

Evaluating Large Language Models in Classical Chinese Poetry Translation: A Case Study of Zang Hua Yin's English Versions

Sicheng Lu

Bozhou University, Bozhou, Anhui, 236800, China

ABSTRACT

This article translates the classical Chinese poem "Burying Flowers" into English using a Large Language Model. We compare it with a human-translated version, examining similarities and differences in translation methods and strategies. We generate suggestions regarding the translation aspects of the Large Language Model, explore its social value, and disseminate outstanding Chinese classical poetry.

KEYWORDS

Translation strategies; Large language Model; Chinese classic poems

1 Introduction

The rapid advancement in AI has spawned large language models (LLMs) like GPT, igniting a global LLM craze. Today, LLMs have achieved deep integration and vertical applications across numerous professional fields and linguistic scenarios. (Xie Wenjun, Guo Zelin, Zhang Yaxi 2025) Since its inception, humanity has continuously explored the linguistic capabilities of artificial intelligence. The emergence of large language models like DeepSeek, ChatGPT, and Wenxin Yiyan has sparked widespread societal discussions about their potential. Classical poetry translation has long faced numerous challenges, including linguistic conciseness, cultural context differences, and the complexity of poetic meter. Traditional translation methods often rely on human translators, resulting in low efficiency and frequent mistranslations. The integration of large language models can leverage their powerful language generation capabilities to provide more accurate and efficient translation solutions. "The Lament for the Fallen Flowers" stands as a quintessential example of classical Chinese poetry. Existing research primarily focuses on analyzing traditional translation methods and strategies, such as comparative studies of its English and French translations. These studies primarily explore strategies employed by different translators during the translation process, such as rhythmic adaptation, imagery transformation, and cultural context reproduction. However, none of these studies mention the application of large language models in the translation process. Therefore, it can be inferred that current translation efforts for "Ode to Burying Flowers" still predominantly rely on traditional translation methods, with no involvement of large language models. This project will analyze textual segments from English translations of "The Lament for the Fallen Flowers" (translated by Yang Xianyi, Xu Yuanhong, and Lin Yutang) alongside translations produced by large language models such as DeepSeek, ChatGPT, and Wenxin Yiyuan. It will examine the translation strategies employed by both human translators and large language models, evaluate their effectiveness, generate translation recommendations leveraging large language models, and explore the societal implications of these models.

2 Large Language Model Translation

2.1 DeepSeek Translation

By the earth's far side at last,
What better sepulchre could be amassed?
Then I'd take a silken bag, and your pure soul hold,
And bury you in clean earth, with a mold.
Now you are pure, when you are underground;
For you were never to the filth in the gutter bound.
So you are now beneath the earth confined,
While I, on my bare shoulders, am left behind.
I tell you, soul of the flower, do not grieve;
I am here, for a little while, to mourn and leave.
Alas! I cannot hold the winged hours fast;
No one can curb the river of time flowing past.

Strange, that I should be the one to mourn,
And come to weep o'er your grave, forlorn.

DeepSeek achieves efficient fusion of dynamic knowledge through streaming data processing, time-series awareness mechanisms, and conflict resolution strategies, breaking through the linear patterns of traditional knowledge transmission. For instance, under "connected mode" search, DeepSeek can deeply read vast web pages to generate comprehensive, accurate knowledge tailored to individual needs, while integrating the latest knowledge and information in real time to better adapt to evolving environments and demands (Zhao Yanan 2025). DeepSeek employs ABAB and AABB rhyming schemes while processing complex imagery throughout the text—such as "willow threads and elm pods" or "swallows between the beams"—to find suitable English equivalents. For culturally specific terms like "maiden in seclusion" or "edge of the sky," it balances fidelity with readability. For instance, "sweet sepulchre" captures both the "fragrant" imagery and the tomb connotation of "sepulchre." Notably, DeepSeek adopts the author's perspective, frequently using subject-driven expressions—a translation approach increasingly resembling human cognition.

