Translation Studies, Summer 2025, Vol. 23, No. 90, pp. 8–21: **Scientific Research Paper**
DOR: 20.1001.1.17350212.1404.23.2.4.6

Evaluating ChatGPT's Performance in Persian-to-English Proverb Translation: The Role of Contextual Information and Prompt Design[1]

Hamideh Nemati Lafmejani[2]
& Seyed Mehdi Mousavian[3]

Abstract

Previous research has documented the challenges translators face when working with idiomatic expressions, particularly proverbs, due to their cultural specificity and linguistic complexity. This study explores the efficacy of integrating AI into the translation process, with a focus on translating Persian proverbs into English. To investigate this, 100 Persian proverbs were selected from Moosavi's (2000) book and supplemented with contextual information from the Daneshchi website, https://www.daneshchi.ir/category/zarbolmasal-irani/, accessed July 2025, to ensure a comprehensive understanding of each proverb's meaning/usage. The Persian proverbs, their figurative meanings, and contextual explanations were provided individually to ChatGPT-4o to generate English translations. These English translations generated by AI were compared with reliable dictionaries for verification of accuracy. The comparison showed that out of 100 proverbs, 70 had accurate English equivalents, 21 had incorrect equivalents, and 9 were paraphrased based on the provided context. These findings suggest that ChatGPT-4o can be a useful tool in translating Persian proverbs when provided with sufficient contextual information, though human oversight remains necessary given the 30% error rate. The study implies that AI-based translation tools, when contextually guided, can support human translators in dealing with culturally bound expressions. However, the need for human translators to critically review and validate AI-generated outputs is still essential.

Keywords: Artificial Intelligence, Prompt design, Proverbs, Translation accuracy

---

1. This paper was received on 23.05.2025 and approved on 21.07.2025.

2. Corresponding Author: Assistant Professor, Department of English Language, SR.C., Islamic Azad University, Tehran, Iran; email: hamideh.nemati@iau.ac.ir

3. MA Student, Department of English Language, SR.C., Islamic Azad University, Tehran, Iran; email: seyedmehdi.mousavian@iau.ir

1. Introduction

Proverbs are commonly used in different cultures and languages. Meider (2004) points out that proverbs although are often linked to the past, they still continue to play a role in contemporary communication. As defined by Meider (1988, p. 119) "a proverb is a short, generally known sentence of the folk which contains wisdom, truth, morals, and traditional views in a metaphorical, fixed and memorizable form and which is handed down from generation to generation." Lauhakangas (2001, p. 20) provided 13 main themes for proverbs that mainly represents basic aspects of human life, i.e., practical knowledge of nature, faith and basic attitudes, basic observations and socio-logic, the world and human life, sense of proportion, concepts of morality, social life, social interaction, communication, social position, agreements and norms, coping and learning, time and sense of time. Proverbs have key characteristics as outlined by Hernadi and Steen (1999, p. 4), regarding phonemic devices, they mention, rhyme, assonance, alliteration, syntactic devices; regarding parallelism/balance, repetition, semantic opposites, ellipsis are listed, and finally for figurative language, metaphor, hyperbole, paradox, and personification are commonly used. Proverbs incorporate particular social traditions that cannot be easily comprehended across different cultures. Moreover, many proverbs appear to be confusing because they can be interpreted both literally and figuratively. Proverbs due to their form and culture-bound meanings are difficult to be translated (Duff, 1989). In these cases, finding a proverb that conveys the desired meaning may take time and effort of the translator. Sometimes, the challenge of finding an appropriate equivalent in the target culture and language is a burdensome. Mastering the target language is not often enough to tackle translation issues including translating proverbs as they are strongly culture-bound (Afrouz, 2022, p. 54).

Duff (1989, p. 11) points out that idiomatic expressions including similes, metaphors, proverbs, and sayings are widely recognized as difficult to translate. When these idioms lack direct equivalents in the target language, Duff (1989)

suggests that translators can adopt several approaches: 1. literal translation, 2. original words in inverted commas, 3. close equivalents, and 4. Non-idiomatic translation. Beekman and Callow (1974, p. 139) suggested three ways to translate a proverb as follows: 1. The words following the proverb could be introduced as the meaning of the proverb; 2. It can be replaced with an equivalent local proverb; and 3. Its non-figurative meaning could be stated straight forwardly.

