© 2026 International Conference on Multidisciplinary Sciences and Innovations . All Rights Reserved.

FROM CHOICE TO PROBABILITY: DOMESTICATION AND DEFAULT LITERALISM IN MACHINE-TRANSLATED UZBEK PROVERBS AND IDIOMS

Bahriddinova Feruza Bahriddin qizi, Husainov Ilyos Jamoliddin o'g'li
Department of Teaching English to Speakers of Other Languages (TESOL),
Webster University in Tashkent, Tashkent, Uzbekistan.
E-mail: rbaxriddinova@webster.edu
Department of Business Management, Turan International University, Namangan, Uzbekistan.
E-mail: husainov_ilyos@turan-edu.uz

Machine translation (MT) systems and chatbots are now widely used to translate culturally loaded language, including proverbs and idioms. Lawrence Venuti's distinction between domesticating and foreignizing translation is the standard framework for describing how a translation treats the "foreign" [7]. However, this framework was developed for human translators who make deliberate choices, not for MT systems. The question is especially relevant for Uzbek, which is classified as a low-resource language in natural language processing, so MT systems tend to behave unevenly and unpredictably with it [6, 4]. In addition, idioms are already a known weak point for neural MT, since such systems often translate them too literally [2]. Venuti's concepts assume a translator who takes a deliberate, ethical stance toward the foreign text. An MT system, however, takes no stance: it simply outputs whatever its training data makes most probable. Therefore, it remains unclear whether "domestication" and "foreignization" can meaningfully describe MT output at all, particularly for a low-resource Turkic language such as Uzbek. The aim of this study is to determine whether Venuti's domestication–foreignization distinction can describe how Google Translate and ChatGPT render Uzbek proverbs and idioms into English, and to propose adjusted concepts where the original distinction does not apply.

Results

1. Method: five Uzbek items (three proverbs, two idioms) were translated once by each system in August 2026. The outputs were classified using Baker's strategies for translating idioms, with an additional "close translation" category [1].

2. The two systems agreed on three of the five items and differed on the remaining two.

3. ChatGPT used substitution, the clearest form of domestication, on three of the five items, while Google Translate used it on only one. In both cases where the systems differed, ChatGPT's output was the more domesticating one.

4. Two factors appear to drive domestication: how well established the English equivalent of an expression is, and the type of system, since a general-purpose chatbot tends to substitute more readily than a dedicated translation service.

5. A new distinction is proposed: strategic domestication, a translator's deliberate choice as described by Venuti, versus retrieval domestication, which occurs when a system merely reproduces an equivalent pair already present in its training data. This is illustrated most clearly by the idiom "kill two birds with one stone," which both systems produced identically.

6. For close, literal MT output, the term default literalism is proposed instead of "foreignization," since such output reflects no genuine stance toward the foreign text [3].

7. Retrieval domestication raises an ethical concern: it is invisible, since there is no translator whose choices can be questioned, cultural images silently disappear, and some Uzbek–English equivalent pairs may in fact be inherited from Russian–English translation resources.

Conclusion

© 2026 International Conference on Multidisciplinary Sciences and Innovations . All Rights Reserved.

Venuti's concepts of domestication and foreignization remain useful for analyzing MT output, but they need adjustment. Both Google Translate and ChatGPT domesticate Uzbek proverbs and idioms, but not uniformly: domestication is strongest when a well-established English equivalent exists, and the general-purpose chatbot substitutes more readily than the dedicated translation service. The two new terms proposed in this study, retrieval domestication and default literalism, are offered as hypotheses for larger studies, given the small sample size of this pilot study.

References

[1] Baker M. (2018). In other words: A coursebook on translation (3rd ed.). Routledge.
[2] Dankers V., Lucas C., Titov I. (2022). Can Transformer be too compositional? Analysing idiom processing in neural machine translation. Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, 3608–3626.
[3] Jones R., Irvine A. (2013). The (un)faithful machine translator. Proceedings of the 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities, 96–101.
[4] Joshi P., Santy S., Budhiraja A., Bali K., Choudhury M. (2020). The state and fate of linguistic diversity and inclusion in the NLP world. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 6282–6293.
[5] Schleiermacher F. (2012). On the different methods of translating (S. Bernofsky, Trans.). In Venuti L. (Ed.), The translation studies reader (3rd ed., pp. 43–63). Routledge. (Original work published 1813)
[6] Veitsman Y., Hartmann M. (2025). Recent advancements and challenges of Turkic Central Asian language processing. Proceedings of the First Workshop on Language Models for Low-Resource Languages, 309–324.
[7] Venuti L. (1995). The translator's invisibility: A history of translation. Routledge.