

UDC 81

IRSTI 16.31.41

<https://doi.org/10.48371/PHILS.2025.4.79.030>

TRANSLATING THE UNTRANSLATABLE IN KAZAKH FAIRY TALES: HUMAN vs AI

*Mizamkhan B.¹, Dyusseimbina G.², Yezmakhunova A.³

^{*1,3} Kazakh Ablai Khan University of International Relations and World
Languages Almaty, Kazakhstan

² Abai Kazakh National Pedagogical University, Almaty, Kazakhstan

Abstract. While the roles of human translators and machines have shifted with the development of such engines as Google Translate and DeepL, the issue of translating the untranslatable or culture-specific items from Kazakh into English still remain challenging. The lack of domain-specific training of machine translation tools as well as the underrepresentation of such languages as Kazakh in diverse textual corpora can lead to misinterpretation of the untranslatable, translation errors, and textual incoherence, and translations that might not feel natural to the target reader. To effectively analyze translation inconsistencies, two prompt-based strategies, using GPT-4o, were developed depending on the results of both human and machine translations.

Analyses demonstrate that when an accurate prompt is generated, the current system available for public use makes fewer critical errors. While shedding light on the integration of human and machine translation strategy, this study argues for well-informed translation prompts incorporated with human translation to improve machine translation quality.

Results also show that, though the translation quality is inconsistent, GPT-4o translation with human assistance enhances the emotional and epistemological perspective of the translation of the untranslatable.

The paper will also argue that closest equivalent of the excerpt should be attempted when translating from the source language.

The research is valuable from both theoretical and practical perspectives as it enhances the understanding of the CSIs as it offers insights into how different translation prompts affect cultural transmission. It will also identify limitations and potential improvements for AI translation tools like Google Translate or ChatGPT in the Kazakh language as a lower resource language in Natural Language Processing.

Keywords: culture-specific items, untranslatable, artificial intelligence, translation equivalence, underrepresented languages, diverse, textual corpora, translation inconsistencies, machine translation

Introduction

Translation is seen as a means of representing other cultures and bridging the gap between two cultural contexts. It is a form of communicative activity designed

to bring speakers of different cultures together and foster mutual understanding. Through translation, readers can access the literary works of other peoples and cultures and facilitate cultural integration, where foreign cultural values are adopted and incorporated into the fabrics of the national culture and literature.

Nevertheless, culture specific items such as realia, names, idioms, and connotative expressions often cause challenges in translation, and some may render certain elements untranslatable [1] influencing the overall quality of the final product.

Kazakh is one of the languages rich in culturally specific items, (CSIs) or the untranslatable that reflects its worldview rooted in nomadic traditions and social values. Not only does this language have complex vocabulary, but it also has difficult-to-translate phrases, which present challenges when transferred to other languages. By focusing on Kazakh language, which is an underrepresented language in the digital pool of vocabulary, this study contributes to broader debates on the use of machine translations and the cultural preservation of minority languages. It compares the quality of translation of Kazakh CSIs to English language created by AI engines, such as ChatGPT, to the one created by a human translator and analyzes how clear-cut prompts improve machine-translation quality. Ultimately, this study contributes to understanding the best ways to ‘teach’ neural machine translation systems and large language models, such as Chat GPT, to translate CSIs by considering context and specific cultural cues.

Theoretical Debates on Translatable and Untranslatable

The debate over translatable and untranslatable centers on the extent to which meaning can be transferred across languages. Some scholars contend that languages share sufficient similarities in terms of meaning, structure and communicative experiences to render translation feasible [2]. For translation studies scholars who argued in favor of translatability, the concept of equivalence held a central role in their theoretical foundations. One of the pioneering scholars in this line, Jakobson, proposed the concept of equivalence in difference — the idea that translation must seek to convey equivalent meaning across languages, even when the grammatical or lexical forms are not directly comparable [2]. Catford further steered this discussion by contending that equivalence exists on a spectrum “rather than [as] a clear-cut dichotomy” [3, p.93]. For Catford, equivalence depends on how much of the meaning can be preserved across different languages. This suggests that there are differences in the actual root meaning of the words depending on the culture, since both language and culture are interconnected.

Expanding on these, Glynn and Hudley [4, p.4] identified three forms of expressions that exhibit “resistance” to translation: conceptual, ideological, and aesthetic or formal. Conceptual refers to ideas, worldviews, or cultural concepts that are deeply embedded in one culture and have no direct equivalent expression in another. For example, “*ар ұят/ar uyať*” is a central concept in Kazakh ethical life. It is often translated in English as “*conscience*,” “*moral integrity*” or “*dignity*”.

