International Journal of Language and Linguistics
2026, Vol. 14, No. 6, pp. 257–265
https://doi.org/10.11648/j.ijll.20261406.13

SciencePG
Science Publishing Group

Research Article

# A Corpus-Based Comparison of Stylistic Features in Google Translate's and ChatGPT's English Translations of Lu Xun's Shexi

**Rui Sun***

School of Foreign Studies, University of Science and Technology Beijing, Beijing, China

## Abstract

The use of AI in translation practice has become increasingly prevalent in recent years, attracting growing scholarly attention. One important research focus in AI translation studies is the comparison of translation quality between neural machine translation models and large language models. However, comparatively less attention has been paid to their stylistic differences. Aiming to explore the two systems' literary translation style, this exploratory corpus-based study investigates Google Translate's and ChatGPT's English translations of Chinese writer Lu Xun's short story *Shexi* (社戏, 'Village Opera'). Using type-token ratio, average sentence length, and lexical density as quantitative stylistic measures, the study examines differences in lexical diversity, sentence splitting, and information richness between the two models' translations. The results show that Google Translate's stylistic indicators are generally closer to those of the source text. Compared with Google Translate, ChatGPT's translation demonstrates a lower type-token ratio, shorter average sentence length, and lower lexical density. Combined with qualitative analysis, these findings suggest that ChatGPT's translation may place greater emphasis on lexical consistency and show a greater tendency towards explicitation, potentially reflecting a more reader-oriented approach. These findings suggest that AI systems may differ in their translation styles. By highlighting style as an important dimension of AI translation analysis, this study complements existing quality-focused evaluations and provides further insights into different AI models' unique ways of translating.

## Keywords

Translator Style, Machine Translation, Corpus-Based Translation Studies, Literary Translation, Google Translate, ChatGPT

## 1. Introduction

The rapid development of technology has significantly transformed the field of translation. From early rule-based and statistical approaches to neural machine translation (NMT), machine translation (MT) systems have undergone noticeable development. The emergence of large language models (LLMs) has introduced a new stage of AI translation. Although general-purpose LLMs are not primarily designed for translation, they are used for translation tasks due to their strong linguistic capabilities. Given these technological developments, a growing body of research has investigated AI translation, particularly comparing the translation quality of NMT models and LLMs. Many scholars primarily use automatic evaluation metrics, such

*Correspondence: Rui Sun (ruisun@ustb.edu.cn)

**Received:** 13 October 2026; **Accepted:** 6 November 2026; **Published:** 26 November 2026

Copyright: © The Author(s), 2026. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

as BLEU [1], TER [2], and ChrF [3] or ChrF++ [4], to compare the translation quality of NMT models and LLMs, sometimes supplemented by human evaluation [5-10]. Some scholars examine translation quality manually [11-15]. These studies have provided valuable insights into the translation performance of the AI systems examined.

In addition to translation quality assessment (TQA), stylistic analysis can also provide an important perspective for understanding AI translation. Style is particularly important in literary translation. While previous research has extensively examined the translation quality of different AI systems, relatively limited attention has been given to the literary style of AI systems. The few studies of this kind are closely tied to human translators' styles: investigating how human-AI collaborative translation affects the manifestation of a human translator's style [16, 17], training machines to reproduce human translators' styles [18-20], or distinguishing translations by human translators from those produced by different AI models [21]. Research that specifically focuses on the translation styles of AI models is rare.

To address this gap, the present study compares the English translations of Lu Xun's *Shexi* (社戏, 'Village Opera') produced by Google Translate and ChatGPT. Google Translate is selected as a representative of NMT models, while ChatGPT represents LLMs. Using a corpus-based method, this study examines their translations in terms of type-token ratio, average sentence length, and lexical density, which are commonly used textual features in corpus-based analyses of translator style [22-25].

### 1.1. Translator Style

Baker defines translator style as 'a kind of thumb-print that is expressed in a range of linguistic — as well as non-linguistic — features' [22]. She compares Peter Bush's and Peter Clark's English translations in terms of the type-token ratio, average sentence length, and the use of say. Adopting Baker's definition, Olohan compares collections of translations from different STs by two translators through the lens of their contraction strategies [26]. Saldanha proposes a revised definition of translator style:

A 'way of translating' which

1) is felt to be recognizable across a range of translations by the same translator, distinguishes the translator's work from that of others,
2) constitutes a coherent pattern of choice,
3) is 'motivated', in the sense that it has a discernible function or functions,
4) cannot be explained purely with reference to the author or source-text style, or as the result of language differences [27].

These widely accepted definitions of translator style conceptualize the concept as a translator's distinctive way of translating that distinguishes their work from that of others. From this perspective, translator style can be investigated by comparing translations produced by different translators to identify their distinctive stylistic patterns. This approach can also be extended to the study of the translation styles exhibited by different AI models.