Some studies have been conducted to check the capability of different Machine translation (MT) systems and artificial intelligence regrading culture-bound and linguistically evolved texts (Bouhadiba, 2016; Al-Khresheh & Almaaytah, 2018; Spencer, 2018; Vukanović, 2025; Hassani, Malekshahi, & Davari, 2025).

Bouhadiba (2016) analyzes MT capability (Systran and Google Translate) in translating linguistically and culturally evolved texts such as proverbs and poetry. His main purpose was to demonstrate how Google Translate and Systran are incapable of rendering syntactic, semantic, stylistic, and cultural nuances in Arabic poetry and proverbs. The results indicate that both systems produce syntactically incorrect, semantically ambiguous, and stylistically substandard outputs culturally-dependent content. Bouhadiba (2016) concludes that examined MT systems are inherently limited since they fail to comprehend context and cultural nuances. As a result, Systran and Google Translate were incapable of producing translations that are both accurate and faithful to the source material, especially for highly structured and culture-bound forms like proverbs.

Al-khresheh and Almaaytah (2018) in their article investigated the challenges of translating English proverbs into Arabic through MT (Google Translate). They selected a set of English proverbs and translated them using Google Translate, then they qualitatively compared the outcomes in terms of lexical, syntactic, and semantic accuracy. The findings revealed some linguistic obstacles: mistranslations, word-for-word translation, incorrect word order, and incorrect lexical choice. They believe that

while MT programs like Google Translate can provide a general impression, they are not yet precise enough to translate subtle, culturally embedded phrases such as proverbs.

Romaniuk-Cholewska (2024) examines the pitfalls and usability of MT systems regarding English-Polish proverbs translations, using tools such as Google Translate, DeepL, PONS and ChatGPT-3.5, and Google Bard. Eighteen commonly known Polish proverbs and their English equivalents were selected from established dictionaries and translated both ways (PL→EN and EN→PL). The results show that AI tools such as ChatGPT and Google Bard obtained better performance than traditional MT systems in handling metaphorical meaning, with ChatGPT achieving 61% accuracy in PL→EN translation. All systems struggled with literal translation and often producing semantically, or culturally-awkward equivalents. While AI tools were able to pick upon meaning over and above a literal translation, they still missed the mark in understanding deeper cultural insight. The results further indicate that adequate contextual information and cultural background are necessary for proficient translation of proverbs, as well as demonstrate significant potential advances in MT with state-of-the-art language models although it is still far from a complete human-level replacement.

In his article, Togaymurodov (2024) explores the difficulties of translating culturally embedded expressions between languages. He highlights that proverbs are not only linguistically unique but also carry deep cultural significance, reflecting societal values, history, and wisdom. He (2024) analyzed a corpus of English proverbs and their Uzbek translations. Using translation theories such as Baker's strategies of equivalence and adaptation, and Venuti's concepts of domestication and foreignization, he investigated common issues like syntactic differences, idiomatic non-equivalence, cultural context dependency, implicit meanings, and loss of poetic elements such as rhyme and rhythm. The results revealed that direct or literal

translation often leads to misinterpretation or loss of meaning, while cultural bias and metaphorical content complicate accurate rendering. Togaymurodov (2024) concluded that translating proverbs demands more than bilingual competence, i.e., it requires intercultural sensitivity and creative problem-solving to preserve both the message and impact of these cultural artifacts.

In his study, Jibreel (2023) investigated the effectiveness of online MT tools in translating proverbs between English and Arabic. The study evaluates five MT engines including Google Translate, Reverso, Yandex, Systran, and Bing. In doing so, 30 proverbs (15 English and 15 Arabic) were selected. Newmark's (1988) taxonomy of translation methods were used for data analysis. Results showed that literal translation is the most common method (52%), followed by semantic (34.7%) and communicative (13.3%). Bing is the most effective, providing the highest percentage of communicative translations, while Yandex is the least effective, often producing literal translations. Google and Bing are comparable in providing semantic equivalents. The study concludes that while MT tools are useful, they struggle with cultural nuances and idiomatic expressions, highlighting the need for human post-editing.

