

The investigation of application related to ChatGPT in foreign language learning

Chen Liu

Department of Instrumentation, University of Science and Technology Beijing, 30
Xueyuan Road, Haidian District, Beijing, 100083, China

1912100420@mail.sit.edu.cn

Abstract. The integration of Chat Generative Pre-trained Transformer (ChatGPT) into the realm of education has garnered significant attention recently. Due to its capability of providing answers to inquiries, the feasibility of ChatGPT's implementation in education has been examined. This review explores the potential benefits of ChatGPT in foreign language learning. Initially, different methods used to research the ChatGPT's plausibility of improving language skills are shown. Furthermore, this review discusses some different application scenarios of ChatGPT, also presenting its drawbacks and limitations. Despite recognized concerns, the viability of ChatGPT as a valuable instrument for foreign language education remains substantiated. The technology exhibits promise in offering personalized and interactive language practice. While acknowledging potential pitfalls, the study underscores the need for further investigation to harness ChatGPT's potential optimally. By emphasizing ongoing research and development, this review envisions a future where ChatGPT contributes significantly to language education, underscoring the necessity for careful exploration of its evolving capabilities.

Keywords: ChatGPT, language learning, large language model.

1. Introduction

In contemporary times, Artificial Intelligence (AI) has penetrated the life of human beings, and its applications in various fields bring individuals great convenience. In particular, AI integrated with education has garnered much attention due to its numerous benefits for learners. Among them, AI in language learning has become increasingly popular concerning its convenience and efficiency, since foreign language proficiency offers several advantages and is frequently necessary for numerous academic and employment prospects [1].

Recent research has been carried out on various topics related to AI and foreign language acquisition. There are some famous AI language learning tools accessible easily on the Internet such as Duolingo, Elsa Talk, and Rosetta Stone. These can facilitate increasing motivation and provide a more personalized learning experience, abundant resources, and real-time feedback [2]. A language learning chatbox named Englishbot was developed and evaluated against a traditional listen-and-repeat interface. It turned out that it improved the students' fluency and showed higher user engagement. Users allocated increased time to interacting with the language learning chatbox, signifying its potential to augment language acquisition outcomes [3].

After being released in November 2022, Chat Generative Pre-trained Transformer (ChatGPT) created a considerable sensation and gained hundreds of millions of users in just a few months [4]. Due to its capability of providing answers to inquiries, the feasibility of the application of it in education appropriately arouses fierce discussion. Given that ChatGPT has been available for no more than a year, the amount of research on ChatGPT's application in language acquisition is still somewhat limited. A paper presented alternative evaluations of ChatGPT's nature, forms, features, benefits, and integration into second language(L2) learning. It developed a theoretical framework that will serve as the basis for additional study once ChatGPT is widely used in L2 contexts around the world [5]. A paper explored the benefits and risks of incorporating ChatGPT into language learning. Students were encouraged to practice critical judgment and teachers were also regularly refresh their information technology understanding [6]. The paper pointed out the plausible scenarios in which ChatGPT may be used in language learning and concluded its significant pedagogical affordances for language learners and teachers [7]. Both advantages and risks of utilizing ChatGPT in higher education language courses were studied in a paper. The paper also gave some suggestions about how to use ChatGPT in higher education [8]. Besides English, ChatGPT's usefulness for learning other languages has also been studied. A paper demonstrated the significant potential of ChatGPT for application in Vietnamese education and educational transformation. It also pointed out the need for a more comprehensive understanding of ChatGPT's capabilities [9].

ChatGPT has also raised concerns among scholars about its impact on academic integrity. A study found that ChatGPT could jeopardize the validity of online exams because it demonstrated critical thinking skills and could generate writing that is remarkably realistic with little input. [10]. An investigation aimed at evaluating experiences in several kinds of educational scenarios was presented in a paper, which revealed various problems including dishonesty caused by ChatGPT [11].

Therefore, this review paper will explore the feasibility and potential benefits of ChatGPT in foreign language learning based on the analysis of the relevant literature in section 2. In section 3, this paper will discuss the plausible application of ChatGPT in foreign language learning. Finally, some suggestions for future research about the implementation of ChatGPT in language learning will be given in this review.

2. Methodology

2.1. Overview of ChatGPT

ChatGPT, developed by OpenAI, is a large-scale generative language model, whose name is partly from the abbreviation of Generative Pre-trained Transformer (GPT). GPT is a series of deep learning models which have undergone iterative enhancements through the introduction of various editions. Reinforcement Learning from Human Feedback (RLHF) was used for model training. To be more specific, ChatGPT, trained on an Azure AI supercomputing infrastructure, is fine-tuned from one model that belongs to the GPT-3.5 series [12].

