

ISTIR 16.21.39

DOI

S.K. ARYSBEK¹, T.A. TAUBYEVA²

Nur-Mubarak Egyptian University of Islamic Culture, Almaty, Kazakhstan

(e-mail: nur-muba19@mail.ru ¹, kadyrarysbek2006@gmail.com ²)

TRANSLATING THE MORPHOLOGICAL SYSTEM OF THE KAZAKH LANGUAGE USING ARTIFICIAL INTELLIGENCE TOOLS: EXPERIENCE AND TESTING

Abstract. The rapid proliferation of Large Language Models (LLMs) and Neural Machine Translation (NMT) architectures has generated considerable optimism regarding the automation of linguistic workflows. This optimism demands rigorous empirical scrutiny when applied to morphologically rich, agglutinative systems such as Kazakh.

This study evaluates how ChatGPT (GPT-4), Google Translate and DeepL widely deployed translation tools and process the morphological density and stylistic nuances of literary Kazakh language while identifying the structural etiologies of their translation artifacts.

The evaluation corpus comprises four textual excerpts systematically sampled from the canonical prose of Mukhtar Magauin and Sherkhan Murtaza. They were selected for their elevated literary registers, high morpheme-to-word ratios and culture-specific lexical depth. Each passage was ingested via zero-shot prompting to isolate baseline capabilities and the resulting outputs were analyzed using a four-tiered philological error taxonomy encompassing morphological, semantic, syntactic and pragmatic-stylistic dimensions. The empirical findings establish a distinct hierarchy of architectural competence.

ChatGPT demonstrates superior contextual modeling, exhibiting sensitivity to literary registers, long-distance syntactic dependencies and language-specific rhetorical devices such as vowel elongation and verbal reduplication.

DeepL occupies an intermediate position, generating syntactically permissible output that consistently erodes cultural and idiomatic depth.

Google Translate performs deficiently manifesting critical semantic reversals, structural disruptions and the erroneous segmentation of transliterated proper nouns.

The theoretical significance of these data indicate that translation errors systematically cluster around specific typological features of Turkic languages that remain underserved by mainstream NLP architectures optimized for Indo-European

structures.

Practical significance deficits pose a concrete sociolinguistic risk to regional digitization initiatives, underscoring an immediate requirement for human-in-the-loop engineering.

The article concludes that active and structural participation from trained philologists is indispensable to refine tokenization strategies, annotate specialized corpora and implement high-fidelity alignment protocols within automated translation systems.

Keywords: Kazakh morphology, machine translation, natural language processing, agglutinative languages, philological error analysis, AI translation evaluation

Introduction

Starting to work seriously with Kazakh literary texts it is noticed something that no translation app seemed able to handle. A single word could carry the meaning of an entire English sentence. A verb form like “*бастырған соң*” encodes not just a completed action but the possessive, temporal, and aspectual context around it. Namely, layers of meaning fused into one orthographic unit. That experience is what motivated this study. The question that makes think over is not whether AI tools are able to translate Kazakh at all, but whether they could do it in a way that preserve what actually matter in the text.

The rapid development of large language models and AI-powered translation tools over the past several years has generated considerable scholarly and public interest in the potential of these systems to transform linguistic work. Tools such as Google Translate, DeepL, and OpenAI's ChatGPT are now consulted millions of times daily by individuals, institutions, and industries across the globe. Their apparent fluency and speed have led many to conclude that translation, once exclusively the province of trained human professionals, is on the verge of full automation [1, pp. 291–295]. This conclusion, however, is premature, and the reasons for that are rooted in the structure of languages like Kazakh.

The performance of AI translation tools varies enormously across different languages, and the gap between high-resource languages such as English, French, or German and low-resource languages such as Kazakh, Uyghur, or Kyrgyz remains structurally significant. The problem is not solely one of insufficient training data, though that is a contributing factor [2, pp. 1715–1716]. More fundamentally, it reflects a deep mismatch between the morphological architecture of languages like Kazakh and the assumptions embedded in the design of most contemporary NLP systems.