Despite the growing interest in AI and its application to culturally rich texts such as proverbs, most current studies tend to speak of AI tools mainly when talking about their translation limitations for idiomatic expressions but do not really delve into the way contextual information or user interaction especially may change or enhance translation quality. Little consideration has also been given to the likely performance of relatively newer AI-driven language models such as ChatGPT in the translation of Persian proverbs.

The current study bridges this gap by studying the applicability of AI, especially ChatGPT-4o, for the translation of Persian proverbs into English. By supplying the model with an in-depth contextual and figurative explanation for each

proverb, the research aims to evaluate ChatGPT-4o's performance in translating Persian-to-English proverbs when the model is provided with contextual information

2. Methodology

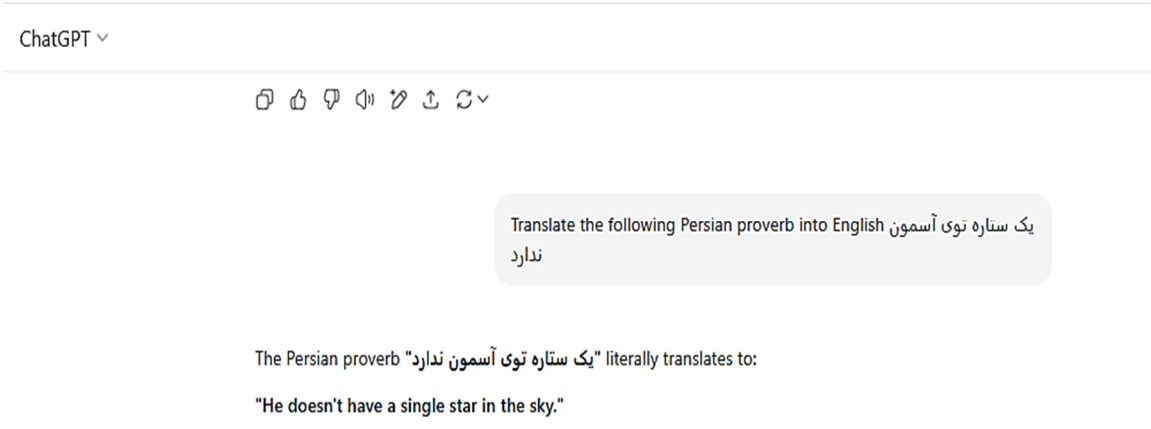
The study employed a systematic approach to examine the potential of artificial intelligence (AI), specifically ChatGPT 4o, in translating culturally embedded idiomatic expressions, with a focus on Persian proverbs.

A total of 100 Persian proverbs were purposefully selected from Moosavi's (2000) comprehensive collection of Persian proverbs, chosen for their rich idiomatic and cultural content. This corpus was deemed suitable due to its wide representation of commonly used Persian proverbs that encapsulate figurative meanings and cultural wisdom.

Then, the contextual definitions were collected from the Daneshchi website, https://www.daneshchi.ir/category/zarbolmasal-irani/, accessed July 2025. For each proverb, its literal meaning, metaphorical sense, and typical usage contexts were noted. This additional information aimed to provide a robust context which would enable the AI model to better understand the deeper senses beneath plain lexical items.

Figure 1 presents a screenshot of the ChatGPT-4o chat interface, showing the prompt for translating a proverb without providing any context.

Figure 1: Screenshot of the ChatGPT-4o Showing the Prompt and the Result for Translating a Persian Proverb



As illustrated in Figure 1, ChatGPT-4o does not accurately translate the proverb into English, offering only a literal translation.

For the purpose of the study, the prompt designed by the researchers was employed for each proverb to provide a context for ChatGPT-4o to translate the proverbs (Figure 2).

Figure 2. Prompt Designed to be Used in Translating the Proverbs

Prompt:

You are an expert phraseologist specializing in idiomatic translation between Persian and English.
Your task is to identify the most accurate English equivalent of a given Persian proverb.