### 1.2. Stylistic Features of AI Translation

In recent years, scholars have begun to investigate the stylistic features of AI translation. Jiang, Lv, and Zhang compare human translations, Google Translate, and ChatGPT translations of spokesperson's remarks using multi-feature analysis to determine whether they can be distinguished by their stylistic features [21]. They find that unsupervised hierarchical clustering fails to separate the three types, while supervised classification performs better. However, supervised methods sometimes struggle to distinguish ChatGPT from Google Translate, whereas human translations are consistently identified.

Kenny and Winters compare Hans-Christian Oeser's English translation of an excerpt of a machine-translated German literary text that he post-edited and his from-scratch translation, and find that the machine-assisted translation is somewhat less representative of Oeser's style [16]. They also compare DeepL's German translation of English literature with Oeser's post-edited version to examine whether the post-edits reflect his usual style [17]. Scholars such as Hansen, Yirmibeşoğlu, and Dallı aim to fine-tune existing pre-trained MT models to reproduce a translator's style using the translator's translation as training data [18-20].

Existing studies of AI translation style have often involved human translations, while few have specifically investigated stylistic differences across AI systems. To address this gap, the present study adopts a corpus-based approach combined with qualitative analysis to examine the stylistic differences between two AI models, Google Translate and ChatGPT. The overarching research question is:

How do Google Translate and ChatGPT differ in their translation styles when translating Chinese writer Lu Xun's *Shexi* into English?

The specific research questions are:

1) How do Google Translate's and ChatGPT's translations differ in terms of type-token ratio, average sentence length, and lexical density?
2) What do these quantitative stylistic indicators reveal about the lexical variety, sentence segmentation, and information richness of the two translations?
3) How does qualitative analysis, combined with the quantitative findings, reveal possible stylistic features of Google Translate's and ChatGPT's translations?

## 2. Materials and Methods

The texts examined in this research consist of:

Chinese ST:

Lu Xun's *Shexi* (社戏, 'Village Opera')

English TT1:

Translation produced by Google Translate

English TT2:

Translation generated by ChatGPT

The ST is a short story written in 1922 by Lu Xun, one of the most influential modern Chinese writers. The story consists of two parts: the first describes the author's experiences of watching operas in Beijing, while the second recounts his childhood memories of watching an opera in the countryside. As one of Lu Xun's representative works distinguished by its literary merit and aesthetic value, *Shexi* has been widely included in textbooks published by major Chinese publishers and translated into multiple languages worldwide.

TT1 was produced by Google Translate, which is one of the most widely recognized and frequently used NMT systems. Developed by Google, it was launched in April 2006 as a statistical MT model. In November 2016, Google announced its transition to an NMT model, marking a major shift in its translation approach. The translation was accessed on 4 September 2026.

TT2 was generated by ChatGPT, which is an advanced LLM developed by OpenAI and released in November 2022. At the time of data collection, GPT-5.6 Luna was the default model provided to users of ChatGPT's free plan. The translation generated by ChatGPT was obtained on 4 September 2026 using the prompt: 'Translate the text into English: [Chinese ST].' Since prior research has shown that prompt engineering can affect the translation output of generative AI [28-31], the prompt in this study was intentionally made simple. This simple prompt aims to help reveal ChatGPT's inherent translation tendencies, making the comparison with Google Translate, which does not offer the same prompt-based interaction, more reflective of their respective characteristics.

The ST and TTs were compiled into a corpus using Sketch Engine [32]. The Chinese ST was automatically segmented into word sequences by Sketch Engine and annotated using the Chinese Penn Treebank POS tagset. The English TTs were POS-tagged using the English TreeTagger POS tagset with Sketch Engine modifications.

Google Translate's and ChatGPT's translations were compared from the perspectives of type-token ratio, average sentence length, and lexical density. The corpus information interface in Sketch Engine directly provides the total number of words, number of unique words, and number of sentences. These measurements were used to calculate type-token ratio and average sentence length. The number of content words required for calculating lexical density is not directly available. The content words were extracted using the Corpus Query Language (CQL) tool. For the Chinese ST, the query [tag="N.*|V.*|AD|JJ|[CO]D"] was used, while the English TTs were queried with [tag="N.*|V.*|J.*|RB.?|CD"]. The queries were designed to retrieve nouns, verbs, adjectives, adverbs, and numerals in the texts.

The quantitative comparison offers an empirical foundation for investigating how Google Translate's and ChatGPT's translations differ in terms of lexical variation, sentence segmentation, and information richness. These quantitative findings are examined alongside a qualitative comparison of the ST and TTs to gain deeper insights into the two models' stylistic characteristics.

