International Journal of English Literature and Social Sciences
Vol-11, Issue-3; May-Jun, 2026

Peer-Reviewed Journal
Journal Home Page Available: https://ijels.com/
Journal DOI: 10.22161/ijels

Translation Errors and Post-editing Strategies in ChatGPT-assisted Translation of Chinese Political Texts: A Case Study of *The Report on China's Right to Development*

Lu Zhou

School of Foreign Languages, Nanjing University of Information Science and Technology, Nanjing, China
Email: 1820661942@qq.com

Received: 29 May 2026; Received in revised form: 23 Jun 2026; Accepted: 27 Jun 2026; Available online: 30 Jun 2026

Abstract**—** *Large language models are increasingly being incorporated into translation workflows, yet their performance in politically sensitive texts remains underexplored. Political discourse poses particular challenges for machine translation because it relies heavily on specialized terminology, ideological meanings, and context-dependent expressions. Drawing on Chapter Three of The Report on China's Right to Development, this paper examines translation errors in ChatGPT-generated output through a comparison with revised human translations. The analysis identifies three recurring problem areas: unnatural linguistic choices, weak information organization, and non-standard rendering of political terms. These problems are traced to limitations in probabilistic text generation, insufficient domain-specific knowledge, and incomplete contextual interpretation. To address them, three post-editing strategies are proposed: register adjustment, structural reorganization, and terminology standardization. The findings suggest that while ChatGPT is capable of producing fluent drafts, high-quality translation of political texts still depends on human expertise. The study highlights the continuing importance of post-editing in ensuring linguistic accuracy, conceptual precision, and discourse appropriateness in AI-assisted political translation.*

Keywords**— *ChatGPT, machine translation, political text translation, post-editing, translation errors***

I. INTRODUCTION

The rapid development of artificial intelligence has significantly transformed translation practices. Among recent advancements, large language models such as ChatGPT have demonstrated strong capabilities in generating fluent multilingual output. However, despite improvements in fluency, concerns remain regarding their performance in specialized discourse domains, particularly political texts characterized by ideological specificity and institutionalized terminology.

Chinese political discourse presents unique challenges for machine translation due to its abstract conceptual structures and culturally embedded expressions. While neural machine translation systems have improved lexical and syntactic processing, their ability to handle pragmatic meaning and discourse-level coherence remains limited.

Against this background, this study aims to examine translation errors in ChatGPT-generated translations of Chinese political texts and explore corresponding post-editing strategies.

II. LITERATURE REVIEW

Machine translation quality has been widely studied from both statistical and neural perspectives. Early evaluation methods mainly relied on automatic metrics such as BLEU to measure translation quality (Papineni et al., 2002; Koehn, 2009). With the development of neural machine translation, research focus has gradually shifted from purely lexical similarity toward fluency, contextual modeling, and overall translation adequacy (Bahdanau et al., 2015; Toral & Sánchez-Cartagena, 2017). Recent studies further highlight that neural-based systems tend to produce more fluent outputs but still face challenges in maintaining semantic and discourse-level consistency.

Studies on neural machine translation suggest that although fluency has improved significantly, issues such as mistranslation, omission, and unnatural expression persist (Castilho et al., 2017). Post-editing research further indicates that machine translation errors impose varying cognitive loads on translators, particularly when structural reformulation is required (Koponen, 2016; Moorkens, 2018).

Recent studies have begun to examine large language models in translation tasks. Research has shown that while ChatGPT performs well in general translation tasks, its performance in domain-specific translation can be further improved through appropriate prompting and domain-specific guidance (Peng et al., 2023; Huang & Zhao, 2026). However, relatively few studies have systematically examined ChatGPT-assisted translation of Chinese political discourse, leaving room for further research in this area..

III. METHODOLOGY

This study adopts a qualitative corpus-based approach. The research corpus is Chapter Three of the Report on China's Right to Development, which is representative of Chinese political discourse with dense terminology and abstract conceptual structures.

The analysis is conducted through a three-text comparison framework involving:

- Source Text (ST)
- ChatGPT-generated Translation (MT)
- Revised Human Translation (HT)

Instances of divergence between MT and HT were identified through close reading and categorized into three types: unnatural expressions, information structure imbalance, and terminological deviation.