You will be provided with:

- The original Persian proverb
- Its figurative (non-literal) meaning
- A brief contextual explanation of how and when it is used

Your translation strategy must follow this priority order:

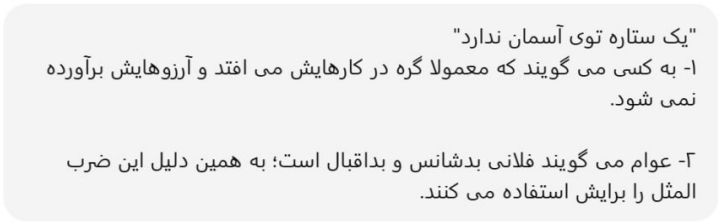
1. Find an English proverb that closely matches the Persian one in both meaning and structure
2. If no such proverb exists, choose an English proverb with a similar meaning (figurative equivalence)
3. If there is no direct proverb, provide a concise idiomatic paraphrase
4. If needed, include a literal or near-literal translation of the Persian proverb as a last resort

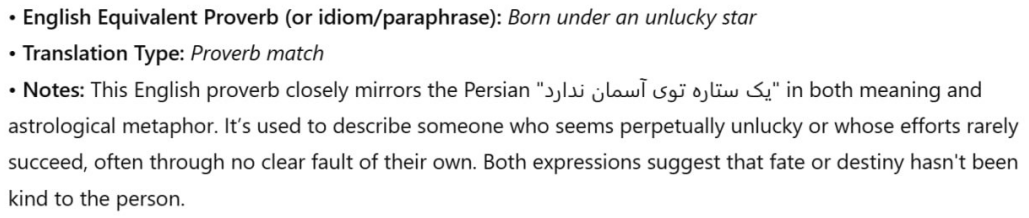
Output Format:

- **English Equivalent Proverb (or idiom/paraphrase):** [Your answer]
- **Translation Type:** [Proverb match / Figurative paraphrase / Literal translation]
- **Notes (optional):** Brief explanation of the reasoning or nuances, if needed

Figure 3 presents a screenshot of the ChatGPT-4o chat interface, showing the prompt for translating a proverb by providing context and designed prompt. It also shows the AI's proposed equivalent for the proverb, based on the context provided.

Figure 3. Screenshot of the ChatGPT, showing the context provided





As can be seen, after being provided with the context and a prompt, ChatGPT suggested an English proverb that corresponds to the Persian proverb.

To determine the accuracy and cultural competency of the translations from AI, all the outputs were cross-checked against standard English language references. Specifically, the Cambridge Dictionary and Merriam Webster Dictionary were used for determining whether AI-selected English translations were actual authentic proverbs or English idiomatic expressions.

Following verification, the results were categorized into three groups:

1. Proper Equivalent: Translations that rightly matched available English proverbs or idiomatic expressions maintaining meaning and stylistic elements.
2. Wrong Equivalent: When the AI produced an English proverb or an idiomatic expression which was not equivalent to the intended meaning of the original Persian proverb.
3. Paraphrased Translations: When the AI produced an explanatory or descriptive paraphrase.

The frequencies in these categories were also recorded to identify the overall efficacy of AI on this technical translation task.

3. Results and Discussion

In order to provide contextual input, Persian proverbs along with their literal and figurative meanings were presented to ChatGPT as prompts. Based on this information, the AI generated English equivalents for the proverbs, were subsequently categorized into three main groups: accurate equivalents, incorrect equivalents, and paraphrased translations.

3.1. Accurate Equivalents

After analyzing ChatGPT-4o's outputs and comparing them against reliable sources such as Merriam-Webster and Cambridge dictionaries, the first group consisted of instances where ChatGPT provided an accurate English equivalent, that is, a proverb or an idiomatic expression, for the corresponding Persian proverb.

Out of the 100 Persian proverbs given to ChatGPT-4o for translation, the AI successfully provided English equivalents for 70 proverbs. These equivalents either took the form of established English proverbs or idiomatic expressions that closely matched the meaning of the original Persian proverbs. This indicates that, in 70% of the cases, ChatGPT-4o was able to accurately translate the figurative meaning of the source proverbs and generate appropriate target-language expressions.

Table 1 presents a selection of these translated proverbs, offering examples from the corpus to illustrate the types of responses generated by the AI system.

Table 1. A Selection of Persian Proverbs and Corresponding AI-generated English Equivalents

No.	Persian Proverb	AI generated English Proverb/idiomatic expression
1	سنگ بزرگ علامت نزدن است	Don't bite off more than you can chew
2	شتر مرد و حاجی خلاص	The dust has settled
3	چغندر گوشت نمیشه، دشمنم دوست نمیشه	A leopard can't change its spots
4	کلاهش پس معرکه است	He's fighting a losing battle

3.2. Incorrect Equivalents

After analyzing ChatGPT-4o's outputs and comparing them against reliable sources such as Merriam-Webster and Cambridge dictionaries, the second group consisted of instances where ChatGPT provided an incorrect English equivalent, that is, a wrong proverb, for the corresponding Persian proverb.

