Original Paper

Research on Cultural Subjectivity Challenges and Strategies for AI Translation in the Era of Large Language Models

Xinxing Li[1]

[1] Henan Normal University, Henan province, China

Abstract

In the context of the accelerated globalization process, cross-language communication is becoming increasingly frequent, involving a wider variety of languages and cultural differences. This has driven a rapid growth in the demand for multilingual translation. This trend has led to the widespread application and in-depth research of large language models based on artificial intelligence in the field of translation. However, the diversity and complexity of languages present numerous challenges for these language models during their development, especially when it comes to accurately addressing cultural differences between source and target languages in translation practice. Specifically, AI translation tools often struggle to fully convey the cultural connotations and values of the source language in English translations, and may even fail to adequately address or downplay cultural differences. This can result in mistranslations or hollow translations at the cultural level.

Keywords

Large language models, AI translation, cultural subjectivity, challenges, strategies

1. Introduction

In recent years, the application of large language models (LLMs) in machine translation has become increasingly widespread. Their breakthroughs in understanding and generating language have brought machine translation closer to the essence of language. For example, Baidu Translate has launched an intelligent translation service based on large model technology, offering a one-stop professional translation solution for all scenarios, addressing issues that traditional machine translation struggles with, such as understanding the original text and incorporating context. In October 2024, Alibaba introduced its newly upgraded AI translation tool, MarcoMT, which was performance-evaluated based on the Flores translation benchmark framework. DeepL, a company specializing in high-precision AI translation, launched new features for its DeepL Pro in 2024. The new version not only offers

higher-quality translation services but also includes industry-specific terminology databases. Additionally, there are many other developments and upgrades in AI translation tools.

With the advancement of AI translation technology, the efficiency of information transmission is often considered the most important goal, while cultural representation is often regarded as secondary. This has sparked new reflections on the development of translation technology and ethical issues, challenging cultural subjectivity. Modern internet technology is the technical cause behind cultural nihilism (Zhang, 2024, p. 28). Furthermore, the use of AI translation tools, especially popular ones like ChatGPT, brings about a series of problems. In the translation process, ChatGPT may exhibit biases on sensitive topics such as gender, race, and religion (Zhang, 2023, p. 80). In addition, ChatGPT may show biases toward certain cultures and languages in its translations, such as overly emphasizing the values of some cultures while ignoring the characteristics and needs of others (Yu & Guo, 2024, p. 119). This can affect the fairness and accuracy of the translation.

This paper will conduct an in-depth analysis of the challenges to cultural subjectivity in the AI translation process and the strategies for addressing them. By examining specific translation cases, it will reveal the cultural subjectivity dilemmas currently faced by AI translation in the context of globalization and propose corresponding optimization strategies. The aim of this research is to enhance translation quality and efficiency, ensuring that AI translation tools can more accurately reflect cultural differences. It also seeks to promote the gradual improvement of cultural adaptability in translation tools during cross-cultural information transmission, thereby achieving more precise, efficient, and culturally sensitive cultural representation.

1.1 The Conflict between Hegemony in AI Translation and Cultural Subjectivity

1.1.1 Cultural Hegemony

The issues in AI translation are primarily manifested in the fact that these models may inherit the viewpoints and biases of the original authors, and human intervention in the translation process can also introduce the subjective biases of the reviewers. Since the pre-training corpora mainly come from Western societies and cultures, the models are prone to being influenced by Western centrism. (Wen & Tian, 2024, p. 32) For example, the political term "Publicity Department of the Communist Party of China Central Committee" has many AI translation tools that translate it as "Propaganda Department of the Communist Party of China". Obviously, although "propaganda" means "advocate", But it also contains the derogatory meaning of "to advocate in a biased or misleading way", which is easy to mislead the target language readers. Large language models (LLMs) are typically trained using vast amounts of internet data, which often contain biases. For example, internet data may primarily be generated by young male authors from developed countries, which could lead to the model exhibiting bias when handling certain topics. (Jurafsky, 2000, p. 207) This can result in the transmission of incorrect values to readers. Translators have found that AI translation frequently translates "XX Leader" as "Mr. XX." Furthermore, on January 16, 2024, International Business Machines Corporation (IBM) released the *2023 Global AI Adoption Index* which surveyed 8,500 global IT professionals

regarding the fastest-growing enterprise use cases for generative AI—IT automation, digital workforce, and customer service. The survey revealed that 42% of companies use AI screening tools, but many experiments have shown that these tools tend to use biased parameters, such as favoring younger candidates or men over women.

