

UC Berkeley

L2 Journal

Title

Using ChatGPT in Learning Chinese Mandarin: An Interactive Dictionary Rather Than a Content Facilitator

Permalink

<https://escholarship.org/uc/item/17d9d1cn>

Journal

L2 Journal, 17(1)

Author

Zhang, Yuxiong

Publication Date

2025-09-26

DOI

10.5070/L2.42018

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, available at <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Peer reviewed

ARTICLE

Using ChatGPT in Learning Chinese Mandarin: An Interactive Dictionary Rather Than a Content Facilitator

YUXIONG ZHANG

Polytechnic Institute of Leiria

E-mail: yuxiong.zhang@ipleiria.pt

As a widely discussed AI tool in recent times, ChatGPT has been the subject of extensive research, particularly in EFL learning contexts where it is often utilized as a sophisticated grammar reviewer. This study shifted its focus to exploring ChatGPT's role in Chinese Mandarin acquisition. Seven senior students from an undergraduate Chinese-Portuguese/Portuguese-Chinese translation program in Portugal participated, with six completing a structured four-task learning activity. Analysis of participant conversations with ChatGPT, a written test conducted immediately after the activity, and a questionnaire assessing satisfaction, perceived learning effectiveness, prior tool usage, and learning outcomes revealed that ChatGPT can function as an interactive, adaptable, and comprehensive real-time dictionary for Chinese Mandarin. It effectively explains word meanings and expressions in specific linguistic contexts and suggests appropriate idiomatic expressions based on learners' prompts. Contrary to previous findings, participants reported lower satisfaction and limited perceived learning effectiveness. This may be due to the role ChatGPT played in the present study, where it was used merely as a chatbot for target-language interaction rather than as a tool facilitating their learning process by providing ready-made responses. Additionally, participants' attitudes and preferences for traditional learning methods may also have contributed to these results. The study suggests that effective integration of ChatGPT into Chinese language learning practices requires time and careful consideration.

INTRODUCTION

Technologies have become an indispensable part of our daily lives. While these technologies may bring advantages to our lives, they also subtly change how we view, learn and interact with the world. It can be explained from a sociocultural perspective: "Incorporating tools like AI does not merely facilitate actions that could occur without them; instead, their inclusion fundamentally alters the entire flow and structure of cognitive functions" (Teng, 2024, p. 54).

Studies reveal notable advantages of the incorporation of innovative technologies for delivering content and motivating learners to engage with learning materials (Bin-Hady et al., 2023; Woo et al., 2024). Recently, researchers have turned their sights towards artificial intelligence, notably generative AI tool such as ChatGPT, and demonstrated an ever-growing curiosity about its utilization for language acquisition (Barrot, 2023; Karataş et al., 2024).

ChatGPT is a public-facing generative artificial intelligence (GenAI) chatbot that allows users to interface with large language models (LLMs). Based on a large language corpus obtained from the Internet, GenAI chatbots can generate responses to users' prompts through statistical prediction and reinforcement learning algorithms (OpenAI, 2023). In the words of Xiao and Zhi (2023), prompts refer to the initial text input by the user to provide context and

request a response, while generated responses are the text output by ChatGPT (2023). Due to their capacities to accurately predict, comprehend natural language, and generate human-like responses, LLM-based chatbots have attracted significant attention (Laato et al., 2023; Liu & Ma, 2024). By January of 2023, ChatGPT gained over 100 million active users within merely two months and became one of the fastest-growing AI tools (Teng, 2024). This number doubled by September of 2024 (Shubham, 2024).

Regarding language learning, ChatGPT can foster activities for interactive learning, provide instant and targeted feedback, and supplement teacher support by cultivating interactions and conversations based on users' prompts (Barrot, 2023; Cai et al., 2023; Yang & Li, 2024). With real-time feedback on their content, learners no longer need to keep waiting for corrections from teachers, which helps to augment their autonomy in learning, build a personalized learning environment, and reduce teachers' workload (Barrot, 2023; Barrot, 2024; Derakhshan & Ghiasvand, 2024).

To explore how ChatGPT can be integrated into Chinese as a Foreign Language (CFL) learning, we start in this article by examining how ChatGPT functions and how chatbots have been used in language education. We then discuss both the potential risks and the benefits associated with using ChatGPT in language learning contexts. Finally, we present five concrete examples of its current application in CFL classrooms, highlighting its pedagogical potential as well as the limitations that warrant further investigation.

Mechanics of ChatGPT

ChatGPT adopts a powerful software system of machine learning called Generative Pre-trained Transformer (GPT-3), created by Open Artificial Intelligence (Open AI) (Bin-Hady et al., 2023, Cai et al., 2023). The large language model (LLM) enables ChatGPT to process and produce human-like interactions with users across different topics, using 175 billion parameters as its foundation (Barrot, 2023; Karataş et al., 2024). However, ChatGPT does not have the ability to truly understand the text; instead, it predicts and associates answers based on probabilities derived from its sophisticated algorithm. While this tool can seamlessly interact with users, it cannot generate any sentences entirely on its own (Sangeethapriya, 2024). Powered by LLM, ChatGPT accesses a significant amount of information and uses Natural Language Processing (NLP) to simulate human-like responses within specific contexts (Agustini, 2023; Barrot, 2024). Although AI tools like ChatGPT cannot understand the texts as humans do in a practical sense, they can accurately extract large volumes of information from natural language and create precise responses to prompts based on this data (Bostrom, 2017).

However, by possessing the capacity to sustain a conversation with users through a question-answer format, ChatGPT differs radically from traditional search engines and enables the individualization of answers for each user (Cai et al., 2023). In other words, ChatGPT does not retrieve information from the Internet in the same way that search engines do; instead, it produces tailored text to respond to the user's prompt through statistical prediction and a powerful algorithm (Escalante et al., 2023). Meanwhile, compared to traditional chatbots, which provide one-size-fits-all feedback and back-and-forth conversations with users, ChatGPT can iterate and facilitate responses, explain the rationale behind its answers, summarize long articles, generate academic papers, conduct logical assessments, debug programming code, and provide translation solutions (Karataş et al., 2024; Xiao & Zhi, 2023; Yang & Li, 2024).

Chatbots and Language Learning

As dialogue agents, chatbots engage human users in dialogue-based interactions and facilitate turn-by-turn language exchanges, providing personalized and interactive language learning experiences (Guo et al., 2022; Xiao & Zhi, 2023). According to sociocultural theory, language learning is a participatory and socially constituted process that cannot be separated from the social context (Vygotsky, 1978). In the words of Pask (1976), learning occurs in an interactive environment through a structured conversation. Thus, even though they were not primarily invented for language pedagogy, as with many other technologies with potential for this purpose (Bin-Hady et al., 2023), chatbots can be viewed as an ideal collaborative and interactive partner for language learning.

Limitations and Risks of Applying ChatGPT to Language Learning

Although ChatGPT is a powerful LLM-based tool capable of generating grammatically correct output, the accuracy of its context related to idiomatic expressions, cultural references, subtleties, and nuances of natural language remains unclear (Mohamed, 2024; Rudolph et al., 2023). In the same vein, ChatGPT does not possess higher-order critical thinking or analytical abilities to understand and contextualize complex concepts, common sense, and emotions, or even to detect writing errors related to profound syntactic structure and pragmatics (Barrot, 2023; Rudolph et al., 2023; Teng, 2024).

In addition, the significant capabilities of easily accessible AI tools like ChatGPT have led to widespread concerns about the misuse, reliability, and ethical use of their output, especially in academic contexts (Escalante et al., 2023; Karataş et al., 2024). Since the responses generated by ChatGPT typically do not contain proper citations, the inappropriate use of ChatGPT by students may result in a loss of fairness, copyright infringement, the spread of misinformation, high-tech plagiarism, and threats to academic integrity (Derakhshan & Ghiasvand, 2024; Xiao & Zhi, 2023). Due to the accurate and coherent output generated by ChatGPT, students can complete writing assignments with minimal effort by using ChatGPT. In this case, the use of LLM-based tools in learning may lead to learning loss, superficial literacy development, and a decrease in creativity, critical thinking, and even linguistic abilities (Barrot, 2023; Derakhshan & Ghiasvand, 2024; Karataş et al., 2024; Mohamed, 2024; Sangeethapriya, 2024; Woo et al., 2024). Additionally, although ChatGPT can provide an interactive learning environment, spoken conversation and emotional support are still irreplaceable (Karataş et al., 2024). Students who excessively rely on AI tools for language learning may miss out on the indispensable dynamic learning experiences that come from interacting with others, which could eventually lead to social isolation (Barrot, 2024).

Even though the emerging application of ChatGPT in learning environments raises concerns, its use should also be contextualized within a broader global perspective. Advanced technologies create a learner-centered environment where students' autonomy prevails; however, they should not be viewed as a threat to the presence of teachers (Sangeethapriya, 2024). Effective teacher-student communication in the classroom is always pivotal and indispensable for a positive learning experience (Mohamed, 2024). Nonetheless, throughout the history of educational technology, ChatGPT is certainly not the first innovation perceived as the end of traditional education. Technologies such as television, computers, the Internet, mobile phones, social media, and augmented reality have all been regarded as revolutionary breakthroughs or panaceas for pedagogical practice. In the words of Rudolph, Tan, and Tan

(2023, p.2), such hopes for total innovation in education were “often exaggerated,” and consideration for the application of technologies in practice was “frequently insufficient.”

