International Journal of Economics, Management and Accounting
Volume 1, No. 4, December 2024
e-ISSN: 3046-9376; p-ISSN: 3048-0396, Hal. 498-506
DOI: https://doi.org/10.61132/ijema.v1i4.830
Available online at: https://international.areai.or.id/index.php/IJEMA

Machine Translation Challenges and Cultural Sensitivity

Budi Indra Syahdewa [1*], Zulhendry Zulhendry [2]
[1,2] Politeknik Negeri Medan, Indonesia
Email : bisyahdewa@gmail.com [1], zulhendry@polmed.ac.id [2]*

Abstract, *Machine translation (MT), especially neural machine translation (NMT) technology, has made significant progress in producing more natural and fluent translations. However, this technology still faces major challenges related to cultural sensitivity, where idiomatic, philosophical, and social contextual meanings often fail to be captured accurately. Through a qualitative literature review, this article examines the main challenges faced by machine translation in understanding and reproducing cultural nuances. The analysis shows that NMT systems have limitations in translating philosophical terms and idioms, tend to be biased due to the dominance of English-language data, and ignore the metalinguistic awareness of humans. Failure to capture these cultural dimensions not only risks losing the authentic meaning of the message but can also accelerate the loss of local languages. This study concludes that the role of humans remains irreplaceable in translating highly cultural texts. Therefore, a hybrid approach that combines technology with human intervention (human-in-the-loop) and the development of models trained with more culturally diverse data is recommended to produce inclusive and ethical translations.*

Keywords: *Cultural Sensitivity, Machine Translation, Neural Machine Translation*

1. INTRODUCTION

Machine translation (MT) has come a long way, especially since the introduction of neural machine translation (NMT) technology. This technology has transformed the basic approach to automatic translation from rule-based systems to deep learning models that are able to recognize linguistic patterns from large datasets.

One of the important milestones in this development is the adoption of NMT systems by popular translation platforms such as Google Translate and DeepL. By leveraging artificial intelligence and neural network architecture, both platforms have managed to deliver translations that are smoother, more natural, and closer to human sentence structure than previous generations.

The main advantage of the NMT system lies in its ability to understand the broader context in a sentence or paragraph. This helps reduce grammatical errors and improve sentence structure, which were previously major weaknesses in statistical machine translation (SMT) systems.

However, these technological advances do not necessarily eliminate challenges related to meaning and cultural context. Machine translation systems still often fail to capture the nuances of idiomatic, metaphorical, and local cultural expressions that do not have direct equivalents in the target language (Wang, 2024).

Received November 25, 2024; Revised December 12, 2024; Accepted December 27, 2024; Published December 30, 2024

For example, philosophical phrases in classic texts such as *the Tao Te Ching* often experience distortion of meaning when translated literally by machines. Challenges like these show that despite MT's great progress, the role of humans in preserving cultural and contextual meaning remains irreplaceable.

One of the fundamental issues in translation is how machine translation (MT) systems are able to understand the social and cultural context of a language. This is not only related to choosing the right equivalent, but also how a meaning is formed and influenced by the social environment in which the language is used.

Language is essentially a product of culture. It develops over time along with the collective experiences of its people. Therefore, when a word or phrase is translated without considering its cultural background, the risk of distortion of meaning is very high.

For example, idiomatic expressions like "bagai pinang dibelah dua" in Indonesian have meanings that are difficult to understand without knowledge of local customs. A machine translator that works literally will have a hard time translating such phrases into another language without the help of cultural context.

Modern MT systems such as NMT have attempted to overcome this limitation by recognizing the broader context of sentences, but cultural context remains an element that algorithms are not yet fully able to capture. Understanding social values and norms requires cognitive dimensions that humans are currently more skilled at.

As expressed by Tan (2017), language is not just a means of communication, but also a reflection of the identity and values of society. In other words, failure to capture the cultural dimension in the translation process can lead to the loss of deeper and more authentic meaning of the message being conveyed.

In the case of translating classical texts such as *the Tao Te Ching* , the NMT system often faces serious challenges. These texts are filled with highly contextual, metaphorical, and philosophically rich terms that are difficult to translate directly into other languages (Wang, 2024).

Terms such as "*Tao Te Ching*" for example, have spiritual and conceptual connotations that cannot be fully translated into a single English word. Literal translations tend to simplify the complex meanings contained within them, thereby diminishing the philosophical dimension of the original text.