Presently, GPT-4 stands as the most advanced system developed by OpenAI, characterized by its capacity to generate safer and more pragmatic responses. GPT-4 acquires broader general knowledge and powerful skills, which help to resolve complex situations more accurately. GPT-4 was improved safer by OpenAI over six months. Regarding internal assessments, GPT-4, compared to GPT-3.5, is 40% more likely to generate accurate responses and also is 82% less likely to answer the prompts which are not allowed.[13].

2.2. Case studies for English learning with ChatGPT

In this section, several applications of ChatGPT in foreign language learning are elaborated. Since the vast majority of Chinese students' first foreign language is English, this section focuses on cases of English learning with ChatGPT's instruction.

2.2.1. ChatGPT for English reading comprehension example. Xiao et al. attempted to develop a system to help English reading comprehension by using ChatGPT to generate relative exercises [14]. Initially, a fine-tuned GPT-2 with PPLM control was used as the baseline method. Subsequently, utilizing ChatGPT, input prompts were designed manually to help generate high-quality reading comprehension passages without additional control. All generated textual content was divided into two kinds, including zero-shot and one-shot. Zero-shot means creating reading passages from scratch and one-shot means creating reading passages based on a referenced passage. Also, questions and corresponding answers for given passages were generated by ChatGPT. Furthermore, extensive evaluations were conducted for both the reading passages and questions. 30 human-written reading passages were randomly selected from textbooks, 30 passages provided by ChatGPT, and 30 passages generated by the baseline model were mixed and shuffled into a sample set. Automatic evaluation metrics and human assessments were both used to evaluate these passages. Each evaluation is composed of 5 scores that represent various aspects of text quality. These factors were carefully chosen based on their significance to the reading comprehension scenario and are frequently employed in human evaluations of text generation experiments. Since no reliable metric for questions is available, only human evaluation was conducted for questions. Evaluation results showed that AI-generated materials were proper for students and even exceeded existing manually developed materials in quality.

2.2.2. ChatGPT for English writing. Kim et al. center on assessing ChatGPT's potential as a supportive tool in English writing [15]. The research process was divided into two steps by testing ChatGPT to complete two tasks. The first task is to design course content for second language learners and the second task is to teach students by using the Task-Based Language Teaching (TBLT) method. The research steps and corresponding test process can be found in Figure 1.

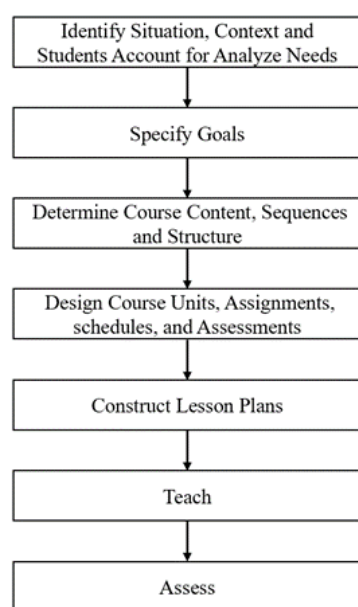


Figure 1. Research steps of course content design.

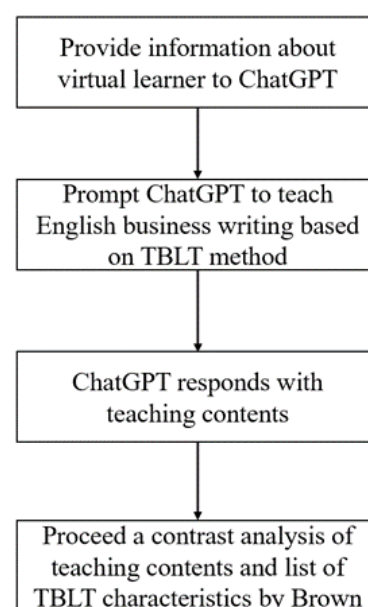


Figure 2. The process of testing ChatGPT by using TBLT.

In terms of Figure 1, at the very beginning, a virtual persona was built to provide its backgrounds and needs to ChatGPT. The figure was defined as a Korean university student whose goal was to develop business English writing skills. Initially, ChatGPT was asked to specify learning objectives based on the information. Subsequently, it was required to set the course framework. Furthermore, it was requested to design specific unit plans for the topic of ‘Writing Effective Emails’, which included assignments and assessments. And then, ChatGPT was asked again to establish lesson plans and teach

common styles and formats that were appropriate for professional email writing. Ultimately, a written work with including wrong answers deliberately was submitted to test its evaluation ability. Looking at the answers given by ChatGPT in every step, every response strictly followed the virtual persona's instruction. In conclusion, the course designed by ChatGPT included several necessary aspects, including relative knowledge, sensible examples, proper exercises, and effective methods.

