AI TRANSLATION OF CULTURE-SPECIFIC AND IDIOMATIC EXPRESSIONS: LINGUISTIC AND CULTURAL CHALLENGES

Sa'dullayeva Shahnoza Ismatullo qizi
Qarshi davlat universiteti Lingvistika kafedrasi o'qituvchisi
DOI:https://doi.org/10.5281/zenodo.23127848

Language functions not only as a means of transmitting information but also as a repository of cultural knowledge, collective experience, and culturally shaped patterns of thought. Consequently, translation involves more than the transfer of lexical and grammatical information between languages. It also requires the interpretation and reconstruction of meanings that are embedded within particular cultural contexts. This issue becomes especially evident in the translation of idioms, proverbs, metaphors, realia, and other culture-specific expressions whose meanings cannot always be derived from their individual lexical components. In Translation Studies, such elements are commonly discussed as culture-specific items (CSIs), which may require particular strategies to preserve their semantic, pragmatic, and cultural dimensions in the target language (Aixelá, 1996).

Idiomatic expressions provide a particularly clear illustration of the relationship between language and culture. For example, the English idiom "kick the bucket" and the Uzbek expression "og'zi qulog'ida" cannot be adequately interpreted through literal translation alone. Their intended meanings depend on conventional usage, contextual information, and shared cultural knowledge. Thus, the successful translation of an idiom requires the translator to identify not only its linguistic form but also its figurative meaning, pragmatic function, and culturally appropriate equivalent. Baker (2011) similarly emphasizes the importance of dealing with non-equivalence and idiomatic expressions at the level of meaning rather than relying exclusively on formal correspondence.

Traditionally, human translators have addressed these challenges by combining linguistic competence with cultural knowledge, contextual interpretation, and pragmatic awareness. However, the rapid development of artificial intelligence (AI) and neural machine translation (NMT) has fundamentally transformed the contemporary translation environment. The introduction of Transformer-based architectures marked a significant technological development in machine translation and natural language processing. Vaswani et al. (2017) demonstrated that attention-based Transformer architectures could achieve strong performance in sequence-to-sequence tasks while enabling more efficient parallel processing. Subsequent models such as BERT and T5 further expanded the capacity of neural systems to process contextual and semantic information. BERT, for example, was designed to represent linguistic information through bidirectional contextual processing, while T5 formulated diverse NLP tasks within a unified text-to-text framework.

Despite these technological advances, the ability of AI systems to process culturally embedded language remains problematic. Recent research demonstrates that idioms and other culture-specific expressions continue to present difficulties for neural translation systems because their meanings frequently extend beyond literal lexical content. A 2024 study of Indonesian cultural idioms, for example, found that contextual understanding is an important factor in the quality of NMT output when translating culturally embedded expressions. More recent research comparing ChatGPT, DeepL, and Google Translate likewise identifies recurring problems involving the preservation of idiomatic meaning and cultural connotations.

These difficulties become particularly significant in the case of low-resource languages. Languages with relatively limited digital corpora may have insufficient representation of idioms, culturally specific

© Ushbu material Creative Commons Attribution 4.0 International (CC BY 4.0) litsenziyasi asosida tarqatiladi.

references, traditional expressions, and context-dependent meanings in the data used to train AI systems. As a consequence, even highly advanced language models may fail to recognize the sociolinguistic and cultural factors that determine how an expression functions in authentic communication. The problem is therefore not restricted to grammatical accuracy or lexical equivalence. Rather, it concerns the ability of AI systems to recognize and reproduce culturally situated meaning.

The challenge can be understood through three closely interconnected dimensions. First, AI systems may have insufficient access to culturally diverse and representative training data. Second, translation systems may produce outputs that prioritize lexical correspondence or surface-level fluency while failing to preserve figurative, pragmatic, or cultural meaning. Third, the development and evaluation of AI translation technologies requires stronger interdisciplinary cooperation between computational linguistics, Translation Studies, cultural linguistics, and related fields. Recent research on culture-aware machine translation similarly indicates that strong general translation performance does not necessarily guarantee equivalent performance in culturally grounded contexts.

The distinction between linguistic adequacy and cultural adequacy is therefore central to the evaluation of AI translation. Conventional evaluation measures have often emphasized dimensions such as lexical similarity, grammaticality, fluency, and semantic correspondence. However, these criteria may not fully capture whether a translated idiom preserves its cultural associations, pragmatic function, emotional connotations, or figurative meaning. Recent benchmark research has consequently begun to propose more explicit frameworks for evaluating culturally grounded expressions and identifying culturally induced meaning deviations.

Comparative studies of contemporary AI translation systems provide further evidence of this problem. Research examining systems such as ChatGPT, DeepL, and Google Translate has identified different approaches to the treatment of idioms and culturally marked expressions, including literal rendering, paraphrase, functional adaptation, and attempts at cultural substitution. Such variation demonstrates that translation quality cannot be assessed solely according to whether an output is grammatically correct or fluent. A culturally appropriate translation must also preserve, where possible, the communicative intention and culturally relevant associations of the source expression.

