The Role of AI in the Cross-Context Communication of Chinese Ancient Poetry Culture--A Comparative Study of ChatGPT & Deepseek's English Translation of Poetry

Xinyi Huang

College of Humanities, Shenzhen University, Shenzhen, China
2022010153@email.szu.edu.cn

Abstract: Under accelerated globalization and frequent cross-cultural exchanges, the English translation of Chinese poetry is of great significance to the dissemination of traditional culture. By comparing the English translations of Chinese ancient poetry by ChatGPT and Deepseek, this study evaluates the cross-contextual communication of AI from the five dimensions. The results show that AI plays a certain role in the cross-context communication of Chinese poetry culture. But simultaneously, the data show that different AIs have different effects on the auxiliary translation of ancient poetry. This means that AI can more accurately convey the cultural connotation of poetry and optimize the effect of cultural transmission by reasonably setting the prompt words. This proves that AI has great potential in promoting the cross-context communication of Chinese ancient poetry culture. Therefore, this paper puts forward the following suggestions: AI developers should tap the potential of models, optimize algorithms, and improve AI's ability to deal with complex cultural connotations and language styles.

Keywords: Machine Translation, Artificial Intelligence, ChatGPT, Deepseek

1. Introduction

With the accelerating process of globalization, cross-cultural exchanges are becoming more and more frequent, and the need for Chinese traditional culture to go global is increasingly urgent. Chinese poetry contains rich historical, philosophical, and literary connotations, and is an important carrier of ancient Chinese culture. However, due to the unique grammatical structure, lexical system, and special cultural connotation of ancient Chinese prose, the English translation of Chinese poetry has always been a major problem in the field of translation [1]. With the development of artificial intelligence technology, the rise of large language models (LLM) has created new possibilities for the English translation of ancient Chinese poetry.

Machine translation (MT) is a translation method that automatically translates text data from one language to another with the help of computers. As one of the key research directions of artificial intelligence, machine translation has attracted high attention in academia and industry [2]. With the development of artificial intelligence technology, many scholars have carried out research on machine translation based on large language models. At present, the research objects of Chinese-English translation include informative, expressive, and operational texts, sentences with cultural elements in articles on people's daily online activities, professional texts in the field of science, and texts related

to the MTI translation course teaching etc. Studies related to Chinese poetry have also focused on the comparison of poetry generation machine translation and human translation, the comparison of traditional translation machine translation and ChatGPT translation, and the construction of data sets and new models. There are few comparative studies on the translation of Chinese classical poetry with different AI. To a large extent, the study of AI translation of Chinese poetry into English is still in a state of exploration.

ChatGPT, developed by OpenAI, has always been a widely concerned language model. It is based on the GPT architecture [3]. After training, it can follow the instructions in the prompt and provide detailed responses. This ability to process natural language is also applicable to machine translation [4]. Deepseek, as a newly emerging large model in China, has performed well in many fields. Some studies have proposed that it has the unique advantages of combining emotion and technology skillfully and being more humanized, and its Chinese processing ability is stronger than that of OpenAI, which is more suitable for Chinese users [5,6].

This study compares the English translation of ancient Chinese poetry by ChatGPT4o and Deepseek. At the same time, it uses different types of prompts to study and explore the effect of different prompts in the English translation of Chinese poetry. It evaluates the translation from five dimensions: language accuracy, fluency, language style, machine translation sense, and overall evaluation, and is committed to filling the gap in the comparative study of different AI software and translation methods, providing case support for improving the quality of English translation of ancient Chinese poetry, and helping the spread of Chinese traditional culture in the world.

2. Method

2.1. Text selection

This study selected 10 poems in the "old style poetry" column of the China Poetry network. The selected Vintage poems will be released in February 2025 to ensure that these poems are not pre-trained by ChatGPT and Deepseek and are random. Poetry consists of seven syllables, four lines, and seven rhythms. The seven-syllable four-line poem is flexible and free in form, while the seven-syllable temperament is rigorous and neat in antithesis. Different genres can provide diverse samples for the study of different AI translation effects.

2.2. Prompt word setting

Wu et al. designed three prompt schemes. Compared with other schemes, ChatGPT has the best translation performance when it requires direct translation of sentences and does not allow the addition of P1 prompts for interpretation [7]. The research of Jiao, W., Wang, W. et al. also shows that direct and concise translation instructions are more referential in evaluating the basic translation ability of the model [4].

Gao et al. in Jiao, W. Based on their research, et al. Proposed that translation tips, including POS tags (TT POS, CD POS), have a positive effect on some translation directions [8]. The POS (part of speech) tag is added to the prompt as auxiliary information, which mainly means that in the process of interaction with ChatGPT, the part of speech information of words is added to the prompt to help ChatGPT better understand and process the input text. Chinese poetry has a unique grammatical structure, so part-of-speech judgment is very important for poetry translation.