Kazakh is an agglutinative Turkic language spoken by approximately 13 million people, the majority residing in Kazakhstan, where it holds the status of state

language. Its morphological system is built on the sequential attachment of suffixes to a root, with each suffix performing a specific grammatical function [3, p. 96]. To appreciate what this means in practice, consider three concrete examples. The word “*қамбамызға*” is a single token meaning “into our granary”: “*қамба*” (granary) + *мыз* (first-person plural possessive) + *ға* (dative case marker). The word “*жазылмаған*” means “not yet written”: “*жаз*” (to write) + *ыл* (passive) + *ма* (negation) + *ған* (past participle). The phrase “*келе жатқанда*” translates as “while they were coming”: *келе* (come, converb form) + *жат* (progressive auxiliary) + *қан* (participle) + *да* (temporal case marker). In each case, a system that fails to correctly segment and interpret the suffix chain will produce a fundamentally wrong translation, and not merely an imprecise one.

The present study sets out to investigate this gap through a small but structured pilot test: four passages from canonical Kazakh literature were submitted to three AI translation tools, and the resulting outputs were analyzed using established categories of philological error analysis. The aim is not to dismiss these tools but to show precisely where and why they fall short, and to make a case for the role that trained philologists must play if AI-assisted translation of low-resource languages is to be trusted in high-stakes contexts. As Kazakhstan pursues digitalization across education, healthcare, legal, and governmental sectors, the quality of Kazakh-language AI tools becomes a matter of real social consequence, and one that philologists are uniquely positioned to address.

Methodology and Methods

The methodological framework of this study combines corpus-based testing with philological error analysis. Rather than relying on automated metrics such as BLEU or TER, which measure statistical proximity to a reference translation without regard for the quality of linguistic reasoning behind the output; this study applies a qualitative, cause-oriented approach that asks not just whether the output differs from the source, but why it differs, and what kind of linguistic knowledge would have been needed to avoid the error [6, pp. 311–312]. For a language like Kazakh, where a morphologically plausible translation can still miss the register, tone, or cultural weight of the original, statistical metrics are simply not sensitive enough to tell us what we need to know.

The test corpus consists of four passages from two major works of contemporary Kazakh literature: Mukhtar Magauin’s “*Kok Munar*” (Blue Tower) [7] and “*Altyn Orda*” (Golden Horde) [9], and Sherkhan Murtaza’s “*Ai men Aisha*” (Ai and Aisha) [8]. These authors represent the canonical core of modern Kazakh prose, published by Qasym Publishing House. The decision to use literary rather than journalistic or administrative text was deliberate. Literary prose places the highest demands on morphological and semantic interpretation. Its register also sits far enough

from the training corpora of current AI systems to expose structural limitations clearly. The four passages were selected with this in mind, as each presents a distinct morphological challenge: verbal suffix chains and military idiom in Passage (1) culturally embedded lexis and rare converb forms in Passage (2) rhetorical elongation and verbal reduplication in Passage (3) and long nominal chains with subordinate clause embedding and proper noun transliteration in Passage (4).

Each passage was submitted in its original Kazakh form to three AI-powered translation systems: Google Translate, DeepL, and ChatGPT (GPT-4). These tools were chosen because they represent distinct technological approaches and together account for a substantial share of real-world AI translation use. Google Translate employs a neural machine translation (NMT) architecture trained on a large multilingual corpus. DeepL is likewise NMT-based but has a reputation for superior syntactic fluency in European languages [1, p. 293]. ChatGPT is a large language model not specifically designed for translation, but capable of performing it through instruction following and in-context learning. All three were accessed through their standard consumer interfaces without specialized prompting or fine-tuning, and outputs were recorded exactly as produced.

Each output was then compared against the Kazakh source text at the word and phrase level. Divergences were identified and classified according to a four-category error taxonomy: morphological errors (incorrect segmentation or interpretation of suffix chains), semantic errors (wrong lexical equivalent due to misread context), syntactic errors (output structure that misrepresents grammatical relationships in the source), and pragmatic-stylistic errors (loss of register, rhetorical force, or culturally specific connotation). Where multiple error types co-occurred in a single instance, the primary cause was recorded. The analysis was conducted by the present author, drawing on training in Kazakh, English and direct familiarity with the literary and historical contexts of the selected passages.

