retrieve literature on ChatGPT for L2 learning. First, based on existing qualitative reviews (e.g., Ansari et al., 2023; Jeon et al., 2023; Lo, 2023), online databases (Fig. 2) were searched by using those Boolean expressions and truncation (*) of key terms pertinent to L2 learning (second language OR foreign language OR L2 OR EFL OR ESL OR language learning OR language teaching OR language acquisition OR speak* OR communicat* OR listen* OR read* OR vocabulary OR writ* OR grammar) AND ChatGPT (ChatGPT OR AI Chatbots OR artificial intelligence agents). Second, through checking references in the identified articles (e.g., Ansari et al., 2023; Cai et al., 2023; Lo, 2023; Sallam et al., 2023), a backward and forward search of the selected studies was launched by adopting snowballing technique (Biernacki & Waldorf, 1981). Third, manual retrieval of language education and educational technology journals (Fig. 2) was done to avoid the incomplete retrieval (Zou et al., 2021, 2022; Li, 2023b, 2023c).

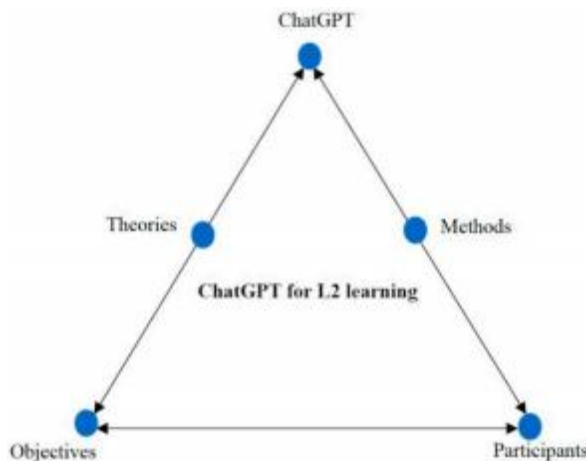


Fig. 1. Proposed model of ChatGPT for L2 learning.

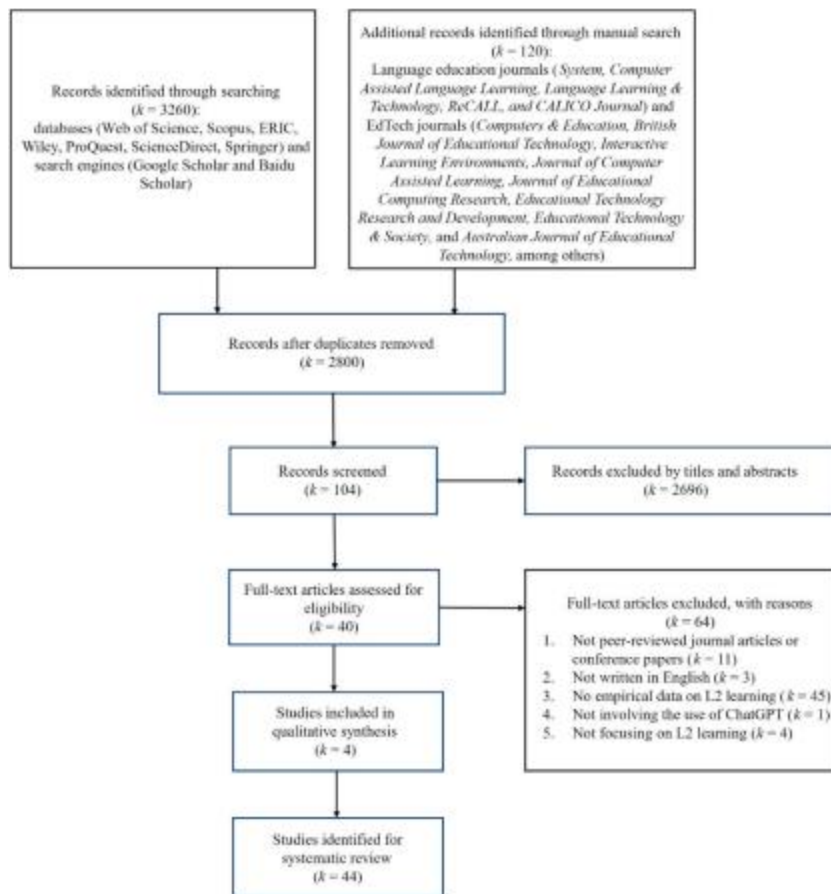


Fig. 2. Data collection flow chart.

3.2. Data trimming criteria

As shown in Fig. 2, the selected studies were further trimmed in relation to the following proposed criteria.

1. Articles written in English should be confined to 2022-2023, because ChatGPT was initially developed in 2022 (Lo, 2023). Studies that are written in any other languages or time range were excluded.
2. Articles selected should be peer-reviewed journal articles or conference proceedings for two considerations: On the one hand, academic rigor of peer-reviewed journal articles and conference proceedings could be warranted, because they reported original findings of ChatGPT for L2 learning based on strict research design (e.g., participants, methods, objectives, and outcomes). On the other hand, the inclusion of both peer-reviewed journal articles and conference proceedings could reduce the risk of publication bias and capture the most up-to-date information of ChatGPT for L2 learning (Huanget al., 2021). Studies in other forms (e.g., pre-prints, dissertations, or non-academic articles) were not considered.
3. Empirical data should be reported in the selected studies. Studies that only propose the development or technology review of ChatGPT without the involvement of L2 learners were excluded.
4. Only studies of ChatGPT for L2 learning should be considered. Those that deal with the use of self-developed chatbots, automated writing evaluation, or other AI tools without any involvement of ChatGPT for second or foreign language learning were not considered.

3.3. Coding scheme

The detailed illustration of coding scheme was presented in Appendix 1, including dimensions, coding types, subtypes, definition of subtypes, and sources.

3.4. Coding procedures

Three coding procedures were performed: First, two coders discussed mutually to improve the consistent understanding of each item. Second, they separately coded all the 44 selected studies related to ChatGPT's roles, ChatGPT's prompt patterns, participants, language skills, learning perceptions, theoretical frameworks, methods, and outcomes with SPSS 24.0. The overall intercoder agreement of $K = 0.950$ ($p < 0.001$) had been obtained at this stage. Third, the discrepancies were addressed through negotiation and analysis of the coding scheme.

4. Results

4.1. ChatGPT

Results of ChatGPT tasks and prompt patterns were obtained as follows: First, the most common ChatGPT tasks reported in the reviewed studies involved content generation ($k = 24$) and feedback ($k = 23$), followed by teaching support ($k = 13$), assessment and grading ($k = 4$), and recommendation ($k = 4$). Second, the most frequent ChatGPT's prompt patterns utilized in the studies were context control ($k = 11$) and output customization ($k = 8$), followed by zero-shot prompting ($k = 2$), error identification ($k = 1$), and interaction ($k = 1$).

