



Sevinch RAVSHANOVA,
O‘zbekiston Milliy universiteti magistranti
E-mail: sevinchravshanova052@gmail.com

TIQXMMI Milliy tadqiqot universiteti dotsenti, PhD K.Alautdinova taqrizi asosida

THE IMPACT OF AI ON IDIOM TRANSLATION IN LITERATURE: A COMPARATIVE STUDY OF CHATGPT, YANDEX, AND MICROSOFT TRANSLATOR

Annotation

This study examines the translation of idioms by three machine translation systems: ChatGPT, Yandex Translate, and Microsoft Translator. The study employs quantitative and comparative methods, using data drawn from the novel “The Prince and the Pauper” by Mark Twain.

Key words: ChatGPT, Yandex Translate, Microsoft Translator, machine translation, idioms, comparative translation, artificial intelligence (AI).

ВЛИЯНИЕ ИИ НА ПЕРЕВОД ИДИОМ В ЛИТЕРАТУРЕ: СРАВНИТЕЛЬНОЕ ИССЛЕДОВАНИЕ CHATGPT, YANDEX И MICROSOFT TRANSLATOR

Аннотация

Данное исследование рассматривает перевод идиом тремя системами машинного перевода: ChatGPT, Yandex Translate и Microsoft Translator. В работе применяются количественные и сравнительные методы, а материалом исследования служат данные, взятые из романа Марка Твена «Принц и нищий».

Ключевые слова: ChatGPT, Yandex Translate, Microsoft Translator, машинный перевод, идиомы, сравнительный перевод, искусственный интеллект (ИИ).

ADABIYOTDA IDIOMALAR TARJIMASIDA SUN'Y INTELLEKTNING TA'SIRI: CHATGPT, YANDEX VA MICROSOFT TRANSLATORNING QIYOSIY TADQIQOTI

Annotatsiya

Ushbu taqdidotda uchta mashina tarjimasi tizimi — ChatGPT, Yandex Translate va Microsoft Translator tomonidan idiomalarning tarjima qilinishini o‘rganadi. Tadqiqotda miqdoriy va qiyosiy usullar qo‘llanilgan bo‘lib, idiomalar Mark Tvenning “Shahzoda va gado” romanidan olingan.

Kalit so‘zlar: ChatGPT, Yandex Translate, Microsoft Translator, mashina tarjimasi, idiomalar, qiyosiy tarjima, sun‘iy intellekt (SI).

Introduction. The field of machine translation (MT) has garnered increasing attention due to the advances made in artificial intelligence (AI), particularly in natural language processing (NLP) and deep learning. Modern MT systems such as Google Translate, Yandex Translate, DeepL, and Microsoft Translator rely heavily on neural networks and transformer models.

AI is not a countable entity; therefore, it cannot be referred to numerically (for example, “20 AIs”). However, it can be categorized by type, approach, or model. Based on the first classification, there are narrow AI, general AI, and superintelligent AI [11]. Narrow AI, also known as weak AI, specializes in a single task and does not match human performance—examples include Google Translate, Siri, and Gemini. General AI represents a theoretical form of intelligence comparable to human cognition, while superintelligent AI is hypothesized to surpass human intelligence, though this remains a speculative and futuristic concept [5].

Regarding the other two classifications, machine learning, deep learning, expert systems, evolutionary algorithms, and reinforcement learning make up the main AI models [3, 6].

Several studies have investigated machine translation from evolutionary and impact perspectives. For instance, one of the key achievements of neural machine translation (NMT) was

the demonstration of near-human translation quality for both Chinese-to-English and English-to-Chinese tasks [7]. This study leveraged the duality of translation by allowing the model to learn bidirectionally—from source to target and vice versa. However, challenges persist in translating languages with complex syntax or limited training data.

In translation pedagogy and literary MT, comparative analyses have been conducted between Google Translate and QTranslate. One study evaluated the error rates of MT in literary texts translated from English into Arabic using J.K. Rowling’s works as source material. The findings revealed that Google Translate outperformed QTranslate, though the difference in accuracy was relatively small [10].

Nevertheless, with the growing use of MT tools by both professional and amateur translators, there is a growing need to identify a relatively accurate system. MT now plays a vital role not only in professional translation but also in language learning and research. However, there remains a lack of MT studies focusing on Yandex Translate, ChatGPT, and Microsoft Translator when translating between English and Russian language pairs.

The present study aims to examine how three MT systems—Yandex Translate, ChatGPT, and Microsoft Translator—handle idiomatic expressions in literary texts. The specific objectives are: a) to evaluate which tool produces the

most accurate translations; b) to analyze the quality of idiom translation from English into Uzbek and Russian.

The study hypothesizes that ChatGPT produces more accurate translations overall than Microsoft Translate, while Yandex Translate demonstrates moderate accuracy in rendering idiomatic expressions in literary texts.

Literature review. This chapter provides the theoretical foundation for machine translation (MT) with particular focus on the translation of idiomatic expressions between Chinese and English. It also synthesizes prior studies and aims to build a conceptual framework for understanding the major challenges, approaches, and research gaps within MT studies.