Discussion and Observation

Before examining the passages themselves, it is worth establishing the theoretical ground on which the analysis rests. Kazakh belongs to the Kypchak branch of the Turkic language family and shares with its relatives a thoroughgoing agglutinative morphology. In agglutinative languages, grammatical relations are expressed by attaching a sequence of affixes to a stem, with each affix performing a discrete and consistent function [4, p. 20]. Unlike inflectional languages such as Latin or Russian, where a single ending may encode several categories simultaneously, Kazakh suffixes are largely one-to-one in their correspondence to grammatical meaning. A suffix chain of five elements therefore encodes five distinct layers of information, all of which must be correctly read for the translation to hold.

The field of natural language processing draws a broad distinction between high-

resource and low-resource languages. High-resource languages benefit from large volumes of digitized text, annotated corpora, and dedicated research communities. Kazakh occupies an intermediate position: not entirely without digital resources, but with existing resources heavily skewed toward contemporary, journalistic, and administrative registers, leaving literary, archaic, and dialectal Kazakh substantially underrepresented [5, p. 52]. This distributional imbalance has direct consequences for the passages tested here, as errors cluster most heavily around lexical items and grammatical structures that simply do not appear in news or web text. This pattern is consistent with broader findings across the field: comprehensive surveys of language representation in NLP research confirm that only a small number of languages command the resources needed for state-of-the-art development, while the majority, including most Turkic languages, remain critically underserved [10, pp. 6282–6283]. With that context established, we can turn to what the passages themselves revealed.

The first passage comes from Magauin's *Kok Munar* [7, pp. 32–33], a scene of military manoeuvre rich in archaic vocabulary and dense verbal suffix chains: “Жаудың көптігінен қара жер қайысатын. Соның арқасында бұларды талай жер соқтырды. Бір-біріне айлап сап, шақ-шақ қылыпта қойды. Селдей қаптап келе жатқанда бетін бұрып жіберіп, басқа жұрттармен қидаластырып қойды” [7, pp. 32–33].

Two features proved particularly resistant to automated translation. The causative phrase “талай жер соқтырды” encodes a military idiom meaning roughly “caused them to cover much ground,” that is, a suggestion of forced march or strategic displacement. Google Translate rendered this as “many earthquakes struck them,” a severe semantic error arising from the misreading of “жер” (earth/land) stripped of its directional and causative context. DeepL produced the more acceptable “forced to march them over many miles” but lost the agglutinative precision of the original. ChatGPT alone correctly identified the military-strategic register. The reduplication “шақ-шақ”, a productive morphological process that intensifies the meaning of the following form [3, p. 98], was rendered by Google Translate as “scattered,” the semantic opposite of the original, which denotes tight formation. DeepL produced “packed in tight ranks,” correctly reading the meaning but losing the iconic intensification of the reduplicate.

Table 1. Comparative analysis of Passage 1 key errors

Tool	талай жер соқтырды	шақ-шақ қылыпта қойды	Error Type
------	--------------------	-----------------------	------------

Google Translate	“many earthquakes struck them”	“scattered” (opposite meaning)	Morphological + Semantic
DeepL	“forced to march over many miles”	“packed in tight ranks”	Pragmatic-Stylistic
ChatGPT	“managed to mislead and send astray”	“keeping them at odds for months”	Pragmatic-Stylistic (minor)

The second passage comes from Sherkhan Murtaza’s “Ai men Aisha” [8, p. 515], which is a philosophical meditation on dreams and human resilience that illustrates a different kind of challenge: culturally embedded lexis that resists statistical translation:

“Адамды Құдай арманынан жерітпесін. Адамның күні арманмен. Арманнан көңіл қалса, ашық аспан да күңгірт болып көрінеді. Өмірде мән-мағына қалмай, кеудең кеуек болып, дүниенің әділетсіздігінен түңіліп, қақпанға түскен қасқырдай құр кіржіңдеп қаласың” [8, p. 515].