4.2. Participants

Results of participants' information were summarized in Table 1. First, regarding target languages, most adopted ChatGPT for English ($k = 40$) learning, followed by Chinese ($k = 1$) and German ($k = 1$) learning. Second, concerning educational levels, ChatGPT was most frequently applied in higher education ($k = 21$), followed by teacher training ($k = 9$), mixed ($k = 6$), and secondary school ($k = 5$). Third, as for sample sizes, the reviewed studies generally utilized small samples ($k = 17$), followed by large ($k = 10$), moderate to large ($k = 7$), and moderate ($k = 6$) sample sizes.

4.3. Objectives

Fig. 3 presented the distribution of research objectives—language skills and learner perceptions—in the selected studies.

Regarding language skills, the most frequently explored language skills were writing ($k = 6$) and general language skills ($k = 4$). When it comes to learner perceptions, most studies focused on general attitudes ($k = 30$), satisfaction ($k = 4$), motivation ($k = 3$), engagement ($k = 3$), perceived ease of use ($k = 3$), perceived usefulness ($k = 3$), perceived learning ($k = 3$), and behavioral intention ($k = 3$), while autonomy ($k = 1$), confidence ($k = 1$), self-regulation ($k = 1$), performance expectancy ($k = 1$), and perceived system usability ($k = 1$) received scant attention.

4.4. Theories

As displayed in Table 2, researchers tended to investigate the use of ChatGPT for L2 learning from the perspective of social ($k = 12$), linguistic ($k = 6$), cognitive ($k = 5$), and other theories ($k = 4$).

For social theories, most studies were conducted based on social constructivism ($k = 3$) and sociocultural theory ($k = 3$), while

Table 1
Participants of the selected studies.

Participants	<i>k</i>
<i>Target languages</i>	
√English	40
√Chinese	1
√German	1
√Not specified	2
<i>Educational levels</i>	
√Higher education	21
√Teacher training	9
√Mixed	6
√Secondary school	5
√Primary school	0
√Pre-school	0
√Not specified	3
<i>Sample sizes</i>	
√Small	17
√Large	10
√Moderate to large	7
√Moderate	6
√Not specified	4

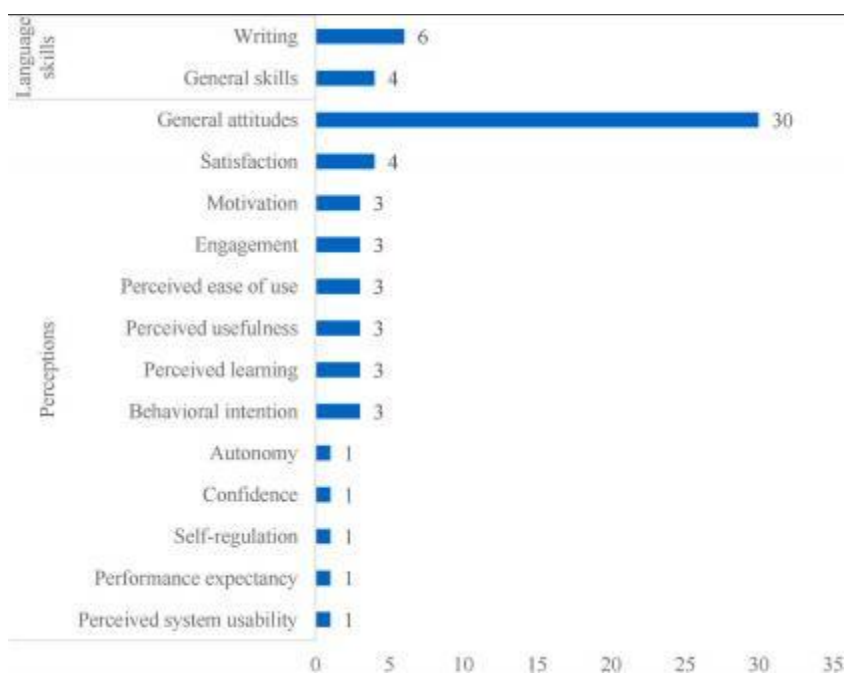


Fig. 3. Distribution of objectives in the selected studies.

Table 2

Theories of the selected studies.

Related theories	Related sources	<i>k</i>
Social theories		12
√Sociocultural theory	Escalante et al., 2023; Rakhmonov & Kurbonova, 2023; Vera, 2023	3
√Social constructivism	Bin-Hady et al., 2023; Rakhmonov & Kurbonova, 2023; Vera, 2023	3
√Activity theory	Cai et al. (2023)	1
√Contextual learning	Lee et al. (2023)	1
√Experiential learning	Yan (2023)	1
√Social cognitive theory	Cai et al. (2023)	1
√Peer comments	Guo and Wang (2023)	1
√Peer scaffolding	Yan (2023)	1
Linguistic theories		6
√Informal digital learning of English	Liu and Ma (2024)	1
√Input hypothesis	Escalante et al. (2023)	1
√Error analysis	Al-Garaady and Mahyoob (2023)	1
√Theory of universal grammar	Imran and Lashari (2023)	1
√Process writing	Barret et al. (2023)	1
√Creative writing	Woo et al. (2023)	1
Cognitive theories		5
√Motivation/self-determination theory	Agustini, 2023; Ali et al., 2023; Cai et al., 2023	3
√Autonomy	Agustini (2023)	1
√Theory of planned behavior	Cai et al. (2023)	1
Other theories		4
√Technology acceptance model	Cai et al., 2023; Liu & Ma, 2024; Zou & Huang, 2023	3
√Reflective learning	Yan (2023)	1
Not specified		28

others, e.g., activity theory ($k = 1$), contextual learning ($k = 1$), experiential learning ($k = 1$), social cognitive theory ($k = 1$), peer comments ($k = 1$), and peer scaffolding ($k = 1$), were less frequently considered.

Besides those social theories, linguistic theories were the second most frequently used ($k = 6$), including informal digital learning of English ($k = 1$), input hypothesis ($k = 1$), error analysis ($k = 1$), theory of universal grammar ($k = 1$), process writing ($k = 1$), and creative writing ($k = 1$).

Researchers also used cognitive theories ($k = 5$) and other theories ($k = 4$). Among those cognitive theories, motivation/self-determination theory ($k = 3$) received considerable attention, followed by autonomy ($k = 1$) and theory of planned behavior ($k = 1$). Other frameworks included technology acceptance model ($k = 3$) and reflective learning ($k = 1$).

4.5. Methods

As displayed in Table 3, all the three general methods were frequently adopted in the selected studies, including quantitative ($k = 16$), qualitative ($k = 15$), and mixed methods ($k = 13$), respectively.

