

Research on the Effectiveness of Large Language Models in the Translation of New Northeast Literature

Zhao Junzhe¹

¹School of Foreign Studies Liaoning University of International Business and Economics, Dalian, Liaoning, China

Correspondence: Zhao Junzhe, School of Foreign Studies Liaoning University of International Business and Economics, Dalian 116052, Liaoning, China.

Received: September 20, 2025 Accepted: October 12, 2025 Online Published: October 15, 2025

Abstract

As the models continue to update and develop, they have also impacted the translation industry. The translation quality of large models has been continuously improving, posing a significant challenge to traditional translation models. Large language models can handle the basic tasks of non-literary translation, but there is still insufficient exploration in the field of translation quality for literary translation. New Northeast Literature presents unique characteristics in vocabulary and narrative style, which will inevitably bring more difficulties to large model translation. This paper will use natural language processing technology to evaluate the effectiveness of machine translation and large language model translation.

Keywords: large language models, new northeast literature, translation effectiveness evaluation

1. Introduction

Currently, the effectiveness of using large language models for literary translation is not ideal, and the background and starting point also increase the difficulty of processing and translating the source language with large language models. Large language models rely on pre-trained data. This paper will use large language models to conduct experiments on the effectiveness of machine translation quality for New Northeast Literature, comparing human translations with large language model translations horizontally, combining a mixed research to fully demonstrate the capabilities and potential of large language models in the field of New Northeast Literature.

2. Machine Translation

Machine Translation (MT) is an important branch of Natural Language Processing (NLP). With the rapid progress of Artificial Intelligence (AI) and neural machine translation services, translation speed has increased, translation costs have decreased, and translation services have become more convenient. In recent years, large language models (LLMs) have rapidly emerged, further amplifying the advantages of MT, and further validating the view that MT has immeasurable value in business communication, customer support, or low-source projects. [2] In contrast, others have pointed out strategies such as data sharing to further reduce mutual interference within multilingual neural translation models. [1] In the translation process, whether it is traditional neural network translation or LLMs, there are cases of meaningless translations or translations that are not faithful to the original text, especially in translations that heavily rely on cultural knowledge or humorous phrases. Additionally, in contexts with complex contexts or frequent shifts in discourse, MT struggles to maintain consistency in longer texts, which is particularly evident in literary translation. [5] Meanwhile, to measure the effectiveness of machine translation, NLP researchers have adopted various evaluation metrics, including BLEU, TER, METEOR, and BERTScore. However, at the nuanced cultural level, it is still difficult to capture such advantages of human translations. [4] Typically, the effectiveness of MT translations still requires human judgment to evaluate. [3]

3. Research Methods and Experimental Design

3.1 Research Questions

This study explores the translation quality and effectiveness of large language models in the field of literary translation, focusing on the more specialized 'New Northeast Literature, aiming to investigate whether the current translation capabilities of large language models can replace human translation.

3.2 Selected Texts

Several works by writers from the New Northeast Literature group have already been translated into multiple languages. Shuang Xuetao's three novellas, *Moses on the Plain*, *The Aeronaut* and *Bright Hall*, have been compiled into a single collection titled *Rouge Street* and published in the spring of 2022.

In this study, the original text selected from *The Aeronaut* serves as the source language material for the large language model, while the comparison translation is selected from Jeremy Tiang's version of *Rouge Street*, specifically *The Aeronaut*. Through the analysis of the original text of *The Aeronaut*, the following characteristics can be observed:

The colloquialism and Northeastern dialect features of the language. The original text contains many colloquial words, loose sentence structures, frequent use of sentences without subjects, and Northeastern dialect phrases.

The text is rich in regional imagery, which is mostly concentrated in Tiexi District and Heping District of Shenyang, forming part of the urban impression, such as 'Red Flag Square' and 'Tiexi Department Store.'

The article is filled with a modern romantic aftertaste and Northeastern-style verbal humor. The novel unfolds its content in a 'nested' mode, revealing to the world the contradiction between the heavy, cruel social reality of workers against the backdrop of the 'layoff wave' and the dramatic nothingness of the story; the verbal humor is further highlighted.

3.3 Research Process

The three experimental paragraphs will be placed into four representative translation platforms: Youdao Translation and Google, ChatGPT and DeepSeek. Large language models rely on pre-trained data, and the more training iterations, the higher the translation quality. Paragraph translation is conducted simultaneously to prevent data quality from affecting subsequent results. At the same time, the prompt was consistently 'You are a professional English translator, and now you are required to translate three selected passages from the new Northeast literary work *The Aeronaut* into English. The passages are not related to each other,' to ensure the reliability and validity of the test.