Unlike its proposed English equivalents, “*ap ұям/ar ууағ*” carries a strong cultural and conceptual weight: “*to lose ap ұям/ar ууағ*” (*ұямтан айырылу*) represents deep social shame, often seen as a loss of humanity or honor. It is often used in public judgements of character and is tied to both individual reputation and family honor. Ideological form involves differences in values, power structures, or political implications between languages and cultures. For instance, kinship terms in Kazakh language reflect not only familial relationships, but also hierarchal social values and obligations within kinship structure. Words like “*моқал/toqal*” (*second or junior wife, often associated with a lower status in the family hierarchy*) convey specific social obligations and gendered expectations that are difficult to fully capture in English, where kinship terminology carries fewer social expectations. Aesthetic untranslatability, positioned within pragmatic effects, concerns those devices in poetic or formal registers – imagery, allusion, syntactic chimes – whose emotive or perceptual force collapses when re-inscribed. Rhyme, pun, or even the insistent rotation of a stressed syllable can be sacrificed or diluted when the translator opts for semantic fidelity.

Recent studies, for instance, have discovered that AI-assisted translators such as ChatGPT, Google Translate, and MS Translator, reinforce gender biases [5]. Such large language models have also demonstrated a tendency to perpetuate raciolinguistic stereotypes that misrecognize linguistic patterns integral to African American English (AAE) [6]. In line with these studies, Tao et al.’s [7] empirical analysis of large language models, such as GPT-4 and GPT-3.5, has shown that these models are inclined to use languages that align with self-expression values common in English-speaking and Protestant European countries.

Behind many of these problems lies the fact that machine translation models are trained on freely available datasets, often chosen from publicly accessible online sources. Unfortunately, current AI-powered language technologies focus on only two to three percent of the world’s most widely spoken languages [8], particularly those with significant financial and political power. As a result, the linguistic pool of AI is either shaped by the dominance of global hegemonic languages or relegated to the marginalization of “low-resource” languages. Consequently, languages with fewer speakers and less political power often have limited representation in the “world system” of digital content [9]. This, in turn, leads to translations of underrepresented languages that are frequently less accurate and culturally insensitive.

Materials and methods

Research material: The research material selected for this study consists of excerpts from a popular Kazakh folktale – “*The adventures of Aldar Köse*”³³. The

³³ It is important to briefly explain the meaning of the character’s name “Aldar Köse”. Literally, it is the combination of two words: “*Al*” + “*dar*” (literally to trick) and “*köse*” (without beard and moustache). According to the *Old Turkic Dictionary*, the root “*al*” conveys meanings beyond “to deceive”; it also encompasses notions of cunning and wisdom. Given these interpretations, it is evident that the name “Aldar Köse” does not solely imply deception; it also conveys meanings of fame, wisdom, and uniqueness.

original Kazakh version of the book, compiled by Älimbaev et al. (2013), consists of a collection of short tales that revolves around the adventures of its main character, *Aldar Köse*, who is sometimes regarded as the Kazakh counterpart to Robin Hood. In these stories, Aldar Köse is often portrayed as witty, humorous, and cunning; a wanderer of the steppe with no wealth or possessions, who helps the poor, tricks the rich, and challenges injustice through intellect rather than force. Due to these personalities of its main character, this folktale is particularly rich in CSIs such as rhymes, humor, wordplay, and idiomatic language that reflect the worldview, social norms, and linguistic creativity of the Kazakh people. These features make *The Adventures of Aldar Köse* an ideal source for studying the translation of CSI from Kazakh into English.

Research method: For the human-translated version, this study utilized the English translation by Mary Lou Masey (1968). For the AI-generated version, the following prompt was used consistently in the GPT-4o interface as the main prompt throughout the analyses: “*Translate the [excerpt] to English language including all the cultural and historical aspects. Also, don’t directly translate the idioms and the set expressions word by word. Use the similar English versions of the idioms*”.

This prompt was designed to encourage the model to attend to both cultural nuance and idiomatic equivalence. Although the uniform use of a single prompt might help reduce output variability and bias, to get more accurate translation, more specific, secondary prompts, follow-up questions, and translation tips were employed. Where necessary, the AI tool was asked to correct the translation. Then, outcomes generated after machine post-editing were compared with the versions translated by a human translator. Further, a comparative textual analysis was conducted to evaluate differences in cultural interpretation, idiomatic usage, and translation strategies. While employing both qualitative and quantitative approaches, this study provides a comparative assessment of both machine and human translation quality. For this purpose, CSI preservation rate was calculated based on three-score criteria where Score 0 – shows omissions, CSI is missing entirely or poorly replaced; Score 1 – more literal, CSIs are translated in a literal sense. Score 2 – adapted idiomatically, CSIs are rendered in a way in which the same tone has been preserved in the target language. After collecting the final data into one table (Table 2), the final statistics are calculated using the following formula:

$$\text{CSI Preservation rate: Preservation \%} \\ = (\text{number of CSIs adapted or retained} / \text{total CSIs}) * 100$$

Then the preservation rate was calculated, and a pie chart was generated to show the results from the data collection.