The popular quantitative methods contained questionnaire ($k = 12$) and tests ($k = 1$). For qualitative methods, qualitative methods including interviews ($k = 10$), class observation ($k = 2$), log data ($k = 1$), audio/video recordings ($k = 1$), and reflection ($k = 1$) were used to evaluate learner perceptions and performance. Regarding mixed methods, combinations of both quantitative and qualitative methods, such as close-ended questionnaires + interviews ($k = 5$), close- and open-ended questionnaires + tests ($k = 1$), tests + reflections ($k = 1$), questionnaires + interviews + log data ($k = 1$), questionnaires + tests + log data ($k = 1$), and close- and open-ended questionnaires ($k = 1$), were frequently used.

4.6. Outcomes

Table 4 summarized the advantages and challenges of adopting ChatGPT for L2 learning.

The most frequently reported benefits included improving perceptions and attitudes towards L2 learning ($k = 19$), improving four language skills ($k = 8$), facilitating autonomous learning and increasing L2 practice ($k = 3$), providing an interactive, enjoyable, and engaging learning environment ($k = 2$), reducing teachers' workload and facilitating flexible teacher roles ($k = 2$), among others.

Those challenges should be cautioned, such as plagiarism or privacy issues ($k = 12$), inaccurate or inappropriate responses ($k = 10$), learners' over-reliance on ChatGPT ($k = 8$), technical issues and learners' unfamiliarity with ChatGPT ($k = 7$), failure to improve language skills ($k = 6$), cultural insensitivity, language biases, and limited personalization ($k = 4$), lack of emotional or empathetic responses ($k = 3$), failure to generate texts with syntactic simplicity and deep cohesion ($k = 3$), failure to integrate ChatGPT into curriculum ($k = 2$), and failure to facilitate advanced L2 skills or offer higher-order feedback ($k = 2$), etc.

5. Discussion

The first research question dealt with ChatGPT's roles and prompt patterns. Regarding ChatGPT's roles, results demonstrated that the primary function of ChatGPT for L2 learning was content generation, feedback, and teaching support. This finding was in apparent contradiction with Ji et al. (2023) who postulated that traditional chatbots were most commonly adopted for conversational practice in language learning. A plausible explanation might rest on the ChatGPT's artificial intelligence generated content (AIGC) feature as compared to traditional chatbots—utilizing AI tools to generate the required content automatically, including standard writing samples (Li et al., 2023a), exercises for L2 use (Lashari et al., 2023), and L2 quizzes (Han et al., 2023). For instance, Yan (2023) explored the influence of ChatGPT's automatic text generation on learner perceptions of EFL writing, and suggested that ChatGPT could improve their writing performance in grammatical accuracy and lexical diversity by generating well-structured writing samples for reference. Besides content generation, ChatGPT could provide L2 learners with personalized feedback, including error detection (Ahmed, 2023), error correction (Escalante et al., 2023), and proofreading suggestions (Faiz et al., 2023), among others. As Guo and Wang (2023) claimed, different from the traditional chatbots that provide one-size-fits-for-all feedback, ChatGPT could not only reiterate feedback in easier terms for learners to understand, but also further explain the rationale for the feedback. Besides, ChatGPT

Table 3
Distribution of research methods in the selected studies.

Methods	k
<i>General methods</i>	
√Quantitative	16
√Qualitative	15
√Mixed	13
<i>Quantitative methods</i>	
√Questionnaires	12
√Tests	1
√Not specified	3
<i>Qualitative methods</i>	
√Interviews	10
√Class observation	2
√Log data	1
√Audio/video recordings	1
√Reflection	1
<i>Mixed methods</i>	
√Close-ended questionnaires + interviews	5
√Close- and open-ended questionnaires + tests	1
√Tests + reflections	1
√Questionnaires + interviews + log data	1
√Questionnaires + tests + log data	1
√Close- and open-ended questionnaires	1
√Not specified	3

Table 4
Outcomes reported in the selected studies.

Outcomes	<i>k</i>
<i>Advantages</i>	
√Improving perceptions and attitudes towards L2 learning	19
√Improving four language skills	8
√Facilitating autonomous learning, and increasing L2 practice	3
√Providing an interactive, enjoyable, and engaging learning environment	2
√Reducing teachers' workload and facilitating flexible teacher roles	2
√Reducing educational inequity and inequality	1
√Outperforming human in the aspects of expressive concreteness, narrativity and syntactic simplicity	1
√Providing more feedback amount than L2 teachers	1
√Generating L2 reading passages with high levels of naturalness	1
√Not specified	15
<i>Challenges</i>	
√Plagiarism or privacy issues	12
√Inaccurate or inappropriate responses	10
√Learners' over-reliance on ChatGPT	8
√Technical issues and learners' unfamiliarity with ChatGPT	7
√Failure to improve language skills	6
√Cultural insensitivity, language biases, and limited personalization	4
√Lack of emotional or empathetic responses	3
√Failure to generate texts with syntactic simplicity and deep cohesion	3
√Failure to integrate ChatGPT into curriculum	2
√Failure to facilitate advanced L2 skills or offer higher-order feedback	2
√Lack of pronunciation and intonation feedback	1
√Different evaluation criteria from L2 teachers	1
√Failure to generate testing items with attractive multiple-choice options and high completion rates	1
√Failure to enhance learner satisfaction and engagement	1
√Failure to internalize ChatGPT output	1
√Distraction to learning tasks	1
√Not specified	16

has been documented to offer L2 learners with necessary teaching support, including conversation agent (Bin-Hady et al., 2023; Jeon & Lee, 2023) and intelligent question-answering (Samarasinghe & Prasangani, 2023). This was because ChatGPT provided learners with unlimited opportunities for L2 practice without temporal and spatial constraints, which could reduce teachers' workload and shift their focus on other important tasks, such as presenting new content, scaffolding activities, and leading discussions (Jeon & Lee, 2023; Ji et al., 2023).

The context control was one of the most frequently adopted prompt patterns, which could be partly due to ChatGPT's failure of understanding EFL learners' intended prompts for specific contexts and the output of irrelevant responses (Kohoke et al., 2023). In other words, ChatGPT could better process instructions and elicit more accurate output with the provision of explicit and specific contextual information. For instance, Escalante et al. (2023) required ChatGPT to focus on the main idea, transitional phrases, and grammatical accuracy in L2 writing, and found that the clarity and specificity of ChatGPT-generated feedback could be improved. Similarly, Barret et al. (2023) required ChatGPT to generate essays on certain topics, such as global warming and climate change, and asserted that through the context control, ChatGPT could serve as a useful tool to brainstorm ideas for the L2 writing process. The second commonly implemented prompt pattern was output customization—allowing learners to tailor the complexity, format, or other properties of ChatGPT output, which could be explained by two possible reasons: On the one hand, learners with different L2 proficiency levels might have distinct preferences for the diversity of ChatGPT's outputs (Jeon & Lee, 2023), and the use of output customization would enable learners to modify the “one-size-fits-all” output to accommodate their personalized needs. On the other hand, the output customization assigning ChatGPT a persona (e.g., waiter and game partner, Jeon & Lee, 2023; hotel receptionist, Javier & Moorhouse, 2023) could take it as a waiter and a game partner to associate learning tasks with the authentic scenarios, which could facilitate learners' L2 use and improve their willingness to communicate (Yang & Li, 2024). Besides the high frequency of context control and output customization, while zero-shot prompting, error identification, and interaction were less frequently adopted, the feasibility of those prompt patterns for L2 learning remains open for debate, warranting further investigation in this regard.