After obtaining the translations, the translations from the four platforms and the human translation were collectively evaluated using Python for BLEU and TER calculations, focusing on surface-level importance, combined with BERTScore to calculate the semantic and literary quality of the translations, with all three used to assess the translation quality.

At the same time, considering that human review can more accurately capture aspects such as translation accuracy, style, and cultural adaptability, the study invited 5 senior translators to score the translations, with a full score of 10, where a higher score indicates greater translation effectiveness. During the process, the 5 translators were unaware of the specific sources of the translations, directly scoring the machine translations and human reference translations to avoid subjective bias affecting the experimental results.

4. Research Results

After using the above prompt to translate the three experimental paragraphs, the corresponding platform translations were obtained. As shown in the table below:

Table 2. Experimental Translation Statistics Table

Paragraph	Paragraph 1	Paragraph 2	Paragraph 3
Reference translation	Turn right at the first hutong on the east side of Tiebai, then turn right again when you see a shoe shop. Proceed up to the third floor of Block 2, where you will find a black security door marked 'Panpan'. Tiebai refers to the Tiexi Department Store, which is the center of Tiexi	She said, "That's okay. Pay me." After I gave her the money, she took it and put it into the top of her stocking. She said, "Your brother is sick now, the money will be used to buy him medicine. Fluose fine hydrochloride." I asked to see him, "Where is he?" Her response is a rejection: "Come on. He can never go back."	Bage Snooker Hall was not large, just a dozen snooker tables, but the soft light made the hall seem warm and spring-like. There are only a few consumers. The balls had been shaped into a triangle, shining like the precious antiques beneath the light. The boss was sitting in front of iMac, playing mahjong. He noticed me looking around, and then stood up. "Are you looking for someone?" "Yes.

	<p>District. I had been there before as a kid. Every Sunday, the place would be crowded with people. Opposite the store, it's a Xinhua Bookstore. Except for the books on the two public shelves, other books that you want to read or buy needed to be passed by the salesmen.</p>		<p>Do you know Li Gang?" "Gnagzi?" "More than thirty, and skinny, his arms filled with tattoos." "I see. That's him. He hasn't been here for a while."</p>
<p>Manual translation</p>	<p>From the east side of TieDep, turn right at the first hutong, right again when you see a shop selling cloth shoes, go up to the third floor of Block 2, and it's the black Panpan security door. TieDep referred to Tiexi Department Store, located in the center of Tiexi District. I'd been there as a child. It got really crowded on Sundays. Across from the shopping center was a Xinhua Bookstore with two racks of books to browse—you had to ask a clerk if you wanted to see the rest behind the counter.</p>	<p>—Good, now pay up. I handed her the cash, and she slipped it into her stocking top. —Your cousin's sick. This money will buy him medicine. Prozac. Now, hurry back to your job in Beijing. It's not like he's your brother. Just say you couldn't find him—no one will blame you. I put my glasses back on. —Never mind getting back to work. You asked me which was more important. I've thought about it, and it's my cousin. I have to see him. We can talk later about whether or not he's going home.</p>	<p>Mynah Snooker Hall wasn't large, just a dozen tables or so, but the lighting was soft and pleasantly warm, like springtime. There weren't many people around, and beneath the lights, the neat triangles of balls looked like precious artifacts in a museum display. The boss sat playing mahjong on a stark white Apple iMac. When he noticed me looking around, he stood up. — Can I help you? —I'm looking for someone called Li Gang. Thirties, skinny, arm tattoos. —Gang-san hasn't been in for a while.</p>
<p>Youdao translation</p>	<p>At Tiebai East, turn right at the first alley. When you see a storefront selling cloth shoes, turn right again. On the third floor of Unit 2, there is a black Panpan security door. Tiebai is the Tiexi Department Store, located in the center of Tiexi District. I went there when I was a child. Every Sunday, it was crowded with people. Opposite it was a Xinhua Bookstore with two open bookshelves. The rest of the books were behind the salespeople. If you wanted to read or buy them, you had to ask the salespeople to throw them over.</p>	<p>She said, "OK. Pay the money." I gave her the money. She stuffed it into the stockings at the root of her thigh and said, "Your brother is ill. You can use these 240 yuan as medicine." Prozac. I said, "I need to meet the person. Where is he?" She said, "Don't see him. He's not going back."</p>	<p>The Bago Billiard hall is not large, with over a dozen tables. However, the lighting is soft and warm like spring. There weren't many people. Under the light, the neatly arranged billiards were in a triangular shape, as if they were precious cultural relics on display in a museum window. The boss was sitting in front of a white Apple all-in-one computer, playing mahjong. When he saw me come in and spread it all over, he stood up and said, "Buddy, looking for someone?" I said, "Li Gang." I'm looking for Li Gang. He said, "Gangzi?" I said, "I have tattoos on both arms. I'm in my thirties and quite thin." He said, "It's Gangzi. I haven't been here recently."</p>