2.2 ChatGPT Translation

The flowers fall, the water carries the red petals,
Countless worries stir in my heart, yet I remain silent, blaming the east wind.
Who could have known, as we parted, that my clear dreams would be so hard to find?
Where could they be?
By the water's edge, in dreams, a gentle breeze brushes softly.
The flowers wither, they fly and fill the sky,
As the redness fades and the fragrance vanishes, who will pity them?
Soft threads of silk drift and settle on the spring pavilion,
The falling petals lightly brush against the embroidered curtain.

ChatGPT's English translation strives to preserve the original imagery. For instance, "花落水流红" is rendered as "The flowers fall, the water carries the red petals," conveying both the visual of petals drifting and evoking a sense of flow and helplessness for English readers. Similarly, "游丝软系飘春榭" is rendered as "Soft threads of silk drift and settle on the spring pavilion," evoking the image of delicate threads floating in the air. Furthermore, since "The Lament for the Fallen Flowers" is poetry, the translation not only prioritizes fidelity to the original meaning but also strives to preserve a certain rhythm and aesthetic appeal in the English version. For instance, the line "Who could know, as we part, that my clear dreams will be so hard to find?" does not follow a strict literal translation but effectively conveys Lin Daiyu's helplessness at parting and her longing for beautiful dreams.

2.3 Literary Heart in a Word Translation

The spring wind blows, and the flowers die,
The summer rain comes, but can't revive.
The autumn frost bites with icy bite,
And the winter snow covers all in white.

Wenxin Yiyan prioritizes creating an overall poetic atmosphere, endowing the translation with a sense of tragic beauty. For instance, the line "Behold the fading spring as flowers fall, for this is when beauty withers and dies" is rendered as: "The spring wind blows, and the flowers die, The summer rain comes, but can't revive. The autumn frost bites with icy bite, And the winter snow covers all in white." This concise yet expressive language vividly portrays the poignant scene of fading spring and fading beauty. As a general-purpose large model, Wenxin Yiyan excels in open-domain text generation and fuzzy semantic processing. (Pan Yuting, Ke Ping 2025)

3 Human Translation

3.1 Yang Xianyi Translation

Fly, Fly, ye faded and broken dream,
Of fragrance, cared for nevermore!
Behold the gossamer entwine the spring flowers,
And the catkins kiss the curtained door.
Here comes the maiden from out her chamber door.
Whose secret no one shall share.
With hoe and basket she treads the fallen blossoms,

And wanders back and forth in prayer.
I smell the scent of elm seeds and the willow.
Where once did blush the peach and pear.

Mr. Lin Yutang's translation theory emphasized "faithfulness," "fluency," and "beauty." In translating "Ode to the Fallen Flowers," he remained faithful to the original text. For instance, the opening employs two instances of "FLY" to correspond with "花谢花飞" (flowers fall and fly), mirroring the rhythm and style of the original poem. The phrases "broken dream" and "of fragrance" aptly convey the original's "红消香断" (red fades, fragrance breaks), capturing the helplessness of spring's impending end. (Huang Ningxia, Yang Ping 2010). In achieving fidelity, fluency, and beauty, Lin Yu-Tang approached the translation from the source text's content, ensuring the recitation flows smoothly while preserving the original's essential hues. His translation captures the poem's rhythmic beauty, cadence, and poetic imagery. Though faithful to linguistic conversion, the original remains inaccessible to foreign readers unfamiliar with Chinese characters. The human-translated English version (Sun Yiqun) handles this challenge with remarkable ingenuity.