In terms of task two, TBLT is regarded as an essential method in language learning because it encourages the development of communication skills and enables students to learn within the context. So ChatGPT's effectiveness in teaching with TBLT was assessed by comparing characteristics of the original TBLT method with the instructions generated by ChatGPT. The assessing process is that ChatGPT was required to teach English business writing based on the TBLT for a redirected virtual persona scenario, which is shown in Figure 2. During the process of comparison analysis, it was illustrated that ChatGPT could utilize the TBLT method to improve students' language skills. What's more, several drawbacks were also mentioned, which included an absence of active interaction, overly simplistic learning content, and limited capability of assessing. This work not only investigated the advantages of ChatGP for English writing but also mentioned that more thorough and methodical research was still required to demonstrate its actual learning impact. Therefore, more studies were advised to properly comprehend ChatGPT's potential for language learning.

2.2.3. ChatGPT for multiple language skills. Vera's work also investigated the effectiveness of using ChatGPT as a tool for learning English for undergraduate students, but it focused on improving various skills including listening, speaking, reading, and writing [16]. 35 participants were randomly selected, then divided into two groups. ChatGPT-based instruction was assigned to people in the experimental group, and the control group received traditional classroom instruction. ChatGPT was required to provide personalized and interactive instruction based on a theoretical framework in Figure 3. To control the variable, both groups received the same instruction during a period of six weeks, but the experimental group received additional ChatGPT-based instruction for 30 minutes per day, five days a week. The control group received traditional classroom instruction for the same amount of time. A pre-test was administered before the instruction began and a post-test was given after receiving instructions to assess the participants' language proficiency. These tests consisted of reading, writing, listening, and speaking sections. Besides, both groups were given a questionnaire to determine how they felt about the instruction. Both open-ended and Likert-scale questions were included in the survey. To assess the pre-and post-test data, independent samples t-tests were used to find out whether the remarkable differences between the two different groups existed. Descriptive statistics were used to analyze the survey data. The study's analysis showed that the performance of the experimental group was much better than the control group on the test in all four language competency domains (reading, writing, listening, and speaking) by a significant margin.

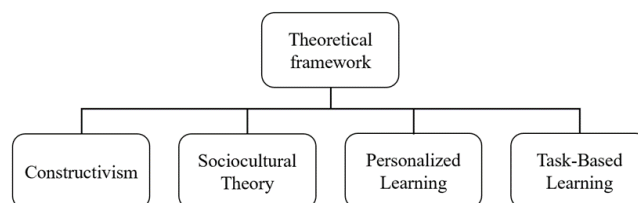


Figure 3. A theoretical framework for instruction generation.

3. Applications and discussion

3.1. Language skills improvement

Some studies have demonstrated that ChatGPT can serve as a valuable tool for users to improve their various language skills. For instance, ChatGPT can be used to refine writing and improve grammatical

accuracy since it has a feature of writing assessment. Based on predetermined criteria, the system rates the written work of users and offers detailed justifications for the score. Therefore, when prompted to give feedback, the tool will highlight the text's advantages as well as drawbacks and make suggestions for how to make the work better. To illustrate, by prompting the tool to edit a text, it can provide corrections on vocabulary, syntax, and grammatical accuracy [17].

Though its great potential as a tool for writing, ChatGPT has some limitations. The developer recognizes that it runs the risk of generating inaccurate and incomprehensible responses. Another limitation is ChatGPT's use of unnecessary statements in the generated text, pointing out it has no personal view as an AI tool. In addition, ChatGPT's capabilities may be constrained while looking for potentially plagiarized texts and tailoring the text to a particular audience. Moreover, ChatGPT is also observed to have some degree of rigorous template usage, which means ChatGPT tends to follow a specific structure even if writing essays about different topics [18]. Since ChatGPT has been improved gradually, the differences between human-written essays and AI-generated essays are harder to tell. Hence, more studies will be needed to cope with these problems to assure academic integrity.

3.2. Cultural understanding

Learning a language entail not only becoming fluent in the language but also becoming aware of the culture it belongs to. ChatGPT can shed light on cultural features including idioms, traditions, or social conventions. Students can get answers to their queries concerning cultural customs and practices, which will help them comprehend the language and its cultural context more fully. The cultural congruence and consistency of the outcomes produced by ChatGPT have been evaluated in a paper. The investigation demonstrates that ChatGPT can more effectively reflect American culture, most likely as a result of a large amount of English training corpus [19]. Therefore, it has the potential to help non-native English speakers to learn about American culture.

However, a cultural adaptation gap between ChatGPT and human civilization has also been discovered in the examined questions. Although ChatGPT can quickly adapt itself to remain socio-culturally congruent, which means it can give different answers when asked the same questions in different languages, it sometimes still tends to accept the injected knowledge which is not real. Also, a massive scope of English training corpus can influence ChatGPT's cultural adaptation to other languages. Therefore, it is sensible that ChatGPT should be trained more by using other languages corpus to enhance its cultural adaptation.

