Ars Artium: An International Refereed Research Journal of English Studies and Culture
ISSN: 2319-7889; E-ISSN: 2395-2423
Vol. 13, January 2025, pp. 127-138
Edited by Vijay Kumar Roy
https://www.arsartium.org

Regional Literature in the Age of Machine Translation: Opportunities and Limitations

*Amrendra Kumar Thakur**

Abstract

Advanced technologies in artificial intelligence field have changed machine translation (MT) into a tool to avail regional literature to a global audience. This research paper discusses the possibilities and difficulties of using MT for regional literary texts, focusing on how MT might assist in preserving and spreading the diversity of language and culture. Beyond the language barriers, the translation of regional literature into general use poses unique difficulties due to dialectical richness, cultural symbolism, and contextual nuances in the process. Using MT case studies, the paper analyses Indian regional text, and the vulnerabilities and inadequacies of the contemporary technology in the current pursuit. It further inquires into the ethical concerns of misrepresentation of culture and loss of regional identity in automated translation. The present paper compares MT outputs with human translations for the purpose of underlining the necessity for a hybrid approach, as the efficiency of MT should be combined with the cultural expertise of a human translator who ensures faithfulness to both the text and the cultural context.

Keywords: Artificial Intelligence, Machine Translation, Regional Literature, Dialectical Richness, Cultural Symbolism

1. Introduction

The fast developments made by artificial intelligence and natural language processing technologies have spurred a rapid pace in MT tool development and usage. According

*PhD Research Scholar, Department of English, B. R. A. Bihar University, Muzaffarpur-842001, Bihar, India. Email: aktmadhubani@gmail.com

© 2025 The Author. Published by Paragon International Publishers. Open Access Under Creative Commons Attribution 4.0 International Licence (CC BY 4.0).

to Sharma, "machine translation often fails to capture idiomatic expressions unique to regional texts" (103). These tools are invaluable in breaking down linguistic barriers, especially in a globally connected world where language can serve the function of either a bridge or a barrier to people. However, as Venuti puts it, "translation is more than a linguistic act; it's a cultural one" since the language is steeped in cultural and contextual meaning. Bose and Kumar argue that "machine translation tools should integrate sentiment analysis to better capture the emotional tone of regional literature" (102). This, in turn, has serious connotations, particularly when dealing with regional literature that "preserves unique cultural, social, and historical perspectives" and cannot be algorithmically translated simply. This means that MT has a tendency to expose in literature those forms of culturally specific elements, which serve as a testimony of what Venuti describes as "a purely algorithmic approach [that] risks erasing cultural specificity" (22).

Regional Indian literature well exemplifies the exact limitations in question. "Regional languages are deeply embedded in local cultures, making them challenging for automated systems" (Verma 22). The Maithili, Angika, Bengali, and Marathi languages are not mere linguistic systems; they carry cultural identity and fill the colloquialisms, cultural symbols, and references deep within the social tapestry of the region and in indigenous knowledge. Roy argues that "regional literature provides a rich ground for evaluating the limitations of machine learning algorithms" (67). For instance, Maithili's idiomatic expressions or metaphors cannot be translated, it is because the MT system has problems dealing with the culturally rich texture of the language. In general, as noted by Wu et al.. Languages with limited data suffer the most, which seriously hampers MT from accurately translating regional dialects. Reddy remarks "Machine translation systems must address the interplay between language and identity in regional texts" (95). As a result of lacking abundant linguistic data, texts may be over digested and inaccurately represented by MT tools, thereby letting them lose their meaning and impact.

Well-translated regional literature serves not only as a medium of expression but also as a cultural heritage preservative. Thus, the regional literature of India is a depository of knowledge and values specific to their culture, and MT can even unwittingly stamp out the nuances of this kind of literature. Rajesh Kumar highlights "the challenges of preserving cultural nuances in regional literature during machine translation" (45). It is against this backdrop that it becomes pertinent to take a closer look at the present scenario of MT in the translation of regional literature and its relative strengths and weaknesses. The actual comparisons between MT-

generated translations and those generated by human translators will help the researcher investigate whether MT is thus armed with the mechanisms to translate culturally nuanced texts while balancing accessibility with accuracy.