Concerning the second research question, most studies focused on using ChatGPT for EFL learning, because English has been included as a compulsory subject in K-12 education and also been considered as an important skill in undergraduate and graduate programs among various non-English speaking countries (Shadiev & Yu, 2022). Results also indicated that ChatGPT was most frequently applied in higher education and teacher training, resonating Lo (2023) who maintained that most ChatGPT studies were conducted among college students. It comes as no surprise to observe that those researchers of ChatGPT for L2 learners who are college teachers would recruit college students as their participants under investigation. Another possible explanation lies in that college students are also easier to gain ethics than those in pre-school, primary school or secondary education levels. The reason why primary and pre-school students were scarcely investigated is that ChatGPT usually generates texts with high syntactic complexity and lexical richness, which might pose challenges for primary students who are still in the early stage of L2 learning to understand the output (Escalante et al., 2023). Most selected studies recruited participants of small sample sizes, probably because ChatGPT was recently released since the year 2022, and large-scale empirical investigations with larger samples might not be currently available.

Regarding the third research question, most ChatGPT studies were conducted under social, linguistic, cognitive, and other theories, since they could provide insights into the interplay between ChatGPT and L2 outcomes. In other words, researchers utilizing ChatGPT attempted to enhance learners' social interactive (e.g., sociocultural theory, Rakhmonov & Kurbonova, 2023; experiential learning, Yan, 2023; social constructivism, Vera, 2023), linguistic (e.g., Escalante et al., 2023; Imran & Lashari, 2023), and cognitive outcomes (e.g., motivation, Ali et al., 2023; autonomy, Agustini, 2023; behavioral intention, Cai et al., 2023), among others. Additionally, twenty-eight studies failed to ground on theoretical underpinnings due to researchers' interest in exploring other aspects of ChatGPT, such as evaluation on ChatGPT output (e.g., Shin & Lee, 2023; Zhou et al., 2023), features of ChatGPT feedback (Guo & Wang, 2023), and learners' ChatGPT use (Han et al., 2023; Javier & Moorhouse, 2023).

The fourth research question investigated objectives of the selected studies. Concerning language skills, ChatGPT was commonly adopted for the development of writing skills, aligning with Su et al. (2023), as it can assist learners in addressing dialogical, structural, and linguistic challenges in writing. Apart from writing skills, learners' general skills were also frequently examined. Vera (2023) reported the use of ChatGPT for the development of learners' L2 performance and perceptions, and posited that due to ChatGPT's personalized, interactive, and engaging features, ChatGPT-based instruction could effectively enhance learners' language skills. Moreover, it indicated that learners' attitudes were most frequently investigated, such as general attitudes (Barret et al., 2023), satisfaction (Ahmed, 2023), motivation (Cai et al., 2023), and engagement (Lee et al., 2023). For instance, Agustini (2023) postulated the argument that ChatGPT enhanced EFL learners' motivation, confidence and autonomy by encouraging self-assessment, affording a non-judgmental learning environment and empowering personalized instruction.

Regarding the fifth research question, for qualitative methods, researchers would adopt open-ended questionnaires (Zulfa et al., 2023), interviews (Agustini, 2023; Ahmed, 2023), or log data (Javier & Moorhouse, 2023; Jeon & Lee, 2023) to explore students' general attitudes towards ChatGPT (Imran & Lashari, 2023), satisfaction (Ahmed, 2023), motivation, autonomy, and confidence (Agustini, 2023). For quantitative methods, researchers would adopt questionnaire surveys (Ali et al., 2023; Barret et al., 2023; Lee et al., 2023) or tests (Vera, 2023) to understand ChatGPT's role in general language skills development (Vera, 2023), attitudes (Barret et al., 2023), perceived learning, satisfaction, and engagement (Lee et al., 2023), and technology acceptance (Liu & Ma, 2024). Researchers also adopted a number of mixed methods, such as close-ended questionnaires + interviews (Nguyen, 2023), tests + questionnaires (Escalante et al., 2023), and questionnaires + interviews + log data (Han et al., 2023), because they used multiple (quantitative + qualitative) sources of data to further triangulate their research findings.

Lastly, the advantages together with challenges of ChatGPT use were obtained. Concerning advantages, enhancing language skills and learner perceptions received considerable attention. It was well-documented that ChatGPT could not only bolster learners' general perceptions (e.g., Liu, 2023; Xiao & Zhi, 2023; Yan, 2023), motivation (Agustini, 2023; Rakhmonov & Kurbonova, 2023), and engagement (Mohamed, 2024; Shahid et al., 2023); but also enhance their general language skills (e.g., Agustini, 2023; Lashari et al., 2023; Vera, 2023) and writing skills (e.g., Athanassopoulos et al., 2023; Marzuki et al., 2023), facilitate autonomous learning and L2 practice (Agustini, 2023; Bin-Hady et al., 2023; Mohamed, 2024), among others. On the other hand, challenges were also reported, such as plagiarism or privacy issues (Barret et al., 2023; Imran & Lashari, 2023), inaccurate or inappropriate responses (e.g., Mohamed, 2024; Nguyen, 2023; Riyadini & Triastuti, 2023), and learners' over-reliance on ChatGPT (e.g., Cai et al., 2023; Harunasari, 2022; Imran & Lashari, 2023). For instance, Yan (2023) reported that learners might copy ChatGPT-generated texts to complete writing assignments, as ChatGPT could avoid plagiarism detection by producing seemingly original content. It was also mentioned that ChatGPT might provide inaccurate and biased contextual information (Javier & Moorhouse, 2023; Mohamed, 2024), because it had limited understandings of cultural context and was pre-trained on data corpuses that contain biased or inaccurate information (Lo, 2023). Marzuki et al. (2023) also criticized that learners might lean too heavily on ChatGPT for proofreading, since its immediate feedback may undermine learners' motivation to check and thoroughly understand their mistakes.