3.3. Mathematics

Some students currently learning English may also need instructions for the standardized tests which may be a compulsory part of application. These tests also include subjects besides English proficiency tests, such as mathematics on SAT or GRE. It is widely accepted that these tests are less demanding, but they are still challenging for those students without a solid foundation in mathematics [6]. Simon et al. tested two iterations of ChatGPT and GPT-4 on publicly accessible datasets as well as manually created ones to examine their mathematical skills. The research demonstrated that ChatGPT functions best as a mathematical assistant for fact-checking queries, serving as a mathematical search engine and knowledge base interface [20].

However, their overall mathematical performance that is well below the level of a graduate student has also been identified. ChatGPT was generally successful in activities that merely involved completing blanks or reciting mathematical facts. However, it does notably poorly on more challenging symbolic computations as well as proof-based questions in the format of graduate-level exercises or mathematical Olympiads. In addition, any kind of uncertainty was hardly conveyed by ChatGPT even if its answers were wrong. Due to the remarkable progress of GPT-4's performance in GRE quantitative compared to GPT-3, it is believed that ChatGPT will have better mathematical skills with the release of a new edition.

4. Conclusion

This review explores the feasibility and potential benefits of ChatGPT in foreign language learning. It starts with illustrating the significant support provided by AI, including ChatGPT, for foreign language learning, but also names the controversies that ChatGPT has stirred up, such as the impact on academic integrity. Although the risks have been acknowledged, some studies still demonstrate that ChatGPT may serve as an effective tool to improve English proficiency through various methods. Furthermore, this review also discusses the other potential applications of ChatGPT and its limitations. Moving forward, as AI will continue to evolve and be used in language learning, both educators and learners must make use of these tools correctly. In addition, future work could focus on better understanding the capabilities of AI and controlling its shortcomings and risks.

References

- [1] Wang H Smyth R and Cheng Z 2017 The economic returns to proficiency in English in China *China Economic Review* 43 p 91-104
- [2] De la Vall R R F and Araya F G 2023 Exploring the benefits and challenges of AI-language learning tools *Int. J. Soc. Sci. Humanit. Invent* 10 p 7569-7576
- [3] Ruan S Jiang L Xu Q et al 2021 Englishbot: An ai-powered conversational system for second language learning 26th international conference on intelligent user interfaces p 434-444
- [4] Openai 2023 ChatGPT <https://openai.com/blog/chatgpt>
- [5] Moqbel M S S and Al-Kadi A M T 2023 Foreign language learning assessment in the age of ChatGPT: A theoretical account *Journal of English Studies in Arabia Felix* 2(1) p 71-84
- [6] Sakai N 2023 Investigating the Feasibility of ChatGPT for Personalized English Language Learning: A Case Study on its Applicability to Japanese Students
- [7] Kohnke L Moorhouse B L and Zou D 2023 ChatGPT for language teaching and learning *RELC Journal* p 00336882231162868
- [8] Baskara F R and Mukarto F X 2023 Exploring the Implications of ChatGPT for Language Learning in Higher Education *IJELTAL (Indonesian Journal of English Language Teaching and Applied Linguistics)* 7(2) p 343-358
- [9] Truong H 2023 Role of ChatGPT in Vietnamese Education
- [10] Susnjak T 2022 ChatGPT: The end of online exam integrity? *arXiv preprint arXiv* p 09292
- [11] Tlili A Shehata B Adarkwah M A et al What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education *Smart Learning Environments* 10(1) p 15
- [12] Ouyang L Wu J Jiang X et al 2022 Training language models to follow instructions with human feedback *Advances in Neural Information Processing Systems* 35 p 27730-27744
- [13] OpenAI R 2023 GPT-4 technical report *arXiv* p 2303.08774
- [14] Xiao C Xu S X Zhang K et al 2023 Evaluating Reading Comprehension Exercises Generated by LLMs: A Showcase of ChatGPT in Education Applications *Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2023)* p 610-625
- [15] Kim S Shim J and Shim J 2023 A Study on the Utilization of OpenAI ChatGPT as a Second Language Learning Tool *Journal of Multimedia Information System* 10(1) p 79-88
- [16] Vera F 2023 Enhancing English language learning in undergraduate students using ChatGPT: A quasi-experimental study *CIAA* p 18
- [17] Klyshbekova M 2023 ChatGPT and Language Learning Available at SSRN p 4488587
- [18] Barrot J Sn 2023 Using ChatGPT for second language writing: Pitfalls and potentials *Assessing Writing* 57 p 100745
- [19] Cao Y Zhou L Lee S et al 2023 Assessing cross-cultural alignment between ChatGPT and human societies: An empirical study *arXiv preprint arXiv* p 2303.17466
- [20] Frieder S Pinchetti L Griffiths R R et al 2023 Mathematical capabilities of ChatGPT *arXiv preprint arXiv* p 2301.13867