Such research also delves into the ethics of placing such reliance on MT while translating literature that demands careful cultural considerations in its translation. It provokes the loss of culture, as Venuti has warned, this might be a risk in the treatment of translation strictly as a technical activity, rather than an interpretive one. Such a model is to say, a hybrid one, putting together both efficiencies coming from MT but, at the same time, cultural sensitivity and ultimately the expertise of human translators. It would thus fill up the gaps of MT by making its application in rapid translation possible, which can then be intervened upon later by humans to correct and maintain cultural sensitivities. The crux of the hybrid model would thus involve paying due respect and retaining the cultural flavour that lies within regional literature while at the same time making it more accessible and inclusive.

Though MT has been able to overcome many barriers of access between languages, it is certainly without severe limitations at least when used in relation to the cultural richness of regional literature. Banerjee suggests that "integrating cultural studies with machine translation research can bridge the gap" (46). It is here that balancing MT capabilities with human expertise can help address these limitations so that the act of translation does not diminish culture but preserves it as an end rather than a means.

2. Literature Review

2.1 History of Machine Translation

Machine translation has come from rule-based to statistical models and most recently Neural Machine Translation (NMT). Recent studies, as noted by Wu et al., have demonstrated the near-human quality of translations through NMT, most effectively in complex syntactic languages. However, despite this progress, "NMT still underperforms with low-resource languages, where limited training data compromises translation quality" (Wu et al. 45). Dialect-rich and culturally saturated texts, as Hassan et al. note, are those which demand familiarity with the culture and the context that MT systems bypass, as they tend to "translate figurative expressions into literal meanings devoid of cultural appropriateness."

2.2 Translation Theory and Cultural Heritage

As translation scholars point out, "preservation of cultural identity in translation" is an essential concern, especially for regional literature, since "translation is more

than a mere substitute for the original text". Here, Venuti's distinction between "domestication" and "foreignization" is applicable because domestication is likely to "dilute cultural identity," a commonly observed effect of MT (Venuti 41). Nida also advocates "dynamic equivalence," something that MT tools can never match up to the standard-preservation level of the content experienced and emotionally perceived within the original text.

2.3 Cultural Constraints of Machine Translation

According to Dingwaney and Maier, the figurative language and idiomatic expressions are simplified by MT, which in turn gives translations something without the one's intended depth. Similarly Mishra points out that while machine translation can manage straightforward texts effectively, it struggles significantly with culturally dense content such as idioms and proverbs, leading to a loss of meaning and coherence (79). For example, "idioms, metaphors, and culturally embedded symbols in Indian regional texts are very often mistranslated" (Singh and Kumar 38).

3.Theoretical Framework

3.1 Translation Theory and Cultural Identity

In analyzing machine translation's effectiveness in translating culturally sensitive texts, this study grounds its workings on the concepts of Lawrence Venuti's foreignization and domestication in his translation theory. First, introduced by Venuti in *The Translator's Invisibility*, these strategies lay the foundations in which the duality of access and cultural preservation in translation should be balanced. Foreignization is a process of trying to keep the source text as it is, with all its cultural and linguistic uniqueness, in order to immerse the reader into the source culture and maintain authenticity. Domestication tries to adapt the text into the foreign culture and language norms, which means it might be more accessible but probably departs from the original cultural depth. As Venuti cites, "a domesticating approach can make culturally rich texts accessible but at the cost of authenticity" (43), which highlights the threat of losing critical elements of culture in the text.