The word күн in the phrase “Адамның күні арманмен” can mean “day”, “sun” or in a literary register, “existence” or “life” [4, p. 34]. Google Translate selected the most statistically common meaning and produced “A person’s day is filled with dreams” reducing a philosophical statement about human existence to a banal observation about daily routine. DeepL produced “A person’s day is with their dream”, which is grammatically closer but similarly unable to access the existential weight the passage carries. The closing word “кіржіңдеп”, a verbal converb form derived from a root meaning to brood with impotent rage, is rare enough to be absent from standard translation resources. Google Translate omitted it entirely; DeepL rendered it as merely “sullen”; ChatGPT alone produced “merely snarling”, an approximation that captures the sense of constrained, futile agitation.

Table 2. Comparative analysis of Passage 2 key errors

Tool	Адамның күні арманмен	құр кіржіңдеп қаласың	Primary Error Type
Google Translate	“A perso’s day is filled with dreams”	Omitted entirely	Semantic + Omission
DeepL	“A person’s day is with their dream”	“as sullen as a wolf”	Semantic (partial)
ChatGPT	“A person lives by dreams”	“merely snarling like a wolf”	Pragmatic-Stylistic (minor)

The third passage, also from Murtaza’s “Ai men Aisha” [8, p. 99] is set during

wartime collectivization and introduces yet another dimension of difficulty, namely morphologically encoded pragmatic meaning that exists not in the words themselves but in the way they are written: “*Біз тары тербеліп тұрған тақтаға да жеттік. Басы иіліп, желмен тербеліп, теңіздей толқиды. Енді орып алмаса, төгіліп қалайын деп тұр. Орып, бастырған соң, біздің қамбамызға түсу қайда-а-а. Бәрін сыпырып-сыпырып, өкіметке өткіземіз*” [8, p. 99].

The elongated rhetorical particle “*қайда-а-а*” is a phonetically extended form of the interrogative “*қайда*” (where), and the elongation itself is the meaning: it signals bitter, resigned recognition that something will never happen. This is not a question but a lament, and any system processing text at the surface level will strip the elongation away and produce a neutral inquiry. That is precisely what Google Translate did, thereby transforming collective despair under Soviet grain requisitioning into the logistical question “where will it go in our barns?” DeepL improved slightly with “where on earth is it going to go in our granary?” ChatGPT went furthest splitting the unit into “would it ever make its way into our own granary? Not at all.” thereby reproducing the pragmatic function through a question-and-denial structure. Similarly, the reduplicated verb form “*сыпырып-сыпырып*” (to sweep up thoroughly, leaving nothing) was rendered with a single verb by both Google Translate and DeepL, losing the thoroughness encoded in morphological repetition [3, p. 98].

Table 3. Comparative analysis of Passage 3 key errors

Tool	қайда-а-а rendering	сыпырып-сыпырып rendering	Error Type
Google Translate	“where will it go?” (neutral)	“sweep it all up”	Pragmatic + Morphological
DeepL	“where on earth is it going?” (partial)	“sweep it all up”	Pragmatic (partial)
ChatGPT	“would it ever...? Not at all.” (split)	“gather up every last bit”	Pragmatic-Stylistic (minor)

The fourth and most syntactically demanding passage comes from Magauin’s “Altyn Orda” [9, p. 190], which is a historiographical reflection on steppe ethno genesis built around the connective “*Tek ..., ... емес, ...*” (“not merely...but rather”): “*Есесіне Алтын Орда іргелене түсті. Бұл кезде барша халықты бұрынғыша қыпшақ-түрік, татар-түрік деп жіктесек, екі тарап теңе-тең жаздайға жетеді, бәлкім, Шығыс өлке тұрғындары басым түсуі де мүмкін. Арада жүз жыл өтпей, манағы әл-Омари айтқандай, өзара туысып, жаппай араласып кетті. Тек татардың қыпшаққа айналуы, немесе керісінше, қыпшақтың татар шығуы, яғни әлденедей ассимиляция емес, үш ғасыр*

бұрын, Шығыс Түрік қағанаты, нәсіл-тегі ортақ жұрт қайтадан тоғысып, біртұтас қауымға айналды” [9, p. 190].