The consequences of inadequate cultural interpretation extend beyond isolated translation errors. AI-based translation technologies are increasingly used in international communication, education, digital services, and multilingual interaction. In diplomatic and international business contexts, culturally inappropriate translations may create misunderstandings or weaken communicative effectiveness. Similar concerns arise in healthcare interpreting and education, where accurate communication can be particularly important. The increasing integration of AI into these domains therefore makes cultural adequacy an important component of responsible language technology.

From this perspective, the development of culturally responsive AI translation should involve the systematic incorporation of culturally diverse linguistic data and culturally informed evaluation procedures. Culture-specific items require translation strategies that take into consideration not only their linguistic form but also their cultural function and contextual meaning. Earlier Translation Studies research has demonstrated the importance of strategies such as retention, substitution, paraphrase, and other forms of cultural adaptation in dealing with culture-specific items. Contemporary AI systems can potentially benefit from these theoretical insights by incorporating culturally informed translation strategies into training, prompting, evaluation, and human-AI interaction.

An interdisciplinary approach is consequently required to address the cultural dimension of AI translation. Computational linguists can contribute expertise in model architecture, corpus construction, machine learning, and evaluation, while linguists and translation scholars can provide theoretical

© Ushbu material Creative Commons Attribution 4.0 International (CC BY 4.0) litsenziyasi asosida tarqatiladi.

frameworks for semantic, pragmatic, idiomatic, and intercultural analysis. Cultural anthropologists and specialists in cultural studies can further contribute knowledge concerning culturally embedded concepts, practices, values, and references. Such collaboration may facilitate the development of AI systems that move beyond surface-level linguistic correspondence toward more context-sensitive and culturally appropriate translation.

Future research should therefore concentrate on three interconnected priorities. First, AI translation systems should be supported by broader and more culturally diverse datasets, particularly for low-resource languages and underrepresented cultural communities. Second, evaluation frameworks should incorporate cultural adequacy alongside conventional measures of fluency, grammaticality, and semantic accuracy. Third, human supervision should remain particularly important when AI systems process literary, artistic, historical, folkloric, or otherwise culturally sensitive material.

Ultimately, the central challenge of AI-assisted translation is not simply whether machines can produce fluent translations, but whether they can adequately interpret and reproduce the cultural meanings embedded in language. The increasing sophistication of models such as BERT, T5, and GPT demonstrates substantial progress in contextual and semantic processing; nevertheless, current research indicates that culturally grounded expressions remain a distinctive challenge. Addressing this challenge requires the integration of linguistic, cultural, computational, and ethical perspectives. Such an approach can contribute to the development of AI translation systems that are not only linguistically accurate but also culturally sensitive, contextually appropriate, inclusive, and respectful of global linguistic diversity.

REFERENCES:

1. Aixelá, J. F. (1996). Culture-specific items in translation. In R. Álvarez & M. C.-A. Vidal (Eds.), Translation, power, subversion (pp. 52–78). Multilingual Matters.
2. Baker, M. (2011). In other words: A coursebook on translation (2nd ed.). Routledge.
3. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of NAACL-HLT 2019, 4171–4186.
4. Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. Journal of Machine Learning Research, 21(140), 1–67.
5. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30.
6. Dewayanti, D. P. S. P., & Margana. (2024). The impact of contextual understanding on neural machine translation accuracy: A case study of Indonesian cultural idioms in English translation. Englisia: Journal of Language, Education, and Humanities, 12(1). https://doi.org/10.22373/ej.v12i1.24433
7. Kolesnikova, M. P. (2025). Interpretation of idioms and culturally marked elements in neural machine translation (based on Russian-English parallels). Litera, 11, 187–216. https://doi.org/10.25136/2409-8698.2025.11.76823
8. Wang, Q., Amini, M., & Tan, D. A. L. (2025). Strategies, errors, and challenges in translating culture-specific items in Chinese-English literary works: A systematic review. Jurnal Arbitrer, 12(2), 259–273. https://doi.org/10.25077/ar.12.2.259-273.2025

© Ushbu material Creative Commons Attribution 4.0 International (CC BY 4.0) litsenziyasi asosida tarqatiladi.

9. Han, B., Wang, Y., Qing, H., Li, T., Yang, F., Ahmat, A., Atawulla, A., Yang, Y., & Zhou, X. (2026). From words to worlds: Benchmarking cross-cultural cultural understanding in machine translation.
10. Yuan, Z., Ye, Y., Feng, X., Li, B., Li, Q., Lu, Y., Tu, D., & Qin, B. (2026). Culture-aware machine translation in large language models: Benchmarking and investigation.
11. Azizov, D. (2024). From idioms to algorithms: Translating culture-specific expressions in AI systems. Kazakh Ablai Khan University, IRE Journals, 7(10).

© Ushbu material Creative Commons Attribution 4.0 International (CC BY 4.0) litsenziyasi asosida tarqatiladi.