Machine Translation (MT) refers to the application of computational methods to render text or speech automatically from one language into another. MT systems analyze the source language and generate corresponding output in the target language, striving to preserve meaning, context, and stylistic nuance. Three principal MT approaches have evolved: Rule-Based MT (RBMT), which depends on predefined linguistic rules and dictionaries; Statistical MT (SMT), which utilizes probabilistic models based on bilingual corpora; and Neural MT (NMT), which relies on deep learning algorithms and neural networks to predict and generate translations [8].

Since idioms seldom have direct linguistic equivalents, translation equivalence becomes a central concern in MT research. Two key dimensions of equivalence are generally discussed: semantic equivalence, which emphasizes the accurate preservation of meaning, and pragmatic equivalence, which considers contextual appropriateness and the effect of the translation on the target audience [2].

Chinese and partner universities have contributed substantially to the theoretical and empirical development of MT. A joint study by Peking University and the University of Edinburgh argued that idioms should be examined through manual analysis despite the existence of automatic evaluation metrics such as BLUE, METEOR, and TER. According to their findings, these metrics provide only a general evaluation and fail to assess specific linguistic or cultural subtleties [12]. The researchers proposed the blacklist method as an alternative approach to detect literal translation errors.

A more recent study conducted at Chengdu University compared the performance of ChatGPT 4.0, DeepL, and StarDesk in translating Chinese idioms into English subtitles [9]. Using BLEU as the evaluation metric, StarDesk achieved the highest score, followed by DeepL and ChatGPT. Nevertheless, manual evaluation produced contrasting results, showing that ChatGPT 4.0 delivered superior performance in linguistic, cognitive, and cultural dimensions, revealing the limitations of relying solely on automatic metrics.

Parallel developments have occurred in Arabic MT research, addressing complex issues at morphological, syntactic, and semantic levels. A study from a Moroccan university evaluated the accuracy of Google Translate, DeepL,

and Microsoft Bing in translating English idioms into Arabic, French, and Spanish. The results revealed an overall accuracy rate of 68.75%, with Bing Microsoft Translator outperforming the others [4]. In Palestine, Alawi critically examined MT errors in Google Translate, emphasizing the cultural and idiomatic challenges that often lead to literal renderings and structural inconsistencies in Arabic output [1].

Despite these contributions, several research gaps remain unaddressed. Firstly, most MT studies are based on short segments, subtitles, or isolated idiomatic phrases rather than extended literary texts such as novels, where contextual continuity shapes idiomatic meaning. Secondly, comparative analyses involving modern neural-based systems—particularly ChatGPT, Yandex, and Microsoft Translator—are still scarce. Third, the dominant use of automated metrics such as BLEU and METEOR fails to capture cultural, pragmatic, and semantic equivalence, especially in idiomatic and figurative contexts. Finally, there is a lack of manual, human-centered evaluations that examine idiomatic accuracy, cultural interpretation, and contextual preservation in narrative translation.

This research seeks to address these gaps by analyzing the translation of idioms within a literary text (novel) using three advanced MT systems: ChatGPT, Yandex, and Microsoft Translator. By integrating automated and manual evaluation methods, the study aims to identify the most accurate system for idiomatic translation. It places particular emphasis on semantic fidelity, pragmatic equivalence, and cultural resonance.

The results of this investigation are expected to contribute to the growing body of MT scholarship by offering a nuanced understanding of idiomatic translation in extended narrative contexts. Moreover, it provides a comparative perspective on contemporary MT systems and highlights the limitations of current evaluation approaches. Ultimately, this study aims to propose a more balanced and context-sensitive framework for assessing machine translation performance in literary discourse, bridging the gap between computational translation technology and human linguistic interpretation.

Research Methodology. This study employed a quantitative research design to numerically demonstrate the accuracy of translation in three AI-based machine translation systems: ChatGPT, Yandex, and Microsoft Translator. A comparative method was also adopted to examine the similarities and differences in idiom translation by three machine translation systems from English into Uzbek and Russian.

13 idioms were selected from “The Prince and the Pauper” by Mark Twain in English, based on archaic features. Each idiom was translated into Russian using three machine translation systems: ChatGPT, Yandex Translate, and Microsoft Translator. The MT outputs were collected manually by entering each idiom into the respective system and recording the resulting translations in a spreadsheet for further analysis.

Table 1. Idioms selected with archaic features in “The Prince and the Pauper” by Mark Twain

№	Idiom (English)	Page number	Notes/Context
1	twist oneself loose	17	The prince twisted himself loose.
2	beyond someone’s wildest dreams	17	He will make thee rich beyond thy wildest dreams.
3	crane someone’s neck	55	The crowd enjoyed this episode prodigiously and pressed forward and craned their necks to see the small rioter.