Benefits of ChatGPT in Language Learning

As an AI chatbot, the most distinctive and relevant characteristic of ChatGPT is its ability to integrate updated knowledge across a wide range of subjects during conversational interactions with users, functioning as an encyclopedic information resource with easy access. In traditional classroom-based learning environments, it is not possible to provide a personalized learning approach tailored to each learner’s level and needs. Due to its repository of vast amounts of continuously updated and current information, ChatGPT can encourage students to reflect on their learning outcomes and biases, take ownership of their learning, and become self-directed learners, thereby bolstering their critical thinking abilities (Agustini, 2023; Barrot, 2023; Laato et al., 2023).

In other words, within a student-centered pedagogical approach, learning can be realized through learners’ experimentation and experience (Rudolph et al., 2023). Nevertheless, in the context of language learning, text-based interactive conversations can be viewed as a supportive and immersive method for practicing theoretical knowledge. Learners are able to apply what they have learned from books in effective language interactions, satisfying individual paces, preferences, and learning objectives, which radically changes the one-size-fits-all approach found in traditional language classrooms (Barrot, 2024).

Since ChatGPT provides a distinctive source of language input, learners can access a vast array of language learning resources and practice their language skills by mimicking conversations and interactions (Barrot, 2024; Kohnke et al., 2023). Additionally, compared to a human language partner, ChatGPT offers 24/7 seamless support, contains a broader stock of language knowledge, and reiterates responses to frequently asked questions (Shaikh et al., 2023). Several authors (Bin-Hady et al., 2023; Karataş et al., 2024; Mohamed, 2024; Shaikh et al., 2023; Yan, 2023) have confirmed that language learners can enhance their conversational, reading, writing, vocabulary, and grammar skills with ChatGPT.

Moreover, as a language learning tool, ChatGPT is distinguished by its accuracy in generating texts, which, according to Barrot (2023, p. 2), are “typically coherent and grammatically correct.” Mizumoto, Shintani, Sasaki, and Teng (2024) compared ChatGPT’s performance to that of human evaluators and a specialized AI-based writing assistant tool, Grammarly, in assessing errors across 232 writing samples. The correlation between assessments by ChatGPT and human evaluators was found to be even stronger than that between Grammarly and humans, suggesting the potential of ChatGPT in enhancing EFL writing accuracy. Yan (2023) examined the effects of texts generated by ChatGPT on EFL learners’ perceptions of writing, and the results suggested that this LLM-based tool can produce grammatically well-structured examples with lexical diversity for reference to improve writing performance.

Regarding the reliability and effectiveness of using ChatGPT as a Chinese grammar-checking tool, Yang and Chen (2025) tested 153 erroneous single-sentence examples and found that both versions 3.5 and 4.0 produced grammatically correct and fluent corrections. Consequently, ChatGPT can be perceived as a powerful tool for helping learners recognize biases and flaws, provide suggestions for improvement, rectify linguistic errors, refine the use of language forms, and enhance writing quality (Barrot, 2023; Qu & Wu, 2024; Yan, 2023).

ChatGPT in Chinese Language Learning

While ChatGPT has been widely studied in the EFL context, its use in Chinese as a Foreign Language (CFL) is gradually gaining attention. Regarding the application of ChatGPT in the context of Chinese language acquisition, Li (2023) examined the experiences of 21 L2 learners using ChatGPT for Chinese writing over 16 weeks. Compared to learners who continued with traditional teaching methods, students in the experimental group demonstrated “significantly better” results in post-writing tests (Li, 2023, p. 67). Similarly, Song (2025) found that university-level L2 Chinese learners consistently achieved higher scores in writing tasks when using ChatGPT for assistance.

Li, Ren, Jiang, and Chen (2023) designed various instructional activities based on students’ learning objectives and linguistic proficiency. Intermediate-low learners were encouraged to use ChatGPT to independently study grammar and engage in conversational exchanges with cultural insights, while intermediate-high learners and near-native speakers in academic writing classes utilized the AI tool to enhance their writing composition. Instructors reported that ChatGPT’s flexibility, real-time responsiveness, and wealth of information provided learners with a valuable discussion platform and enriched their learning experience.

Within the same Chinese educational institution, Jiang, Li, and Chen (2024) examined the impact of AI-powered activities on critical thinking skills in an introductory and an intermediate course for second-language learners, as well as an academic writing course for near-native Chinese speakers. The results indicated that key cognitive skills—including remembering, understanding, applying, analyzing, evaluating, and creating—were consistently applied across all three language courses at different proficiency levels.

Additionally, Ge (2024) explored the potential of using ChatGPT to teach Chinese characters and enhance reading skills for heritage learners. The study concluded that the AI chatbot effectively provided detailed explanations of Chinese characters, facilitated reading comprehension, fostered deep cultural exchanges, and offered personalized learning assistance.

Nevertheless, compared to its widespread application in EFL contexts, research on the use of ChatGPT in Chinese as a Foreign Language (CFL) remains notably limited. Moreover, existing CFL studies tend to highlight general benefits, such as “enriching the learning experience,” “facilitating reading comprehension,” or “fostering cultural exchange,” without clearly articulating the mechanisms behind these outcomes or providing detailed examples of learner-AI interaction. As a result, the specific ways in which ChatGPT contributes to CFL learning remain unclear.

A common limitation is the narrow focus on writing. Few studies examine how learners actually engage with the AI tool during writing tasks, in what aspects their performance improves, or whether learners critically interact with ChatGPT or simply extract ready-made content without deeper reflection. In addition, most studies assess outcomes through pre- and post-test comparisons and neglect the interactive, process-oriented nature of AI-mediated learning, failing to explore how learners prompt ChatGPT, evaluate its responses, or integrate its feedback into their work.

These gaps highlight the need to closely examine learners’ real-time interactions with ChatGPT in CFL contexts. Key questions remain, such as how learners engage with the AI chatbot during specific tasks, what kinds of learning improvements are observable, and which instructional designs most effectively support meaningful engagement with AI tools. A deeper understanding of learner-AI interaction, especially how learners formulate prompts, interpret

AI responses, and incorporate feedback, may provide clearer insight into the mechanisms behind language development in CFL environments.

According to Xia, Ge, Shen, and Najjar (2024), several Western countries have shown a growing interest in studying Chinese Mandarin among youth. For instance, in Portugal, three municipalities—São João de Madeira, Estarreja, and Espinho—have implemented Chinese Mandarin as a compulsory subject in primary education. According to the Portuguese Ministry of Education (Direção-Geral da Educação, n.d.), by the 2021/2022 school year, 11 secondary school groups from different regions offered Chinese language courses as an optional subject. In higher education, the University of Lisbon, the University of Minho, the University of Aveiro, and the Polytechnic Institute of Leiria offer undergraduate programs related to the Chinese language. Additionally, there is strong demand for Chinese Mandarin learning beyond these educational stages; currently, there are five Confucius Institutes in Portugal.

Accordingly, the following three research questions are proposed:

1. What are the observable ways in which CFL learners use ChatGPT as an AI chatbot to obtain information they need for Chinese language learning activities?
2. In what ways do CFL learners replicate and integrate texts generated by ChatGPT into their own Chinese language production?
3. How do CFL learners perceive the effectiveness of using ChatGPT as an AI chatbot in language learning when it does not provide direct answers?

METHODOLOGY

Study Design

To explore the impact of ChatGPT on Chinese Mandarin acquisition, we prepared a 70-minute exploratory learning activity designed to encourage students to interact more with ChatGPT in Chinese. Instead of introducing the AI tool during the writing process for direct error correction, we designed four different exercises aimed at helping students use the target language to communicate with ChatGPT in order to find possible solutions for the tasks, leveraging its original function as a chatbot.

Previous studies have indicated that text-based interactive conversations with ChatGPT can be an immersive method for practicing theoretical linguistic knowledge and can encourage students to take ownership of their learning and become self-directed learners (Agustini, 2023; Barrot, 2023; Barrot, 2024). Therefore, we first aimed to assess whether ChatGPT could be efficiently utilized by CFL learners to comprehend texts within a limited time frame. For this purpose, we included an exercise in which participants were asked to summarize a long passage. Our objective was to explore Chinese language acquisition through written communication in a variety of contextual situations.

The remaining three tasks were designed to encourage participants to use the target language to describe the information they needed and to analyze relevant responses during their interactions with the AI tool. In particular, the second, third and final exercises focused on identifying the names of specific objects or places in Chinese and determining the appropriate idiomatic expression for certain contexts. This approach aligns with the belief that chatbot-led learning can enhance micro-learning effectiveness and facilitate information acquisition (Cai et al., 2023). Our goal is for students not only to gain new linguistic insights into Mandarin Chinese in terms of grammar and vocabulary but also to enhance their knowledge of Chinese culture during their interactions with the chatbot.

Tasks

In the first exercise, students are asked to summarize the main idea of an approximately 700-character text that includes several idiomatic expressions and relatively difficult words not covered in their curriculum. The objective is to familiarize students with the tool's functionality and response mechanism.