IV. TRANSLATION ERRORS AND THEIR POST-EDITING STRATEGIES

4.1 Unnatural Expressions

Source Text (ST):

20 世纪 70 年代，在全球性问题的加剧以及"能源危机"的冲击下，人们的环境意识如大梦觉醒，各种环保组织和人士将环境问题的严峻性揭示给公众和国际社会，有关"增长的极限"开始在世界范围内展开，并掀起了环保运动的浪潮。

ChatGPT Output (MT):

In the 1970s, driven by the intensification of global issues and the shock of the "energy crisis," environmental awareness suddenly awakened among people.

Revised Translation (HT):

In the 1970s, against the backdrop of growing global challenges and the impact of the "energy crisis," public awareness of environmental issues began to rise.

Although the machine-generated translation successfully conveys the basic meaning of the source text, it exhibits a lack of naturalness in target-language expression. The phrase "environmental awareness suddenly awakened" is a literal rendering of the metaphor "如大梦觉醒," which leads to an unnatural collocation in English.

While the verb awaken is grammatically acceptable, it is rarely used with abstract nouns such as awareness in this semantic configuration. In English, collocations such as raise awareness, increase awareness, or develop awareness are more conventional. The machine translation therefore reflects a direct mapping of source-language metaphorical structure without sufficient adaptation to target-language usage norms.

In addition, the metaphor "如大梦觉醒" functions rhetorically in Chinese political discourse to emphasize a sudden collective realization. However, English academic and policy-oriented discourse tends to favor neutral and non-personified expressions. The machine translation retains the metaphorical structure but fails to adjust its stylistic realization, resulting in a text that appears stylistically marked and less idiomatic.

By contrast, the revised translation replaces the metaphorical expression with the collocation "public awareness… began to rise," which preserves the semantic content while aligning with conventional English usage in academic and policy contexts.

4.2 Information Structure Imbalance

Source Text (ST):

生态发展权是新时期发展权的一项子权利，是在遵循人与自然与社会三者和谐发展规律的基础上，获得物质与精神成果的权利。

ChatGPT Output (MT):

Ecological development rights are a sub-right of development rights in the new era. Based on the principles of harmonious development among humanity, nature, and society, it refers to the right to obtain both material and spiritual achievements.

Revised Translation (HT):

The right to ecological development, a component of the right to development in the new era, refers to the right to obtain both material prosperity and spiritual enrichment based on the harmonious development of humanity, nature, and society.

The machine translation divides the original sentence into two independent clauses, reflecting a tendency to segment Chinese syntactic structures into shorter English sentences. While this strategy improves readability at a superficial level, it disrupts the logical hierarchy of the definition.

In the source text, the sentence functions as a unified conceptual definition. However, the machine output separates background information from the definitional core, weakening syntactic cohesion. As a result, the conceptual relationship between"sub-right status" and"definition of ecological development rights" becomes less integrated.

The revised translation restructures the sentence into a single complex sentence. The phrase "a component of the right to development in the new era" is embedded as an appositive, allowing the main subject"the right to ecological development" to remain syntactically and semantically central. This structure better reflects English academic conventions, which prioritize hierarchical integration of information.

4.3 Terminological Deviation

Source Text (ST):

中国虽然在经济发展上取得了巨大的成就，但贫富差距，城乡差距不断拉大，区域发展不平衡日益严重。基于这一发展现状，"五位一体"发展权要求在政治、经济、文化、社会、生态等领域向弱势群体进行倾斜保护。

ChatGPT Output (MT):

Although China has achieved remarkable economic success, the wealth gap, urban-rural disparity, and regional development imbalance are continuously widening. Based on this development reality, the"five-in-one" development rights framework requires inclined protection of vulnerable groups in the political, economic, cultural, social, and ecological domains.

Revised Translation (HT):

Although China has made significant progress in economic development, the gap between the rich and the poor, urban-rural disparities, and regional imbalances continue to widen. In response, the"Five-Sphere Integrated Plan" for the right to development calls for preferential protection of vulnerable groups across the political, economic, cultural, social, and ecological domains.

The term"五位一体" is a highly institutionalized concept in Chinese political discourse and is conventionally translated as"Five-Sphere Integrated Plan." The machine translation"five-in-one" represents a literal word-for-word rendering that fails to activate its established political equivalence in English discourse. As a result, the expression lacks interpretive clarity for international readers unfamiliar with the Chinese political terminology system.