6. Implications

6.1. Implications for teachers

First, since ChatGPT has potentially innovated L2 learning, teachers should integrate ChatGPT into their teaching practice and utilize it to enhance L2 teaching processes. By using ChatGPT-powered tools and strategies, teachers can personalize L2 learning, improve L2 learning outcomes, and better prepare learners for success in the digital age. Second, since prompt patterns could help learners customize outputs and interactions with ChatGPT to meet their personalized needs (White et al., 2023), teachers should train learners how to adjust prompts through the use of context control and output customization (Jeon & Lee, 2023). Teachers could develop a customized plan for implementing ChatGPT that aligns with their L2 educational goals and train their learners regarding how to use a variety of ChatGPT prompts with the aid of the existing prompt library. Third, to tackle with student plagiarism via ChatGPT (Li et al., 2023b; Liu, 2023), teachers should emphasize the importance of academic integrity, establish guidelines for appropriate ChatGPT use, and assign higher-order tasks that highlight the application, analysis, and creation of knowledge (Ansari et al., 2023; Lo, 2023). Fourth, despite ChatGPT's pedagogical affordances for L2 education, teachers should not only mitigate learners' potential risks of becoming over-reliant on ChatGPT, but also minimize the negative impacts on their problem-solving skills and critical awareness. One solution is to raise learners' critical awareness of ChatGPT's limitations (e.g., inaccurate or inappropriate output) and encourage them to use authoritative sources, evaluate and validate the factual correctness of ChatGPT-generated content (Barrot, 2023).

6.2. Implications for designers

First, since ChatGPT had limited background knowledge about learners' personalized needs (Guo & Wang, 2023), designers could consider incorporating ChatGPT into student information systems, analyze student data, and deliver personalized materials tailored to their individual differences, such as L2 proficiency, cognitive styles, and learning preferences (Li, 2022b, 2022c, 2023a; Yang & Li, 2024). Second, given that ChatGPT might generate inaccurate or biased output (Cai et al., 2023; Javier & Moorhouse, 2023), designers should guarantee the accuracy of ChatGPT responses and develop it with cultural sensitivity to avoid language biases and stereotypes (Mohamed, 2024; Zou & Huang, 2023). Third, designers should offer user-friendly interface with clear navigation, simplify ChatGPT operation, and integrate gamified elements (Zou et al., 2021) to facilitate human-computer interaction or collaboration with ChatGPT.

6.3. Implications for researchers

First, since learning theories could inform what variables tend to affect ChatGPT's effectiveness and explicate how ChatGPT could be integrated into L2 curriculum (Li et al., 2021; Zou et al., 2022), researchers should explicitly state the adopted social (sociocultural, constructivism, and experiential learning), linguistic (input hypothesis, informal digital learning), and cognitive (self-determination theory, autonomy) theories in their future studies. Second, while a predominant focus has been paid on students of higher education, little is known about the effectiveness of ChatGPT for L2 learning among K-12 students. As such, researchers should not only investigate how ChatGPT could bolster K-12 students' perceptions (e.g., motivation, satisfaction and engagement) and facilitate their language skills (Meng & Li, 2023), but also develop their AI competencies regarding the ethical use, potential benefits and risks of utilizing ChatGPT for L2 learning. Third, given that most studies employed questionnaire survey and interviews to explore learners' attitudes towards ChatGPT, future study may adopt pre-/posttests experimental designs or other mixed methods to triangulate findings of ChatGPT for L2 learning. Fourth, as L2 learners are likely to be over-reliant on ChatGPT's desired outputs generation that would pose challenges to their higher-order thinking skills (Kohnke et al., 2023), future research could investigate ChatGPT's long-term influence in L2 learners' critical awareness, creativity, and problem-solving skills.

7. Conclusion

This study indicated that ChatGPT was frequently adopted for content generation, feedback, and teaching support in L2 learning. The majority of learners utilized such prompt patterns as context control and output customization to elicit effective output. Most studies employed ChatGPT for EFL learners in higher education institutions and adopted small sample sizes. The majority of studies focused on learners' general attitudes/perceptions and writing skills. Sociocultural theory, social constructivism, and technology acceptance model were more frequently adopted than other theories. Most studies used qualitative, quantitative and mixed methods, with a particular eye on questionnaire surveys, interviews, log data and written texts. Lastly, detailed illustrations of advantages and challenges were reported as well. The current study could not only contribute to presenting a panoramic vision about ChatGPT for L2 learning, but also inform future research and pedagogy in relation to ChatGPT roles, prompt patterns, learning theories, methodology, advantages and challenges.

Despite those promising findings, there were several limitations as well. First, this study only considered journal articles and conference proceedings for analysis. Future research may include other types of literature, such as PhD dissertations and monograph chapters. Second, since this study only involved such dimensions as ChatGPT's roles and prompt patterns, participants, objectives, theories, methodology, and outcomes, future endeavor could consider other important coding categories, such as role of learners' individual differences and application effectiveness.

Funding

This research was supported by the Scientific Research Funds from Hunan Provincial Ministry of Education (Grant No. 23A0040), China.

8. Data availability

The data that support the results of this study will be offered.

Declaration of competing interest

There is no conflict of interest.

Conflict of interest

There is no conflict of interest.

CRedit authorship contribution statement

Lu Yang: Writing – review & editing, Writing – original draft, Investigation, Conceptualization. **Rui Li:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Acknowledgements

The author would like to thank the anonymous reviewers for their constructive and insightful comments and suggestions.

Appendix 1

Coding scheme of the selected studies.