The second exercise requires students to identify the origins of six types of lanterns from various regions in Asia and find their names in Chinese. These lanterns show notable differences in their shape and material, and include, for instance; a traditional Chinese wooden lantern, a Korean lotus lantern with Buddhist origins, a Chinese bamboo lantern with two characters on it, Japanese lanterns in white and red colors featuring multiple characters, a Korean *cheongsachorong* in red and blue within a cuboid shape, and colorful Vietnamese lanterns in a diamond shape. This task makes use of ChatGPT's ability to provide users with up-to-date information across various subjects, while students are encouraged to describe each image in the target language to discover the answers, using the AI chatbot to consult specific words.

In the third exercise, students describe several images of a Chinese city to ChatGPT in order to determine its location within China. The first picture features a multistory traditional pagoda with flying eaves on each floor and a statue of two cranes, indicating its name—Yellow Crane Tower. The second picture depicts a *sakura* park with numerous trees in full bloom with pink flowers and a Japanese-style tower in the background. The last picture shows a double-deck road and rail bridge spanning a vast river, with two gazebos on either side. Compared to the previous exercise, this requires students to provide more detailed information, demonstrating a more advanced capacity for language organization and expression. Additionally, our test showed that by providing key information, even without detailed descriptions, ChatGPT is capable of deducing the city depicted in the three images (Figure 1).

In the final exercise, students are given three groups of illustrations and must use *chengyu*, a kind of traditional Chinese idiomatic expression, to describe the situations depicted in each group. This exercise has no standard answers, allowing each student to express their own perspective and understanding.

Parameters

The activity was conducted under the guidance and monitoring of the author, with participants required to interact with ChatGPT exclusively in Chinese and without the use of dictionaries. Each exercise was timed—25 minutes for the first, 15 minutes for the second, 10 minutes for the third and 20 minutes for the last. We printed the exercises on separate sheets of paper and distributed them after all participants completed the last task. The instructions for each exercise were written in Chinese to prevent the use of any other language during the activity. To assist students in effectively initiating their interactions with ChatGPT, we provided suggested approaches at the end of each exercise prompt.

Students were asked to document their conversations with ChatGPT by taking screenshots or sharing the output link provided by the tool, then submit these records to the author for further analysis. Before the activity, participants received a brief presentation on ChatGPT's key characteristics, functions, and potential applications for learning. Any questions about its use were addressed during the session. Additionally, we explained the learning activity process and the main objectives of the study.

After completing the learning activity, participants undertook a written test consisting of four questions designed to evaluate the learning outcomes achieved through their interactions with ChatGPT. The first question comprised five multiple-choice sub-questions aimed at assessing the students' comprehension of idiomatic expressions and words that they were presumably unfamiliar with to determine whether they fully understood the text from the first task. The second and third questions required students to identify the main characteristics of traditional Chinese lanterns and describe an image of the same landscape from a different angle, with the goal of examining whether they imitated or applied any idiomatic expressions or syntactic structures used by ChatGPT during their interactions. Finally, the last question asked participants to specify which idiomatic expression they used and its meaning in Chinese Mandarin, allowing us to gauge their understanding, ability to explain the expressions in the target language and potential learning outcomes with the AI tool.

Figure 1

Testing Chatgpt's Ability to Deduce the Name of a Chinese City

请根据我的描述推测出我所说的中国城市：1.有一个古楼，楼前面有鸟的雕像。2.一大片树林开着粉色的花，树后面有一个日式的楼。3.在一条江上有座双层桥，上面有车，下面有铁路。

根据您的描述，我推测您说的城市可能是武汉。

1. **古楼和鸟的雕像**：武汉的黄鹤楼是著名的古楼，黄鹤楼前的景区内有不少雕像和纪念性建筑，可能符合“古楼”和“鸟的雕像”这个描述。
2. **树林开粉色花和日式楼**：武汉有不少花园和公园，其中武汉大学的樱花非常有名，每年春天，樱花盛开，形成一片粉色的花海。另外，武汉大学内也有一些日式风格的建筑。
3. **双层桥**：武汉的长江大桥是一个非常著名的双层桥，上面有公路，下面有铁路，连接着武汉的汉口与武昌。

这些特征都符合武汉的景观特色，尤其是黄鹤楼、武汉大学的樱花和长江大桥。

Translation in English:

Author: Please guess which Chinese city I am describing: 1. There is an ancient tower with carvings of birds in front of it; 2. A large tree is blooming with pink flowers, and there is a Japanese-style pavilion under the tree; 3. There is a double-deck bridge across a river. The upper level has cars, and the lower level has trains.

ChatGPT: Based on your description, the city you are referring to is likely Wuhan.

1. **Ancient tower and bird carvings:** Wuhan's Yellow Crane Tower is a famous landmark. The scenic area around Yellow Crane Tower contains carvings and statues of birds, which may match the "ancient tower with bird carvings" you described.
2. **Trees with pink blossoms and pavilions:** Wuhan has many parks and gardens. The cherry blossoms at Wuhan University are especially famous; every spring the cherry trees bloom magnificently, forming a sea of pink flowers. In addition, some of Wuhan's parks have Japanese style pavilions, which fits your description.
3. **Double-deck bridge:** Wuhan's Yangtze River Bridge is a very well-known double-deck bridge, with a road on the upper level and a railway on the lower level, connecting the two

sides of Wuhan across the Yangtze River. These features all align with Wuhan's landscape, especially Yellow Crane Tower, Wuhan University's cherry blossoms, and the Yangtze River Bridge.)

Questionnaire

Additionally, we prepared a questionnaire to investigate participants' opinions regarding their learning outcomes. In technology-assisted learning practices, satisfaction plays a vital role and serves as a fundamental measure of how participants respond to the learning process. While satisfaction may not guarantee learning outcomes, dissatisfaction can significantly hinder learning effectiveness (Cai et al., 2023; Woo et al., 2024). We adopted the questionnaire used by Woo, Wang, Guo, and Susanto (2024), which included 14 satisfaction items adapted from Fisher, Schumaker, Culbertson, and Deshler (2010). In their study, the authors reported a Cronbach's alpha value of 0.98 for the adapted satisfaction questionnaire.

Since the effectiveness of ChatGPT in L2 learning practices has yet to be evaluated (Mohamed, 2024; Yang & Li, 2024), and the overwhelming majority of research focuses on English learning experiences (Jeon et al., 2023; Liu & Ma, 2024; Yang & Li, 2024), we included five items related to learning effectiveness based on Cai, Lin, and Yu's (2023) investigation of the influencing factors on 458 learners' attitudes towards ChatGPT-assisted language learning. This allows us to evaluate participants' perceptions of ChatGPT's effectiveness in learning Chinese Mandarin.

All items were first translated into Portuguese by two experienced translators and then sent to three experts in the Portuguese linguistic field for their feedback. After incorporating their suggestions, we presented the final version of the questionnaire in European Portuguese to collect participants' opinions on their satisfaction and the effectiveness of ChatGPT.

In the last section of the questionnaire, we included an open-ended question about participants' previous experiences using ChatGPT to learn Chinese Mandarin. Additionally, participants were directly asked about their learning outcomes, with the hope that they would share their experiences from a first-person perspective as participants in the learning activity.

Ethical Principles

During data collection, no private information was sought; only gender and learning time were requested. According to Oppenheim (2000, p. 84), "the respondent's right to privacy and the right to refuse to answer certain questions, or to be interviewed at all, should always be respected, and no undue pressure should be brought to bear." All students voluntarily participated in the learning activity. Before the presentation on ChatGPT and prior to completing the questionnaire, they were informed about the data to be collected and assured that no private information would be disclosed. Data collection only commenced after obtaining consent from all participants.

Participants

This study employed a purposive sample consisting of seven senior students (six female and one male) from an undergraduate Chinese-Portuguese/Portuguese-Chinese translation program at a public higher education institution in Portugal. Over the past three years, these students have completed more than 1,300 hours of Chinese Mandarin coursework aimed at developing a range of linguistic skills. One participant (participant G) is of Chinese descent

and was raised in Portugal. During the second task, one participant (participant A) opted to withdraw from the activity, and in accordance with ethical research principles, his/her decision was respected, leaving six participants to complete the activity and the written test.

Data Analysis

With the objective of exploring the effects of applying ChatGPT in CFL learning practices through written interactions as an AI chatbot, we first analyzed the conversations between learners and ChatGPT submitted by the participants in order to understand how they interacted with the AI tool to achieve their objectives. Since the activities were designed differently from traditional writing exercises, participants had to interact with ChatGPT to find answers for each task. Therefore, the way they utilized ChatGPT may differ from previous studies, which have typically limited its use to more passive functions.

To evaluate the learning outcomes of these four learning activities, we analyzed not only the responses provided during the learning activities but also those from the post-written test. The evaluation methods mainly included two aspects: the accuracy of the provided answers and the task completion rate, which was measured by directly counting the total responses and the number of correct answers, since participants were not obligated to answer all the questions in the written test due to ethical considerations. Except for the first task, which required participants to summarize a 700-character text, we opted to assess whether they could produce a summary of approximately 30 characters. The second and third tasks featured fixed responses. Although the last task was open-ended, the *chengyu* provided by learners still needed to align with the linguistic context described in the images, based on the definitions provided in the *Modern Chinese Dictionary (Xiandai Hanyu Cidian, 7th edition)*.