Research procedure: As analyzing the entire story would be too lengthy for a paper, by adapting Glynn and Hudley’s framework, the following criteria

to select CSIs was applied: 1) proper and common names, 2) idiomatic phrases and metaphors specific to the cultural context, 3) references to local customs, traditions, and 4) religious or kinship terminology with layered cultural expressions. They are culturally unique, vivid and expressive. The selected CSIs and their explanations are given in Table 1.

Results

This section analyzes how CSIs are rendered in translation, focusing on the challenges and nuances involved in preserving the cultural uniqueness and meaning embedded in the original passage. It also demonstrates how ChatGPT-4o translation was improved when Mary Lou Masey’s translation was given as a prompt and what translation errors occurred after each prompt.

Excerpt 1. (Älimbaev et al., 2013, p.133-134)

Ерте заманда, құйрығы келте заманда Шықбермес Шығайбай дейтін бай болыпты. Төрт түлігі сай болыпты. Қөңілі жай болыпты. Сол Шықбермес Шығайбай үйіне ешкімді қондырмайды екен, қондырса, дәм бермейді екен. Тіпті тастан да қатты сараң, өзі тойса да, көзі тоймайтын адам екен. Үйіне кісі келсе оның жауабы: — Шық, әй! — екен. Осыдан жұрт оны Шығайбай атап кеткен екен. Шықбермес Шығайбайдың аты естілмеген ел болмайды екен. Оның «Шық-әйін» естілмеген жан болмайды екен. Қайтсем де, қалай болса да Шығайбайдан дәм татам деген талайлардың тауы шағылған болады екен. (Älimbaev et al., 2013, p.133-134).

Table 1. CSIs from the selected excerpts and their explanations

	Culturally Specific Items	Explanations
1.	<i>Ерте, ерте, ертеде, құйрығы келтеде/erte, erte, ertede, quiryǵy keltede</i>	Almost all Kazakh fairy tales start with this figurative expression. It means that the story/action happened in time immemorial. It is a mythical, distant past. The literal translation of this part of figurative expression is “with a short tail.”
2.	<i>Шықбермес/Shyqbermes</i>	It is a noun-forming adjective, which is literally translated as following: “shyq” – get out, “bermes” – the one who does not give anything. Literally it means “someone who knows only one word: get out! Figuratively: a tight-fisted person.
3.	<i>Шығайбай/Shygaibai</i>	The first passage of the fairy tale explains the origin of this name quite well. It’s an abrupt, informal way of telling someone to leave. The verb “shyq” in Kazakh means “get out,” and “ai” is an exclamation used in commanding speech. In addition, bai is the Kazakh term for the wealthy person. “Shyq” and “ai” are combined to create a proper name “Shygai” (the sound -q changed to -g because of regressive assimilation).
4.	<i>Төрт түлігі сай болыпты/Tört tülgi sai bolypty</i>	In Kazakh culture ‘ <i>mәpm мүлік/tört tülk</i> ’ refers to abundance, endless wealth meaning that the man possesses all types of livestock such as horses, cows, camels, and sheep. It is an important cultural symbol of wealth and prosperity.
5.	<i>Тастан да қатты сараң/ tastañ da qatty sarañ</i>	<i>It is a widely used idiom to describe a greedy person, which can literally be translated as “harder than a stone.”</i>

Translating the untranslatable in Kazakh fairy tales: Human vs AI

6.	Өзі тойса да, көзі тоймайтын адам/ özi toisa da, közi toimaityn adam	It is another commonly used idiom to describe a greedy person. It can literally be translated “even when his stomach is full, his eyes are not.”
7.	Шық-әй/Shyq-ai	The verb “shyq” in Kazakh means “get out,” and “ai” is an exclamation used in commanding speech.
8.	Тауы шағылған/ тауы шағылған	This is an idiomatic expression which can literally be translated as “one’s mountain has been shattered.” It is a Kazakh idiom meaning to be deeply disappointed, to suffer a major setback, or to lose hope.

Prompt 1: Translate the [excerpt] to English language including all the cultural and historical aspects. Also, don’t directly translate the idioms and the set expressions word by word. Use the similar English versions of the idioms.

*Once upon a time, in the distant past, (1) there was a rich man named Shyghaybay (2), who was known as Shykbermes Shyghaybay (3). He had all his livestock in order (4), and his heart was content. However, Shyghaybay never allowed anyone to stay at his house; even if he did, he wouldn’t offer them any food. He was even **harder than stone in his stinginess (5)—he himself would feast, but his eyes were never satisfied.** (6) If anyone came to his house, his response was:*

*- «Leave, hey!» (7) And so, the people began to call him Shyghaybay. The name of Shykbermes Shyghaybay was known throughout the land. No one could say they hadn’t heard his «Shy, hey!» Whenever anyone tried, by any means possible, to get a taste of his food, **they were left disappointed (8).***

Analysis:

CSI 1. Ерме, ерте, ертеде, құйрығы келтеде/erte, erte, ertede, quiryğy keltede

ChatGPT: *Once upon a time, in the distant past*

Here, readers can see that ChatGPT simply omits the “quiryğy keltede” part of the expression and goes straight to “once upon a time.” In fact, the replacement of “quiryğy keltede” to equivalents like “once upon a time” compromises the cultural nuance and stylistic uniqueness in the original text. However, it erases the rhythmic flavor of traditional Kazakh storytelling, where the use of similarly sounding words such as “erte” and “kelte” enhance the narrative’s musicality and oral appeal.