This construction requires the reader to hold an opening negative clause in suspension across several lines before resolving it with a contrastive conclusion, precisely the kind of long-distance dependency that challenges transformer-based models [2, p. 1717]. Google Translate disrupted the logical structure entirely, presenting the argument as loosely connected statements rather than a carefully built contrast. It also rendered the Arabic proper noun “әл-Омари”, the name of the medieval Egyptian scholar “al-Umari”, transliterated into Kazakh script, as “Manaq al-Omari” splitting the Arabic definite article (әл-) and treating it as a separate Kazakh word. Misidentifying a cited scholarly source is not a minor stylistic slip; it has direct implications for any reader relying on the translation for research. DeepL preserved more of the structure but truncated mid-sentence at the most complex subordinate clause. ChatGPT alone produced a complete and structurally coherent translation, correctly identifying the “Tek ..., ... емес, ” construction as a contrastive frame.

Table 4. Comparative analysis of Passage 4 key errors

Tool	“Tek ..., ... емес,.....”	әл-Омари	Completeness	Error Types
Google Translate	Lost: presented as list	“Manaq al-Omari” (fragmented)	Complete	Syntactic + Proper Noun
DeepL	Partially preserved	“al-Omari” (correct)	Truncated mid-clause	Syntactic + Omission
ChatGPT	“not simply... rather” (correct)	“al-Omari” (correct)	Complete	None significant

Results

Taking together, the four passages support a clear and consistent ranking of the three tools. ChatGPT demonstrates the strongest overall performance, showing sensitivity to literary register, rhetorical devices, and long-range syntactic dependencies. Its failures are primarily of the pragmatic-stylistic type. These include occasional paraphrasing where more precise rendering would be preferable. Importantly, such failures are qualitatively different from the structural failures observed in the other two tools.

DeepL occupies an intermediate position. It is grammatically more acceptable than Google Translate in most cases and handles proper nouns more accurately. However, it exhibits a truncation failure in the longest passage and consistently loses cultural and literary depth.

Google Translate performs worst across all four passages. Its errors are both the most frequent and the most severe. These include semantic reversals, structural disruptions, and proper noun fragmentation.

Table 5. Summary of tool performance across analytical criteria

Criterion	Google Translate	DeepL	ChatGPT
Morphological accuracy	Weak	Moderate	Strong
Semantic accuracy	Weak	Moderate	Strong
Syntactic structure	Moderate	Moderate	Strong
Literary register	Weak	Moderate	Strong
Rhetorical devices	Weak	Partial	Moderate
Proper noun handling	Weak	Strong	Strong
Completeness	Strong	Partial	Strong

What is striking about these results is that the errors are not random. They cluster around three structural features of Kazakh that are particularly resistant to current AI processing: Firstly, agglutinative suffix chains involving causative, aspectual, and directional morphemes stacked onto a single root [3, p. 96]. Secondly, morphologically encoded pragmatic information such as elongated vowels, reduplicated forms, and zero-marked illocutionary functions that cannot be recovered from the surface string alone. Thirdly, long-distance syntactic dependencies in complex subordinate clause structures. All three features require a processing system to integrate multiple levels of linguistic information simultaneously, including grammatical competence, register awareness, and cultural familiarity with the literary tradition.

These results also carry broader social consequences. In the medical domain, semantic reversals of the type produced by Google Translate could cause direct patient harm if used in clinical communication. In the legal domain, fragmentation of proper nouns and loss of syntactic structure in complex clauses can alter the interpretation of documentary evidence. In the educational domain, inaccurate AI translation of literary texts can distort students' understanding of the Kazakh literary tradition. Kazakhstan's ongoing digitalization program, which has elevated Kazakh to the status of state language, makes the quality of these tools a matter of national significance, not merely academic interest.