Dimensions	Coding types	Subtypes	Definition of subtypes	Sources
ChatGPT	ChatGPT's roles	1. Content generation	Generating learning materials, e.g., writing samples, exercises, and quizzes	Yan et al. (2024)
		2. Feedback	Providing feedback on learners' performance	
		3. Teaching support	Serving as a conversation agent and intelligent question-answering	
		4. Assessment and grading	Assessing learners' language performance	
		5. Recommendation	Recommending learners with personalized learning materials	
		6. Not specified	No report of specific ChatGPT's roles involved	
ChatGPT	Prompt patterns	1. Context control	Controlling the contextual information in which ChatGPT operates, e.g., requiring ChatGPT to focus on grammatical accuracy when providing feedback	White et al. (2023)
		2. Output customization	Tailoring the complexity, styles or other properties of the output generated by ChatGPT, e.g., requiring ChatGPT to generate texts with simple words	
		3. Zero-shot prompting	Prompting ChatGPT with no additional training, e.g., requiring ChatGPT to help translate or summarize the given texts	
		4. Error identification	Identifying and resolving the errors of the output generated by ChatGPT, e.g., requiring ChatGPT to help identify errors of the output	
		5. Interaction	Adjusting the interaction between learners and ChatGPT, e.g., requiring ChatGPT to generate a quick quiz or automatically ask questions	
		6. Not specified	No report of specific prompt patterns involved	
Participants	Target languages	1. English	Studies that used ChatGPT for English learning	Jeon, Lee, and Choi (2023)
		2. Chinese	Studies that used ChatGPT for Chinese learning	
		3. German	Studies that used ChatGPT for German learning	
	Educational levels	1. Pre-school	Preschool or kindergarten students	Xie et al. (2019)
		2. Primary	Elementary school students	
		3. Secondary	Junior middle school or senior high school students	
		4. Higher education	College or graduate students	
		5. Teacher training	Pre- or in-service teachers	
		6. Mixed	Mixed educational levels	
	Sample sizes	1. Small	Fewer than 30 participants	Hwang and Fu (2019)
		2. Moderate	30–50 participants	
		3. Moderate to large	51–100 participants	
4. Large		Over 100 participants		
Objectives	Language skills	1. Writing	Studies that used ChatGPT for L2 writing development	Zou et al. (2022)
		2. General language skills	Studies that used ChatGPT for domain-general L2 skill development, e.g., listening, speaking, reading, writing, vocabulary, and grammar, etc.	
	Learning perceptions	1. General attitudes	Participants' overall attitudes towards using ChatGPT for L2 learning	Xie et al. (2019)
		2. Satisfaction	Participants' satisfaction towards using ChatGPT for L2 learning	
		3. Motivation	Participants' motivation towards using ChatGPT for L2 learning	
		4. Engagement	Participants' social, cognitive and behavioral engagement in ChatGPT for L2 learning	
		5. Perceived ease of use	Participants' perceived ease of use in ChatGPT for L2 learning	
		6. Perceived usefulness	Participants' perceived usefulness in ChatGPT for L2 learning	
		7. Perceived learning	Participants' perceived learning gains in ChatGPT for L2 learning	
		8. Behavioral intention	Participants' intention to use ChatGPT for L2 learning	
		9. Autonomy	Participants' autonomy in ChatGPT for L2 learning	
		10. Confidence	Participants' confidence in ChatGPT for L2 learning	
		11. Self-regulation	Participants' ability to self-regulate their L2 learning with ChatGPT	
12. Performance expectancy	Participants' expectancy of ChatGPT as a valuable tool for L2 learning			
13. Perceived system usability	Participants' evaluation of the usability of using ChatGPT for L2 learning			
Theories	Learning theories	1. Social	Studies that reported social learning theories, e.g., sociocultural theory, social constructivism, social cognitive theory, etc.	Shadiev and Yu (2022)

(continued on next page)

(continued)

Dimensions	Coding types	Subtypes	Definition of subtypes	Sources
Methods	General methods	2. Linguistic	Studies that reported language learning theories, e.g., informal digital learning of English, input hypothesis, error analysis, etc.	Zou et al. (2022)
		3. Cognitive	Studies that reported cognitive theories, e.g., motivation, self-determination theory, autonomy, etc.	
		4. Others	Studies that reported other theories, e.g., technology acceptance model	
		5. Not specified	No report of any theories involved	
		1. Quantitative	Studies that utilized quantitative methods, including questionnaire surveys and test	
Outcomes	Application outcomes	2. Qualitative	Studies that utilized qualitative methods, e.g., interviews, class observation, log data, etc.	Shadiev and Yu (2022)
		3. Mixed	Studies that utilized both quantitative and qualitative methods, e.g., Close-ended questionnaires + interviews	
		1. Advantages	Studies that reported pedagogical benefits of applying ChatGPT for L2 learning	
		2. Challenges	Studies that reported drawbacks of applying ChatGPT for L2 learning	

References

- Agustini, N. P. O. (2023). Examining the role of ChatGPT as a learning tool in promoting students' English language learning autonomy relevant to Kurikulum Merdeka Belajar. *Edukasia: Jurnal Pendidikan Dan Pembelajaran*, 4(2), 921-934. <http://jurnaledukasia.org/index.php/edukasia/article/view/373>.
- Ahmed, M. A. (2023). ChatGPT and the EFL classroom: Supplement or substitute in Saudi Arabia's eastern region. *Information Sciences Letters*, 12(7), 2727-2734. <https://doi.org/10.18576/isl/120704>
- Al-Garaady, J., & Mahyoob, M. (2023). ChatGPT's capabilities in spotting and analyzing writing errors experienced by EFL learners. *Arab World English Journals*, 9, 3-17. <https://doi.org/10.24093/awej/call9.1>
- Ali, J. K. M., Shamsan, M. A. A., Hezam, T. A., & Mohammed, A. A. (2023). Impact of ChatGPT on learning motivation: Teachers and students' voices. *Journal of English Studies in Arabia Felix*, 2(1), 41-49. <https://doi.org/10.56540/jesaf.v2i1.51>
- Ansari, A. N., Ahmad, S., & Bhutta, S. M. (2023). Mapping the global evidence around the use of ChatGPT in higher education: A systematicscoping review. *Education and Information Technologies*, 1-41. <https://doi.org/10.1007/s10639-023-12223-4>
- Athanassopoulos, S., Manoli, P., Gouvi, M., Lavidas, K., & Komis, V. (2023). The use of ChatGPT as a learning tool to improve foreign language writing in a multilingual and multicultural classroom. *Advances in Mobile Learning Educational Research*, 3(2), 818-824. <https://doi.org/10.25082/AMLER.2023.02.009>
- Barrett, A., & Pack, A. (2023). Not quite eye to AI: Student and teacher perspectives on the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education*, 20, 1-24. <https://doi.org/10.1186/s41239-023-00427-0>
- Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, 1-6. <https://doi.org/10.1016/j.asw.2023.100745>
- Biernacki, P., & Waldorf, D. (1981). Snowball sampling: Problems and techniques of chain referral sampling. *Sociological Methods & Research*, 10(2), 141-163. <https://doi.org/10.1177/004912418101000205>
- Bin-Hady, W. R. A., Al-Kadi, A., Hazaea, A., & Ali, J. K. M. (2023). Exploring the dimensions of ChatGPT in English language learning: A global perspective. *Library Hi Tech*, forthcoming. <https://doi.org/10.1108/LHT-05-2023-0200>
- 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*. <https://doi.org/10.1080/10447318.2023.2261725>. forthcoming.
- Chapelle, C. A. (2005). Interactionist SLA theory in CALL research. In J. Egbert, & G. Petrie (Eds.), *CALL research perspectives* (pp. 53-64). Mahwah, NJ: Laurence Erlbaum Associates.
- Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: Insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education*, 20, 1-20. <https://doi.org/10.1186/s41239-023-00425-2>
- Faiz, R., Bilal, H. A. B., Asghar, I., & Safdar, A. (2023). Optimizing ChatGPT as a writing aid for EFL learners: Balancing assistance and skill development in writing proficiency. *Linguistic Forum*, 5(3), 24-37. <https://doi.org/10.53057/linfo/2023.5.3.3>
- Fang, J. W., Hwang, G. J., & Chang, C. Y. (2022). Advancement and the foci of investigation of MOOCs and open online courses for language learning: A review of journal publications from 2009 to 2018. *Interactive Learning Environments*, 30(7), 1351-1369. <https://doi.org/10.1080/10494820.2019.1703011>
- Guo, K., & Wang, D. (2023). To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing. *Education and Information Technologies*, 1-29. <https://doi.org/10.1007/s10639-023-12146-0>
- Han, J., Yoo, H., Kim, Y., Myung, J., Kim, M., Lim, H., Kim, J., Lee, T. Y., Hong, H., Ahn, S.-Y., & Oh, A. (2023). Recipe: How to integrate ChatGPT into EFL writing education. In *Proceedings of the tenth ACM conference on learning* (pp. 1-8). <https://doi.org/10.1145/3573051.3596200>
- Harunasari, S. Y. (2022). Examining the effectiveness of AI-integrated approach in EFL writing: A case of ChatGPT. *International Journal of Progressive Sciences and Technologies*, 39(2), 357-368. <https://doi.org/10.52155/ijpsat.v39.2.5516>
- Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, 38(1), 237-257. <https://doi.org/10.1111/jcal.12610>
- Hwang, G. J., & Chang, C. Y. (2023). A review of opportunities and challenges of chatbots in education. *Interactive Learning Environments*, 31(7), 4099-4112. <https://doi.org/10.1080/10494820.2021.1952615>
- Hwang, G. J., & Fu, Q. K. (2019). Trends in the research design and application of mobile language learning: A review of 2007-2016 publications in selected SSCI journals. *Interactive Learning Environments*, 27(4), 567-581. <https://doi.org/10.1080/10494820.2018.1486861>
- Imran, A. A., & Lashari, A. A. (2023). Exploring the world of artificial intelligence: The perception of the university students about ChatGPT for academic purpose. *Global Social Sciences Review*, 8(1), 375-384. [https://doi.org/10.31703/gssr.2023\(VIII-I](https://doi.org/10.31703/gssr.2023(VIII-I)
- Javier, D. R. C., & Moorhouse, B. L. (2023). Developing secondary school English language learners' productive and critical use of ChatGPT. *TESOL Journal*, 1-9. <https://doi.org/10.1002/tesj.755>
- 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, 15873-15892. <https://doi.org/10.1007/s10639-023-11834-1>
- Jeon, J., Lee, S., & Choe, H. (2023b). Beyond ChatGPT: A conceptual framework and systematic review of speech-recognition chatbots for language learning. *Computers & Education*, 206, 1-21. <https://doi.org/10.1016/j.compedu.2023.104898>
- 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*, 1-19. <https://doi.org/10.1080/10494820.2023.2204343>
- Ji, H., Han, I., & Ko, Y. (2023). A systematic review of conversational AI in language education: Focusing on the collaboration with human teachers. *Journal of Research on Technology in Education*, 55(1), 48-63. <https://doi.org/10.1080/15391523.2022.2142873>

- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537–550. <https://doi.org/10.1177/00336882231162868>
- Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, 28(1), 973–1018. <https://doi.org/10.1007/s10639-022-11177-3>
- Lashari, A. A., Munawar, I., Mastoi, M., Niaz, P., Buriro, S. A., & Golo, M. A. (2023). Unlocking the potentials of ChatGPT: The efficacy of ChatGPT in ESL learning outcomes. *PalArch's Journal of Archaeology of Egypt/Egyptology*, 20(1), 1135–1143.
- Lee, H., Hsia, C. C., Tsou, A., Choi, S., Hou, H., & Ni, S. (2023). VisionARY: Exploratory research on contextual language learning using AR glasses with ChatGPT. In *Proceedings of the 15th biannual conference of the Italian SIGCHI chapter* (pp. 1–6). <https://doi.org/10.1145/3605390.3605400>
- Li, R. (2022a). Research trends of blended language learning: A bibliometric synthesis of SSCI-indexed journal articles during 2000–2019. *ReCALL*, 34(3), 309–326. <https://doi.org/10.1017/S0958344021000343>
- Li, R. (2022b). Effects of blended language learning on EFL learners' language performance: An Activity Theory approach. *Journal of Computer Assisted Learning*, 38, 1273–1285. <https://doi.org/10.1111/jcal.12697>
- Li, R. (2022c). Effects of mobile-assisted language learning on EFL/ESL reading comprehension. *Educational Technology & Society*, 25(3), 15–29. <https://www.jstor.org/stable/48673721>.
- Li, R. (2023a). Still a fallible tool? Revisiting effects of automated writing evaluation from activity theory perspective. *British Journal of Educational Technology*, 54(3), 773–789. <https://doi.org/10.1111/bjet.13294>
- Li, R. (2023b). Effects of mobile-assisted language learning on EFL learners' listening skill development. *Educational Technology & Society*, 26(2), 36–49. https://www.j-ets.net/collection/forthcoming-articles/26_2.
- Li, R. (2023c). Investigating effects of computer-mediated feedback on L2 vocabulary learning. *Computers & Education*, 198, 1–11. <https://doi.org/10.1016/j.compedu.2023.104763>
- Li, R. (2024). Effects of mobile-assisted language learning on foreign language learners' speaking skill development. *Language, Learning and Technology*, 28(1), 1–26. <https://www.iltjournal.org/item/10125-73553/>.
- Li, B., Kou, X., & Bonk, C. J. (2023a). Embracing the disrupted language teaching and learning field: Analyzing YouTube content creation related to ChatGPT. *Languages*, 8(3), 197–215. <https://doi.org/10.3390/languages8030197>
- Li, X., Li, B., & Cho, S. J. (2023b). Empowering Chinese language learners from low-income families to improve their Chinese writing with ChatGPT's assistance afterschool. *Languages*, 8(4), 238–254. <https://doi.org/10.3390/languages8040238>
- Li, R., Meng, Z., Tian, M., Zhang, Z., & Xiao, W. (2021). Modelling Chinese EFL learners' flow experiences in digital game-based vocabulary learning: The roles of learner and contextual factors. *Computer Assisted Language Learning*, 34(4), 483–505. <https://doi.org/10.1080/09588221.2019.1619585>
- Liu, B. (2023). Chinese university students' attitudes and perceptions in learning English using ChatGPT. *International Journal of Education and Humanities*, 3(2), 132–140. <http://i-jeh.com/index.php/ijeh/index>.
- Liu, C., & Hwang, G. J. (2023). Roles and research trends of touchscreen mobile devices in early childhood education: Review of journal publications from 2010 to 2019 based on the technology-enhanced learning model. *Interactive Learning Environments*, 31(3), 1683–1702. <https://doi.org/10.1080/10494820.2020.1855210>
- Liu, G., & Ma, C. (2024). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*, 18(2), 125–138. <https://doi.org/10.1080/17501229.2023.2240316>
- Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), 410–425. <https://doi.org/10.3390/educsci13040410>
- Marzuki, Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education*, 10, 1–17. <https://doi.org/10.1080/2331186X.2023.2236469>
- Meng, Z., & Li, R. (2023). Understanding Chinese teachers' informal online learning continuance in a mobile learning community: An intrinsic–extrinsic motivation perspective. *Journal of Computing in Higher Education*, forthcoming. <https://doi.org/10.1007/s12528-023-09352-7>
- Mohamed, A. M. (2024). Exploring the potential of an AI-based chatbot (ChatGPT) in enhancing English as a foreign language (EFL) teaching: Perceptions of EFL faculty members. *Education and Information Technologies*, 29, 3195–3217. <https://doi.org/10.1007/s10639-023-11917-z>
- Nguyen, T. T. H. (2023). EFL teachers' perspectives toward the use of ChatGPT in writing classes: A case study at van lang university. *International Journal of Language Instruction*, 2(3), 1–47. <https://doi.org/10.54855/ijli.23231>
- Rakhmonov, I. U., & Kurbonova, R. S. (2023). The pedagogical principles and effectiveness of utilizing ChatGPT for language learning. *Research in Education*, 2(9), 226–243. <http://sjifactor.com/passport.php?id=22258>.
- Riyadini, M. V., & Triastuti, A. (2023). Teachers' perspectives on ChatGPT as a language teaching resource: Benefits, challenges, and pedagogical considerations. In *Conference on English Language Teaching* (pp. 1105–1115). <https://proceedings.uinsaiizu.ac.id/index.php/celti/article/view/614>.
- Sallam, M. (2023). ChatGPT utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns. *Healthcare*, 11(6), 887–907. <https://doi.org/10.3390/healthcare11212819>
- Samarasinghe, K., & Prasangani, K. S. N. (2023). Reliance on AI tools and fostering creativity among Sri Lankan ESL learners: Special focus to ChatGPT. In *3rd International Conference on Educational Technology and Online Learning* (pp. 97–101). <http://repository.lib.vpa.ac.lk/handle/123456789/1844>.
- Shadiev, R., & Yu, J. (2022). Review of research on computer-assisted language learning with a focus on intercultural education. *Computer Assisted Language Learning*, 1–31. <https://doi.org/10.1080/09588221.2022.2056616>
- Shahid, A., Hayat, K., & Iqbal, Z. (2023). Comparative analysis: ChatGPT vs traditional teaching methods. *Pakistan Journal of Society, Education and Language*, 9(2), 585–593. <https://pjsel.jehanf.com/index.php/journal/article/view/1236>.
- Shin, D., & Lee, J. H. (2023). Can ChatGPT make reading comprehension testing items on par with human experts? *Language, Learning and Technology*, 27(3), 27–40. <https://hdl.handle.net/10125/73530>.
- Smutny, P., & Schreiberova, P. (2020). Chatbots for learning: A review of educational chatbots for the facebook messenger. *Computers & Education*, 151, 1–11. <https://doi.org/10.1016/j.compedu.2020.103862>
- Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, 57, 1–11. <https://doi.org/10.1016/j.asw.2023.100752>
- Vera, F. (2023). Enhancing English language learning in undergraduate students using ChatGPT: A quasi-experimental study. In *Congreso Internacional de Aprendizaje Activo* (pp. 18–21). <https://apolo.unab.edu.co/ws/portalfiles/portal/27240222/Libro-de-actas-CIAA-2023.pdf#page=18>.
- White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C. (2023). A prompt pattern catalog to enhance prompt engineering with ChatGPT. <https://doi.org/10.48550/arXiv.2302.11382>. arXiv:2302.11382.
- Woo, D. J., Wang, Y., Susanto, H., & Guo, K. (2023). Understanding English as a foreign language students' idea generation strategies for creative writing with natural language generation tools. *Journal of Educational Computing Research*, 61(7), 1464–1482. <https://doi.org/10.1177/07356331231175999>
- Wu, X. Q., & Li, R. (2024a). Investigating effects of robot-assisted language learning on EFL skill development. *Journal of Educational Computing Research*, 62(4), 1010–1034. <https://doi.org/10.1177/07356331231226171>
- Wu, X. Q., & Li, R. (2024b). Unraveling effects of AI chatbots on EFL learners' language skill development: A meta-analysis. *Asia-Pacific Education Researcher*, forthcoming. <https://doi.org/10.1007/s40299-024-00853-2>
- Xiao, Y., & Zhi, Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions. *Languages*, 8(3), 212–224. <https://doi.org/10.3390/languages8030212>
- Xie, H., Chu, H. C., Hwang, G. J., & Wang, C. C. (2019). Trends and development in technology-enhanced adaptive/personalized learning: A systematic review of journal publications from 2007 to 2017. *Computers & Education*, 140, 1–16. <https://doi.org/10.1016/j.compedu.2019.103599>
- Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, 28, 13943–13967. <https://doi.org/10.1007/s10639-023-11742-4>

- Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2024). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology*, *55*, 90–112. <https://doi.org/10.1111/bjet.13370>
- Yang, L., & Li, R. (2024). Contextualized game-based language learning: Retrospect and prospect. *Journal of Educational Computing Research*, *62*(1), 357–375. <https://doi.org/10.1177/07356331231189292>
- Yang, L., Li, R., & Zhou, Y. (2024). Research trends of game-based language learning in K-12 education: A systematic review of SSCI articles during 2009–2022. *Journal of Computer Assisted Learning*, *40*(3), 1218–1230. <https://doi.org/10.1111/jcal.12944>
- Zhou, T., Cao, S., Zhou, S., Zhang, Y., & He, A. (2023). Chinese Intermediate English Learners outdid ChatGPT in deep cohesion: Evidence from English narrative writing. *System*, *118*, 1–10. <https://doi.org/10.1016/j.system.2023.103141>
- Zou, M., & Huang, L. (2023). To use or not to use? Understanding doctoral students' acceptance of ChatGPT in writing through technology acceptance model. *Frontiers in Psychology*, *14*, 1–9. <https://doi.org/10.3389/fpsyg.2023.1259531>
- Zou, D., Huang, Y., & Xie, H. (2021). Digital game-based vocabulary learning: Where are we and where are we going? *Computer Assisted Language Learning*, *34*(5–6), 751–777. <https://doi.org/10.1080/09588221.2019.1640745>
- Zou, D., Luo, S., Xie, H., & Hwang, G. J. (2022). A systematic review of research on flipped language classrooms: Theoretical foundations, learning activities, tools, research topics and findings. *Computer Assisted Language Learning*, *35*(8), 1811–1837. <https://doi.org/10.1080/09588221.2020.1839502>