CSI 2: Шықбермес/Shyqbermes

ChatGPT: *...who was known as Shykbermes Shyghaybay*

ChatGPT translates the original noun-forming adjective using a noun clause beginning with “who...”. This strategy involves a shift in word order. Such structural adjustments are normal in English translations. What is problematic, however, is that ChatGPT does not translate the main character’s name, instead just

transcribes it word-for-word as “*Shykbermes Shyghaybay*”. As it was mentioned before, the name “*Shyqbermes Shygaibai*” in Kazakh language implies that the character is a very stingy (greedy) person. However, ChatGPT’s translation does not reflect that meaning.

CSI 3: *Төрт түлігі сай болынты/Tört tülіgі sai bolypty*

ChatGPT: *He had all his livestock in order.*

ChatGPT rendering still lacks the cultural depth and nuance conveyed in the original text.

CSI 4: *тастан да қатты сараң/ tastañ da qatty sarañ*

ChatGPT: *harder than stone in his stinginess*

This is another commonly used idiom that conveys an extreme level of miserliness. ChatGPT’s translation “*harder than stone in his stinginess*” is more literal and closer to the original meaning. But it could be slightly improved if more familiar idiomatic expression (e.g., “*as stingy as stone*”) is used.

CSI 5: *өзі мойса да, көзі тоймайтын адам / özi toisa da, közi toimaityn adam*

ChatGPT: *he himself would feast, but his eyes were never satisfied*

The idiom ‘*Özi toisa da, közi toimaityn adam*’ literally translates as “*even when his stomach is full, his eyes are not,*” describing insatiable greed. ChatGPT translates it more literally: “*he himself would feast, but his eyes were never satisfied*”.

CSI 6: *Шық-әй/Shyq-ai*

ChatGPT: *Leave-hey.*

ChatGPT rendered “*Shygai*” as “*Leave-hey.*” This means ChatGPT misses the figurative meaning of “*Shyq-ai*” and its cultural connotation to Shygaibai’s name in the original text.

CSI 7: *тауы шағылған/ tauy şağylğan*

ChatGPT: *.... they were left disappointed*

ChatGPT rendered the idiom “*тауы шағылған/tauу şağylğan*” (lit.: *their mountain shattered*) as “*they were left disappointed.*” Although ChatGPT does not necessarily use an idiomatic expression, it conveys the contextual meaning of the original idiom. However, it is worth noting that the idiom “*тауы шағылған/tauу şağylğan*” in Kazakh carries a deeper emotional weight; it is more than disappointment, rather a great sense of emotional collapse. Although, ChatGPT’s translation “*they were left disappointed*” captures the general sentiment, it does not fully reflect the emotional intensity.

After this analysis, a more specific prompt will be proposed based on Mary

Lou Masey's translation. This part can be considered as a post-editing of the raw machine translation and aims to determine whether machine post-editing will influence the outcome and can generate a high-quality translation. Such automatic post-editing will 'reduce repeated manual corrections and improve translation productivity' [10]. This way, the study will help the ChatGPT decode new CSIs, retrieve contextual information and train the data.

Prompt 2. Improve the quality of translation of highlighted word expressions relying on Mary Lou Masey's translation. Pay attention to the untranslatable, phonetic peculiarities of Kazakh personal names as well as Mary Lou Masey's omissions of CSIs. Offer a tone that better matches the original folktale.

Mary Lou Masey's translation: *Once upon a time (1) there lived a rich man named Shigaibai (2) who was famed far and wide for his incredible greed (3). In the whole vast steppe, there was no one more miserly than he. Not once in his life had Shigaibai given anything away, not even so much as a drop of water (5 and 6), nor had he ever invited anyone to be his guest. (7) – no translation. (8) – no translation.*

ChatGPT: Long, long ago, back when *even the foxes had short tails*, there lived a wealthy man named Shygaibai, who came to be known as *Shyqbermes Shygaibai* – Shygaibai the Ungiving. His herds were plentiful, and his fortune complete. *He lived in comfort*, yet not once did he invite a soul into his home. And if by chance someone did step inside, he wouldn't offer them so much as a sip of tea. He was *stingier than stone itself* — even when *his belly was full, his greedy eyes were never content*. If a traveler came knocking, all they'd get was a sharp bark: — “*Out, you!*” And so, the people began calling him Shygaibai. His name, Shyqbermes Shygaibai, was known across the steppe, and there wasn't a soul who hadn't heard his bitter “*Out, you!*” echo in the wind. Many tried — by hook or by crook — to get a taste of his food, but each one left with *their hopes dashed and their pride bruised*.