Machine translation, particularly with new neural models, domestication has transformed into a sort of approach by default due to the fact that algorithms are more inclined toward fluency and readability of the target language. Chauhan observes that "machine translation tools often prioritize dominant languages, marginalizing regional ones" (15). Such an approach can dilute cultural context, leading to translations losing the nuances that are indispensable to identifying source material.

That is a profound result for Indian regional literature, where linguistic diversity and cultural symbolism are prominent. According to Desai, "regional literature often challenges the linear processing approach of machine translation" (62). A foreignizing approach, while challenging for MT with its narrow cultural knowledge, is simply unavoidable when local texts are full of idiomatic expressions, and colloquial and community-specific references. It is through the foreignization-domestication framework that an assessment of the performance of MT has been provided to establish whether it is either a facilitator of the preservation or simplification of the cultural identities as seen in the regional literature.

3.2 Cognitive Linguistics and Neural Machine Translation

Another theoretical dimension of this research is cognitive linguistics, especially its application to NMT. Indeed, NMT represents a big step in the development of MT technology, as it simulates human cognitive processes via deep neural networks that can elaborate coherent translation and are aware of the context. As cognitive linguistics understands language as one of the cognitive functions in the lives of human beings, translation does not seem to be a purely technical task, but rather a profoundly interpretive one. It is definitely limited since algorithm processes don't have experiential knowledge and intuitive awareness of cultural differences which humans perceive while processing the same language.

This gap is reflected when MT is translated with regional literature that embodies unique dialects and culturally specific idioms, among them references to indigenous knowledge systems. As observed by Wu et al, the MT models have a critical weakness in handling languages with limited resources of datasets, where "limited comprehensive datasets for dialects and culturally specific vocabulary" (46) has resulted in being a liability for NMT models in the management of nuanced language patterns . Cognitive linguistics highlights that NMT models, while being extremely strong in finding general patterns of the language, could not encapsulate deep human cognition, most particularly with contextually rich and culturally sensitive texts. Thus, even more advanced NMT systems face what Wu et al. call an impossible-to-bridge "cognitive gap" between human translators and machines that usually brings about a flattened or distorted translation when it comes to regional literature.

The gap becomes more relevant in the regional Indian context, where different languages like Maithili, Angika, etc., bring in specific syntactic structures, culturally embedded metaphors, and phrases that do not have one-to-one translations from universally supported languages like English. The issue is that this demands more

significant datasets with large amounts of linguistic corpora to train the NMT system models. Hence, without rich linguistic data, MT is likely to deliver translations lacking in contextual meanings and cultural depths that human translators, with their culturally enlightened glasses, are capable of retaining. It is to this end that the present paper, having already brought to light some technical weaknesses in MT, does not fail to discuss ethical considerations since translations that fail to perceive the culture are likely to devaluate the regional text.

4. Methodology

4.1 Selection of Texts

The research starts with the selection of text excerpts from Maithili and Angika literature that are, in particular, culturally dense and idiomatic in terms of usage of idiomatic expressions, cultural symbols, and references indigenous to that culture. These were selected owing to their importance in regional literature of India and complexity in the language feature, and because they could be difficult for machine translation systems to translate. Excerpts include folk tale selections, poetry, and oral narratives, genres which are particularly hard to translate because they depend heavily on regional metaphors, idioms, and colloquial expressions, among others, which can be a challenge to determine proper translation. Such kinds of folk tales, for example, from Maithili, are steeped in proverbs, mythological allusions, and dialect-related phrasing that require an acute sense of culture to allow for an accurate translation. Such "culturally dense passages" are a good indicator of MT's weakness and strength in handling texts deeply rooted in regional identity.