Peng, K., Ding, L., et al. proposed task-specific cue words (TSP) and domain-specific cue words (DSP) [9]. The research emphasizes that task information prompts can improve translation performance in complex tasks; The introduction of correct domain information can improve the translation performance of ChatGPT.

Based on the above research on the cue words, this study designed the following three kinds of cue words:

BT: basic translation task: please provide the translation from Chinese to English for the following materials: [specific ancient poetry sentences].

DSC: domain-specific and cultural notation: you are a translator in the field of translating ancient Chinese culture. Please provide English translations of these Chinese poetry sentences, paying attention to the cultural details, grammatical structure, and special meaning in the sentences: [specific ancient poetry sentences].

Part of speech tag assistance (PTA): Please provide English translation for these materials with the help of the given part of speech tag: [specific ancient poetry sentences], and the part of speech tag is as follows: [list specific part of speech tags].

2.3. Translation evaluation perspective

The study will comprehensively evaluate translation from five perspectives, including accuracy, fluency, language style, naturalness, and overall evaluation. In the evaluation system of machine translation, accuracy and fluency have long been important standards in early machine translation evaluation [10]. It is also the evaluation angle contained in the human evaluation criteria in most studies [11,12]. Accuracy measures the level of accuracy in conveying the original meaning of the original text based on presenting the key information completely and without distorting it. Fluency mainly measures the degree to which the translation conforms to the rules and norms of the target language, focusing on the smoothness, coherence, and naturalness of the translation.

In literary translation, language style is also a key consideration [13]. This dimension evaluates the adaptability and continuity of the language style of the source text, including mood, intonation, rhythm, poetic expression, format, rhyme, and cultural details.

Some scholars pointed out that there is a lack of "humanization" in AI translation, which will make the translated text give people a sense of alienation [14]. Therefore, this study also takes translation naturalness as a criterion for evaluation.

Finally, referring to the evaluation framework of human poetry proposed by Wang et al., the evaluation standard of "overall evaluation" is added to evaluate whether the translation successfully captures the essence and tone of the original [15].

In this experiment, five professionals with a master's or doctorate in English translation and native Chinese were invited to form an evaluation team. The translated content is displayed on the Wenjuanxing online survey platform. Using the 5-point Likert scale, the evaluators scored each translation result according to the criteria of each dimension.

3. Results

3.1. AI plays a positive role in the dissemination of chinese culture

To explore whether ChatGPT and Deepseek have differences in translation performance, this study uses the linear mixed effect regression method of the lme4 package in R language to analyze the data by referring to the data analysis method of Gao et al. In the analysis, the translator (ChatGPT/Deepseek) is set as a fixed effect, the rater and poetry are taken as random effects, the scores of each scoring standard are taken as dependent variables for analysis, and the significance is calculated with the help of the lmertest package.

From the performance of Deepseek and ChatGPT's English translation of ancient poems, both can play a positive role in the external dissemination of Chinese culture.

From the perspective of language accuracy, ChatGPT score was 3.65, and Deepseek score was 3.74, with no significant difference (β=0.0875, SE=0.09277, t=0.943, DF=79, P=0.3484). This means

that both AI can accurately grasp the key information of ancient Chinese poetry and convey the core semantics to the target audience, laying a solid foundation for the dissemination of ancient Chinese poetry culture.

In terms of fluency, ChatGPT scored 3.40 points, and Deepseek scored 3.77 points, with no significant difference (β=0.375, SE=0.243, t=1.546, DF=7, P=0.1660). Their translations are similar in following the grammar rules and expression habits of the target language, which can ensure that the translated readers can smoothly understand the content of poetry, reduce the difficulties in cultural understanding caused by language barriers, and promote the widespread spread of Chinese ancient poetry culture in the world.

In the dimension of naturalness, ChatGPT scored 3.48 points, and Deepseek scored 3.88 points, with no significant difference (β=0.4, SE=0.237, t=1.689, DF=7, P=0.1350). This shows that the two AIs have a good performance in translating, getting rid of the stiff feeling of machine translation and closer to human natural expression, which can enhance the applicability of Chinese ancient poetry translation in literary scenes and enhance the affinity of cultural transmission.

3.2. Automation transformation of AI technology in cultural communication

This study also studied the differences in translation performance of Deepseek under different cue words and also used the linear mixed effects regression of the LME package in R language to analyze the data. At this time, the fixed effect is different cue words (Deepseek1/Deepseek2/Deepseek3), and the random effect is the rater and poetry. The score of each scoring standard is used as the dependent variable to analyze, and the significance calculation is also carried out with the help of the lmertest package.