# 3. Results

## 3.1. Type-Token Ratio

The type-token ratio measures the diversity of vocabulary employed in a text. It is calculated using the number of unique words (types) divided by the total number of words (tokens). The higher the type-token ratio, the richer the lexis used in the text. As displayed in Table 1, Google Translate's translation has a higher type-token ratio than ChatGPT's.

***Table 1.*** *Type-token ratio of the ST, TT1 and TT2.*

| Text | ST | TT1 | TT2 |
|---|---|---|---|
| Total number of words | 3,681 | 4,018 | 4,625 |
| Number of unique words | 1,257 | 1,208 | 1,285 |
| Type-token ratio | 0.34 | 0.30 | 0.28 |

ChatGPT's lower type-token ratio compared with Google Translate may partly be attributable to the greater length of its output. The type-token ratio is sensitive to text length. Generally, the longer the text, the lower the type-token ratio. One commonly used method to minimize the influence of text length is the standardized type-token ratio (STTR), which involves dividing a text into successive segments of 1,000 tokens, calculating the type-token ratio for each segment, and then taking the average of these values. Any final segment containing fewer than 1,000 tokens is excluded from the calculation. However, given the relatively limited length of the texts examined in this study, applying STTR would result in a considerable proportion of the material being excluded from the analysis. The final 18.5% of the source text and 13.5% of ChatGPT's translation would be omitted from the calculation. An alternative measure is the moving-average type–token ratio (MATTR), which calculates type-token ratio across successive, overlapping windows of a fixed number of tokens and averages the resulting values, reducing its sensitivity to overall text length and allowing the entire text to contribute to the analysis [33]. However, implementing MATTR is more technically demanding and requires careful methodological decisions, particularly the selection of an appropriate window size.

In light of these considerations, the present study adopts an equal-length sampling approach to facilitate a straightforward comparison between the two models' outputs. It compares the first 4,018 words of both TTs; 4,018 is the total word count of

the shorter TT, which was produced by Google Translate. This approach ensures equal sample lengths while allowing the entire Google Translate output to be included in the comparison. It should be noted, however, that ChatGPT's translation is not included in its entirety. The first 4,018 words of ChatGPT's translation contain 1,182 unique words, yielding a type-token ratio of 0.29. Although there is a slight increase from the previously calculated value of 0.28, it remains marginally lower than Google Translate's type-token ratio of 0.30. This suggests that Google Translate's translation uses slightly more varied vocabulary and has less repetition. Nevertheless, the two values are very close. This small numerical difference, on its own, does not provide sufficient evidence to conclude that the two models differ in their translation styles in terms of lexical variety.

Although the quantitative results reveal only a marginal difference between the two models' translations, a close qualitative comparison of the ST and the TTs suggests that ChatGPT may place greater emphasis on consistency and explicitation, i.e., a translation technique that makes implicit information in a ST explicit in the TT [34-39]. Example 1 offers a representative illustration of this tendency.

Example 1

*ST* 在小村里，一家的客，几乎也就是公共的。

('In small village, one family's guest, was almost also the public's.')

*Google Translate* In this small village, a guest from one family was practically a member of the public.

*ChatGPT* In such a small village, a guest belonging to one family was practically a guest belonging to the whole community.

When translating 客 (kè 'guest'), Google Translate uses two different expressions, *guest* and *member*, while ChatGPT consistently translates it as *guest* in both cases. This suggests that ChatGPT may place greater emphasis on lexical consistency and maintaining the same semantic reference throughout the translation. ChatGPT also adopts a parallel structure, repeating the pattern *a guest belonging to*, whereas Google Translate avoids such structural repetition by using a more varied expression. Moreover, ChatGPT makes the implicit belonging relationship in the ST more explicit by clearly showing that a guest of one family is also considered a guest of the whole community. Compared with Google Translate, which appears to follow the surface meaning of the ST more closely, ChatGPT provides a clearer explanation of the social relationship implied in the ST, making the meaning easier for English readers to understand and better capturing the interpersonal bonds in a small, traditional Chinese village.