4.2 Machine Translation Tools

To check the performance of the developed machine translation systems, this research makes use of Google Translate and Microsoft Translator. They are famous technologies with advanced capabilities in Neural Machine Translation (NMT). Wu et al. describe those platforms as "among the most sophisticated for general language translation" (47), thus suitable tools for this comparative study. Google Translate and Microsoft Translator are highly accessible and most frequently used for translating general language content, thus having put large amounts into neural network models intended to improve contextual and semantic understanding. However, these tools may fail to realize the required cultural sensitivity for regional texts with specific dialects and expressions. From this consideration, this study selects these two tools to probe whether MT systems, despite advanced NMT technology, can produce a translation that does justice to culturally rich content. The human translations provide

the base to proceed with further comparison toward judging the correctness, cultural interpretation, and the overall comprehension in the human translation compared to the one translation by the machine.

4.3 Evaluation Criteria

Evaluations of MT outturns are framed within three core criteria: linguistic adequacy, cultural closeness, and reader understanding. Such criteria allow a holistic judgment about whether or not the MT system successfully manages both the linguistic and cultural complexities of selected regional texts for processing.

4.3.1 Linguistic Accuracy

This is an assessment of the grammar, vocabulary, and syntax of the translations. Since meaning relies extensively on the lexical meaning of the source text, linguistic accuracy is significantly crucial to prevent structural or semantic errors. Assessments of the MT system are conducted in terms of whether the system is able to deal with unique syntactical structures frequently present in Maithili and Angika and its specific vocabulary.

4.3.2 Cultural Fidelity

This criterion would check whether the translation maintains cultural markers, idioms, and symbols that are essential for holding up the identity of the source text. According to Singh and Kumar, "cultural fidelity evaluates the preservation of cultural markers and idioms" (39), and this aspect is very pertinent if the translations are being done for regional literature as it contains more culture-specific things. The cultural fidelity is, here, tested by analyzing the outcome of MT outputs in retaining cultural references and symbols, i.e. local expressions and culturally important terms. This measure would seek to identify places where MT may have domesticated culturally specific elements, thereby reducing its original meaning.

4.3.3 Reader Comprehension

How readable and interpretable is the translated text? A reader comprehension test is adopted for determining the clarity of a translation's output and coherence in structure for an average reader. This test would ensure whether the output conveys the meaning the original text intends. The feedback from the readers who were also familiar with source and target languages would allow understanding as to whether the translation was clear, culturally correct, and close to its original essence.

This methodology will apply these three evaluation criteria to the machine-translated outputs against a human translation and try to give a nuanced understanding

of how MT stands with regard to handling regional literature in terms of its current capabilities and limitations. It will come out highlighting in which areas the simple content is actually possible to be translated by the MT systems and in what arenas human expertise remains unsurpassed. Bassnett emphasizes that "translators are cultural mediators, shaping how one culture understands another" (87). The results of this study can be useful in the development of a hybrid model of translation, which combines the effectiveness of MT with the cultural sensitivity of human translators, especially for culturally complex regional texts.

5 Analysis and Discussion

5.1 Linguistic Accuracy

Linguistic accuracy is perhaps the greatest challenge posed by these machine translation (MT) systems especially when the regional language happens to be Maithili or Angika. The former has grammatical structures that are far more different and complex than that of the latter and both happen to permit idiomatic expressions that are hard for algorithms to process properly. For instance, "translating Maithili idioms and proverbs into English poses unique challenges due to their cultural specificity and figurative nature" (Mishra 79). Proverbs are, indeed, an integral constituent of the language and form part of the richness of its literature as well as of its culture. These proverbial sayings mainly comprise short phrases representing wisdom and social norms and do not have corresponding English words. "Regional literature embodies unique rhythms and sounds that are difficult to translate" (Ahuja 11). When these sentences pass through the machine translation process, the translations often lack meaning, which may lead to syntax errors or nonsensical sentences. This difference is very critical in literature because the accuracy of the language used affects not only the understanding of the text but also the cultural content of the text. MT systems tend to overgeneralize dialect-specific structures in a translation, thereby making the output grammatically incorrect or without subtle meanings that a human translator intuitively preserves.