<p>Google Translation</p>	<p>Tiebai East, turn right at the first alley. You'll see a shop selling cloth shoes, then turn right again. Unit 2, third floor, black Panpan security door. Tiebai is the Tiexi Department Store, located in the heart of Tiexi District. I used to go there as a child, and it was packed every Sunday. Across from it is the Xinhua Bookstore, with two open bookshelves. The rest of the books are behind the salesperson; if you want to read or buy them, you have to ask the salesperson to toss them to you.</p>	<p>She said, "Okay, pay me." I gave her the money, and she stuffed it into her stocking at the base of her thigh. She said, "Your brother is sick. Think of these two hundred and forty yuan as medicine. Prozac." I said, "I need to see him. Where is he?" She said, "No, he's not coming back."</p>	<p>The Bage Billiard Hall is small, with a dozen or so tables, but the lighting is soft and warm, like spring. There are few people there. Under the light, the billiard balls are arranged in a triangle, like precious artifacts on display in a museum window. The boss was sitting at a pristine white Apple all-in-one computer, playing mahjong. He saw me come in, scrambling around, then stood up and said, "Hey, looking for someone?" I said, "Li Gang." Li Gang. He said, "Gangzi?" I said, "He has tattoos on both arms, is in his thirties, and is quite thin." He said, "It's Gangzi. He hasn't been here lately."</p>
<p>ChatGPT Translation</p>	<p>East of Tie-Bai: take the first alley and turn right; when you see a storefront selling cloth shoes, turn right again. Unit 2, third floor, black Panpan security door. "Tie-Bai" means the Tiexi Department Store, in the heart of Tiexi District. I went there as a kid; Sundays it was packed with people. Across the street was a Xinhua Bookstore with two open shelves; the rest of the books were behind the clerks. If you wanted to browse or buy, you had to have a clerk toss one over to you.</p>	<p>She said, "Okay—pay up." I handed her the money. She stuffed it into the waistband of her stockings at the top of her thigh and said, "Your brother's sick. Think of this two hundred and forty yuan as medicine money. Paxil." I said, "I need to see him—where is he?" She said, "Don't. He's not coming back."</p>	<p>The Eight-Bro Billiards Hall wasn't big—maybe a dozen tables—but the lighting was soft, warm as spring. Hardly anyone there. Under the lamps, the racked balls formed a triangle, like precious artifacts on display in a museum case. The owner sat at a sleek white Apple all-in-one, playing mahjong. When he saw me come in, he wiped his hands around and stood up. "Buddy, you looking for someone?" "Li Gang," I said. "I'm looking for Li Gang." "Gangzi?" he said. "Two tattooed arms, thirty-something, pretty skinny," I said. "That's Gangzi," he said. "He hasn't been in lately."</p>
<p>DeepSeek Translation</p>	<p>Tiexi Department Store. Take the first right, then right again when you see a cloth shoe vendor. Third floor, unit two, black Panda security door. Tiexi Department Store, or Tie Bai for short, is located in the heart of Tiexi District. I went there as a kid—every Sunday, it was packed with people. Across the street was a</p>	<p>She said, "Alright, pay up." I handed her the money, and she stuffed it into her stocking at the top of her thigh. "Your brother's sick," she said. "Consider this 240 yuan your contribution for his medicine. Prozac." I said, "I need to see him. Where is he?" She said, "Don't bother. He's not coming back."</p>	<p>The Parrot Billiards Hall wasn't very big, with just over a dozen tables, but the lighting was soft and the atmosphere warm as spring. There were only a few people there. Under the lights, the racked billiard balls formed a triangle, like precious artifacts displayed in a museum case. The owner was sitting in front of a pristine white Apple iMac, playing mahjong. When he saw me come in and start scanning the</p>