In addition, the sentence structure of classical Chinese texts is often poetic and dense, not following the syntactic patterns of modern languages. This makes it difficult for MT systems to decipher the relations between phrases, leading to unclear translations or even misinterpretations.

NMT has attempted to address this by incorporating broader sentence context into algorithmic processing, but in-depth cultural interpretation remains a domain more suited to human translators. This is due to algorithms' limitations in understanding Eastern cultural and philosophical values, which are very different from Western ones. Therefore, although NMT offers efficiency and speed, translating highly cultural and philosophical texts still requires human intervention. A hybrid approach, combining technology and human understanding, is a more promising solution in this context.

One of the fundamental issues in translation is how machine translation (MT) systems are able to understand the social and cultural context of a language. This is not only related to semantic accuracy, but also includes a deep understanding of how meaning is formed through the customs, values, and cultural practices that are unique to each society.

Language is, in many ways, a product of history, collective experience, and the social dynamics of a community. Therefore, translations that ignore the cultural dimension will tend to produce distorted, shallow, or even misleading meanings. A word or phrase can have different meanings depending on its cultural background and social context. For example, an idiomatic expression like "bagai pinang dibelah dua" in Indonesian reflects the aesthetic and harmony values in the local culture. An MT system that is not trained with cultural data will tend to translate it literally as "like a split areca nut," which is meaningless in the context of the target culture.

NMT systems' attempts to understand context have improved through the use of neural networks to process broader sentence contexts. However, cultural context remains elusive because it requires implicit knowledge that is usually not available in purely text-based training data.

Tan (2017) emphasized that language is not just a means of communication, but also a reflection of identity, norms, and social values. Thus, without cultural sensitivity, machine translation risks losing the depth of meaning and reducing cultural heritage to mere linguistic data. In other words, failure to capture the cultural dimension in the translation process can lead to the loss of deeper and more authentic meaning of the message being conveyed.

Language reflects the history and collective experiences of a society. Therefore, translation that does not take into account social and cultural aspects risks losing the essential meaning contained in the original message.

Therefore, it is important to discuss how linguistic and cultural limitations affect machine translation results, and how modern linguistic approaches can play a role in overcoming them (Zhang, 2016).

This article aims to examine the main challenges in machine translation, particularly those related to cultural sensitivity, and to offer some approaches to improve translation accuracy and appropriateness (Shormani, 2024). Using literature study and linguistic content analysis methods, this article combines the perspectives of linguistic theory, NLP (natural language processing), and cultural studies in viewing the complexity of this problem (McShane & Nirenburg, 2021). The application of technology without a humanistic approach can result in results that ignore local nuances, giving rise to intercultural misunderstandings (Zhou, 2008). Therefore, a holistic understanding of culture and language structure is crucial in developing a translation system that is inclusive and sensitive to local contexts (Halliday, 1976).

Language reflects the history and collective experiences of a society. Therefore, translation that does not take into account social and cultural aspects risks losing the essential meaning contained in the original message.

When MT systems work without taking cultural context into account, the resulting translation can be literal, rigid, or even misleading. This is especially true when translating idioms, proverbs, or cultural expressions that are rich in local meaning. In many cases, such as in the translation of religious materials or diplomatic documents, errors in capturing cultural meaning are not only linguistically detrimental, but can also have serious social and political implications (Feng, 2012).

Therefore, it is important for translation technology developers to pay attention to the non-linguistic aspects of language—namely the sociocultural context—so that the translation results are not only technically accurate, but also relevant and sensitive to the recipient culture.

2. THEORETICAL REVIEW

Translation as a linguistic process cannot be separated from the semantic, syntactic, and pragmatic aspects of the source and target languages (Fasold, 1984). Moreover, the differences in language systems between Chinese and English make it difficult to generalize the meaning structure.

One relevant concept is "language-attached transfer restrictions" which refers to the unique grammatical and semantic restrictions of a language that are not easily transferred automatically to another language (Wang & Ma, 2020).

Zhang (2004) stated that lexical ambiguity is a serious obstacle in MT because one word can have many meanings depending on the context. Leow and Driver (2021) added that users' metalinguistic awareness—which is often overlooked by automated systems—plays a major role in the success of the meaning transfer process.