Out of the 100 Persian proverbs given to ChatGPT-4o for translation, the AI was not able to successfully find a correct equivalent for 21 Persian proverbs. This represents approximately 21% of the total data. In these cases, ChatGPT-4o either failed to provide a meaningful equivalent, generated an unrelated proverb, or produced a literal translation that did not convey the figurative meaning of the original Persian saying.

Table 2 presents a selection of these translated proverbs, offering examples from the corpus to illustrate the types of responses generated by the AI system.

Table 2. A Selection of Persian Proverbs and Wrong Corresponding AI-generated English Equivalents

No.	Persian Proverb	AI generated English Proverb/idiomatic expression
1	مشت نمونه خروار است	The evil is in the details
2	گدا به گدا رحمت به خدا	It's the blind leading the blind
3	آدم شاخ در می‌آورد	He/she has horns
4	تا کور شود هر آنکه نتواند دید	Be blind to one's own faults

3.3. Paraphrased Translations

After analyzing ChatGPT-4o's outputs and comparing them against reliable sources such as Merriam-Webster and Cambridge dictionaries, the third group consisted of instances where ChatGPT-4o provided a paraphrased English equivalent for the corresponding Persian proverb.

Out of the 100 Persian proverbs provided to ChatGPT for translation, the AI generated paraphrased English equivalents for 9 proverbs, representing approximately 9% of the total dataset. In these instances, instead of identifying an established English proverb or idiomatic expression with a similar figurative meaning, ChatGPT-4o responded with descriptive rewordings that attempted to convey the general idea or moral of the original Persian proverbs. While these paraphrased translations succeeded in capturing the core message to some degree, they lacked the concise, culturally embedded nature of traditional proverbs.

Table 3 presents a selection of these translated proverbs, offering examples from the corpus to illustrate the types of responses generated by the AI system.

Table 3. A Selection of Persian Proverbs and Corresponding AI-generated Paraphrased English Equivalents

No.	Persian Proverb	AI generated English Proverb/idiomatic expression
1	دود اگر بالا نشیند کسر شان شعله نیست	A diamond doesn't lose its value because it's buried in the mud
2	خدا در و تخته رو با هم جور می کنه	God brings together those who are compatible or similar to each other
3	از مال پس است و از جان عاصی	Nothing to lose and nothing to live for
4	هر چه دیدی از چشم خودت دیدی	You have no one to blame but yourself

These findings suggest that while the model demonstrates a reasonable ability to interpret and translate many culturally embedded expressions, it still faces challenges when dealing with certain idiomatic or metaphorical structures unique to the Persian language.

This suggests that when faced with expressions that do not have a direct counterpart in English, ChatGPT-4o tends to rely on explanatory phrasing rather than searching for or generating equivalent figurative language.

4. Conclusion

The present study aimed to evaluate ChatGPT-4o's performance in translating Persian proverb into English, with an aim to explore how contextual information and prompt design influence the retrieval of accurate and culturally appropriate translations. Out of 100 Persian proverbs in the corpus of the study, ChatGPT-4o successfully translated 70 proverbs (70%) of the Persian proverbs into equivalent English proverbs or idiomatic expressions. However, ChatGPT-4o failed to translate 21 proverbs by generating unrelated proverbs, or producing a literal translation that did not convey the figurative meaning of the original Persian.

The findings of this study are consistent with those of Romaniuk-Cholewska (2024), who reported a 61% accuracy rate in the translation of Polish proverbs into English. Nonetheless, as in her study, certain inaccuracies were observed, especially in relation to cultural and idiomatic equivalence. Vukanović (2025)'s study also showed high level of translation accuracy of ChatGPT in translating expressions that contain metaphorical language for instance metaphors embedded in idioms and proverbs.

The results also reinforce the importance of prompt design in fine-tuning the performance of AI models in translation. This is in line with the finding by Yamada (2023), Gao, Wang, and Hou (2023), and Romaniuk-Cholewska (2024) who indicated that AI systems like ChatGPT perform better when they are supported by rich contextual input.