4	goad someone into something	55	Presently they began to taunt him and mock at him, purposely to goad him into a higher and still more entertaining fury.
5	stood someone's ground	55	He stood his ground and defied the mob right royally.
6	lay about	56	The mob closed in on the warrior, who backed himself against a wall and began to lay about him with his long weapon like a madman.
7	as fast as someone's leg could carry someone	56	A troop of horsemen came charging down upon the mob, who fled out of harm's reach as fast as their legs could carry them.
8	with one accord	57	The great assemblage bent their heads upon their breasts with one accord.
9	to say someone nay	57	None rise up to say me nay?
10	strike a chill to someone's heart	57	The tidings struck a chill to the heart of the poor little waif, and sent a shudder through his frame.
11	send a shudder through his frame	58	The tidings struck a chill to the heart of the poor little waif, and sent a shudder through his frame.
12	wax and wane	66	But in my last battle I was taken captive, and during the seven years that have waxed and waned me since then
13	raise the hue and cry	117	In a moment he was up and off like the wind, the gentlemen following after and raising the hue and cry, lustily as he went.

To evaluate the accuracy of idiom translation, the collected data were analyzed using the blacklist method, which identifies incorrect or unacceptable translations based on literal translation. In this approach, each idiom output from ChatGPT, Yandex Translate, and Microsoft Translator was checked against a predefined list of incorrect renderings. If an idiom was translated word-for-word or failed to convey its figurative meaning, it was classified as incorrect. The proportion of

correct and incorrect translations was presented in numbers and calculated in percentages to compare the performance of the three systems.

Analysis and results. This section depicts the findings of the blacklist analysis of idiom translations produced by ChatGPT, Yandex Translate, and Microsoft Translator. The idioms were from the source language (English) into the target languages (Uzbek and Russian).

Table 2. Translation accuracy of idioms from English into Uzbek by three MT systems.

MT system	Total Idioms	Correct Translations	Incorrect Translations	Accuracy (%)
ChatGPT	13	12	1	92.30
Yandex Translate	13	5	8	38.46
Microsoft Translator	13	4	9	30.76

Table 3. Translation accuracy of idioms from English into Russian by three MT systems

MT system	Total Idioms	Correct Translations	Incorrect Translations	Accuracy (%)
ChatGPT	13	12	1	92.30
Yandex Translate	13	8	5	61.53
Microsoft Translator	13	7	6	53.84

Overall, ChatGPT produced the highest number of correct translations, followed by Yandex Translate, while Microsoft Translate had the lowest accuracy. The findings also indicate that translations into Uzbek showed lower accuracy in Yandex Translate and Microsoft Translator compared to translations into Russian. This could be due to ChatGPT's contextual understanding, which aligns with earlier studies.

Conclusion and recommendations. These results suggest large language models (LLMs) like the future ChatGPT

5.0 version may handle figurative language better than traditional MT systems. Translators, teachers, researchers, and students may consider using LLM tools when working with idioms. Nevertheless, the study had limitations, including a small database and a focus on one type of figurative language. Future research should expand the corpus and include more languages. BNC or COCA should be used to derive the source materials for similar research.

REFERENCES

1. Alawi, N. Machine Translation: The Cultural and Idiomatic Challenge, 2017, <https://digitalcommons.aaru.edu.jo/alazhar/vol19/iss2/9/>
2. Almubark, A. Equivalence and Non-equivalence in Translation: The Concept of Translation Equivalence. Lap Lambert Academic Publishing, 2015, <https://www.bol.com/nl/nl/p/equivalence-and-non-equivalence-intranslation/1001004010212349>
3. Alpaydin, E. Machine Learning: The New AI (2nd ed.). MIT Press, 2021 – 105-159 p
4. Ben-Yahia, M. Examining the Efficiency of Machine Translation in Translating English idioms used in American Media, 2024, <https://share.google/5YeCe8tKxWOISZ96h>
5. Bostrom, N. Superintelligence: Paths, Dangers, Strategies. Oxford University Press, 2014 – 23 p
6. Goodfellow, I., Bengio, Y., Courville. Deep Learning. MIT Press, 2016 – 6 p
7. Hassan, H., Aue, A., Chen, Ch. Achieving Human Parity on Automatic Chinese to English Translation, 2018, <https://share.google/SjEJ0jDvbbxWUUS3U>
8. Koehn, P. Neural Machine Translation. Cambridge University Press, 2020 – 33 p
9. Liu, H., Li, X. Evaluating ChatGPT's Translation of Chinese Idioms in Subtitles: A Cognitive Approach Through the Case Study of Minglan. Journal of Arts, Humanities and Social Sciences, 2024
10. <https://share.google/TIMimSBilOo6M86yv>
11. Omar, A., Gomaa, Y. The Machine Translation of Literature: Implications for Translation Pedagogy. International Journal of Emerging Technologies in Learning, 2020, <https://share.google/qEa7wBuf1HKVqGnDI>
12. Russel, S. J., Norvig, P. Artificial Intelligence: A Modern Approach (4th ed.). Pearson, 2021 – 4 p
13. Shao Ch., Greedy Search with Probabilistic N-gram Matching for Neural Machine Translation, 2018, <https://share.google/AwnbVImpQeTUb0d9k>
14. Stahlberg, F. Neural Machine Translation: A Review and Survey, 2020, <https://www.sciencedirect.com/science/article/pii/S2666651020300024>
15. Twain, M. The Prince and The Pauper. Collins Classics. Harper Press, 2011 – 17-117 p