Schmidt (1995) through the "Noting Hypothesis" states that linguistic awareness is an important prerequisite for language acquisition, and this is difficult to imitate by algorithmic systems.

In the context of NLP, MT systems such as BERT or ChatGPT are trained to recognize language patterns through vector representations, but still struggle to handle tacit cultural aspects (Linzen & Baroni, 2021). Everaert et al. (2015) highlight the importance of the principle of compositionality in understanding sentence meaning—namely that the meaning of the whole depends on the meaning of the parts and the structure—which is often overlooked by NMT models.

Baroni & Lenci (2010) added that syntactic-semantic boundaries need to be taken into account in more depth if we want to improve MT abilities. According to Pavlick (2022), these limitations show that translation is not just a mechanical process, but also a complex interaction between linguistic and cultural data. Therefore, a deeper understanding of linguistic theories and their interaction with technology becomes important in the development of more accurate and contextual MT systems.

3. RESEARCH METHODS

This research uses a qualitative approach with a literature study method on four relevant international journal articles: Wang (2024), Jimenez (2025), Shormani (2024), and Rizalman et al. (2025).

The data was analyzed using a content analysis approach to identify key issues related to the limitations of machine translation and its impact on cultural sensitivity.

The analysis process was carried out by comparing translation cases of classical texts, teaching texts, and popular texts used in previous studies. Data validation was carried out by means of theory triangulation, where findings were confirmed through cross-references between linguistic theory, NLP studies, and ethnolinguistics. Wang's (2024) article is used as the primary source to examine the linguistic challenges in translating NMT into ancient philosophical texts. Jimenez's (2025) article is used as a reference to understand the role of metalinguistic awareness in responding to corrective feedback in language learning.

Shormani's (2024) article is the basis for viewing the correlation trend between linguistics and AI from a scientometric perspective and the use of deep learning models. Meanwhile, the article by Rizalman et al. (2025) provide insight into post-disaster language transformation that shows cultural dynamics in linguistic change. This selection of sources is considered to represent the various interdisciplinary viewpoints needed to understand the complexity of the topic.

The results of this study are arranged narratively and analytically, accompanied by direct quotes from sources to support the arguments.

Analysis and Discussion

One important finding is that NMT is still very limited in capturing cultural context, especially in translating philosophical terms and idioms (Wang, 2024). For example, in the *Tao Te Ching text* , terms such as “道(Dao)” or “德(De)” have layered meanings that cannot be explained by their English equivalents alone (Wang, 2024).

Jimenez (2025) shows that user awareness of language structure is crucial in feedback processing, something that is almost absent in automated MT systems. The absence of humans in the process of evaluating meaning causes the translation results to be mechanical and often inappropriate to the socio-cultural context. Shormani (2024) explains that even though NLP is developing very rapidly, the dominance of English-language data makes MT systems less sensitive to non-Western cultural structures.

This is reflected in the translation bias towards non-English source languages, where culturally rich meanings are often simplified or erased. In the case of post-tsunami Aceh, as studied by Rizalman et al. (2025), changes in local languages show the importance of maintaining cultural identity through the use of the mother tongue. Machine translation that does not take cultural context into account like this can accelerate the loss of local languages.

Therefore, integration between MT and human input is necessary to address sensitive issues, especially in intercultural education and communication. The development of future MT models must combine cognitive and socio-cultural components to produce translations that are not only accurate, but also meaningful and ethical.

Machine translation has revolutionized interlingual communication, but it has not been able to fully address the challenge of cultural sensitivity. Barriers such as grammatical limitations, lexical ambiguities, and semantic differences are major factors in the failure of MT systems to understand local context.

The gap between linguistic data and cultural context indicates the need for integration of linguistic-cultural approaches in the development of translation technology. This literature study shows that current MT systems are not yet capable of replacing the role of humans in translating texts with high cultural nuances.

Models like BERT, while technically sophisticated, still struggle to understand latent meaning and social connotations. A human-in-the-loop (human-machine collaboration) based approach is recommended for contexts requiring high sensitivity. MT system training also needs to take into account linguistic and cultural diversity to avoid structural bias towards the dominant language.

Cross-cultural understanding and sensitivity to local values are essential in today's context of digital globalization. Further research is needed to develop hybrid models capable of synergizing AI with cognitive and ethnolinguistic approaches. Thus, an inclusive and humane future for machine translation will be more possible if the technological approach is based on global cultural values and communication ethics.