In the written test, the answers to the first two tasks were fixed. The third exercise required learners to describe an image of the same city from a different angle. Here, the evaluation criteria focused on vocabulary choice and grammatical accuracy in Mandarin Chinese. For the final question, which directly asked for an idiomatic expression used by the participant and its meaning, we evaluated whether the *chengyu* was written correctly and whether the accompanying explanation demonstrated understanding. We assumed that participants would demonstrate their learning process by replicating words, idiomatic expressions, and linguistic structures generated by the AI chatbot during the post-test (Kohnke et al, 2023; Barrot, 2024).

It was also our objective to assess learners' satisfaction and self-perceived learning effectiveness throughout the four learning tasks involving the integration of ChatGPT. We adopted two questionnaires previously used in similar studies (Cai et al., 2023; Woo et al., 2024), and the results are also discussed in this section. Additionally, we included a section in the questionnaire to directly inquire about participants' previous experience using ChatGPT for Chinese language learning.

Therefore, the findings section of the present study is organized into three main parts: 1) learning process and learning outcomes, based on the four tasks and the analysis of conversations with the AI tool and post-written tests; 2) motivation and perceived learning effectiveness, based on the questionnaire responses; 3) previous use of ChatGPT before the learning activity, also assessed through the questionnaire.

These methods were chosen to capture both qualitative and quantitative aspects of ChatGPT's role in language learning: interaction logs reveal students' autonomous use of the chatbot as a conversational partner, rather than merely as a content facilitator that provides

ready-made answers; post-tests offer evidence of language learning outcomes; and questionnaires provide insights into learners' attitudes and self-perceived progress.

FINDINGS

Learning Process and Learning Outcomes

Task 1

In the first task, participants were asked to summarize a text containing approximately 700 Chinese characters. No specific instructions were given on how ChatGPT should be used; the goal was to allow participants to become familiar with the AI tool and observe whether they could leverage ChatGPT to address challenges in their Chinese Mandarin learning process. Upon analyzing the conversations between students and ChatGPT, we observed that participants primarily used the tool as a dictionary. For example, participant D asked for the pronunciation of a character with “立” (which means “set up” or “erect” in English) as the upper component and “早” (“early”) as the lower component. After receiving the pronunciation and several examples illustrating the character's meaning, the participant then inquired about the meaning of the word “文章,” (“article”) which contained a character the participant was unfamiliar with. However, when ChatGPT's initial explanation was unclear to her, she requested a simplified version (Figure 2).

Even though the interaction consisted of only a brief explanation requested by the learner, it suggests that ChatGPT can function as a Chinese-language dictionary, enabling users to specify and refine their questions to obtain the desired information in an autonomous learning environment directed by the learner him/herself. Learners can leverage learning objectives. Based on the limited information provided by the participants regarding their conversations with ChatGPT, we found that all participants utilized the chatbot as a dictionary to explore the meaning of certain characters, words, or even clauses, especially participants A and B (see Table 1).

Figure 2

Conversation between Participant D and ChatGPT

有一个子我不知道怎么读，上面是“立”，下面是“早”

您说的字是“章”，读作 **zhāng**。这是一个常见的汉字，有多种意思，例如：

- 文章的一部分（章节）
- 规章制度（章程）
- 章子、印章（用于盖印的工具）

如果您有其他问题，随时问我！

那“文章”是什么意思？

“文章”指的是通过语言文字表达思想、情感或传递信息的书面作品。它可以涵盖各种形式的书写内容，如散文、小说、诗歌、论文、报告等。具体来说：

1. **书面作品**：文章是一种用文字写成的表达形式，通常有一定的结构和主题。
2. **表达思想或情感**：文章常用来传递作者的观点、感受、经验等。
3. **不同体裁**：文章可以是长篇大论的论文，也可以是简短的随笔或诗歌。

例如，一篇小说是用故事情节表达作者的想法，而一篇评论文章则是分析或评价某个主题。

我比太明白，能更容易地告诉我妈

当然！我可以用更简单的方式解释：

文章就是用文字写出来的内容，用来表达想法或告诉别人一些事情。比如：

- **故事**是文章，用来讲一个有趣的情节。
- **作文**是文章，小朋友在学校写的。
- **新闻**也是文章，用来告诉我们发生了什么事。

所以，文章就是用文字写出来的一段话，帮助别人了解或明白某些东西。

Translation in English:

Participant: There's a character (the user mistakenly typed “子” instead of “字”, meaning “character”) I don't know how to read. The top is “立,” the bottom is “早.”

ChatGPT: The character you mentioned is “章,” pronounced “zhāng”. This is a common Chinese character with several meanings, for example:

- A part of a written work (chapter)
- Rules or regulations (regulation)
- Seal or stamp (a tool used for stamping)

If you have any other questions, feel free to ask me anytime!

Participant: Then what does “文章” mean?

ChatGPT: “文章” refers to a written work that expresses ideas, thoughts, feelings, or conveys information through language. It can cover many types of written content, such as essays, novels, poems, theses, reports, etc. Specifically, it can mean:

1. Written composition: An article is a form of text composition, usually with a certain structure and theme.
2. Expression of ideas: An article is often used to convey the author’s opinions, feelings, experiences, etc.
3. Different forms: An article can be a serious academic essay, or it can be a short and casual piece.

For example, a novel is a piece of writing that tells a story created by the author, while a review article is used to analyze and evaluate a certain topic.

Participant: I don’t quite understand (the participant mistakenly typed “比” instead of “不”, meaning “not”). Can you explain it in a simpler way?

ChatGPT: Of course! I can explain it in a simpler way:

“文章” is a piece of writing made up of words. It is used to express ideas or to tell others something. For example:

- A story is an article, used to describe what happened that day.
- A short essay is an article, where a child writes about their friends.
- A news report is also an article, used to tell us what has happened.

So, an article is just a passage written with words to help others understand certain things.)

Table 1

Use Of Chatgpt To Consult The Meaning Of Characters, Words, Or Clauses

| | Character | Word | Clause |
|---------------|-----------|------|--------|
| Participant A | 1 | 7 | - |
| Participant B | 1 | 6 | - |
| Participant C | 2 | 2 | 1 |
| Participant D | 1 | 3 | 1 |
| Participant E | - | 1 | - |
| Participant F | - | - | 1 |
| Participant G | - | 1 | - |

However, ChatGPT often provides detailed responses to participants' brief and straightforward prompts, indicating a notable contrast in response length and complexity (Table 2). A strategy that may enhance comprehension, as demonstrated by participant A, is to request explanations in simpler terms. This approach led to a significant difference in the response length generated for participant A compared to the other two participants (Figure 3). In fact, as recorded in Figure 2, after Participant D asked the chatbot to explain its answer in a simpler way, the length of its output temporarily decreased from 203 characters to 135 characters (Table 2).

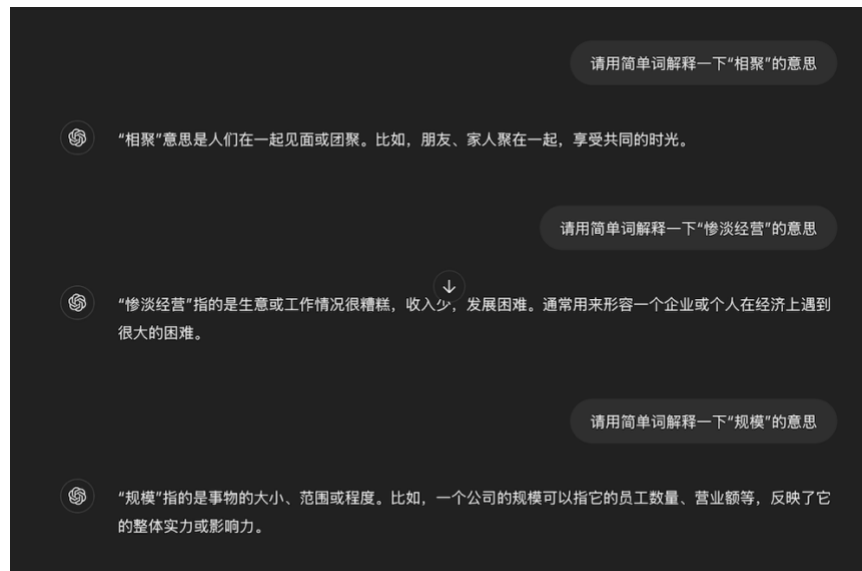
Table 2

Number of Characters (Including Punctuation) Exchanged between Participants A, B, and D and ChatGPT During Task 1

| | | | | | | | | | |
|----------------------|----------------------|-----|-----|-----|-----|-----|-----|-----|----|
| Participant A | Input by participant | 18 | 15 | 16 | 15 | 10 | 16 | 18 | 16 |
| | Output by ChatGPT | 66 | 48 | 56 | 46 | 73 | 39 | 53 | 57 |
| Participant B | Input by participant | 9 | 7 | 8 | 7 | 7 | 7 | 15 | 6 |
| | Output by ChatGPT | 169 | 166 | 160 | 158 | 158 | 225 | 67 | 67 |
| Participant D | Input by participant | 25 | 11 | 15 | 8 | 8 | 5 | 14 | |
| | Output by ChatGPT | 80 | 203 | 135 | 138 | 187 | 162 | 214 | |

Figure 2

Conversation between Participant A and ChatGPT



Translation in English:

Participant: Please explain the meaning of “相聚” in simple words.