The difference in the two models' translation consistency is also illustrated by how they render the repeated expression 秩秩斯干 in the ST, which appears twice in the ST. The first occurrence appears as part of the full line 秩秩斯干幽幽南山, the opening line of *Sigan* (斯干, '*The Stream*') in *The Classic of Poetry*, which evokes the image of a flowing stream and secluded southern mountains. Google Translate renders this line as *The stream flows gently, the southern mountains are secluded*. However, when 秩秩斯干 appears again later in the ST without the following phrase 幽幽南山, Google Translate adopts a different strategy, retaining the original Chinese expression and adding the parenthetical explanation *a phrase from a poem*. This addition makes the expression's literary origin explicit. However, the shift from semantic translation to ST retention is inconsistent across the two occurrences. By contrast, ChatGPT translates both occurrences as *The winding stream flows through the secluded southern hills*. The two models differ in how they organise the imagery of the original line. In the ST, the stream and the mountains are depicted separately as two juxtaposed elements, a structure that Google Translate preserves in its translation. ChatGPT, in contrast, integrates the two elements into a single, interconnected scene, depicting the stream as flowing through the secluded southern hills. Furthermore, ChatGPT provides the fuller semantic rendering even when the second occurrence in the ST consists only of 秩秩斯干. In doing so, it draws on the broader textual context to restore the imagery conveyed by the full line, introducing information that is omitted in the second occurrence itself. This suggests that ChatGPT not only makes contextual information more explicit but also prioritises consistency in rendering recurring expressions.

A similar pattern can be observed in their translations of 叫天 (Jiaotian), which appears six times in the ST and refers to Tan Xinpei (谭鑫培), a renowned Peking opera actor. The first occurrence of 叫天 appears within 谭叫天, which combines his surname 谭 with a shortened form of his stage name, 小叫天. The stage name derives from that of his father, Tan Zhidao (谭志道). Because Tan Zhidao had a high-pitched, resonant voice, audiences likened him to the Jiaotianzi (叫天子), a pipit known for its far-carrying call, and thus gave him the stage name Jiaotian. His son, Tan Xinpei, inherited the family's artistic tradition and became known as Xiao Jiaotian (小叫天, 'Little Jiaotian'). The second occurrence of 叫天 appears within his full stage name 小叫天. The remaining four occurrences in the ST use 叫天 alone to refer to Tan Xinpei. Google Translate handles the six occurrences of the stage-name reference inconsistently. It renders 叫天 as *Jiaotian* in the first two occurrences, omits it in one of the subsequent instances, and translates the other three as *"Calling Heaven"*, *the call to heaven*, and *Jiao Tian*, respectively. By contrast, ChatGPT renders the first occurrence as *Xinpei*, identifying the actor by his name rather than strictly following the ST. This may be regarded as a form of referential explicitation, although it also involves replacing the original stage-name reference. ChatGPT then consistently renders the remaining five occurrences as *Jiaotian*. This suggests that ChatGPT may draw on contextual and historical knowledge to identify the referent while maintaining a consistent rendering of the recurring name in the subsequent instances. Compared with Google Translate, ChatGPT demonstrates greater consistency in handling repeated references to the same individual.

Taken together, it seems that ChatGPT tends to render recurring expressions more consistently than Google Translate, while also making certain implicit information more explicit. Greater consistency in lexical choice may contribute to ChatGPT's marginally lower type-token ratio, as the repeated use of the same lexical items can reduce the number of unique words in a text. However, given the small numerical difference between the two type-token ratio values, the findings suggest that the models demonstrate similar extents of lexical diversity overall. The qualitative examples offer a possible explanation for the slight numerical difference.

## 3.2. Average Sentence Length

The average sentence length of a text is determined by dividing the total number of words by the number of sentences. As shown in Table 2, the average sentence length of both translations is noticeably shorter than that of the ST, particularly in ChatGPT's translation.

***Table 2.*** *Average sentence length of the ST, TT1 and TT2.*

| Text | ST | TT1 | TT2 |
|---|---|---|---|
| Total number of words | 3,681 | 4,018 | 4,625 |
| Number of sentences | 172 | 222 | 353 |
| Average sentence length | 21.4 | 18.1 | 13.1 |

Although both translations have a greater total length than the ST, they also contain a noticeably larger number of sentences, indicating that both models adopt a sentence-splitting strategy. Such a strategy may reflect the two models' effort to improve readability by breaking down complex information into smaller units. Compared with Google Translate, ChatGPT has a longer total length but a much higher number of sentences. This suggests that ChatGPT's translation involves more extensive sentence segmentation, resulting in shorter average sentence length.

A comparison of the ST and its TTs reveals that ChatGPT frequently splits a single ST sentence into two or more sentences, whereas Google Translate is more inclined to retain the original sentence structure. Example 2 provides a typical case.

Example 2

*ST* 我先是没有爬上去的勇气，接着便联想到私刑拷打的刑具，不由的毛骨悚然的走出了。

('I at first didn't have the courage to climb up, then thought of instruments of private punishment torture, could not help but walking out with body hair standing up, spine bone turning cold. ')

*Google Translate* At first, I didn't have the courage to climb up, then I thought of instruments of torture, and a chill ran down my spine as I walked away.