5.2 Cultural Fidelity

MT systems often remain shortsighted in cultural fidelity, especially in unique cultural markers buried within regional literature. Bose identifies "how cultural context shapes the interpretation of regional literature in translation" (17). Regional texts usually refer to items, practices, or symbols unique to a particular culture or dialect. To state the example, Singh and Kumar, using the term "madua" in the Angika, use a staple

crop that goes with deep cultural roots. MT systems tend to simplify such terms to generic words like "crop" or "grain," stripping away the contextual aspect that brings it nearer to its meaning in Angika culture (Singh and Kumar 39). Such oversimplification dilute the symbolic and emotional value attached to the term in the source text, and the translation does not represent the source text. The MT tools also mostly break down on cultural metaphors or idiomatic expressions that carry meanings specific to regional languages. For example, the myths and folkloric personalities or activities that are used as idioms and colloquial phrases in Maithili literature carry cultural meaning. Without knowledge and appreciation, the core components of culture behind the MT output come to lose the flavour of the text and decrease the cultural depth and authenticity of its contents.

5.3 Ethical and Cultural Issues

There are ethical issues revolving around the possibility of cultural erasure and misrepresentation in using MT to translate regional literature. According to Venuti, there exists a danger of "cultural erasure and misrepresentation" (42), especially for those culturally deep texts reduced to accessibility. These oversimplifications lead to what scholars call "cultural homogenization" within regional and indigenous literature, wherein the unique elements of the marginalized voices are either muted or lost in translation. Gupta's research also explores "the ethical implications of machine translation in literature" (54). Such is specifically disconcerting for regional texts that would detail indigenous knowledge, social norms, and historical narratives because poor translations could equally dilute the preservation of cultural identity enshrined in the literature.

The cultural constraints of MT also raise an ethical question regarding the accountability of an accurate representation of the marginalized communities. It may inadvertently become part of linguistic imperialism in that the dominant languages and cultures drown out the finer nuances of regional dialects. Here, cultural authenticity plays an important role when portraying artistic flavours of diverse literary traditions before audiences. MT technology offers broader access but only at the cost of losing richness on the cultural complexities and reduces it to more generic simple texts. It points out the urgent need for a hybrid translation model that can balance MT speeds with human translator sensitivity in regional literature translations deserving respect and, above all, accuracy.

6. Conclusion

This research paper highlights that, though machine translation tools bring some undeniable benefits in the aspect of access as well as efficiency, they totally fail to carry proper cultural insight and sensitivity so that it can capture all regional literature nuances, especially apparent while transcribing regional texts like those found in Maithili and Angika, which carry structures of dialect, idioms, and culturally based symbols with which standard algorithms are not equipped.

In critically appraising MT tools such as Google Translate and Microsoft Translator, this study discovers that they do just about all right for general content but are not really up to par in tackling the dilemmas of regional dialects. The problem of accuracy in the language is specifically a concern; there is likely to be much syntactic and lexical and structural wrongness where the source language's grammar is radically different or where the expressions are dialectal. Thus, the translations will be grammatically faulty or deficient in contextual content of the original. The cultural fidelity criterion, furthermore, shows how MT tools fail to capture a lot of culturally significant words and idiomatic expressions, representing them in a simplifying or generalizing manner, which happens to cancel out their essential features of regional identity. For instance, terms like "madua," in Angika, are selected because of their cultural weight and historical context. Such terms are often translated into generic terms like "crop," thus depriving the words of their crucial cultural resonance for which they are constituted as important.

However, the study raises ethical concerns related to cultural erasure and homogeneous voices of the marginalized. Inadequate translation of regional literature contributes to what Venuti describes as "cultural homogenization, whereby unique characteristics of the culture are lost while making texts more accessible to a broader public" (42). This phenomenon is more destructive to regional and indigenous literature as such literature texts frequently operate as essential records of cultural information, social practice, and the narratives of history.