ChatGPT: “相聚” means people meeting or gathering together. For example, friends or family members getting together to enjoy shared time.

Participant: Please explain the meaning of “惨淡经营” in simple words.

ChatGPT: “惨淡经营” refers to a situation where a business or work is doing poorly, with low income and difficult development. It is often used to describe an enterprise or individual facing significant financial difficulties.

Participant: Please explain the meaning of “规模” in simple words.

ChatGPT: “规模” refers to the size, scope, or extent of something. For example, the scale of a company can refer to its number of employees, turnover, etc., reflecting its overall strength or influence.)

With respect to the completion of the first task, only Participants D and G successfully summarized the text in the required 30 characters. These two participants also correctly answered all the questions regarding the meanings of five words and idiomatic expressions not covered in their curriculum. Meanwhile, Participant F provided two correct answers, while the remaining participants each had only one correct answer.

Task 2

Based on participants' answers, the second task appeared to be the most challenging (Table 3). Participants generally used two kinds of strategies to identify the names of the lanterns. The first approach involved describing the primary features of each lantern in the images. However, when participants provided only general or vague descriptions to ChatGPT the tool sometimes struggled to accurately identify the intended item. Analysis of ChatGPT's responses revealed several simplified and incorrect answers on this topic. For example, when Participant D described one lantern as “made of wood” and another as resembling a flower in red and yellow, ChatGPT responded with the general terms “wooden Chinese traditional lantern (lamp)” and “flower-shaped yellow and red Chinese traditional lantern (lamp)” (Figure 4).

Table 3

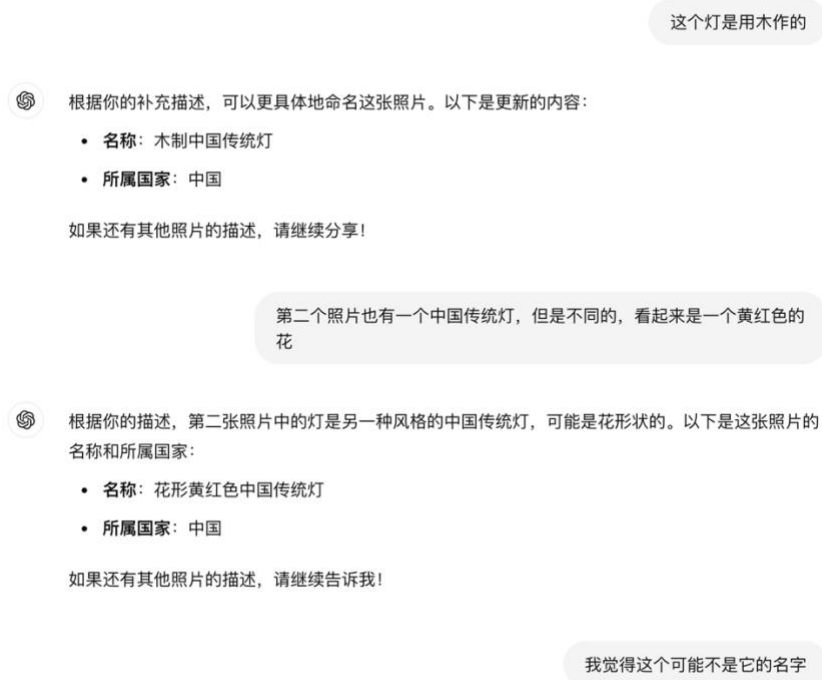
Answers Provided by Participants in the Second Exercise of the Written Test

| | Number of answers given for the names of lanterns | Number of correct answers | Number of answers given for the origin of lanterns | Number of correct answers |
|---------------|---|---------------------------|--|---------------------------|
| Participant B | 4 | 0 | 4 | 2 |
| Participant C | 3 | 0 | 0 | - |
| Participant D | 1 | 1 | 0 | - |
| Participant E | 1 | 0 | 0 | - |
| Participant F | 2 | 0 | 3 | 0 |
| Participant G | 1 | 1 | 0 | - |

Another strategy was utilized by Participant G. As a fluent bilingual, the participant asked ChatGPT to identify traditional lanterns across Asia and received a response containing 937 Chinese characters, which requires a relatively advanced level of reading comprehension in Mandarin Chinese. Nonetheless, the information provided by ChatGPT might not contain all the information of the lanterns of the task and the Participant G ended up answering only one of the names.

In the written test, participants were asked to describe the main characteristics of traditional Chinese lanterns. We hoped to see students replicate the information generated by ChatGPT in their responses. However, the participants primarily chose to either describe the images of lanterns provided in the learning activity (Participants B, C, and E) or to offer a general overview of Chinese lanterns, which possess various meanings and colors (Participant F).

Figure 3

Conversation between Participant D and ChatGPT

Translation in English:

Participant: This lamp is made (the participant mistakenly typed “作” instead of “做”, meaning “made”) of wood.

ChatGPT: Based on your additional description, we can name this photo more specifically. Here is the updated content:

- Name: Wooden Traditional Chinese Lamp
- Country: China

If you have descriptions for other photos, please continue sharing!

Participant: The second photo also has a Traditional Chinese lamp, but it's a different one; it looks like a yellow-red flower.

ChatGPT: Based on your description, the lamp in the second photo is another style of Traditional Chinese lamp, possibly flower-shaped. Here is the name and country for this photo:

- Name: Yellow-Red Flower-Shaped Traditional Chinese Lamp
- Country: China

If there are no descriptions for other photos, please let me know!

Participant: I think this might not be its correct name.)

Task 3

In a similar way, participants were required to describe three images of a Chinese city to ChatGPT and identify the city to which the images belonged in the third task. With the exception of Participant G, all participants completed the task; however, Participant E provided only the name of the city without supplying the additional required information regarding its location. Notably, none of the participants correctly identified the city in question. Furthermore, general descriptions such as “old building,” “park with *sakura*,” or “big bridge” were insufficient for ChatGPT to accurately locate the city depicted in the images, as these characteristics could apply to numerous large-scale cities in China. Overall, participants tended to accept the answers generated by ChatGPT without seeking further information to validate them.

In the written test, four participants described the city and provided examples of its tourist attractions as required by the exercise. Participants B and E referenced several characteristics from the images presented in the learning activity, while Participant D articulated a response in his/her own words, which were genuinely original and not derived from the output generated by ChatGPT. In contrast, Participant F reiterated the name of the city suggested by the AI tool, despite the actual name having already been provided in the title of the exercise.

Task 4

The final task requires participants to use fixed Chinese idiomatic expressions, known as *chengyu*, to describe each situation depicted in three sets of illustrations. Given the subjective nature of interpreting the images, multiple acceptable answers may exist. While all participants provided responses to varying degrees, only a small portion of these answers could be considered correct (see Table 4). The incorrect answers were primarily attributed to either inappropriate usage or the expressions not being recognized as *chengyu*.

Table 4

Answers Provided By Participants In The Last Task

| | Number of answers given | Number of correct answers |
|---------------|-------------------------|---------------------------|
| Participant B | 3 | 1 |
| Participant C | 2 | 0 |
| Participant D | 3 | 0 |
| Participant E | 3 | 1 |
| Participant F | 3 | 1 |
| Participant G | 1 | 1 |

Some participants (E and G) chose to describe the illustrations to ChatGPT and relied on the tool to generate appropriate answers for them, while others (C, D, and F) sought solutions from ChatGPT for the situations they described. In both cases, ChatGPT offered multiple solutions for consideration; however, participants may not have comprehended the explanations generated by the AI tool, leading them to select inadequate *chengyu* for given situations, despite the presence of correct answers in ChatGPT's output. Additionally, we noted that not all the responses provided by ChatGPT were *chengyu*, as some were common

idiomatic expressions in modern Chinese, which resulted in incorrect answers for the task (see Figure 5).

Figure 5

Some Answers do not Correspond to Chengyu Provided by ChatGPT

🌀 这种情况可以用成语“羡慕嫉妒恨”来形容。这个成语表达的是人们对他人所拥有的优点或优势的羡慕和嫉妒。在你的描述中，动物们彼此看着，希望能像其他动物一样大或有某种特质，正体现了对其他动物的羡慕之情。

你描述的情况可以用成语“群起而攻之”来形容。这个成语指的是一群人齐心协力，共同对付某个目标或问题。在你的描述中，七个人互相指责、鼓励，这种集体行为可以与“群起而攻之”对应。

2. 旧瓶装新酒

- **意思：**指用旧的形式装载新的内容。形容表面上有变化，但本质上没有改变。
- **用法：**可以说：“苹果的手机设计虽有更新，但实际上只是旧瓶装新酒。”

针对这个情况，可以使用以下成语：

1. 羡慕嫉妒恨

- **意思：**形容对别人拥有的东西感到羡慕、妒忌，并因此产生怨恨。
- **用法：**可以说：“我对那些人的努力感到羡慕嫉妒恨，但他们却没有对其他人产生这种情绪。”

Translation in English:

This situation can be described using the idiom “羡慕嫉妒恨”. This *chengyu* expresses feelings of envy and jealousy towards the advantages or strengths that others possess. In your description, the animals looking at each other, wishing to be as big or to have certain traits like the others, perfectly illustrates this feeling of envy towards other animals.