A study of the effects of different cue words on the translation performance of Deepseek shows that there are no significant differences in the translation performance based on basic translation task cue (BT), domain-specific and cultural connotation cue (DSC) and part of speech tag auxiliary cue (PTA) in terms of language accuracy, language style, fluency and naturalness. In terms of language accuracy, there was no significant difference between Deepseek1 (BT prompt word) and Deepseek2 (DSC prompt word), Deepseek1 and Deepseek3 (PTA prompt word), Deepseek2 and Deepseek3; The same is true of language style, fluency, and naturalness. This shows that Deepseek can better understand the semantics of the original poetry and translate it relatively accurately no matter what kind of cue words are used.

Data analysis found that only in the overall evaluation, there were significant differences in the translation of different cue words. The score of Deepseek1 was significantly lower than that of Deepseek2 (3.71 vs. 3.94, β=-0.2250, SE=0.0795, t=-2.830, DF=158, P=0.0145), that is, the translation performance of DSC was significantly better than that of BT. This shows that specific cue words can guide Deepseek to better grasp the overall artistic conception and cultural connotation of poetry to improve the overall translation quality.

Compared with previous studies in which more detailed domain or part-of-speech tagging cues have a greater impact on ChatGPT translation, Deepseek is relatively less affected by different cues. This may be because Deepseek can quickly integrate and absorb the latest information for "deep thinking" in the real-time networking state [16]. When researchers give it the simplest clue word, it can carry out in-depth thinking and analysis. In this study, Deepseek's logical chain of "deep thinking" on BT prompts has self-prompted specific fields and cultural connotations, and also conducted part of speech analysis on keywords. Therefore, different cue words may have less influence on the translation effect of Deepseek. This shows that with the continuous improvement of technology, the degree of automation of AI is gradually improving, and the effect of human intervention on translation results is declining. Deepseek may be able to maintain a stable translation level under different

prompts by its strong learning and understanding ability, reducing the dependence on specific prompts.

3.3. English translation differences between Deepseek and ChatGPT

Deepseek and ChatGPT are similar in multiple evaluation dimensions, but there are significant differences in language style (β=0.237, SE=0.104, t=2.283, DF=79, P=0.0251). After the data analysis, this study interviewed some evaluators. Combined with the translation results, this difference is reflected in the use of words, sentence patterns, and image construction.

In terms of wording, taking "I would like to be a wildflower whose fragrance fills my sleeve" as an example, ChatGPT translates it as "I wish to be a wildflower, its fragrance filling my sleeves", which is relatively straightforward and plain, and can clearly convey the basic semantics. Deepseek translated this sentence as "Let wild perfume drench sleeves in twilight's hue" and used the word "drench" to emphasize the intensity and penetration of fragrance compared with "fill". "twilight's hue" creates a poetic atmosphere with the color of twilight, making the translation more literary.

In the sentence patterns, when translating "Bucket handle back Yin New Year order, spirit snake plate Rui trillion double spring", the sentence patterns of ChatGPT translation are relatively regular, and the expression is stable and direct. While Deepseek's translation "hopper's handle returns to tiger's celestial sphere, serpent coils imperial, twin springs now appear." combines the length of sentences, such as the expression of "tiger's celestial sphere", which is full of imagination, making the rhythm of sentences more varied and rhythmic.

In terms of image construction, Deepseek used expressions such as "cradle", "Silk mist" and "Snow - Heron stands etched" to vividly depict the relationship between the stream and the golden carp, and between the cold gauze and the snow goose, creating a hazy and poetic picture, skillfully integrating the image of Chinese classical culture. In contrast, ChatGPT may prefer to convert images in a straightforward and concise way and pay attention to the accurate transmission of semantics.

In addition, there is the evaluator's feedback on the actual translation performance. Deepseek sometimes pays attention to rhyme and the expression of cultural images, which makes the translation difficult to understand. The translation of ChatGPT is relatively more concise and easy to understand, which is also an important manifestation of the differences in translation characteristics between the two.

4. Conclusion

This study focuses on ChatGPT and Deepseek's English translation of ancient Chinese poetry and analyzes it from multiple dimensions. The results are of great significance in the field of translation, the application of AI technology, and cultural communication.