*ChatGPT* At first I did not have the courage to climb onto it. Then I suddenly thought of the instruments used in torture and corporal punishment. I shuddered and walked out.

Compared with Google Translate, which maintains the ST structure by combining the three actions into one sentence, ChatGPT divides the ST into three separate sentences, making the sequence of events and the psychological progression easier for English readers to follow. ChatGPT's more frequent sentence splitting possibly reflects the model's preference for clarity and readability. The tendency is also evident in its translation of 毛骨悚然, which is a Chinese idiom whose literal meaning is 'the hair on one's body stands up and one's spine turns cold', describing a state of extreme fear or horror. Google Translate conveys it as *a chill ran down my spine*, which corresponds to the physical imagery embedded in the idiom more closely. In contrast, ChatGPT renders it simply as *I shuddered*, prioritizing accessibility for English readers over preserving the figurative features of the ST. This choice can be understood as a form of reader-oriented normalisation, whereby a culturally and figuratively marked expression is rendered in a more conventional and accessible form in the TT. Although this strategy may enhance readability, it may also be regarded as a loss in translation quality under an evaluation framework that prioritizes semantic correspondence. This highlights the distinction between style and quality as two separate dimensions of translation assessment.

In some cases, both ChatGPT and Google Translate split a ST sentence into shorter sentences, but ChatGPT does so to a greater extent. Example 3 presents one such case.

Example 3

*ST* 我忍耐的等着，许多工夫，只见那老旦将手一抬，我以为就要站起来了，不料他却又慢慢的放下在原地方，仍旧唱。

('I waited patiently, for a long time, just saw the old female character raise hand, I thought about to stand up, unexpectedly the character slowly lowered back to original place, continued singing.')

*Google Translate* I waited patiently, and after a long while, I saw the old woman raise her hand. I thought she was about to stand up, but instead, she slowly lowered it back to its original position and continued singing.

*ChatGPT* I waited patiently.

After a very long time, the old female performer finally raised one hand.

I thought she was about to stand up.

But no.

She slowly lowered her hand back to the same place and continued singing.

Whereas Google Translate renders the ST sentence in two sentences, ChatGPT produces five, each presented as a separate paragraph. This tendency to restructure the ST is frequent in ChatGPT's translation. Such restructuring may reflect an attempt to enhance readability by presenting information in shorter, more easily digestible units. Overall, ChatGPT appears more inclined than Google Translate to depart from the

ST's original structural boundaries.

### 3.3. Lexical Density

Lexical density, which estimates the information richness of a text, can be calculated using the number of content words divided by the total number of words in a text. The analysis shows that Google Translate's translation has a very slightly higher lexical density than ChatGPT's (see Table 3), indicating that Google Translate's translation is marginally richer in information content.

***Table 3.*** *Lexical density of the ST, TT1 and TT2.*

| Text | ST | TT1 | TT2 |
|---|---|---|---|
| Total number of words | 3,681 | 4,018 | 4,625 |
| Number of content words | 2,606 | 2,414 | 2,718 |
| Lexical density | 0.71 | 0.60 | 0.59 |

The results should be interpreted with caution, particularly when comparing lexical density across Chinese and English. The Chinese Penn Treebank part-of-speech tagset does not provide dedicated POS tags for auxiliary and modal verbs. Therefore, these items were grouped with other verbs and included in the content-word count. The English TreeTagger PoS tagset with Sketch Engine modifications assigns modal verbs a dedicated tag, allowing them to be excluded from the count. However, auxiliary verbs could not be consistently excluded, so auxiliary verbs were included in the content-word count. Although this procedure does not allow for the complete exclusion of all function words, the same query was applied across both English TTs, providing a consistent basis for comparison within the English TTs. The resulting lexical-density figures remain informative as approximate measures for comparative analysis, although differences between Chinese and English should be interpreted with greater caution.

The difference in lexical density is negligible and cannot be interpreted as evidence of a meaningful stylistic difference between the two translations. Nevertheless, the values are consistent with the results regarding total length, average sentence length, and type-token ratio. ChatGPT's translation has a longer total length but shorter average sentence length, suggesting that the model tends to unpack the information in the ST and present it through more separate but semantically explicit units. Meanwhile, its lower type-token ratio and lexical density suggest that ChatGPT does not prioritize lexical variation or a highly condensed style. Instead, it tends to adopt consistent and explicit lexical choices to make the meaning of the ST clearer and more accessible to readers.

Overall, these features suggest that ChatGPT tends to prioritize consistency, explicitation, and readability. In contrast, Google Translate shows a preference for staying closer to the ST.