1.1.2 Information Technology Hegemony

John Zysman (2006) introduced the concept of the "algorithmic revolution" in his article (Zysman, 2006), which became a core element driving the transformation of the entire digital economy. (Li, Y., & Li, J., 2023, pp. 73-84) Subsequently, American scholar Frank Luciano explicitly proposed "digital imperialism" in 2003. (Frank & Huang, 2003) Comprehensive algorithms are forming a global surveillance system, leading to digital hegemony. (Liu, 2021) Information technology, as a derivative of "algorithms," along with large language models and AI translation tools, represents a new era of information technology. These technologies, in their data collection, algorithm analysis, and text processing, reflect the values and culture of Western countries. Moreover, according to data released by the China Academy of Information and Communications Technology, the Americas region has the largest AI market, projected to reach $288.6 billion in 2023, accounting for 64.3% of the AI software market. In contrast, the Asia-Pacific market is smaller, estimated at $55 billion in 2023, which shows that production and capital are increasingly concentrated in the hands of large digital multinational corporations. This concentration gives rise to a series of issues. AI translation systems, when processing information, may replicate mainstream viewpoints and marginalize alternative knowledge, thereby reinforcing existing power structures and knowledge systems, resulting in a form of informational hegemony. (Lindemann, 2024, pp. 1-14) Furthermore, OpenAI's release of the trillion-parameter multimodal pre-trained model GPT-4, along with currently circulating AI chips and artificial large models such as AIGC, still see Western countries dominate the market. These developments demonstrate their absolute control over the global digital market.

In this context, AI translation tools, whether in data input, text analysis, translation generation, or output processes, are inevitably influenced by Western culture. The text analysis in AI translation tools primarily involves natural language processing (NLP). As one of the key tools in the field of natural language processing, large-scale language models describe patterns mathematically and learn from massive amounts of textual data for text selection. The corpus, as the foundation of large-scale language models, directly affects AI translation in terms of its size and diversity. However, not all data retrieved from large datasets are necessarily accurate. (Veljko, 2020, p. 206) Issues in the algorithm design of foreign large language models can lead to translation results that tend to reflect Western cultural perspectives and expressions. (Wang & Li, 2024) For example, when translating the political term "make China a favored destination for foreign investment", large language models may simplistically translate it as "Invest in China," Such translation methods fail to fully consider China's unique historical context, development process, and cultural connotations, resulting in an inability to accurately convey the political and cultural significance contained in the original text, which could

potentially affect students' correct understanding of China's political system and development path.

2. The Impact of Hegemony in AI Translation on Cultural Subjectivity

2.1 The Marginalization of Traditional Culture

With the widespread use of AI translation tools, many local cultural traditions and folk arts in translated texts may be influenced by hegemonic cultures and thus marginalized. The rapid development and technological progress of modern society promote a "more efficient" and "universal" cultural model, often neglecting the uniqueness and depth of traditional cultures. Cultural values rooted in history, religion, language, and customs may be regarded as outdated or inconsistent with modern trends, and even gradually disappear. (Wang, Li, & Liang, 2024) Under the influence of cultural hegemony, the subjectivity of traditional cultures may be weakened, and many indigenous values may be seen as "irrational" or "useless," accelerating their decline in the context of globalization.

2.2 Cultural Homogenization

The development of AI translation technology may exacerbate the erosion of cultural subjectivity by hegemonistic forces, leading to the marginalization or neglect of certain cultures, thereby weakening the value of cultural diversity in global society. Driven by this cultural hegemony, different cultural traits and traditions are easily replaced by a unified global cultural standard, resulting in the disappearance of minority cultures, non-mainstream cultures, or local cultures. During the process of globalization, the uniqueness of specific regions, ethnic groups, and belief systems may be overlooked, making cultural diversity more vulnerable. This uniformity of cultural hegemony not only erases cultural pluralism but may also lead to the loss of traditional knowledge and cultural creativity, and even trigger the dissolution of national character and identity.