Similarly, the phrase"倾斜保护" is translated as"inclined protection," which reflects a direct lexical transfer without consideration of semantic convention in English policy discourse. The term inclined in English primarily refers to physical orientation or personal tendency and is not conventionally used to describe policy-based preferential treatment. The revised translation adopts "preferential protection," which is widely used in political and policy contexts and thus ensures terminological accuracy and communicative adequacy.

4.4 Summary

Overall, the main issues in ChatGPT-generated translations of political texts are not limited to basic semantic transfer, but are more prominently reflected in stylistic naturalness, discourse organization, and terminological standardization. These findings indicate that as machine translation systems improve in fluency, their limitations shift from sentence-level translatability to discourse-level adequacy. Consequently, post-editing remains an indispensable stage in political text translation, particularly in ensuring stylistic appropriateness, conceptual precision, and terminological consistency.

V. CAUSES of TRANSLATION ERRORS AND POST-EDITING STRATEGIES

The analysis of machine-translated outputs shows that the identified errors are not random but follow systematic patterns. Based on both the corpus data and relevant literature, the causes of translation errors in ChatGPT-generated political texts can be explained from three perspectives: the language generation mechanism, the role of training data, and limitations in contextual understanding.

5.1 Limitations of the Language Generation Mechanism

Neural machine translation systems are primarily trained on large-scale corpora and rely on probabilistic prediction of target-language sequences based on contextual probability distributions. As a result, the generated output tends to favor expressions with higher statistical frequency rather than those that are pragmatically or stylistically optimal.

Castilho et al. (2017), in a comparative study of neural machine translation (NMT) and phrase-based statistical machine translation (PBSMT), found that NMT significantly improves fluency and readability, with fewer word order and morphological errors. However, issues such as mistranslation, omission, and unnatural expression remain persistent. The study further highlights that different language pairs exhibit different error profiles, and that word order errors, mistranslations, and structural imbalance continue to impose significant challenges for post-editing quality.

Moreover, previous research indicates that different types of MT errors impose varying cognitive loads on post-editors. Compared with surface-level edits such as lexical substitution or minor adjustments, structural reformulation, mistranslation correction, and idiomatic adaptation require substantially higher cognitive effort (Koponen, 2016). Importantly, even seemingly acceptable outputs may conceal deeper pragmatic or structural issues, which increases the complexity of post-editing tasks.

This pattern is evident in the present corpus. For instance, "人们的环境意识如大梦觉醒" is translated as "environmental awareness suddenly awakened." Although the meaning is broadly preserved, the collocation is unnatural in English usage. The verb awaken is rarely used with abstract nouns such as awareness, making the expression stylistically marked and syntactically rigid.

This reflects a reliance on statistical co-occurrence patterns rather than constraints of natural collocation in the target language, leading to what may be described as "semantically accurate but pragmatically distorted" translation.

In response to such issues, post-editing should prioritize register adjustment and collocational refinement. Expressions such as awareness awakened should be revised into more conventional forms such as awareness increased or awareness began to grow, thereby improving target-language naturalness and acceptability.

5.2 Corpus-driven Terminological Deviations

The quality of machine translation output is heavily dependent on its training data. However, political discourse contains highly institutionalized terms with fixed ideological and policy meanings, which are often underrepresented or inconsistently represented in general corpora.

Recent studies show that although neural machine translation systems have achieved substantial improvements in fluency, they still struggle with terminological consistency in specialized domains, particularly when domain adaptation is limited (Dinu et al., 2019). Such issues are not always"errors" in a strict sense, but rather non-standard expressions that are interpretable yet not conventionally acceptable.

In the present data, the term"五位一体" is translated as"five-in-one," which illustrates this problem. While the expression is semantically transparent, it fails to reflect its institutionalized meaning within Chinese political discourse. Instead of adopting the established equivalent, the system resorts to a literal compositional strategy.

This indicates that, in the absence of domain-specific constraints, machine translation systems tend to prioritize surface form equivalence over terminological standardization, resulting in outputs that are understandable but not norm-compliant.

Therefore, post-editing should focus on terminology alignment. Translators should consult authoritative sources such as official government documents and bilingual policy texts to ensure consistency. For instance, five-in-one should be revised as The Five-Sphere Integrated Plan, and inclined protection should be replaced by preferential protection. Such standardization significantly improves both professional quality and discursive authority.