## 4. Discussion

Quantitative analysis suggests that Google Translate's output is closer to the Chinese ST. The comparison should, however, be interpreted cautiously. The Chinese ST serves only as a tentative reference point given language differences in word segmentation, sentence-boundary conventions, and the conveyance of information. Still, the qualitative analysis also suggests that Google Translate tends towards more conservative renderings, whereas ChatGPT's translation demonstrates greater flexibility in reformulating and adapting the ST. This tendency reflects the characteristics commonly associated with NMT systems, which are trained to learn correspondences between source and target languages based on large-scale parallel data. As a result, NMT models may maintain relatively strong alignment with ST structures and linguistic patterns. In contrast, ChatGPT's translation exhibits a more reader-oriented style characterized by greater consistency, explicitness, and restructuring of information. This may be associated with the generative nature of LLMs, which could draw on contextual information when producing TTs. Such an approach may allow LLMs to produce translations that are more coherent and accessible for target readers.

This study has several limitations. First, as an exploratory comparison, the analysis is based on a single literary ST and one output from each model. Therefore, the stylistic differences identified may be specific to the text and outputs examined and may not generalize to other texts, genres, or translation contexts. Future research could investigate translation styles across a broader range of STs representing different genres, collect multiple outputs from LLMs, and examine the effects of different prompting strategies on LLMs' translations. Where appropriate, statistical tests could also be used to assess the significance and robustness of the observed differences across texts and outputs. Second, Google Translate and ChatGPT were selected as representatives of NMT's and LLMs' translations; however, the findings may not be fully generalizable to other models within the same categories, as individual models may differ in their training data and optimization strategies. Third, the study relies on a limited number of corpus-based stylistic indicators. Future research could expand the comparison by examining multiple NMT models and LLMs across different genres and language pairs, while incorporating a wider range of stylistic features, such as reporting verbs, keywords, and key clusters, to provide a more comprehensive analysis of AI models' translation styles.

## 5. Conclusions

This exploratory study examined stylistic differences between Google Translate's and ChatGPT's English translations of Lu Xun's *Shexi* through a corpus-based method. The

comparison of the two models' translations in terms of type-token ratio, average sentence length, and lexical density, alongside qualitative analysis, reveals that, while Google Translate tends to stay closer to the ST, ChatGPT demonstrates a stronger tendency toward explicit and reader-oriented expression. These findings suggest that different AI models may have different stylistic orientations in translation. The stylistic analysis offers insights into how different systems organize information and represent meanings in the target language. By highlighting style as an important dimension of AI translation analysis, this study contributes to a deeper understanding of translations produced by Google Translate and ChatGPT.

## Abbreviations

| | |
|---|---|
| AI | Artificial Intelligence |
| CQL | Corpus Query Language |
| LLMs | Large Language Models |
| MATTR | Moving-Average Type–Token Ratio |
| MT | Machine Translation |
| NMT | Neural Machine Translation |
| POS | Part-of-Speech |
| ST | Source Text |
| STTR | Standardized Type-Token Ratio |
| TQA | Translation Quality Assessment |
| TT | Target Text |

## Supplementary Material

The supplementary material can be accessed at https://doi.org/10.11648/j.ijll.20261406.13

## Author Contributions

**Rui Sun:** Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Validation, Visualization, Writing – original draft, Writing – review & editing

## Funding

This work is supported by the Fundamental Research Funds for the Central Universities (Grant No. FRF-TP-26-072).

## Data Availability Statement

The data supporting the outcome of this research work have been reported in this manuscript.

## Conflicts of Interest

The author declares no conflicts of interest.

## References

[1] Papineni, K., Roukos, S., Ward, T., Zhu, W.-J. BLEU: A Method for Automatic Evaluation of Machine Translation. In Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, Philadelphia, USA, 2002; pp. 311–318. https://doi.org/10.3115/1073083.1073135

[2] Snover, M., Dorr, B., Schwartz, R., Micciulla, L., Makhoul, J. A Study of Translation Edit Rate with Targeted Human Annotation. In Proceedings of the 7th Conference of the Association for Machine Translation in the Americas: Technical Papers, Massachusetts, USA, 2006; pp. 223–231.