In practice, these findings suggest that AI can serve as a helpful assistant to human translators. However, due to its current incapability of grasping cultural

nuances, AI cannot be regarded as a complete replacement for human specialists in literary and idiomatic translation (Vukanović, 2025).

Although the results are worth noting, this study also faced limitations. There is inherent limitation in having only used a single AI model, ChatGPT 4o, when it could have been compared to other models. Contextual input improved translation accuracy substantially, but the study did not systematically investigate the effect of the kinds, or the degrees of context on the translations. In addition, the corpus was limited to a curated selection of Persian proverbs. While it was representative of many proverbs, it did not represent the idiomatic expression available in each regional or dialectical variation. These limitations reduce generalizability and indicate a need for more extensive and diverse datasets and a more disciplined evaluation means.

Future research could benefit from dedicated proverb databases and multilingual culture specific corpora where more relevant idiomatic and metaphorical content could be explored and used for the training of language models.

Incorporating culture-based databases would help language models learn not just linguistic patterns but also the socio-cultural underpinnings of proverbs and idioms. Furthermore, the fine-tuning of large language models on particular annotated datasets that account for metaphorical use and cultural context could possibly improve translation quality. In addition, we will need comparative working papers on various AI models and prompts designs.

References:

Afrouz, M. (2011). Translating provebs in the Gulistan of sa'di: Developing a new taxonomy based on Baker's (2011) model. *Hikma, 21*(1), 53–83.

Al-khresheh, M. H., & Almaaytah, S. A. (2018). English proverbs into Arabic through machine translation. *International Journal of Applied Linguistics & English Literature, 7*(5), 158–166.

Beekman, J., & Callow, L. (1974). *Translating the word of God.* The Zondervan Corporation for Summer Institute of Linguistics.

Bouhadiba, M. (2016). The deficiencies of machine translation of proverbs and poetry. *Revue Maghrébine des Langues, 10*(1), 222–240.

Duff, A. (1989). *Translation*. Oxford University Press.

Gao, Y., Wang, R., & Hou, F. (2023). How to Design Translation Prompts for ChatGPT: An Empirical Study. *Proceedings of the 6th ACM International Conference on Multimedia in Asia Workshops*, 26–37.

Hassani, G., Malekshahi, M., & Davari, H. (2025). AI-Powered transcreation in global marketing: Insights from Iran. *English Language Overseas Perspectives and Enquiries, 22*(1), 203–221.

Hernadi, P., & Steen, F. (1999). The tropical landscapes of proverbia: A crossdisciplinary travelogue. *Style, 33*, 1–20.

Jibreel, I. (2023). Online machine translation efficiency in translating fixed expressions between English and Arabic (Proverbs as a case-in-point). *Theory and Practice in Language Studies, 13*(5), 1148–1158. https://doi.org/10.17507/tpls.1305.07

Lauhakangas, O. (2001). *The Matti Kuusi international type system of proverbs.* ASF and Outi Lauhakangas.

Meider, W. (2004). *Proverbs: A handbook*. Greenwood Press.

Meider, W. (1988). *English proverbs.* Philipp Reclam.

Moosavi, M. (2000). *A versified dictionary of Farsi proverbs and their English equivalents.* Jahan Rayaneh.

Romaniuk-Cholewska, D. (2024). The machine translation (MT) of proverbs in the Eng-PL language pair. *Studies in Logic, Grammar, and Rhetoric, 69*(82), 553–570.

Spencer, T. (2018). *Proverbial machine translation: Translating proverbs between Spanish and English using phrased based statistical machine translation*. (Master's thesis, University of Southern Mississippi).

Togaymurodov, D. D. (2024). The translation challenges of proverbs, idioms, and cultural references. *Academic Research in Educational Sciences, 5*(5), 468–475.

Vukanović, M. B. (2025). Translating (metaphors) in the age of AI: Opportunities, challenges, and implications for the EFL classroom. *English Language Overseas Perspectives and Enquiries, 22*(1), 185–201.

Yamada, M. (2023). Optimizing machine translation through prompt engineering: An investigation into ChatGPT's customizability. *Proceedings of Machine Translation Summit XIX, 2*, 195–204.