Analysis: When a correct prompt is given, even the machine can perform cognitive effort and can identify translation inconsistencies and correct them. This time ChatGPT used a more evocative and resonant language. The dramatic effect is more appropriate. CSIs are not translated literally, instead, it tried to find a parallel to a source language CSIs.

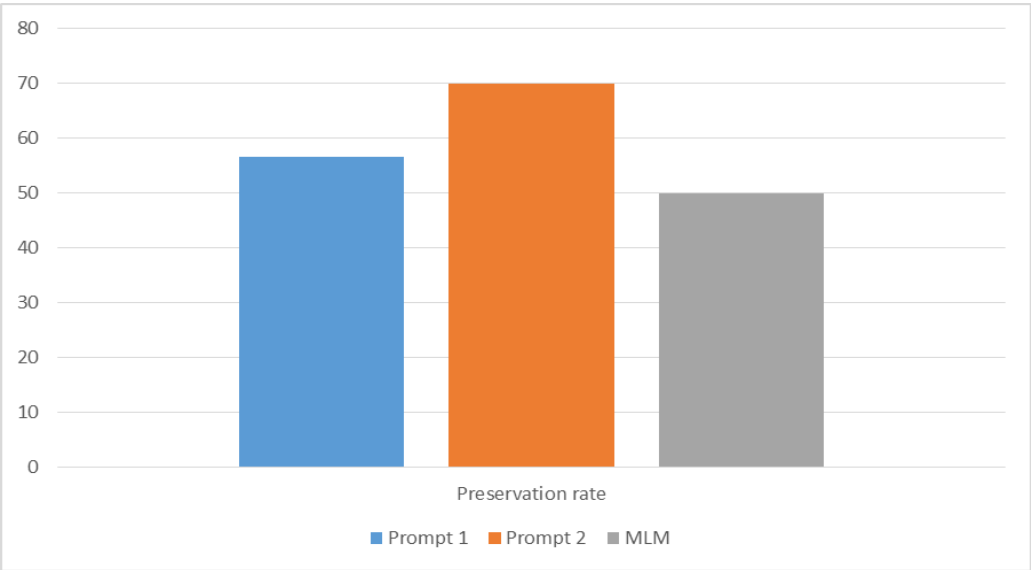
Table 2. Summaries of translation problems identified

CSI Phrase (Kazakh)	Prompt 1	Prompt 2	M a s e y Translation	Score: Prompt 1	S c o r e : Prompt 2	Score: Masey
“Ерте заманда”	Adapted	Adapted	Adapted	2	2	2
“Шықбермес”	Transcribed	Adapted	Adapted	2	2	2
“Төрт түлігі сай болыпты”	Literal	Adapted	Omitted	1	2	0
“Тастан да қатты сараң”	Adapted	Adapted	Adapted	2	2	2

“Көзі тоймайтын адам”	Literal	Adapted	Omitted	1	2	0
“Шық-эй”	Literal	Literal	Omitted	1	1	0
“Тауы шағылған”	Adapted	Adapted	Omitted	2	2	0

Altogether, 30 CSIs (tables1 and 2) were evaluated. By the end of the experiment, Prompt 1 scored 34, Prompt 2 scored 42, and MLM scored 29, respectively. This analysis shows that translation inconsistencies occurred in all three cases as these CSIs are not customary in the target language. As a result, the assumption that human mind and creativity excel the machine translation in many ways needs to be revisited. ChatGPT could cope with complex CSIs and adapt the database when an accurate prompt incorporated with human translation was given.

The outcome is as follows:



Picture 1 – Preservation rate table

ChatGPT with Prompt 1 has earned a quality score of 56.67%, which is higher than human translation quality by 6.67%. When the second prompt was given, it could generate even more coherent and readable output while efficiently preserving CSIs. The total quality score it earned was 70% compared to 50% for a human translator. Mary Lou Masey failed to translate most of the CSIs, many of the untranslatable were omitted which negatively impacted the target text quality. These experimental results support the claim of this paper that the human-AI synergy will generate a more high-quality translation of the CSIs of underrepresented languages such as Kazakh.

Discussion

This study based on qualitative-quantitative analyses of CSIs in human versus ChatGPT translations reveals persistent challenges in translating

“untranslatables”—words or expressions deeply rooted in culture or philosophy that resist straightforward equivalence [11]. These challenges manifest as domestication, oversimplification, omission, and insufficient contextualization.

The human translation by Mary Lou Masey tends to simplify or omit CSIs to maintain narrative cohesion and align with moral values deemed appropriate for the target audience. For instance, Mary Lou Masey omits several CSIs such as “*eşki quiryğy keltede*” (1) and “*тәрт мүлігі сай болынты/төрт түлігі сай бөліпті*” (4) and reduces “*toqal*” to “*his servant girl*.” As indicated in the previous section, in the traditional Kazakh culture, *toqal* is the second or junior wife, often associated with a lower status in the family hierarchy. It appears that the author chooses to omit first two idioms because they are deeply embedded in the source culture and lack direct English equivalents, while the last term is modified to protect Western readers, especially children, from culturally sensitive content. Although these strategies enhance readability and reduce cognitive load, they risk erasing vital cultural nuances, stylistic richness, and epistemological depth. Such preference for domestication and adjustment in translation can reinforce hegemonic narratives by prioritizing the target culture’s norms over the source culture’s unique voice.