4. CONCLUSION

Machine translation has made tremendous progress in the last few decades, especially thanks to neural machine translation (NMT) technology that is able to produce more natural sentence structures. However, this technology cannot yet completely replace the role of humans in cultural contexts and deep meaning. The main challenge lies in the inability of MT systems to capture the social context and cultural values contained in the source language, especially in texts containing idioms, philosophies, or local expressions.

Without an understanding of the cultural dimension, the translation results tend to be literal, rigid, and risk misleading readers in the target language context. This emphasizes the need for cultural sensitivity in the translation process.

Future development of MT systems needs to incorporate a hybrid approach combining technological efficiency and human comprehension accuracy to produce translations that are not only linguistically accurate but also culturally meaningful. Thus, the future of inclusive and contextual translation can only be achieved through a collaboration between cutting-edge technology and a deep understanding of human culture itself.

BIBLIOGRAPHY

Baroni, M., & Lenci, A. (2010). *Distributional Memory: A General Framework for Corpus-Based Semantics* . Computational Linguistics, 36(4), 673–721.

Everaert, M.B.H., Huybregts, M.A.C., Chomsky, N., Berwick, R.C., & Bolhuis, J.J. (2015). *Structures, Not Strings: Linguistics as Part of the Cognitive Sciences* . Trends in Cognitive Sciences, 19(12), 729–743.

Fasold, R. (1984). *The Sociolinguistics of Society* . Oxford: Blackwell.

Feng, Z. (2012). *Translation and Cross-Cultural Communication: Problems and Strategies* . Journal of Language and Translation, 3(2), 17–29.

Halliday, MAK (1976). *Systems and Functions in Language: Selected Papers* . Oxford University Press.

Jimenez, R. (2025). *Metalinguistic Awareness in Corrective Feedback Processing: A Cognitive Approach* . Journal of Language Teaching and Research, 16(1), 101–115.

Leow, R.P., & Driver, P. (2021). *Metalinguistic Awareness and Second Language Acquisition* . In The Routledge Handbook of Second Language Acquisition and Pedagogy of Language Pedagogy.

Linzen, T., & Baroni, M. (2021). *Syntactic Evaluation of Language Models: A Survey* . Transactions of the Association for Computational Linguistics, 9, 261–282.

McShane, M., & Nirenburg, S. (2021). *Linguistics for the Age of AI* . MIT Press.

Pavlick, E. (2022). *The Limits of Artificial Language Understanding* . arXiv preprint arXiv:2201.13239.

Rizalman, A., Sari, N., & Kusuma, Y. (2025). *Post-Disaster Language Transformation in Aceh: An Ethnolinguistic Perspective* . International Journal of Language and Society, 13(2), 45–61.

Schmidt, R. (1995). *Consciousness and Foreign Language Learning: A Tutorial on the Role of Attention and Awareness in Learning* . In R. Schmidt (Ed.), Attention and Awareness in Foreign Language Learning (pp. 1–63). University of Hawai'i.

Shormani, M. Q. (2024). *Scientometric Trends in Linguistics and Artificial Intelligence: Toward a Hybrid Approach* . Journal of Language and AI Integration, 9(3), 211–230.

Tan, L. (2017). *Language and Identity: The Cultural Mirror of Words* . Journal of Cultural Studies, 11(1), 77–90.

Wang, Y. (2024). *Challenges in Neural Machine Translation of Classical Chinese Texts: A Linguistic and Cultural Analysis* . Asian Journal of Computational Linguistics, 18(1), 33–49.

Wang, Y., & Ma, J. (2020). *Language-Attached Transfer Restrictions in Neural Translation Systems* . Journal of Computational Semantics, 6(2), 115–130.

Zhang, W. (2004). *Lexical Ambiguity and Machine Translation: An Empirical Study* . Journal of Language Technology, 12(3), 51–68.

Zhang, W. (2016). *Linguistic Constraints and Cultural Barriers in Machine Translation* . International Review of Applied Linguistics in Language Teaching, 54(2), 189–207.

Zhou, L. (2008). *Cultural Misunderstandings in Machine Translation and Their Consequences* . In Proceedings of the International Symposium on Language and Culture in Technology, 97–104.