<p>Xinhua Bookstore with two open bookshelves; the rest of the books were kept behind the counter. If you wanted to look at one or buy it, you had to ask the clerk to toss it over.</p>	<p>place, he stood up and said, "Looking for someone, pal?" I said, "Li Gang. I'm looking for Li Gang." He said, "Gangzi?" I said, "Tattoos on both arms, in his thirties, pretty thin." He said, "That's Gangzi. Hasn't been around lately."</p>
--	---

By utilizing natural language processing technology to evaluate the translation quality of three experimental paragraphs, the following results can be obtained:

Table 3. Automated Evaluation Numerical Statistics Table

Paragraph	Type	BLEU	TER	BERTScore (F1)
1	Manual	0.6432	0.215	0.911
	Youdao	0.4827	0.412	0.849
	Google	0.5376	0.367	0.866
	ChatGPT	0.5893	0.298	0.882
	DeepSeek	0.5612	0.321	0.875
2	Manual	0.5987	0.243	0.902
	Youdao	0.4215	0.487	0.829
	Google	0.4689	0.438	0.842
	ChatGPT	0.5032	0.392	0.865
	DeepSeek	0.4876	0.405	0.858
3	Manual	0.6125	0.228	0.908
	Youdao	0.4468	0.453	0.836
	Google	0.4721	0.419	0.85
	ChatGPT	0.5384	0.361	0.875
	DeepSeek	0.5197	0.374	0.868

ChatGPT and DeepSeek followed closely, with average scores of 0.5436 and 0.5228, respectively, outperforming Google and Youdao.

The average TER score for human translation is 0.229, significantly lower than other platforms, indicating that the translation requires almost no modification to match the reference translation. ChatGPT's TER score is 0.350, performing best among machine platforms, with DeepSeek following closely, while Google and Youdao scored 0.408 and 0.451, respectively.

The average F1 score for human translation is 0.907, significantly higher than other platforms. ChatGPT's F1 score is 0.874, performing closest to human translation, with DeepSeek following closely (0.867), while Google and Youdao scored 0.853 and 0.838, respectively.

Combining the three metrics, human translation performed best in all dimensions, verifying its reliability as a reference translation. Among machine platforms, ChatGPT performed excellently in BLEU, TER, and BERTScore, demonstrating its comprehensive capabilities in literary translation. DeepSeek's performance was also relatively stable, approaching ChatGPT in semantic restoration.

In the translator's scoring, human translation received the highest score of 41 points, while Google Translate scored the lowest with 31 points.

Table 4. Translation Evaluation Table

	Translator 1	Translator 2	Translator 3	Translator 4	Translator 5	Overall
Manual	8	8	9	8	8	41
Youdao	7	6	6	7	7	33
Google	6	6	6	7	6	31
ChatGPT	7	8	7	7	8	37
DeepSeek	8	9	8	7	7	39

5. Conclusion

This study focused on literary texts, comprehensively using automated evaluation metrics and subjective scoring by translators to systematically analyze the translation quality of four translation platforms and human translators. The results show that automated metrics can effectively capture the performance of translations in terms of lexical matching, structural restoration, and semantic fidelity, especially BERTScore, which demonstrated high applicability in semantic-level evaluation. However, translator scoring revealed dimensions that automated metrics struggle to cover, such as literary style, cultural imagery, and emotional conveyance, suggesting that literary translation quality assessment should adopt a multi-dimensional approach. The results show that automated metrics and human scoring have a certain consistency in overall trends, but at the detailed level, they also reveal differences among platforms in terms of semantic restoration, style conveyance, and translator acceptance.

References

- [1] Qorbani, A., Ramezani, R., Baraani, A., & Kazemi, A. (2025). Multilingual neural machine translation by cascading computational graphs. *Expert Systems with Applications*, 128722. <https://doi.org/10.1016/j.eswa.2025.128722>
- [2] Hatim, B., & Munday, J. (2004). *Translation: An advanced resource book*. Routledge.
- [3] Khudaverdiyeva, T. (2024). Translation problems and solutions: bridging linguistic and cultural differences. In Proceedings of the Republican Scientific Conference on Contemporary Problems of Translation Studies ("Tərcüməşünaslığın müasir problemləri"), Azerbaijan University of Languages and Nakhchivan State University (pp. 1–139).
- [4] Khudaverdiyeva, T. (2025). Machine and human translation: An Applied Linguistics Evaluation. *Journal of Foreign Language Teaching and Applied Linguistics*, 11(1), 98–104.
- [5] Proietti, L., Perrella, S., & Navigli, R. (2025). Has Machine Translation Evaluation Achieved Human Parity? The Human Reference and the Limits of Progress. arXiv preprint arXiv:2506.19571.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).