The situation you described can be described using the idiom “群起而攻之”. This *chengyu* refers to a group of people uniting to jointly confront a target or problem. In your description, the seven people accusing and encouraging each other represents this kind of collective action, which corresponds to “群起而攻之”.

2. “旧瓶装新酒”

- **Meaning:** It means putting new wine in old bottles. It describes a situation where the form appears old but the content is new, or superficially there is change, but the essence remains the same.
- **Usage:** One could say: “Although Apple's phone designs have been updated, in reality, it's just putting new wine in old bottles.”

For this situation, the following idiom can be used:

1. “羡慕嫉妒恨”

- **Meaning:** Describes feeling envy, jealousy, and even resentment towards what others have.
- **Usage:** One could say: “I feel envy and jealousy (“羡慕嫉妒恨”) towards those people's efforts, but they themselves don't have these feelings towards others.”

In the written test, only three participants (B, C, and D) completed the final exercise, and two of them (C and D) shared their interactions with ChatGPT in the fourth task. Notably, we observed an instance of the learning process in Participant C, who replicated the explanation of the *chengyu* generated by ChatGPT verbatim (see Figure 6). In this sentence, the learner used the idiomatic Chinese structure “虽然……但是……”, which means “although...but...” in English. While this construction is grammatically correct in Chinese, it is considered incorrect in both English and Portuguese, where combining “although” with “but”, or “*embora*” with “*mas*” in the same clause is redundant and ungrammatical. This suggests that the learner likely recalled and applied the syntactic structure generated by the AI tool, rather than producing it independently or relying on his/her first language, which does not permit such usage.

Figure 6

Output Generated by ChatGPT and Answer of Participant C

可以用的成语是***“表里如一”***。这个成语指的是外表和内心一致，形容事物的表象与本质相同。在你的例子中，虽然看起来不同，但其实它们本质上是一样的。如果你还有其他问题，随时问我！

意思是“虽然看起来不同，但其实它们本质上是一样的。”

Translation in English:

ChatGPT: “The *chengyu* that can be used is “表里如一”. This *chengyu* means that the outside appearance and the inner reality are consistent, describing a situation where the surface appearance is the same as the essential nature. In your example, even though they look different, they are essentially the same. If you have other questions, feel free to ask me anytime!”

Participant C: “The meaning is “even though they look different, they are essentially the same”.)”

Satisfaction and Learning Effectiveness

This study also aims to assess participants’ satisfaction and perceived effectiveness of using ChatGPT as a chatbot for completing activities in learning Chinese Mandarin. This was evaluated through a questionnaire containing 13 items on satisfaction and 5 items on learning effectiveness. The Cronbach’s alpha for the 13 satisfaction items was 0.938, indicating a high reliability of the results (Table 5). Overall, participants expressed limited satisfaction with the learning experience using ChatGPT, with a mean satisfaction score of approximately 2.67 on a 5-point scale, falling below the midpoint of 3. According to individual mean scores, participant A, who did not complete the full learning activity, was the least satisfied, while participants D and F also reported lower satisfaction levels. In contrast, the remaining three participants showed neutral to slightly positive satisfaction with the activity (Table 6).

Table 5

Reliability Statistics for Satisfaction

| Cronbach's alpha | Cronbach's alpha based on standardized items | N of items |
|------------------|--|------------|
| .938 | .943 | 13 |

Table 6

Mean Scores for Each Participant

| | Mean | Std. Deviation |
|---------------|------|----------------|
| Participant A | 1.36 | 0.72 |
| Participant B | 3.29 | 0.96 |
| Participant C | 3.64 | 0.61 |
| Participant D | 2.07 | 0.80 |
| Participant E | 3.14 | 0.91 |
| Participant F | 2.21 | 0.77 |
| Participant G | 3.14 | 0.91 |

Regarding participants' attitudes toward learning effectiveness, we adopted the definition put forth by Cai, Lin, and Yu (2023: 4), which defines it as "a holistic construct that includes increasing levels of learning/subject mastery in ChatGPT-assisted language learning." In the context of the present study, this refers specifically to learners' experiences in accomplishing the objectives of each of the four tasks.

The Cronbach's alpha test also indicated high reliability, with an alpha value of 0.901 (Table 7). While participant A showed a negative attitude toward the learning effectiveness of the activity, the lowest score was recorded by participant F, who completely disagreed with all statements about ChatGPT's effectiveness in learning. Additionally, participants D, C, and G also expressed doubts about the tool's effectiveness in the learning activity (Table 8). Among the five items related to learning effectiveness, participants expressed the most disagreement with the idea that ChatGPT could enhance their learning motivation (Table 9).

The overall relatively negative feedback regarding learning effectiveness when using ChatGPT in the activities we designed may correspond to the unsatisfactory results in the post-written tests. For instance, only two participants managed to summarize the text within the required character limit in the first exercise, five out of six participants failed to provide the correct names of the lanterns in the second exercise, none of the participants identified the correct city in the third exercise, and only three participants completed the final exercise.

Table 7

Reliability Statistics for Learning Effectiveness

| Cronbach's alpha | Cronbach's alpha based on standardized items | N of items |
|------------------|--|------------|
| .901 | .891 | 5 |

Table 8

Mean Scores For Each Participant

| | Mean | Std. Deviation |
|---------------|------|----------------|
| Participant A | 1.4 | 0.8 |
| Participant B | 3.6 | 0.8 |
| Participant C | 2.2 | 0.4 |
| Participant D | 1.8 | 0.4 |
| Participant E | 3.2 | 0.75 |
| Participant F | 1 | 0 |
| Participant G | 2.2 | 0.4 |

Table 9

Mean Scores for Each Learning Effectiveness Item

| | Mean | Std. Deviation |
|--|--------|----------------|
| ChatGPT can improve my language learning efficiency. | 2.2857 | 1.25357 |
| ChatGPT can improve my learning performance. | 2.2857 | 1.25357 |
| ChatGPT can enhance my learning motivation. | 1.8571 | .69007 |
| ChatGPT can enhance my learning efficacy. | 2.1429 | 1.06904 |
| ChatGPT can provide more learning resources. | 2.4286 | 1.13389 |

To gain further insights into participants' learning experience with ChatGPT, they were also asked to reflect on their learning outcomes from the activity. All six participants generally felt that ChatGPT was not sufficiently effective for learning Chinese Mandarin (Table 10). Participant C remarked that "using ChatGPT intelligently and effectively is harder than it looks," and Participant B emphasized that it requires time and practice to make the tool significantly effective for learning Chinese. Even as a bilingual student, Participant G commented on the challenges of having adequate linguistic ability to interact with ChatGPT effectively in completing learning tasks.

Table 10

Participant's Opinions on the Learning Effectiveness of ChatGPT (English Translation)

| | |
|---------------|---|
| Participant A | None, since I gave up halfway through. |
| Participant B | I believe that ChatGPT can benefit language learning when used correctly and with ample time. However, in this context, with limited time and little experience using it, I don't think it has offered any of the benefits that my usual traditional methods provide. |
| Participant C | Using ChatGPT intelligently and effectively is harder than it looks. It's still an underdeveloped tool compared to others I've used, and it lacks the resources to improve its usability. It has hardly helped me understand the content. |
| Participant D | It has hardly helped me understand the content. |
| Participant E | I now have a better understanding of how ChatGPT can assist with learning Chinese. However, I still feel that, in general, ChatGPT is not useful for my study methods or my learning. |
| Participant F | In general, I still feel that ChatGPT is not useful for my study methods or learning. I prefer other resources, such as physical or online dictionaries (like Pleco), because they are much more accessible and better meet my needs. |

| | |
|---------------|--|
| Participant G | I don't think it helped me much because, to complete the tasks, I needed a high level of Chinese. This meant I didn't have the language skills required to explain to ChatGPT what I wanted. |
|---------------|--|

Previous Use of ChatGPT Before Learning Activity

Finally, we asked participants whether they had prior experience with ChatGPT or used it in their daily learning. Four respondents reported no prior use of the AI tool, preferring to acquire language skills by interacting with native speakers or consulting dictionaries. Participant D, however, mentioned hesitating to use ChatGPT due to a lack of confidence in his/her Chinese language proficiency. Participant C had previously experimented with ChatGPT for refining an article written in Chinese, but found the experience unsatisfactory, preferring human feedback. Participant G used ChatGPT as a dictionary to understand word usage in specific contexts. In contrast, Participant E could be considered an advanced user, employing ChatGPT to summarize articles, answer questions, and even generate short compositions based on prompts from him/herself or teachers (Table 11).