In contrast, ChatGPT handles complex linguistic structures and processes large volumes of texts with notable accuracy [12]. Several empirical studies have confirmed that ChatGPT’s commendable capabilities in translating complex texts are sometimes almost equal in quality to those of human translators with advanced expertise. Similarly, Hernández-Guerra [13] showed that machine-generated and human-produced subtitles were closely similar in quality; only individuals with advanced translation expertise could notice the subtle differences. Consistent with these studies, the present study observed that ChatGPT tends to generate a more accurate translation, often prioritizing fidelity to the original text and minimizing omissions as compared to human translation when a clear and appropriate prompt is given.

Nonetheless, machine translation often fails to comprehend or elaborate on the cultural and emotional significance of the original text. Consequently, the outcomes generated by ChatGPT are often diluted or misinterpreted. For instance, after the first prompt, when translating the imperative “*Shyq-ai*,” ChatGPT failed to recognize the cultural connotation embedded in the phrase, particularly its association with the main character’s stinginess. To overcome such limitations, Li et al. [14], introduce a scalable, context-aware multilingual idiom knowledge base (IDIOMKB) developed in assistance with large language models.

Conclusion

Although advancements in language modelling enhance AI tools’ ability to decode idiomatic expressions, training remains heavily biased toward high-resource languages with extensive corpora. Consequently, assistive models still fall short when handling culturally specific items from languages that lack

expansive datasets, such as Kazakh. The system's inability to retrieve accurate meaning demonstrates the need for human assistance to faithfully convey the layered cultural, emotional, and lexical shape of Kazakh CSIs into English.

Findings of the current study demonstrate that translation is far from being a mere linguistic task. It involves the cultural elements that highly affect translation quality. Cultural divergence along with implicit information embedded in source and target languages obstruct a complete correspondence of an utterance with its translation. Even in an era of rapidly deploying neural machine translation systems, large language frameworks struggle to mediate culturally specific items. Recommended as a time-saving measure, tools like ChatGPT process whole prompts promptly yet miss the contextual, conceptual, and emotional gestures integral to the source text. Failure remains conspicuous when the source language is Kazakh, a tongue marginal to the digital ecosystems of databases and corpora, whose vocabulary is still held in the periphery of world linguistic systems. Systematic overuse of automated translation, therefore, depreciates quality while quietly erasing the specific textures of marginal cultures and affording uneven prominence in international debate to already dominant linguistic communities.

Therefore, while advocating for an alliance between AI and human translations, the study emphasizes the culturally centric translation practices that do not omit, weaken, or simplify the source text overly. Such a collaborative approach helps to attain both the fidelity to the original text and its cultural significance.

REFERENCES

- [1] Cassin B., Apter E., Lezra J. and Wood M. (eds.). Dictionary of untranslatables: A philosophical lexicon. – Princeton: Princeton University Press, 2014. – P. 13-44.
- [2] Jakobson R. On linguistic aspects of translation // On translation. – Cambridge, MA: Harvard University Press, 1959. – P. 232–239.
- [3] Catford J. C. A linguistic theory of translation. Vol. 31. – London: Oxford University Press, 1965. – 110 p.
- [4] Glynn D., Hadley J. Theorizing (un)performability and (un)translatability // Perspectives. – 2021. – №29(1). – P. 20–32.
- [5] Ghosh S. and Caliskan A. ChatGPT perpetuates gender bias in machine translation and ignores non-gendered pronouns: Findings across Bengali and five other low-resource languages // Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society. Association for Computing Machinery. – 2023. – P. 901–912.
- [6] Hofmann V., Kalluri P. R., Jurafsky D. and King S. AI generates covertly racist decisions about people based on their dialect // Nature. – 2024. – № 633(8028). – P. 147–154.
- [7] Tao Y., Viberg O. Cultural bias and cultural alignment of large language models // Cognitive Systems Research. – 2024. – № 83. –P. 101–174.
- [8] Helm P. Gábor B., Gertraud K. Fausto G. Diversity and Language

Technology: How Techno-Linguistic Bias Can Cause Epistemic Injustice. – 2023. <https://arxiv.org/abs/2307.13714>

[9] Apter E. Afterword: Towards a theory of reparative translation / Giusti F. and Robinson B. L. (eds.) The work of world literature. – ICI Berlin Press, 2021. – P. 209–228. doi:10.37050/ci-19_09.

[10] Chatterjee R. Automatic post-editing for machine translation. – 2019. <https://arxiv.org/abs/1910.08592>

[11] Large D., Akashi M., Józwikowska W. and Rose, E. (eds.). Untranslatability: Interdisciplinary perspectives. – London: Routledge, 2018. – 236 p.