[3] Popović, M. ChrF: Character N-gram F-score for Automatic MT Evaluation. In Proceedings of the Tenth Workshop on Statistical Machine Translation, Lisbon, Portugal, 2015; pp. 392–395. https://doi.org/10.18653/v1/W15-3049

[4] Popović, M. ChrF++: Words Helping Character N-grams. In Proceedings of the Second Conference on Machine Translation, Copenhagen, Denmark, 2017, pp. 612–618. https://doi.org/10.18653/v1/W17-4770

[5] Son, J., Kim, B. Translation Performance from the User's Perspective of Large Language Models and Neural Machine Translation Systems. Information. 2023, 14, 574. https://doi.org/10.3390/info14100574

[6] Jiao, W., Wang, W., Huang, J., Wang, X., Shi, S., Tu, J. Is ChatGPT a Good Translator? Yes with GPT-4 as The Engine. arXiv. 2023. https://doi.org/10.48550/arXiv.2301.08745

[7] Hendy, A., Abdelrehim, M., Sharaf, A., Raunak, V., Gabr, M., Matsushita, H., Kim, Y. J., Afify, M., Awadalla, H. H. How Good Are GPT Models at Machine Translation? A Comprehensive Evaluation. arXiv. 2023. https://doi.org/10.48550/arXiv.2302.09210

[8] Alkhawaja, L. Unveiling the New Frontier: ChatGPT-3 Powered Translation for Arabic–English Language Pairs. Theory and Practice in Language Studies. 2024, 14(2), 347–357. https://doi.org/10.17507/tpls.1402.05

[9] Cetin, E. D., Sahin, C. Multi-Metric Evaluation of Translation-Based Cross-Lingual Sentiment Consistency Using Large Language Models and Neural Machine Translation. Applied Sciences. 2026, 16(16), 7878. https://doi.org/10.3390/app16167878

[10] Castilho, S., Mallon, C. Q., Meister, R., Yue, S. Do Online Machine Translation Systems Care for Context? What About a GPT Model? In Proceedings of the 24th Annual Conference of the European Association for Machine Translation, Tampere, Finland, 2023; pp. 393–417.

[11] Wang, Q., Wang, Y., Amini, M., Xian, M., Fu, Z. Quality Assessment of Tibetan–Chinese Poetry Translation: Integrating Automated Metrics and Qualitative Insights Through a Cross-System Comparison of Dedicated NMT Engines and a Prompted LLM. Language Resources and Evaluation. 2026, 60. https://doi.org/10.1007/s10579-026-09918-8

[12] Zayed, A. B. Evaluating the Fidelity and Accuracy of ChatGPT 4 and Google Translate in Translating Legal English Documents into Arabic–and Vice Versa. Journal of the Faculty of Languages, University of Tripoli, Libya. 2024, 1(29), 64-87.

[13] Sun, R. Evaluating the Translation Accuracy of ChatGPT and DeepL Through the Lens of Implied Subjects. AWEJ for Translation & Literary Studies. 2024, 8(4), 41-53. https://doi.org/10.24093/awejtls/vol8no4.5

[14] AlAfnan, M. A. Large Language Models as Computational Linguistics Tools: A Comparative Analysis of ChatGPT and Google Machine Translations. Journal of Artificial Intelligence and Technology. 2025, 5, 20–32. https://doi.org/10.37965/jait.2024.0549

[15] Dong, L. A Comparative Study on the Translation Quality of Chinese Diplomatic Discourse by NMT and LLMs Based on Multidimensional Quality Metrics. International Journal of Applied Linguistics and Translation. 2025, 11(4), 107–115. https://doi.org/10.11648/j.ijalt.20251104.12

[16] Kenny, D., Winters, M. Machine Translation, Ethics and the Literary Translator's Voice. Translation Spaces. 2020, 9(1), 123–49. https://doi.org/10.1075/ts.00024.ken

[17] Kenny, D., Winters, M. Customisation, Personalisation and Style in Literary Machine Translation. In Translation, Interpreting and Technological Change: Innovations in Research, Practice and Training, Böser, U., Deane-Cox, S., Winters, M., Eds., London: Bloomsbury Academic; 2024, pp. 59-79. https://doi.org/10.5040/9781350212978

[18] Hansen, D., Esperança-Rodier, E., Blanchon, H., Bada, V. 2022. La traduction littéraire automatique: Adapter la machine à la traduction humaine individualisée [Automatic literary translation: Adapting the machine to individualised human translation]. In Vers une robotique du traduire [Towards a robotics of translation], Baillot, A., Carter, E., Grass, T., Fabo, P. R., Ed., special issue of Journal of Data Mining and Digital Humanities. https://doi.org/10.46298/jdmdh.9114

[19] Yirmibeşoğlu, Z., Dursun, O., Dalli, H., Şahin, M., Hodzik, E., Gürses, S., Güngör, T. Incorporating Human Translator Style into English–Turkish Literary Machine Translation. In Proceedings of the 24th Annual Conference of the European Association for Machine Translation, Tampere, Finland, 2023; pp. 419–428. https://aclanthology.org/2023.eamt-1.40

[20] Dallı, H., Dursun, O., Güngör, T., Gürses, S., Hodzik, E., Şahin, M., Yirmibeşoğlu, Z. Giving a Translator's Touch to the Machine: Reproducing Translator Style in Literary Machine Translation. Palimpsestes. 2024, 38, 15–36. https://journals.openedition.org/palimpsestes/8696

