G.Zh. Baishukurova[1*], A.B. Irgebayeva[2], N.N. Aitova[3*]

[1, 2]Abai Kazakh National Pedagogical University, Almaty, Kazakhstan;
[3] Maksut Narikbayev University, Astana, Kazakhstan, e-mail:
(e-mail: [1]baishukurova@mail.ru, [2]kaldanov70@mail.ru ,[3]nurlykhan.an@gmail.com)

Akhmet Baitursynuly's Parallel Corpus: Lexical transformations in translation

This article explores various types of lexical transformations applied in the translation of Akhmet Baitursynuly's selected works into Russian. The research is based on the developing Kazakh-Russian parallel corpus of Baitursynuly's texts, which serves as a valuable linguistic resource. The study is motivated by the authors' dual role as both translators of Baitursynuly's works and developers of the parallel corpus. The objective of the research is to analyze the characteristics of lexical transformations in translation and to outline the principles of working with the Kazakh-Russian parallel corpus in translation practice. The advantages of using a parallel corpus as a tool for translation research and practice are also discussed. The study identifies several types of lexical transformations, including transcription, transliteration, concretization, generalization, lexical addition, omission, semantic development, and holistic transformation. A significant contribution of this study is the development of software that enables the automatic comparison of Russian and Kazakh texts. This tool enhances translation efficiency by reducing the time required to identify equivalents, facilitates translation analysis, and enables an assessment of the adequacy of applied translation strategies. The findings contribute to the advancement of translation studies and corpus linguistics, demonstrating the effectiveness of parallel corpora in translation research and practice. Particular attention is given to custom-built software developed to automatically compare Kazakh and Russian texts. This tool speeds up the process of finding equivalents, increases the accuracy of translation analysis, and enables the assessment of translation strategies. The results presented make a valuable contribution to the fields of translation studies and corpus linguistics, creating new opportunities for researching Baitursynov's legacy. The relevance of this study is supported by current trends in the use of corpus linguistics in translation studies. The parallel corpus does not perform automatic translation functions but supports in-depth analysis to improve translation quality. This article is the first to provide a detailed description of the development stages and application principles of the Kazakh-Russian parallel corpus of A. Baitursynov. The article also presents examples of translators' detailed work to ensure accurate transfer of meaning and form, along with average statistics on the types of lexical transformations found in the analyzed texts.

Keywords: A. Baitursynuly, corpus linguistics, parallel corpus, concordance, translation, lexical transformations.

Introduction

The development of parallel corpora represents one of the most significant and rapidly advancing areas in contemporary Kazakhstani linguistics. Considerable attention is devoted to the creation and utilization of parallel corpora, which consist of texts in different languages that maintain identical content. These corpora serve as essential tools for linguistic research, comparative language analysis, and the advancement of machine translation technologies [1].

This study aims to analyze equivalent lexical units within the Kazakh-Russian parallel corpus of Akhmet Baitursynuly. The research objectives include tracing translation logic, evaluating the necessity and effectiveness of applied translation strategies, and assessing the adequacy of translations through a comparative analysis of source and target texts.

The study hypothesizes that various types and techniques of lexical transformations are integral to achieving high-quality translations. Lexical transformations are unavoidable in translation, as they enhance the accuracy and adequacy of both individual textual elements and the translated text as a whole. The selection of transformation techniques depends on multiple factors, particularly in addressing culturally specific concepts and terms. To preserve Baitursynuly's linguistic style and imagery, translators must employ appropriate strategies that align with the lexical, cultural, and stylistic features of the original language. Certain types of lexical transformations serve as key mechanisms in ensuring a precise and contextually appropriate rendition of meaning [1; 18].

---

* Corresponding author's e-mail: *baishukurova@mail.ru*

The significance of this research lies in the application of corpus linguistic tools, including concordances and parallel corpora, for conducting comparative and contrastive translation analysis. The parallel corpus functions as an invaluable research instrument, facilitating a deeper understanding of literary works for linguists, philologists, and researchers. Rather than being designed for automatic translation, the corpus provides a framework for comprehensive analysis, enabling the exploration of translation features and the refinement of translation quality.

For the first time, this study describes the principles of working with the Kazakh-Russian parallel corpus of Akhmet Baitursynuly in the context of translation. The developed software allows for an automated comparison of Russian and Kazakh texts, significantly reducing the time required to identify lexical equivalents. Furthermore, it enables systematic translation analysis and assessment of the effectiveness of specific translation strategies. The empirical and statistical data obtained contribute to both theoretical research and practical applications in translation studies.