This is not a problem that can be fixed simply by adding more training data. It is a structural challenge that requires human expertise at the point of evaluation, and that is where the trained philologist's role becomes indispensable. Even ChatGPT,

which substantially outperforms its competitors, does not produce translations acceptable for scholarly, legal, or medical use without human review. The errors it makes are precisely those hardest for non-specialists to detect: they do not produce obviously nonsensical output, but rather subtly flatten or distort the meaning of the source. The philologist's role is therefore threefold: as evaluator, identifying errors that statistical metrics would score highly because they produce fluent English while misrepresenting the Kazakh source [6, p. 311]; as annotator, contributing higher-quality training data from literary and specialized registers; and as system architect, specifying the morphological and stylistic features that NLP systems must correctly handle.

Conclusion

This study set out to answer a specific question: how well do three widely used AI translation tools, namely ChatGPT (GPT-4), Google Translate, and DeepL, and how they handle the morphological and literary complexity of Kazakh? The answer, based on a philological analysis of four literary passages, is that none of them handles it well enough for high-stakes use without expert human oversight. The comparison was made precisely because these three tools currently dominate real-world AI translation use, and understanding their relative strengths and failures is essential for anyone working with Kazakh language materials in a professional or academic context. The results reveal a consistent hierarchy: ChatGPT leads, DeepL occupies the middle ground, and Google Translate trails significantly, yet none of them clears the bar for independent use in contexts where accuracy matters.

The novelty of this study lies in its approach. Rather than applying automated metrics that measure surface-level similarity to a reference translation, it uses a cause-oriented philological taxonomy that identifies not just what went wrong but why, and what kind of linguistic knowledge would have been needed to avoid the error. This approach reveals something that automated metrics cannot: the failures of these tools are structurally predictable, concentrated at the exact points where Kazakh morphological complexity, including agglutinative suffix chains, pragmatically encoded forms, and long-distance clause dependencies, diverges most sharply from the analytic language patterns that dominate AI training data.

The significance of this finding is both theoretical and practical. Theoretically, it demonstrates that the gap between AI capability and philological need is not a temporary data deficit but a structural mismatch rooted in how these systems are designed. Practically, it points toward a clear path forward: the development of philologist-in-the-loop translation quality assurance systems for Kazakh, where AI-generated drafts are reviewed by trained specialists at the points of highest

error risk. Students of philology are well positioned to drive this process, not by replacing AI tools, but by becoming the expert human layer that makes those tools trustworthy. As Kazakhstan advances its digitalization agenda and as the Kazakh language asserts itself more fully in public and professional life, building that layer of expertise is not optional. It is a necessity.

References

1. Forcada, M. L. Making sense of neural machine translation. *Translation Spaces* vol. 6, no. – Amsterdam, The Netherlands: John Benjamins Publishing Company, 2017. - pp. 291–309.
2. Sennrich, R., Haddow, B., Birch, A. Neural machine translation of rare words with subword units. *Proceedings of ACL*. – Berlin: Association for Computational Linguistics, 2016. – pp. 1715–1725.
3. Kessikbayeva, G., Cicekli, I. A rule-based morphological analyzer and a morphological disambiguator for Kazakh language. *Linguistics and Literature Studies*, vol. 4, no. – Los Angeles, USA: Horizon Research Publishing, 2016. - p. 104.
4. Comrie, B. *The Languages of the Soviet Union*. - Cambridge: Cambridge University Press, 1981. – p. 317.
5. Kirchhoff, K., Turner, A. M. Unsupervised resolution of acronyms and abbreviations in nursing notes using document-level context models. *Proceedings of the 7th International Workshop on Health Text Mining and Information Analysis*. – Texas, USA: Association for Computational Linguistics, 2016. - pp. 52–60.
6. Papineni, K., Roukos, S., Ward, T., Zhu, W.-J. BLEU: A method for automatic evaluation of machine translation. *Proceedings of ACL*. – Philadelphia USA: Association for Computational Linguistics, 2002. - pp. 311–318.
7. Magauin, M. *Kok Munar (Blue Tower)*. - Almaty: Qasym Publishing House, 2019. – p 318.
8. Murtaza, Sh. *Ai men Aisha (Ai and Aisha)*, vol.: 2nd collection. - Almaty: Qasym Publishing House, 2019. – p. 686.
9. Magauin, M. *Altin Orda (Golden Horde)*. - Almaty: Qasym Publishing House, 2019. – p. 446.
10. Joshi, P., Santy, S., Budhiraja, A., Bali, K., Choudhury, M. 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*. – Online: Association for Computational Linguistics, 2020. - pp. 6282–6293.