This implication calls for an accommodation approach where the dimensions of speed and accessibility associated with MT could be married to the paramount cultural knowledge that human translators are construed to enjoy. Here, in such an approach, initial translation work undertaken by an MT system can be finalized by human translators who, with such knowledge and sensibility about the culture, will hold onto the authenticity of the source text.

Though MT tools are very helpful for raising the scope and accessibility of regional literature, they can never replace the depth of understanding and cultural acquaintance brought by human translators into the exercise. It is only with the teamwork between the strengths of MT and human excellence that a guarantee on the consistency of accuracy and appreciation for culture in translation from regional literature is sure to come about, as suggested by Venuti. Similarly Mehta said, "hybrid approaches combining machine and human translation yield better results" (33). This method holds much potential for creating translations that honour a regional community's unique voices while making them accessible around the world.

Works Cited

Ahuja, Kavita. *The Soundscapes of Regional Literature in Translation*. HarperCollins, 2018.

Bassnett, Susan. *Translation Studies*. 4th ed., Routledge, 2013.

Banerjee, Tuhin. "Bridging Cultural Studies and Machine Translation." *Translation Studies Review*, vol. 19, no. 2, 2021, pp. 44–50.

Bose, Anindita, and Arjun Kumar. "Sentiment Analysis in Regional Machine Translation." *Translation Technology Journal*, vol. 14, no. 2, 2021, pp. 100–05.

Bose, Sudipta. "Cultural Context in Translating Regional Texts." *Translation Quarterly*, vol. 16, no. 2, 2022, pp. 15–20.

Chauhan, Arvind. "Dominance in Machine Translation: A Regional Perspective." *International Journal of Translation Studies*, vol. 18, no. 1, 2020, pp. 12–20.

Desai, Vikram. "Linear and Nonlinear Approaches in Machine Translation." *Journal of Translation Technology*, vol. 10, no. 1, 2021, pp. 60–70.

Dingwaney, Anuradha, and Carol Maier, editors. *Between Languages and Cultures: Translation and Cross-Cultural Texts*. U of Pittsburgh P, 1995.

Gupta, Neha. "Ethics in Machine Translation: A Literary Approach." *Digital Humanities Quarterly*, vol. 10, no. 2, 2019, pp. 52–59.

Hassan, Hany, et al. *Achieving Human Parity on Automatic Chinese to English News Translation*. Microsoft Research, 2018.

Kumar, Rajesh. *Cultural Nuances in Machine Translation: A Study of Regional Literature*. Oxford UP, 2019.

Mehta, Arjun. "Hybrid Models in Translation Studies." *Computational Linguistics Review*, vol. 27, no. 3, 2018, pp. 30–40.

Mishra, Shailendra Kumar. "Translation Strategies and Challenges for Translating Maithili Idioms and Proverbs into English." *Research Review: International Journal of Multidisciplinary*, vol. 9, no. 4, 2024, pp. 79–88.

Nida, Eugene A. *Toward a Science of Translating: With Special Reference to Principles and Procedures Involved in Bible Translating*. E. J. Brill, 1964.

Reddy, Sandeep. *Language and Identity in Machine Translation*. HarperCollins, 2018.

Roy, Aniket. "Regional Literature as a Benchmark for Machine Translation." *Journal of Linguistic Studies*, vol. 12, no. 3, 2018, pp. 64–75.

Sharma, Anita. Idiomatic *Challenges in Machine Translation*. Routledge, 2020.

Singh, Anupam, and Rajesh Kumar. "Challenges in Machine Translation of Indian Regional Languages: A Case Study of Maithili and Angika." *Asian Journal of Translation Studies*, vol. 9, no. 1, 2021, pp. 112–32.

Venuti, Lawrence. *The Translator's Invisibility: A History of Translation*. Routledge, 1995.

Wu, Yonghui, et al. *Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation*. Arxiv, 2016, https://doi.org/10.48550/arXiv.1609.08144.

❑❑❑