Table 11

Participant's Previous Experience Using ChatGPT (English Translation)

| | |
|---------------|---|
| Participant A | No, I don't like using tools like these to learn languages, especially one as difficult as Chinese—or even for other things. To me, it's a form of cheating, as I don't need to make any effort. I prefer practicing my Chinese by talking with friends, and if I have any doubts, I can just call them or send a message. I also practice my listening comprehension, which is quite weak, by communicating with them. |
| Participant B | I've never used it before, as I prefer methods like dictionaries or talking to native speakers. |
| Participant C | Yes, but it didn't work, it didn't help me as much in correcting a text as a native speaker would. |
| Participant D | No, because I feel that my level of Chinese is not yet sufficient to express my doubts clearly. |
| Participant E | Yes, I have used ChatGPT to summarize articles and other texts, to ask questions, and to create short compositions based on prompts from myself or from teachers. |
| Participant F | I had never used it until now because I didn't think it was useful for my study and learning. I had never been interested in using ChatGPT. |
| Participant G | Yes, I have used ChatGPT before to learn how to use vocabulary in specific contexts. |

DISCUSSION

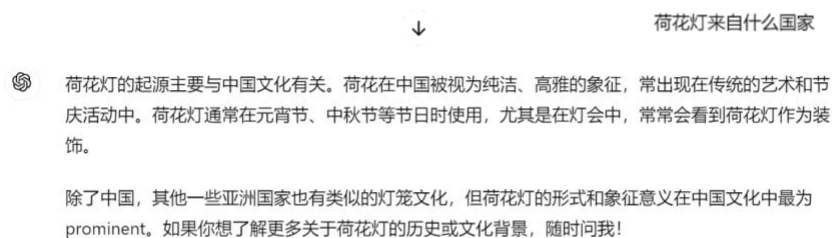
To integrate ChatGPT as a reliable language learning tool, it is essential to validate the quality of Chinese language usage in the outputs generated by the model. As a less commonly trained language compared to English, which has a significant amount of resources available, Chinese is not considered a strength of ChatGPT, as noted by Rudolph, Tan, and Tan (2023). Upon analyzing all interactions between participants and the LLM-based tool, we identified only one error: when Participant B inquired about the origin of the lotus lantern, the response generated by ChatGPT included the English word "prominent," despite the expectation that the answer should have been solely in Mandarin Chinese (see Figure 7).

While ChatGPT predominantly produces grammatically correct texts in Mandarin, the reliability of the content it generates remains questionable. In the second task, students were

asked to identify the names of specific lanterns illustrated in the images; however, ChatGPT fabricated nonexistent names based on the information provided by the participants. This limitation arises because, while ChatGPT can generate responses based on vast databases, it lacks true understanding of texts and their conveyed meanings (Agustini, 2023; Barrot, 2024; Sangeethapriya, 2024). Consequently, it may fail to link to accurate information resources when users do not express their queries clearly. As a synchronous chatbot, ChatGPT generally prioritizes providing an immediate response over seeking further clarification, even when it lacks sufficient and efficient data. For learners without advanced language skills or deep topic knowledge, this can result in acquiring inaccurate information during the learning process. This might explain why one participant remarked that ChatGPT "lacks the resources to improve its usability" in the questionnaire. Additionally, we observed that ChatGPT sometimes failed to distinguish between Chinese fixed idiomatic expressions—*chengyu*—and more common phrases, leading to confusion among participants.

Figure 4

Output by ChatGPT in Chinese Mandarin Containing an English Word



Translation in English:

Participant: What country is the lotus lantern from?

ChatGPT: The origin of the lotus lantern is primarily associated with Chinese culture. In China, the lotus flower is regarded as a symbol of purity and elegance, often featured in traditional art and festive activities. Lotus lanterns are commonly used during holidays such as the Lantern Festival and the Mid-Autumn Festival, especially in lantern fairs, where they are frequently seen as decorations.

Besides China, some other Asian countries also have similar lantern traditions, but the form and symbolic meaning of the lotus lantern are most “prominent” in Chinese culture. If you would like to learn more about the history or cultural background of lotus lanterns, feel free to ask me!

Nonetheless, in contrast to the results of a previous study conducted by Woo, Wang, Guo, and Susanto (2024), which reported a high level of satisfaction among participants after utilizing ChatGPT in an EFL writing classroom, this study yields different findings, revealing a low level of satisfaction among participants. In the present study, we aim to revive the function of ChatGPT as a chatbot rather than as a grammar reviewer in the students’ learning

process, hoping that participants can interact with the tool in the target language to complete the tasks we designed. This may explain the dissimilar results obtained. As the authors noted, students generally express satisfaction with ChatGPT because the tool generates responses aligned with their preferences and facilitates their writing process; however, in our study, participants' written production in Chinese was not similarly assisted. According to research conducted by Cai, Lin, and Yu (2023), hedonic motivation is a crucial factor influencing performance expectancy and perceived satisfaction in ChatGPT-assisted language learning.

Several studies (Agustini, 2023; Karataş et al., 2024; Qu & Wu, 2024) have highlighted that ChatGPT can bolster learners' motivation; however, participants in this study primarily disagreed that ChatGPT enhances learning motivation. One reason for this may be that the students involved in this study each have their own pace and learning methods for acquiring Chinese Mandarin. Three participants stated that they had not used ChatGPT for learning the Chinese language because the tool did not align with their preferred learning strategies. Furthermore, students may feel more comfortable with the presence of human interaction and emotional connection in their language learning (Agustini, 2023). As Mohamed (2024) stated, AI-powered language learning tools may not be equally suitable for everyone, particularly for those who prefer more interaction with humans.

Additionally, two participants noted their lack of language proficiency when interacting with ChatGPT in the target language. Although previous studies have indicated that ChatGPT is capable of customizing the outputs it generates to accommodate learners' personalized needs and pace (Barrot, 2024; Yang & Li, 2024), students may still lack confidence in their proficiency with the target language and harbor reservations about interacting with the tool in Chinese Mandarin.

However, despite participants reporting a low level of satisfaction and negative attitudes towards the learning effectiveness of ChatGPT in Chinese language learning, we observed an instance of one participant who appears to have unconsciously integrated the output generated by the AI tool into their subsequent written test. While it remains questionable whether this was merely a product of short-term memory, it is evident that the learner read the response carefully and understood the information it conveyed. This suggests that it is possible for learners to acquire information during interactions with ChatGPT. However, it is advisable to always verify the reliability of the information provided by the AI tool through other available resources.

CONCLUSION

To explore the actual effects of ChatGPT as an AI-based chatbot rather than a content facilitator during the process of Chinese language acquisition, we designed a learning activity comprising four different tasks. Rather than relying on pre-generated answers from the AI tool as a mere content facilitator, these tasks aimed to encourage students to communicate with ChatGPT in the target language to obtain the desired responses. The tasks simulated potential real-life scenarios in which learners might use ChatGPT for language learning purposes and resolve language-related questions in their daily routines, such as summarizing a long text, identifying the names of specific objects or places in the target language, and finding the appropriate idiomatic expression for a given context.

In order to answer the question: "What are the observable ways in which CFL learners use ChatGPT as an AI chatbot to obtain information for Chinese language learning activities?", we analyzed learners' interactions with the AI tool and found that students used

ChatGPT not only to understand the meaning of unfamiliar words encountered in tasks but also to clarify terms found in the chatbot's responses during the first task, which required learners to summarize a long text in the target language. There was no specific instruction or hint provided regarding how to use ChatGPT in this activity. Learners engaged with ChatGPT to resolve their questions about characters or vocabulary, primarily due to the tool's inherent interactive features. In task 4, they described specific situations or illustrated content to explore the most appropriate *chengyu* for each case. ChatGPT generated different responses in real time based on each learner's prompts. Unlike a traditional paper dictionary, the AI tool provided personalized answers adapted to learners' unique situational descriptions, effectively functioning as an "interactive dictionary." Notably, identifying the correct usage of a Chinese idiomatic expression solely through a situational description is not feasible with a traditional paper dictionary. Moreover, although four learners provided only one correct *chengyu* in the last task, all their responses were different, suggesting that ChatGPT may offer a personalized learning environment.

Additionally, as for the question "In what ways do CFL learners replicate and integrate texts generated by ChatGPT into their own Chinese language production?", we observed that one participant replicated ChatGPT's explanation verbatim in the written test. Notably, the structure used does not align with the syntactic norms of the learner's first language, suggesting that the learner had understood and internalized the explanation provided by the chatbot. This suggests a tangible effect of incorporating this tool into L2 learning practice. The second and third exercises in the post-written test were designed to assess whether learners had acquired any idiomatic expressions or syntactic structures from their interactions with ChatGPT. However, this expectation was not confirmed by the participants' answers.

Finally, the answer we reached to the third research question—"How do CFL learners perceive the effectiveness of using ChatGPT as an AI chatbot in language learning when it does not provide direct answers?"—was generally negative. Unlike learning practices where ChatGPT is used as a content facilitator or grammar reviewer that provides ready-made, grammatically correct answers, participants in this study interacted with the tool solely as a chatbot to complete predefined tasks. Consequently, students may not have experienced the "miraculous powers" often attributed to the AI tool in facilitating language learning, which could explain their lack of satisfaction and negative attitudes toward its effectiveness—findings that diverge from previous studies. Another potential factor shaping participants' views could be their perceived attitudes regarding ChatGPT and preference for traditional learning methods, suggesting that not every student is inherently prepared for ChatGPT's integration into their learning process. Successfully and efficiently incorporating this AI tool may thus require time and thoughtful adaptation.