[21] Jiang, Z., Lv, Q., Zhang, Z. Distinguishing Translations by Human, NMT, and ChatGPT: A Linguistic and Statistical Approach. arXiv. 2023. https://doi.org/10.48550/arXiv.2312.10750

[22] Baker, M. Towards a Methodology for Investigating the Style of a Literary Translator. Target. 2000, 12(2), 241–266. https://doi.org/10.1075/target.12.2.04bak

[23] Bosseaux, C. A Study of the Translator's Voice and Style in the French Translations of Virginia Woolf's The Waves. In CTIS Occasional Papers, Olohan, M., Ed., Manchester: Centre for Translation and Intercultural Studies; 2001, pp. 55–75.

[24] Li, D., Zhang, C., Liu, K. Translation Style and Ideology: A Corpus-Assisted Analysis of Two English Translations of Hongloumeng. Literary and Linguistic Computing. 2011, 26(2), 153–166. https://doi.org/10.1093/llc/fqr001

[25] Chen, M., Qu, X. A Corpus-Based Study of the Translator's Style of Two Versions of Fu Sheng Liu Ji—Take Pratt & Chiang Su-Hui's and Graham Sanders's Versions as Examples. Theory and Practice in Language Studies. 2023, 13(3), 617–623. https://doi.org/10.17507/tpls.1303.10

[26] Olohan, M. How Frequent Are the Contractions? A Study of Contracted Forms in the Translational English Corpus. Target. 2003, 15(1), 59–89. https://doi.org/10.1075/target.15.1.04olo

[27] Saldanha, G. Translator Style: Methodological Considerations. The Translator. 2011, 17(1), 25–50. https://doi.org/10.1080/13556509.2011.10799478

[28] Gao, Y., Wang, R., Hou, F. How to Design Translation Prompts for ChatGPT: An Empirical Study. arXiv. 2023. https://doi.org/10.48550/arXiv.2304.02182

[29] Peng, K., Ding, L., Zhong, Q., Shen, L., Liu, X., Zhang, M., Ouyang, Y., Tao, D. Towards Making the Most of ChatGPT for Machine Translation. arXiv. 2023. https://doi.org/10.48550/arXiv.2303.13780

[30] Yamada, M. Optimizing Machine Translation through Prompt Engineering: An Investigation into ChatGPT's Customizability. arXiv. 2023. https://doi.org/10.48550/arxiv.2308.01391

[31] Aldosari, L. A., Altuwairesh, N. Assessing the Effects of Translation Prompts on the Translation Quality of GPT-4 Turbo Using Automated and Human Evaluation Metrics: A Case Study. Perspectives (published online first on 08 May 2025). 1–25. https://doi.org/10.1080/0907676X.2025.2464120

[32] Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., Suchomel, V. The Sketch Engine: Ten Years On. 2014, Lexicography. (1), 7–36. https://doi.org/10.1007/s40607-014-0009

[33] Covington, M. A., McFall, J. D. Cutting the Gordian Knot: The Moving-Average Type–Token Ratio (MATTR), Journal of Quantitative Linguistics. 2010, 17(2), 94–100. https://doi.org/10.1080/09296171003643098

[34] Vinay, J.-P., Darbelnet, J. Comparative Stylistics of French and English: A Methodology for Translation; Translated by Sager, J. C. from Stylistique comparée du français et de l'anglais: Méthode de traduction published in 1958, Amsterdam: John Benjamins Publishing Company; 1995, p. 342.

[35] Blum-Kulka, S. Shifts of Cohesion and Coherence in Translation. In Interlingual and Intercultural Communication: Discourse and Cognition in Translation and Second Language Acquisition Studies, Blum-Kulka, S., House, J., Eds., Tübingen: Gunter Narr Verlag; 1986, pp. 17–36.

[36] Séguinot, C. Pragmatics and the Explicitation Hypothesis. TTR. 1988, 1(2), 106–113. https://doi.org/10.7202/037024ar

[37] Øverås, L. In Search of the Third Code: An Investigation of Norms in Literary Translation. Meta. 1998, 43(4), 557–570. https://doi.org/10.7202/003775ar

[38] Klaudy, K. Explicitation. In Routledge Encyclopedia of Translation Studies, Baker, M., Saldanha, G. Eds., London: Routledge; 2008, pp. 104–108.

[39] Baker, M. Corpus-Based Translation Studies: The Challenges That Lie Ahead. In Terminology, LSP and Translation: Studies in Language Engineering in Honour of Juan C. Sager, Somers, H., Ed., Amsterdam: John Benjamins Publishing Company; 1996, pp. 175–186.