[12] Fan R., Zhang Y. Strengths and limitations of machine translation and human translation' / Role of AI in translation and interpretation. – IGI Global Scientific Publishing, 2025. – P. 95–118.

[13] Hernández-Guerra E. Towards Inclusive Machine Translation: The Intersection of Artificial Intelligence, LGTBIQ+ Terminology, and Gender Bias // ELUA: Estudios de Lingüística, – 2024. – №39(1). – P. 125-145

[14] Li S., Chen J., Yuan S., Wu X., Yang H., Tao S. and Xiao Y. Translate meanings, not just words: IdiomKB's role in optimizing idiomatic translation with language models // Proceedings of the AAAI Conference on Artificial Intelligence. – 2024. – № 38(17). – P. 18554–18563.

ҚАЗАҚ ЕРТЕГІЛЕРІНДЕГІ БАЛАМАСЫЗ ЛЕКСИКАНЫ АУДАРУ: КӘСІБИ АУДАРМАШЫ VS ЖИ

*Мизамхан Б.¹, Дюсембина Г.², Езмахунова А.³

^{*1,3} Абылай хан атындағы Қазақ Халықаралық Қатынастар және Әлем
Тілдері Университеті, Алматы, Қазақстан

² Абай атындағы Қазақ Ұлттық Педагогикалық Университеті
Алматы, Қазақстан

Андатпа. Google Translate және DeepL сияқты жүйелердің дамуына байланысты кәсіби аудармашылар мен машиналық аударма құралдары атқаратын рөлдер айтарлықтай өзгеріске ұшырап жатыр. Соған қарамастан, мәдени реңттегі мәтіндерді қазақ тілінен ағылшын тіліне аудару мәселесі әлі де өзекті және күрделі күйінде қалып отыр. Машиналық аударма құралдарының қазақ салалық ерекшеліктерге бейімделмеуі, сондай-ақ қазақ тілі сияқты тілдердің әртүрлі мәтіндік корпустарда жеткілікті дәрежеде қамтылмауы тікелей аударылмайтын ұғымдарды бұрмалап түсіндіруге, аударма қателіктеріне, мәтіндегі сөздер, ұғымдар мен сөйлемдердің арасында логикалық байланыстың болмауына және аударма тілде сөйлейтін аудитория үшін мәтіннің жасанды болып естілуіне жол ашып отыр.

Аудармадағы осындай сәйкессіздіктерді тиімді талдау үшін GPT-4o жүйесін пайдалана отырып, сондай-ақ, кәсіби аудармашы мен машиналық аударма нәтижелеріне сүйене отырып, екі түрлі нұсқауға (prompt) негізделген стратегия жасалды.

Талдау көрсеткендей, нұсқау дұрыс берілген жағдайда, қазіргі көпшілікке қолжетімді жасанды интеллект (GPT-4o) жүйесі жіберетін қателер саны бұдан айтарлықтай аз болмақ. Адам және машиналық аударма стратегияларын үйлестіру мәселесіне назар аударып, бұл зерттеу машиналық аударманың сапасын арттыру үшін кәсіби аудармашымен бірлесе отырып, мазмұнды нұсқау жасап қолдануды ұсынады.

Зерттеу нәтижелері, сонымен қатар, аударма сапасы әрқилы болғанымен, GPT-4o жүйесі мен кәсіби аудармашының бірлесе жұмыс істеуі баламасыз элементтердің эмоциялық және танымдық (эпистемологиялық) мәнін жақсартуға түсетінін де көрсетеді.

Бұл мақала, сондай-ақ, аудармада түпнұсқа мәтіннің мазмұнына барынша сай келетін балама табуға ұмтылу қажет екені де атап өтеді.

Зерттеу теориялық тұрғыдан да, практикалық тұрғыдан да құнды, өйткені ол мәдени-реңкке ие элементтерді тереңірек түсінуге мүмкіндік береді және әртүрлі аударма тәсілдерінің мәдени мазмұнды жеткізуге қалай әсер ететінін көрсетеді. Сонымен қатар, бұл жұмыс қазақ тілі табиғи тілді өңдеу (Natural Language Processing) саласында аз қамтылған тіл ретінде қарастырылған жағдайда, Google Translate немесе ChatGPT сияқты жасанды интеллекттің көмегімен аударудың кемшілігін анықтауға және оларды жетілдіру жолдарын ұсынуға мүмкіндік береді.