3.2 Yang Xianyi's Translation

Alone, her hoe in hand, her secret tears
Falling like drops of blood one by one bare bough.
Dusk falls and the cuckoo is silent;
Her hoe brought back, the lodge is locked and still;
A green lamp lights the wall as sleep enfolds her,
Cold rain pelts the casement and her quilt is chill
What causes my two-fold anguish?
Love for spring and resentment of spring;

Mr. Yang Xianyi's translation theory prioritizes utmost fidelity to the original text. When translating a classic like *Dream of the Red Chamber*, he adopted a domestication strategy. For a classic poem like "Burying Flowers," drawn from a renowned Chinese novel, he deeply understood its cultural context and thus emphasized its cultural characteristics in translation. For instance, "Burying Flowers" contains two allusions: One appears in line 24: "Falling like drops of blood on each bare bough." The other is in line 25: "Dusk falls and the cuckoo is silent." (Li Rongrong, 2010) It was precisely Mr. Yang Xianyi's profound understanding of Chinese culture that enabled his unique English translations of these two allusions.

3.3 Xu Yuanchong's Translation

O in the three hundred and sixty days each year.
The cutting wind and biting frost make flowers sear!
How long can their fragrant blossoms last fresh and fair?
Once blown away, they can be found nowhere. It's harder to find fallen blossoms than those in bloom;
Before the steps their grave-digger is filled with gloom.
Alone with hoe in hand, my tears secretly shed,

Mr. Xu Yuanchong's "Three Beauties Theory" is his own translation philosophy, inheriting from Wen Yiduo's New Poetry Movement principles of "musical beauty, pictorial beauty, and architectural beauty." It blends Western truth-seeking with Chinese aesthetic traditions: - Poetic Beauty: Preserving the original's imagery and philosophical depth; - Rhythmic Beauty: Recreating the poem's cadence; - Structural Beauty: Mirroring the source's sentence patterns. For instance, "The Lament for the Fallen Flowers" brims with "aesthetic beauty" of sorrow, grief, desolation, and melancholic splendor. The text conveying the heroine's resentment and yearning is fully rendered in the translation. The line "O in the three hundred and sixty days each year. The cutting wind and biting frost make flowers sear!" The original "wind knives" and "frost swords" clearly employ metaphorical imagery. While translating them as "the knives of wind and swords of frost" might seem more "faithful" to the original, a problem arises: would such a "faithful" translation be acceptable to English readers? Does it align with their linguistic habits? Such parallel imagery not only makes the sentence cumbersome and lifeless but also strikes English readers as stiff and peculiar. In fact, David Hawes' translation of the same line employs only one metaphor—"swords of frost"—while rendering the wind knife as "slaughtering gale." This approach somewhat diminishes the inherent formal beauty of the language, falling short of the clarity and expressive power achieved by the concise, equally potent pair of -ing verbs in "cutting wind and biting frost." (Kong Xiangli, 2008) Mr. Xu Yuanchong, while preserving Chinese cultural elements, applied the "Three Beauties Theory" to this poem, endowing it not only with aesthetic appeal but also with the unique charm of Chinese culture.

4 Comparison Between Large Language Models and Human Translation

Translations of "The Lament for the Fallen Flowers" by three large language models—DeepSeek, Wenxin Yiyuan, and ChatGPT—reveal that while these models bear traces of human translation, they are also deepening their own interpretations by mimicking human translation habits, thought processes, and approaches. However, compared to human translation, it is evident that when handling literary texts—especially poetry—LLMs tend to produce translations that lean toward singular poetic construction. They lack the holistic consideration of human translators and the support of unique translation theories. Consequently, the core of their translations often leans toward mass appeal, lacking distinctive translation characteristics. Human translation, having emerged earlier than large language models, has developed distinctive translation theories and produced numerous classic works, thereby expanding dissemination channels for cultural creations. However, human translation still faces significant time constraints compared to large language models, and the historical examples it draws upon are relatively dated. Unlike large language models, it cannot keep pace with evolving translation practices or adapt to the latest cultural interpretations, thereby failing to meet the demands of cultural development. More critically, there is concern about potential bias in the transmission of cultural ideas through large language models. Within the context of semiotics, "thought" and "expression" correspond to the signified and signifier, respectively. Thus, both are fluid entities. When certain content becomes frequently used and difficult to avoid, it cannot be privatized. Its frequent usage increases the likelihood of its transformation into a thought. (Yao Ye 2025)