С.Қ. АРЫСБЕК¹, Т.Ә.ТӘУБЕЕВА²

ЖАСАНДЫ ИНТЕЛЛЕКТ ҚҰРАЛДАРЫ АРҚЫЛЫ ҚАЗАҚ ТІЛІНІҢ МОРФОЛОГИЯЛЫҚ ЖҮЙЕСІН АУДАРУ: ТӘЖІРИБЕ ЖӘНЕ СЫНАУ

Аңдатпа. Үлкен тілдік модельдер мен нейрондық машиналық аударма архитектураларының қарқынды дамуы лингвистикалық жұмыс процестерін автоматтандыруға қатысты үлкен оптимизм тудырды. Оптимизмді бейсызық жұрнақтар қазақ тілі секілді морфологиялық жағынан бай, жалғамалы (агглютинативті) жүйелерге қолдану барысында қатаң эмпирикалық тексеруден өткізу талап етіледі.

Мақаланың зерттеу мақсаты кеңінен қолданылатын ChatGPT (GPT-4), Google Translate, DeepL үш аударма құралдардың көркем әдебиеттегі морфологиялық және стилистикалық реңктерін өңдеу қабілеттерін бағалау және аударма дефектілерінің құрылымдық этиологиясын анықтау. Зерттеудің бағалау жүйесін құруда Мұхтар Мағауин мен Шерхан Мұрта-заның канондық прозаларының сөз құрамындағы морфемалардың жоғары арақатынасы және мәдени маркерленген лексикалық тереңдігімен ерекшеленетін төрт мәтіндік үзінді іріктеліп талқыланады.

Базалық архитектуралық мүмкіндіктерді оқшаулау мақсатында әрбір мәтін нөлдік іріктеу әдісімен енгізілді. Алынған нәтижелер морфологиялық, семантикалық, синтаксистік және прагматикалық-стилистикалық өлшемдерді қамтитын төрт деңгейлі филологиялық қателер таксономиясы бойынша талданады.

Эмпирикалық нәтижелер құралдардың архитектуралық құзыреттілігінің айқын иерархиясын айқындайды. ChatGPT дауыстылар мен етістіктердің қосарлануы сияқты тілге тән риторикалық тәсілдерге сезімталдық танытып, жоғары мәнмәтіндік модельдеу қабілетін көрсетеді.

Google Translate құралы маңызды семантикалық инверсияларды, құрылымдық қателерді және транслитерацияланған жалқы есімдердің қате сегментациясын көрсететінін байқатады.

DeepL синтаксистік тұрғыдан дұрыс аударма жасағанымен, бірақ мәтіннің мәдени және идиомалық тереңдігін үнемі әлсіретіп отырады.

Мақаланың теориялық маңыздылығы үндіеуропалық құрылымдарға оңтайландырылғанын атап өтіп, жеткілікті түрде ескерілмейтін түркі тілдеріне аударма жасау қателері кездесетінін көрсету.

Тәжірибелік маңыздылығы цифрландыру бастамалары үшін нақты әлеуметтік-лингвистикалық тәуекел тудырады және бұл процеске адамның тікелей қатысуын қамтамасыз ету қажеттілігін негіздейді.

Зерттеу барысында автоматтандырылған аударма жүйелеріндегі токени-

зация стратегияларын жетілдіру және соған филологтардың құрылымдық тұрғыдан белсенді қатысуы міндетті екені туралы қорытынды жасалады.