Тірек сөздер: мәдени реңкке ие сөздер, баламасыз лексика, жасанды интеллект, аударма баламасы, аз зерттелген тіл, әр түрлі, мәтін корпусы, аударманың әрқилылығы, машиналық аударма

ПЕРЕВОД НЕПЕРЕВОДИМОГО В КАЗАХСКИХ СКАЗКАХ: ПРОФЕССИОНАЛЬНЫЙ ПЕРЕВОДЧИК vs ИИ

*Мизамхан Б.¹, Дюсембина Г.², Езмахунова А.³

^{*1,3} Казахский Университет Международных Отношений и Мировых
Языков имени Абылай хана, Алматы, Казахстан

² Казахский Национальный Педагогический Университет имени Абая

Аннотация. Несмотря на то, что роль профессиональных переводчиков и машинного перевода изменилась с развитием таких систем как Google Translate и DeepL, проблема перевода непереводаемых или культурно-специфических терминов с казахского на английский язык по-прежнему остается сложной. Недостаток специализированного обучения по использованию инструментов машинного перевода, а также недостаточная представленность таких языков, как казахский, в разнообразных текстовых корпусах могут привести к неверному толкованию непереводаемых элементов, ошибкам в переводе и текстовой непоследовательности, а также к переводам, которые могут показаться неестественными для целевого читателя. Две стратегии для эффективного анализа переводческих несоответствий анонимно разработаны на основе GPT-4o-промптов, учитывающего и человеческий, и машинный перевод. Проведенный

анализ показывает, что при корректном составлении промпта нынешняя публично доступная система делает меньше критических ошибок. Освещая интеграцию стратегий человеческого и машинного перевода, данное исследование выступает за использование тщательно проработанных переводческих промптов, дополненных человеческим переводом, для повышения качества машинного перевода.

Результаты также демонстрируют, что, несмотря на нестабильное качество перевода, перевод GPT-4o с участием человека способствует улучшению эмоционального и эпистемологического аспектов передачи непереводаемых элементов.

В статье также отмечено, что при переводе на язык оригинала необходимо стремиться к наиболее близкой к оригинальному эквивалентности элементов перевода (эквиваленту элемента).

Эксперимент полезен как в теоретическом, так и в прикладном плане, расширяя представление о культурно-специфических элементах и позволяя понять, как перевод на разные языки требует применения различных подходов к дополнению их передачи. Кроме того, он поможет лучше понять ограничения и возможные направления для улучшения переводческих инструментов на базе искусственного интеллекта, а таких как Google Translate, и ChatGPT, в частности, при работе с малоизученным языком в обработке текстов на естественном языке казахским.

Ключевые слова: культурно-специфические элементы, непереводаемое, искусственный интеллект, переводческая эквивалентность, малоизученные языки, различный, корпус текстов, несоответствия перевода, машинный перевод

Мақала түсті / Стаття поступила / Received: 15.10.2025.

Жариялауға қабылданды / Принята к публикации / Accepted: 25.12.2025.

Information about the authors:

Mizamkhan Baglan – Candidate of Philological Sciences, Kazakh Ablai Khan University of International Relations and World Languages, Almaty, Kazakhstan, e-mail: balapane@gmail.com, <https://orcid.org/0000-0001-5581-7147>

Dyussembina Gulsamal – PhD, Abai Kazakh National Pedagogical University, Almaty, Kazakhstan, e-mail: gulsamal86@gmail.com, <https://orcid.org/0000-0002-0121-9069>

Yezmakhunova Arzigul – MA, Kazakh Ablai Khan University of International Relations and World Languages, Almaty, Kazakhstan, e-mail: arzusha85@mail.ru, <https://orcid.org/0009-0007-6709-2906>

Авторлар туралы мәлімет:

Мизамхан Бағлан – филология ғылымдарының кандидаты, Абылай хан атындағы Қазақ Халықаралық Қатынастар және Әлем Тілдері Университеті, Series “PHILOLOGICAL SCIENCES” No. 4 (79) 2025

Mizamkhan B., Dyussembina G., Yezmakhunova A.

Алматы, Қазақстан, e-mail: balapane@gmail.com, <https://orcid.org/0000-0001-5581-7147>

Дюсембина Гүлсамал – PhD, Абай атындағы Қазақ Ұлттық Педагогикалық Университеті, Алматы, Қазақстан, e-mail: gulsamal86@gmail.com, <https://orcid.org/0000-0002-0121-9069>

Езмахунова Арзигуль – магистр, Абылай хан атындағы Қазақ Халықаралық Қатынастар және Әлем Тілдері Университеті, Алматы, Қазақстан, e-mail: arzusha85@mail.ru, <https://orcid.org/0009-0007-6709-2906>

Информация об авторах:

Мизамхан Бағлан – кандидат филологических наук, Казахский Университет Международных Отношений и Мировых Языков имени Абылай хана, Алматы, Казахстан, e-mail: balapane@gmail.com, <https://orcid.org/0000-0001-5581-7147>

Dyussembina Gulsamal – PhD, Казахский Национальный Педагогический Университет имени Абая, Алматы, Казахстан, e-mail: gulsamal86@gmail.com, <https://orcid.org/0000-0002-0121-9069>

Езмахунова Арзигуль – магистр, Казахский Университет Международных Отношений и Мировых Языков имени Абылай хана, Алматы, Казахстан, e-mail: arzusha85@mail.ru, <https://orcid.org/0009-0007-6709-2906>