From the perspective of translation performance, ChatGPT and Deepseek are similar in language accuracy, fluency, and naturalness, indicating that the current mainstream AI translation technology has been relatively reliable at the basic level of language conversion. They can accurately convey the key information of poetry, follow the rules of the target language, and make the translation have a certain degree of naturalness, which provides an effective way for the preliminary cross-language communication of ancient Chinese poetry. In terms of language style, Deepseek can create a more poetic atmosphere through the use of words, sentence patterns, and image construction, which is related to its possible combination of emotion and technology, and more humanized characteristics. This also shows that in the process of literary translation, it is important to choose AI tools that are more suitable for the style of literary creation to retain the charm of the original text.

As for the impact of different cue words on the translation performance of Deepseek, the research results show that, compared with ChatGPT, Deepseek is less affected by different cue words, which

may be because Deepseek has strong autonomous learning and understanding ability, as well as its characteristics of real-time online integration of information. This not only reflects the progress of AI technology but also implies that AI translation may reduce human intervention and achieve more automated and efficient translation in the future.

The current research also has some limitations. In terms of sample selection, only 10 old-style poems were selected, which makes it difficult to fully cover the rich genres, styles, and cultural connotations of ancient Chinese poetry. In addition, the evaluation subject is only composed of professionals with specific backgrounds, which may have limitations in the evaluation perspective. Future research can expand the sample range to cover ancient poems of different dynasties, genres, and themes, and enhance the universality of the research results. In addition, it can optimize the evaluation system, introduce more diverse evaluation subjects, such as native English readers and literary critics, and improve the evaluation dimensions and standards. It can also further explore new types of cue words, and make more targeted designs based on the specific cultural background and translation purpose of poetry to improve the research level and translation quality of AI translation of ancient Chinese poetry.

References

[1] Cheng, F. (2017). Chinese Poetic Writing. New York Review of Books, New York.

[2] Yang, S., Wang, Y., & Chu, X. (2020). A Survey of Deep Learning Techniques for Neural Machine Translation. arXiv preprint arXiv: 2002.07526.

[3] Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving Language Understanding by Generative Pre-Training.

[4] Jiao, W., Wang, W., Huang, J. T., Wang, X., & Tu, Z. (2023). Is ChatGPT a Good Translator? A Preliminary Study. arXiv preprint arXiv: 2301.08745, 1(10).

[5] Qin, S. (2025). Deepseek: Change and Thinking. China Automotive News, (002).

[6] Wei, Y., Jia, K., Zeng, R., He, Z., Qiu, L., Yu, W., & Jiang, Y. (2025). Artificial Intelligence Innovation, Development and Governance Reform under the Breakthrough Effect of Deepseek. E-government 1-38.

[7] Wu, Y., & Hu, G. (2023). Exploring Prompt Engineering with GPT Language Models for Document-Level Machine Translation: Insights and Findings. In Proceedings of the Eighth Conference on Machine Translation, 166-169.

[8] Gao, Y., Wang, R., & Hou, F. (2024). How to Design Translation Prompts for ChatGPT: An Empirical Study. In Proceedings of the 6th ACM International Conference on Multimedia in Asia Workshops, 1-7.

[9] Peng, K., Ding, L., Zhong, Q., Shen, L., Liu, X., Zhang, M., & Tao, D. (2023). Towards Making the Most of ChatGPT for Machine Translation. arXiv preprint arXiv: 2303. 13780.

[10] Freitag, M., Foster, G., Grangier, D., Ratnakar, V., Tan, Q., & Macherey, W. (2021). Experts, Errors, and Context: A Large-Scale Study of Human Evaluation for Machine Translation. Transactions of the Association for Computational Linguistics, 9, 1460-1474.

[11] Gao, R., Lin, Y., Zhao, N., & Cai, Z. G. (2024). Machine Translation of Chinese Classical Poetry: a Comparison among ChatGPT, Google Translate, and DeepL Translator. Humanities and Social Sciences Communications, 11(1), 1-10.

[12] Seljan, S., Dunđer, I., & Pavlovski, M. (2020, September). Human Quality Evaluation of Machine-Translated Poetry. In 2020, 43rd International Convention on Information, Communication and Electronic Technology (MIPRO), 1040-1045. IEEE.

[13] Hutchins, W. J., & Somers, H. L. (1992). An Introduction to Machine Translation. (No Title).

[14] Wang, H., & Liu, W. (2021). In the Era of Artificial Intelligence, Translation Technology Turns to Research. Foreign language teaching, (05), 87-92.

[15] Wang, S., Wong, D. F., Yao, J., & Chao, L. S. (2024). What is the Best Way for ChatGPT to Translate Poetry? arXiv preprint arXiv: 2406.03450.

[16] Dai, J., & Qin, Y. (2025). Research on the Development and Dissemination of Chinese Civilization Enabled by Deepseek Type Generative Artificial Intelligence. Journal of Chongqing University (Social Sciences Edition), 1-14.