Like other advanced technologies, ChatGPT was not initially designed specifically for educational purposes. While it possesses significant capabilities for generating and refining text in multiple languages, its use in learning contexts requires careful consideration. Identifying effective ways to integrate ChatGPT into students' learning processes, rather than using it solely as a content facilitator, remains an important area for exploration.

Limitation of the Present Study

As exploratory research utilizing a purposive sample, this study involved only seven participants. Additionally, in accordance with ethical principles, we respected the decision of one participant who chose to discontinue the learning activity halfway through. Consequently,

our analysis was based solely on the interaction records we received from the remaining participants, which did not encompass the entirety of the conversations they had with ChatGPT. The incomplete data collection ultimately impacted our analysis of potential learning outcomes during the activity. Furthermore, three participants explicitly expressed a preference for traditional learning methods, which may have influenced the study's results. We recommend that future research expand the participant pool and explore whether prior attitudes towards using ChatGPT in language learning affect satisfaction with its usage in learning activities. Moreover, since over half of the participants had no previous experience using ChatGPT, a more comprehensive training session may be required than the one-hour presentation as we provided. A well-structured orientation could help learners engage more effectively with AI tools during the learning process. There is no way to guarantee participants' familiarity with the AI tool, and inefficient usage or insufficient knowledge about it may result in limited learning gains and, consequently, low satisfaction. This could prevent an accurate demonstration of ChatGPT's true capabilities in Chinese language acquisition. It would also be beneficial for future studies to focus on the application of ChatGPT in learning Chinese Mandarin over a more extended period.

Compared to previous studies that reported high levels of satisfaction with using ChatGPT in EFL writing practice, our study focused on its role in developing learners' linguistic capacities as a real-time conversational peer, providing fluent and grammatically correct interactions in the target language. However, future research should also consider other influential factors that may differentiate the activities in the present study from those in previous research, such as differences in task genres, disparities between the target language and the learner's first language, variations in writing proficiency, and the teacher's role in interactions with AI tools.

ACKNOWLEDGEMENTS

This work is funded by National Funds through the FCT - Foundation for Science and Technology, I.P., within the scope of the project Ref^a UIDB/05507/2020 and DOI identifier <https://doi.org/10.54499/UIDB/05507/2020>. Furthermore, we would like to thank the Centre for Studies in Education and Innovation (Ci&DEI) and the Polytechnic of Leiria for their support.

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. <https://doi.org/10.62775/edukasia.v4i2.373>
- Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, 100745. <https://doi.org/10.1016/j.asw.2023.100745>
- Barrot, J. S. (2024). ChatGPT as a language learning tool: An emerging technology report. *Technology, Knowledge and Learning*, 29(2), 1151–1156. <https://doi.org/10.1007/s10758-023-09711-4>
- 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*. <https://doi.org/10.1108/LHT-05-2023-0200>
- Bostrom, N. (2017). *Superintelligence: Paths, dangers, strategies* (Reprint with corrections). Oxford University Press.
- 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>

- Derakhshan, A., & Ghiasvand, F. (2024). Is ChatGPT an evil or an angel for second language education and research? A phenomenographic study of research-active EFL teachers' perceptions. *International Journal of Applied Linguistics*. <https://doi.org/10.1111/ijal.12561>
- Direção-Geral da Educação. (n.d.). *Projeto-piloto de oferta do Mandarim no ensino secundário*. Retrieved from <https://www.dge.mec.pt/projeto-piloto-de-oferta-do-mandarim-no-ensino-secundario>
- 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), 57. <https://doi.org/10.1186/s41239-023-00425-2>
- Fisher, J. B., Schumaker, J. B., Culbertson, J., & Deshler, D. D. (2010). Effects of a computerized professional development program on teacher and student outcomes. *Journal of Teacher Education*, 61(4), 302–312. <https://doi.org/10.1177/0022487110369556>
- Ge, H. (2024). The role of ChatGPT in Chinese reading education for Chinese as a heritage language (CHL) learners. In *Proceedings of the 2024 7th International Conference on Big Data and Education* (pp. 17–21). <https://doi.org/10.1145/3704289.3704298>
- Guo, K., Wang, J., & Chu, S. K. W. (2022). Using chatbots to scaffold EFL students' argumentative writing. *Assessing Writing*, 54, 100666. <https://doi.org/10.1016/j.asw.2022.100666>
- 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>
- Jiang, X., Li, J., & Chen, C. H. (2024). Enhancing critical thinking skills with ChatGPT-powered activities in Chinese language classrooms. *International Journal of Chinese Language Teaching*, 5(1). <https://doi.org/10.46451/ijclt.20240105>
- 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*, 1–24. <https://doi.org/10.1007/s10639-024-12574-6>
- 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>
- Laato, S., Morschheuser, B., Hamari, J., & Björne, J. (2023). AI-assisted learning with ChatGPT and large language models: Implications for higher education. In *2023 IEEE International Conference on Advanced Learning Technologies (ICALT)* (pp. 226–230). IEEE. <https://doi.org/10.1109/ICALT58122.2023.00072>
- Lexicography Department of the Institute of Linguistics, Chinese Academy of Social Sciences. (2016). *Xiandai Hanyu Cidian* [Modern Chinese Dictionary] (7th ed.). The Commercial Press.
- Li, J., Ren, X., Jiang, X., & Chen, C. H. (2023). Exploring the use of ChatGPT in Chinese language classrooms. *International Journal of Chinese Language Teaching*, 4(3), 36–55. <https://doi.org/10.46451/ijclt.20230303>
- Li, G. (2023). Personalized international Chinese education based on ChatGPT. In *2023 13th International Conference on Information Technology in Medicine and Education (ITME)* (pp. 63–67). IEEE. <https://doi.org/10.1109/ITME60234.2023.00024>
- 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>
- Mizumoto, A., Shintani, N., Sasaki, M., & Teng, M. F. (2024). Testing the viability of ChatGPT as a companion in L2 writing accuracy assessment. *Research Methods in Applied Linguistics*. <https://doi.org/10.1016/j.rmal.2024.100116>
- Mohamed, A. M. (2024). Exploring the potential of an AI-based Chatbot (ChatGPT) in enhancing English as a Foreign Language (EFL) teaching: perceptions of EFL Faculty Members. *Education and Information Technologies*, 29(3), 3195–3217. <https://doi.org/10.1007/s10639-023-11917-z>
- OpenAI. (2023). *GPT-4 system card*. Retrieved from <https://cdn.openai.com/papers/gpt-4-system-card.pdf>
- Oppenheim, A. N. (2000). *Questionnaire design, interviewing and attitude measurement*. Continuum.
- Pask, G. (1976). *Conversation theory: Applications in education and epistemology*. Elsevier.
- 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>
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of applied learning and teaching*, 6(1), 342–363. <https://doi.org/10.37074/jalt.2023.6.1.9>

- Sangeethapriya, S. (2024). AI based ChatGPT impacts on second language learners. In *2024 10th International Conference on Communication and Signal Processing (ICCSP)* (pp. 790-794). IEEE.
<https://doi.org/10.1109/ICCSP60870.2024.10543703>
- Shaikh, S., Yayilgan, S. Y., Klimova, B., & Pikhart, M. (2023). Assessing the usability of ChatGPT for formal English language learning. *European Journal of Investigative Health, Psychology and Education*, *13*, 1937–1960.
<https://doi.org/10.3390/ejihpe13090140>
- Shubham, S. (2024, September 2). ChatGPT statistics (Sep. 2024) – 200 million active users. *Demandsage*.
<https://www.demandsage.com/chatgpt-statistics/>
- Song, D., & Tang, A. F. (2025). Exploring AI-assistance in L2 Chinese writing with standardized assessment tasks. *Language Learning & Technology*, *29*(2), 161–189. <https://hdl.handle.net/10125/73613>
- Teng, M. F. (2024). A systematic review of ChatGPT for English as a foreign language writing: Opportunities, challenges, and recommendations. *International Journal of TESOL Studies*, *6*(3), 36-57.
<https://doi.org/10.58304/ijts.20240304>
- Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Woo, D. J., Wang, D., Guo, K., & Susanto, H. (2024). Teaching EFL students to write with ChatGPT: Students' motivation to learn, cognitive load, and satisfaction with the learning process. *Education and Information Technologies*, 1-28. <https://doi.org/10.1007/s10639-024-12819-4>
- Xia, J., Ge, Y., Shen, Z., & Najjar, M. R. (2024). The Auxiliary Role of Artificial Intelligence Applications in Mitigating the Linguistic, Psychological, and Educational Challenges of Teaching and Learning Chinese Language by non-Chinese Students. *International Review of Research in Open and Distributed Learning*, *25*(3), 116-133. <https://doi.org/10.19173/irrodl.v25i3.7680>
- 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. <https://doi.org/10.3390/languages8030212>
- Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, *28*, 1–25. <https://doi.org/10.1007/s10639-023-11742-4>
- Yang, C. T. Y., & Chen, H. H. J. (2025). ChatGPT and L2 Chinese writing: evaluating the impact of model version and prompt language on automated corrective feedback. *Computer Assisted Language Learning*, 1-29. <https://doi.org/10.1080/09588221.2025.2453205>
- 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>