2.3 The Diminution of Cultural Identity

AI translation tools often assume that all cultural differences are relative, and that the boundaries between cultures can be blurred or even dissolved. This hegemonic way of thinking may gradually weaken people's sense of identification with their own original culture. Traditional cultural values, customs, and historical memories are increasingly viewed as outdated and unnecessary. Especially under the influence of globalization and information technology, many local cultural features and traditions are replaced by standardized global culture, making cultural identity increasingly vague. Cultural hegemony, by diminishing cultural differences, may reduce individuals' or groups' attachment to and pride in traditional culture, thereby affecting the formation of cultural subjectivity.

3. Strategies to Counteract Hegemony in AI Translation

3.1 At the National Level

Currently, bilingual parallel corpora internationally are largely centered around English, and the multilingual service industry has become the main arena for the United States to seek its discourse influence. Therefore, it is imperative for the country to accelerate the construction of multilingual

Chinese-centered parallel corpora. (Wang, 2022) The government should actively promote the development of large-scale localized training datasets to ensure the diversity and compatibility of these datasets, reducing biases and mistranslations caused by data imbalance. This not only helps break the Western-centric model of corpus construction but also allows for the proactive control of cultural dissemination. (Wang, 2024) In addition, the government should actively support the research and development of large language models led by Chinese, enhancing the ability to process Chinese information and improve Chinese-to-foreign translation capabilities. Through policy support, financial investment, and talent cultivation, the country can gradually reduce dependence on foreign models, thus avoiding infiltration or manipulation by hegemonic national ideologies. The development of AI translation tools requires the state to strengthen the governance of artificial intelligence, actively promote international cooperation in the field of AI development, and form consensus to prevent certain countries or regions from using AI technology to promote unilateralism and hegemonism. Of course, this requires the international community to work together to build a fair and reasonable AI governance framework.

3.2 At Social Level

In today's rapidly developing global society, it is important for society to focus on cultivating the public's cultural critical awareness, guiding people to recognize the limitations of AI translation tools, and encouraging the use of language expressions that align with local cultural values. This approach helps avoid the risks of cultural homogenization and unification, prevents political conflicts arising from cultural clashes, and highlights the cultural tension unique to China. (Sun, 2008) Especially when translating texts related to values, moral norms, and social ethics, it is essential to deeply understand the cultural background behind them. When applying large language models, the human-centered approach should be adhered to, strengthening critical thinking abilities, enhancing awareness of technological ethics, and clarifying the boundaries and relationships in human-machine collaboration. This ensures that translation teaching is not only humane but also deep and innovative, effectively addressing the challenges brought about by over-reliance on large language models. (Wang & Xie, 2024)

Society should strengthen the ethical and legal regulation of AI translation to ensure that AI translation tools respect user privacy, protect intellectual property, and avoid generating misleading information or biases. (Li, Du, & Ma, 2022) Furthermore, AI translation systems need to have cultural awareness, enabling them to understand and address language differences caused by cultural disparities; otherwise, misunderstandings or even offenses may occur.

3.3 At the Individual Level

Individuals should first enhance their cultural awareness and cultural confidence, cultivate cross-cultural sensitivity, and fully recognize that language is not just a conversion of words but a carrier of culture. When using AI translation tools, they should remain vigilant and avoid blindly adopting Western cultural expressions. In addition, individuals should continually improve their

translation skills, especially when translating content involving complex cultural and ideological topics, and try to avoid relying solely on AI translation tools. By employing human editing and proofreading, translation quality can be enhanced to ensure both accuracy and cultural appropriateness. In daily life, individuals can participate in translation practices, such as literary translation and academic translation, to hone their translation abilities, while also reflecting on and improving the limitations of AI translation tools. By comparing the effects of AI translation and human translation, individuals can better understand the limitations of algorithms and the creativity inherent in human translation.