5 Conclusion

Against the backdrop of large language model development, their translation capabilities continue to advance with the times. However, it is also evident that compared to human translation, large language models lack their own cognitive framework—what we refer to as "understanding the changes throughout history and forming one's own unique perspective." Their mechanical translation approach results in a lack of understanding regarding the "humanity" inherent in translation. Instead, they mechanically reference translation theories from various translators and broadly synthesize them. None of these models demonstrate their own unique translation theories. This highlights that the mechanical translation of large language models functions not as a "human" in cultural exchange but as a "tool." Nevertheless, their summarization capabilities prove to be of paramount importance. When confronted with obscure, elusive language that can only be grasped intuitively, they can collect vast amounts of data, synthesize it, and deliver relatively satisfactory results.

This paper's limitations lie in selecting only three of the most widely used large language models, failing to broadly consider the potential of other large language models for translation. It also lacks deeper textual analysis: the chosen text is a classical Chinese poem, and the large language models were not sufficiently trained on such texts. Furthermore, the comparison with human translation only highlights the humanities aspect, leaving other dimensions unexplored.

References

- [1] Qianxi Li, Bao Pang, Yong Song, Hongze Fu, Qingyang Xu, Xianfeng Yuan... & Chengjin Zhang. (2026). Large language model assisted hierarchical reinforcement learning training. *Information Sciences*, 723, 122688-122688.
- [2] Zhou Zisen, Huang Qi, Tan Zehong, Liu Rui, Cao Ziheng, Mu Fangman... & Qin Shaozheng. Emotional Competence Assessment in Dynamic Social Interaction Scenarios Using Multimodal Large Language Models. *Acta Psychologica Sinica*, 1-24.
- [3] Chen Guoliang, Zhang Lin & Gou Zhenyu. Research on an Intelligent Generation Framework for Peer Review Elements Based on Large Language Models. *Theory and Practice of Information Science*, 1-12.
- [4] Giulia Polverini & Bor Gregorcic. (2025). Multimodal large language models and physics visual tasks: comparative analysis of performance and costs. *European Journal of Physics*, 46(5), 055708-055708.
- [5] Chen Han, Xu Zefeng, Jiang Jiu, Fan Fan, Zhang Junjian, He Chu & Wang Wenwei. Automatic Diagnosis of Cognitive Assessment Scales Based on Large Language Models and Deep Networks. *Computer Science*, 1-15.
- [6] Cui, L. F., Wang, Z. S., Bao, Y. X., & Zhao, H. (2025). A review of automatic question-answering systems and evaluation frameworks in the era of large language models. *Computer Engineering and Applications*, 1-17.
- [7] Jing Wang, Jiahua Zhang, Kaida Yang, Beth Bradshaw Ghavidel, Benyamin Khajetash, Abolfazl Sarikhani... & Meysam Tavakoli. (2025). Enhancing auto-contouring with large language model in high-dose rate brachytherapy for cervical cancers. *Medical Physics*, 52(10), e70034.
- [8] Maikel Leon. (2026). GPT-5 and open-weight large language models: Advances in reasoning, transparency, and control. *Information Systems*, 136, 102620-102620.
- [9] Peijin Li, Zhijian Hu, Yichen Shen, Xinyan Cheng & Mohannad Alhazmi. (2025). Short-term electricity load forecasting based on large language models and weighted external factor optimization. *Sustainable Energy Technologies and Assessments*, 82, 104449-104449.
- [10] Kibum Moon, Adam E. Green & Kostadin Kushlev. (2025). Homogenizing effect of large language models (LLMs) on creative diversity: An empirical comparison of human and ChatGPT writing. *Computers in Human Behavior: Artificial Humans*, 6, 100207-100207.