5.3 Lack of Contextual Understanding

Another major source of translation errors lies in insufficient contextual modeling. Political texts are deeply embedded in specific policy frameworks and discourse systems. Their meaning is constructed not only at the lexical level but also through pragmatic intent and discourse function.

Previous studies suggest that while neural machine translation systems have improved sentence-level semantic processing, they remain limited in handling pragmatic meaning and contextual inference (Toral et al., 2018). In particular, when dealing with abstract concepts or implicit logical relations, models tend to perform surface-level lexical substitution rather than context-sensitive interpretation.

For example, "倾斜保护" is translated as inclined protection. Although lexically aligned, the term inclined in English typically refers to physical orientation or personal tendency, and does not convey the policy meaning of preferential resource allocation. This results in a pragmatic shift rather than a purely lexical error.

It is important to distinguish this type of error from terminological deviation discussed in Section 4.2. The issue here is not the absence of a standardized equivalent, but the failure to extend meaning based on contextual constraints, leading to what can be described as"lexically correct but pragmatically inadequate" translation.

This limitation is also evident in long and complex sentences. In definitions such as that of "ecological development rights", machine translation often segments sentences to reduce processing difficulty. However, such segmentation weakens logical cohesion and disrupts the hierarchical structure of information.

To address this issue, post-editing should emphasize structural reconstruction. Translators need to identify core and subordinate information and reorganize them using subordination, embedding, or nominalization strategies, thereby restoring textual coherence and discourse hierarchy.

VI. CONCLUSION

This study examined translation errors in ChatGPT-generated translations of Chinese political texts, based on Chapter Three of The Report on China's Right to Development. Three main types of errors were identified: unnatural expressions, information structure imbalance, and terminological deviation.

The findings show that ChatGPT can generally convey the basic meaning of the source text with fluent output. However, problems arise at the level of stylistic naturalness, discourse organization, and terminological accuracy, especially in context-dependent political discourse.

These errors suggest that the limitations of large language models are no longer mainly related to basic comprehension, but to discourse-level and pragmatic adequacy. Machine translation may be fluent, but not always appropriate for political communication.

Therefore, post-editing remains essential. It functions not only as error correction, but also as a process of refining style, restoring structure, and ensuring terminological consistency in AI-assisted translation.

Future research may expand the dataset and apply quantitative methods to further validate these findings.

REFERENCES

[1] Castilho, S., Moorkens, J., Gaspari, F., et al. (2017). A comparative quality evaluation of PBSMT and NMT using professional translators. *Frontiers in Psychology*, 8, 1282.

[2] Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. (2002). BLEU: A method for automatic evaluation of machine translation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics* (pp. 311–318).

[3] Koehn, P. (2009). *Statistical machine translation*. Cambridge University Press.

[4] Bahdanau, D., Cho, K., & Bengio, Y. (2015). Neural machine translation by jointly learning to align and translate. In *Proceedings of ICLR*.

[5] Toral, A., & Sánchez-Cartagena, V. M. (2017). A multifaceted evaluation of neural versus statistical machine translation for 9 language directions. *In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics.*

[6] Moorkens, J. (2018). Eye tracking as a measure of cognitive effort for post-editing of machine translation. In *Eye Tracking and Multidisciplinary Studies on Translation* (pp. 55–70).

[7] Koponen, M. (2016). Is machine translation post-editing worth the effort? A survey of research into post-editing and effort. *The Journal of Specialised Translation*, 25, 131–148.

[8] 黄协安，赵善江. (2026). 国内外主流大语言模型中译英翻译质量的实证对比分析. *语言与翻译*, (1), 50–59, 66.

[9] Dinu, G., Mathur, P., Federico, M., et al. (2019). Training neural machine translation to apply terminology constraints. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 3063–3068).

[10] Toral, A., Castilho, S., Hu, K., et al. (2018). Attaining the unattainable? Reassessing claims of human parity in neural machine translation. In *Proceedings of the Third Conference on Machine Translation: Research Papers* (pp. 113–123).

[11] Peng, K., Ding, L., Zhong, Q., Shen, L., Liu, X., Zhang, M., Ouyang, Y., & Tao, D. (2023). Towards making the most of ChatGPT for machine translation. *Findings of the Association for Computational Linguistics: EMNLP* 2023, 8402–8418.