Түйін сөздер: қазақ морфологиясы, машиналық аударма, табиғи тілді өңдеу, жалғамалы тілдер, филологиялық қателерді талдау, ЖИ аудармасын бағалау.

С.К. АРЫСБЕК¹, Т.Ә.ТӘУБЕЕВА²

Египетский университет исламской культуры Нур-Мубарак

ПЕРЕВОД MORFOЛОГИЧЕСКОЙ СИСТЕМЫ КАЗАХСКОГО ЯЗЫКА С ПОМОЩЬЮ ИНСТРУМЕНТОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: ОПЫТ И ТЕСТИРОВАНИЕ

Аннотация. Стремительное развитие языковых моделей и архитектур нейронного машинного перевода породило значительный оптимизм в отношении автоматизации лингвистических рабочих процессов. Однако применение дан-ного оптимизма к морфологически богатым агглютинативным системам, та-ким как казахский язык, требует строгой эмпирической проверки.

Целью статьи является оценка способности трёх широко используемых переводческих инструментов – ChatGPT (GPT-4), Google Translate, DeepL передавать морфологические и стилистические оттенки художественного текста, а также выявление структурной этиологии переводческих дефектов.

Для построения системы оценки были отобраны и проанализированы четыре текстовых фрагмента из канонической прозы Мухтара Магауина и Шерхана Муртазы, отличающиеся высокой концентрацией морфемных структур и культурно маркированной лексической глубиной.

С целью изоляции базовых архитектурных возможностей каждый текст вводился методом zero-short prompting. Полученные результаты анализируются в рамках четырёхуровневой таксономии филологических ошибок, охватывающей морфологические, семантические, синтаксические и прагматико-стилистические параметры.

Эмпирические результаты выявляют чёткую иерархию архитектурной компетентности инструментов. ChatGPT демонстрирует высокий уровень контекстуального моделирования, проявляя чувствительность к языковым риторическим приёмам, таким как удвоение гласных и глагольных форм.

Google Translate показывает наличие существенных семантических инверсий, структурных нарушений и ошибочной сегментации транслитерированных имён собственных.

DeepL несмотря на синтаксически корректный перевод, систематически ослабляет культурную и идиоматическую глубину текста.

Теоретическая значимость статьи заключается в демонстрации того, что современные системы машинного перевода преимущественно оптимизированы под индоевропейские структуры, тогда как при переводе тюркских языков, остающихся недостаточно представленными в обучающих корпусах, возникают устойчивые ошибки.

Практическая значимость исследования состоит в том, что подобные ограничения создают конкретные социолингвистические риски для инициатив цифровизации и обосновывают необходимость непосредственного участия человека в данном процессе.

В ходе исследования сделан вывод о необходимости совершенствования стратегий токенизации в автоматизированных системах перевода, а также о структурно активном участии филологов в разработке и оценке подобных технологий.

Ключевые слова: казахская морфология, машинный перевод, обработка естественного языка, агглютинативные языки, филологический анализ ошибок, оценка ИИ-перевода

Information about the author:

Arysbeek Serikuly Kadyr – 3rd year student, Foreign Philology (Arabic-English), Nur-Mubarak Egyptian University of Islamic Culture, Almaty, Kazakhstan;

Арысбек Серікұлы Қадыр – Шетел филологиясының 3-курс студенті, (Араб-Ағылшын), Нұр-Мұбарак Египет ислам мәдениеті университеті, Алматы, Қазақстан;

Арысбек Серикулы Кадыр – студент 3 курса, Иностранная Филология (Арабский-Английский), Египетский университет исламской культуры Нур-Мубарак, Алматы, Казахстан.

Taubeyeva Aliya Taubekyzy – senior teacher, Nur-Mubarak Egyptian University of Islamic Culture, Almaty, Kazakhstan;

Тәубеева Әлия Тәубеқызы – аға оқытушы, Нұр-Мұбарак Египет ислам мәдениеті университеті, Алматы, Қазақстан;

Тәубеева Әлия Тәубеқызы – старший преподаватель, Египетский университет исламской культуры Нур-Мубарак, Алматы, Казахстан.