4. Conclusion

With the rapid development of artificial intelligence technology, especially the widespread application of advanced large language models like ChatGPT in the field of translation, translation teaching and practice are undergoing profound changes. However, the use of these technologies has also raised issues of cultural hegemony, particularly in areas such as data sources, algorithm design, values, and discourse structures. For example, foreign large language models, which rely primarily on data trained from Western societies, often carry Western-centric biases, which are particularly evident when translating texts related to Chinese history, culture, and politics. To address these challenges, research has proposed several strategies. First, it is necessary to build diverse and inclusive corpora to reduce biases and mistranslations caused by data imbalance, thus breaking the Western-centric model of corpus construction. Second, the logic of AI translation must be carefully examined to avoid over-reliance on technology. Students should be encouraged to combine their professional knowledge to verify and optimize AI translation results, thus preventing cultural misinterpretation and other fundamental errors. In addition, it is essential to strengthen cultural critical awareness, cultivate students' cultural consciousness and confidence, safeguard cultural diversity, and promote the healthy development of ideology in translation education. Through these strategies, the cultural hegemony traps posed by large language models can be effectively avoided, fostering the sustainable development of translation education and contributing to the protection of national cultural security and the enhancement of cultural soft power.

References

Frank, L. C., & Huang, H. L. (2003). Digital Imperialism and Cultural Imperialism. *Marxism and Reality*, *5*.

Jurafsky, D. (2000). *Speech and language processing*.

Li, J., Du, S. Y., Ma, W. Y. et al. (2022). Research on Translation Ethics in the Age of Artificial Intelligence. *Science and Technology Think Tank*, *2022*(11), 62-69.

Li, Y., & Li, J. (2023). Political Economy Critique of Digital Imperialism under the Background of the "Algorithm Revolution". *Teaching and Research*, (12), 73-84.

Lindemann, N. F. (2024). Chatbots, search engines, and the sealing of knowledges. *AI & SOCIETY*,

1-14. https://doi.org/10.1007/s00146-024-01944-w

Liu, H. Y. (2021). Data Hegemony and the New Predation of Digital Imperialism. *Contemporary Economic Research*, *2*.

Source: China Academy of Information and Communications Technology: New Landscape of the Global Digital Economy. (2023). China Academy of Information and Communications Technology. *Global Digital Economy White Paper*.

Sun, Y. F. (2008). Cultural Translation and Global Localization. *Chinese Translators Journal*, (01), 5-11 + 95.

Veljko, K. (2020). *Succeeding with AI: How to make AI work for your business* (p. 206). Manning Publications Co.

Wang, H. S., & Xie, F. (2024). A Study on the Innovation of Translation Education Practice Models Driven by Large Language Model Technology. *Chinese Translators Journal*, *45*(02), 70-78.

Wang, H. S., Li, D., & Liang, X. R. (2024). Cultural Traps and Pathways to Breakthrough: A Study on Resistance Strategies Against Cultural Hegemony in Translation Teaching in the Era of Large Language Models. *Foreign Language Education Research*, *12*(04), 2-10.

Wang, H. S., Li, D., & Liang, X. R. (2024). Cultural Traps and Pathways to Breakthrough: A Study on Resistance Strategies Against Cultural Hegemony in Translation Teaching in the Era of Large Language Models. *Foreign Language Education Research*, *12*(04), 210.

Wang, K. F. (2022). Development and Application of a Multilingual Chinese-English Parallel Corpus Cluster Centered on Chinese. *Foreign Language Teaching*, *43*(06), 1-7.

Wen, X., & Tian, Y. L. (2024). A Study on the Effectiveness of ChatGPT in Translating Discourse with Chinese Characteristics. *Shanghai Translation*, *2024*(2), 32.

Yu, H., & Guo, Y. Y. (2024). Risk and Transcendence: An Ethical Analysis of ChatGPT Empowering Translation. *Chinese Translators Journal*, *45*(04), 115-122.

Zhang, Y. P. (2024). Critique and Correction of Contemporary Chinese Cultural Nihilism Thought. *School Party Building and Ideological Education*, (17), 27-30.

Zhang, Y. X. (2023). "Design Protection" of Personal Information in Human-Computer Dialogue—A Case Study of the ChatGPT Model. *Library Forum*, *2023*(8).

Zysman, J., & Newman, A. (2006). *How Revolutionary was the Digital Revolution?: National Responses, Market Transitions, and Global Technology*. Stanford University Press. https://doi.org/10.1515/9781503625730