Methods and materials

The primary material for this study is the one-volume edition of *Selected Works* by Akhmet Baitursynuly [2], [3], which was specifically prepared for translation into Russian, English, and Turkish. This edition comprises scholarly and academic texts, along with selected journalistic works by the author. The original texts and their Russian translations have been incorporated into the database of the Kazakh-Russian parallel corpus, allowing for their comprehensive use in the research. The necessary data were extracted using corpus-based methods and subsequently analyzed.

A combination of linguistic and computational methods was employed to examine lexical transformations in translation. The methodological framework consists of the following key approaches:

Data collection: The first stage of the research involved the collection and preparation of materials. A continuous sampling method was applied to extract linguistic data from Baitursynuly's texts and their Russian translations. To ensure precise segmentation and alignment of texts, corpus tools were used for text structuring, allowing for an accurate comparison of source and target language units.

Data analysis: Following the data collection, segmentation and comparison were conducted to divide the original and translated texts into corresponding units. This facilitated a comparative and contrastive analysis, which systematically identified similarities and differences between the Kazakh originals and their Russian translations.

Lexical and semantic analysis: The study also employed lexical-semantic analysis to examine the transformation of lexical units during translation. A semantic analysis method was used to investigate shifts in meaning, while componential analysis helped identify structural changes in vocabulary. Additionally, statistical analysis was applied to quantify the frequency of lexical transformations, revealing patterns in translation strategies.

Classification of transformations: To further systematize the findings, the research focused on the classification of lexical transformations. This process involved assessing the degree of correspondence between lexical units in the original text and their equivalents in the Russian translation. Based on the identified patterns, different types of lexical transformations were categorized to provide a structured overview of translation strategies.

Parallel corpus formation and data processing: A crucial aspect of the research was the formation and processing of the Kazakh-Russian parallel corpus of Baitursynuly's works. This bilingual database was designed to facilitate the identification of lexical transformations in translation by aligning the original Kazakh texts with their Russian translations. The compilation of the corpus followed several key stages. First, text data were collected and curated for inclusion in the corpus. Next, the texts underwent markup and preparation, during which all materials were standardized to a uniform format, extraneous symbols were removed, and spelling differences were normalized. Sentences were then segmented into smaller units that corresponded to semantic structures.

After the automated alignment process, a verification and correction stage was carried out. Manual review ensured the accuracy of segment matches, and adjustments were made when discrepancies were found by refining segment boundaries or adding new aligned pairs. Once the corpus was finalized, the data underwent further processing. Statistical processing was conducted to measure the frequency of lexical units and their translation equivalents, as well as to identify recurring translation strategies. Additionally, automated processing was applied, utilizing N-gram analysis to compare segments and detect patterns in lexical transformations.

Through the integration of these methodological approaches, the study provides a structured and empirical basis for analyzing translation strategies within the Kazakh-Russian parallel corpus, offering valuable insights into the dynamics of lexical transformations in translation.

The methodological approach adopted in this study provides a systematic and empirical basis for analyzing translation strategies within the Kazakh-Russian parallel corpus, offering valuable insights into the dynamics of lexical transformations in translation.

In modern translation studies, comparative translation analysis is increasingly conducted using corpus linguistics tools, particularly concordances and parallel corpora. However, a parallel corpus does not serve as a tool for direct text translation; rather, it functions as an instrument for an in-depth and comprehensive analysis of translated texts. This approach enhances translation quality by providing empirical and statistical data that can be applied in both translation theory and practice.

The development of the Kazakh-Russian parallel corpus of A. Baitursynuly [4] was driven by the necessity of conducting a scientific analysis of his vocabulary and grammatical structures, as well as examining historical linguistic changes that took place during the formation of the Kazakh script. Many of Baitursynuly's manuscripts have been preserved in different versions of the Arabic script (tote zhazu) and reflect the spelling norms of their time. These texts were transliterated from the Arabic script into Cyrillic to facilitate linguistic analysis. The foundation of the parallel corpus consists of two primary texts: the original works of Baitursynuly in Kazakh and their Russian translations.

The use of a parallel corpus enables researchers to identify lexical correspondences and explore the principles, methods, and techniques employed in translating words and concepts from one language to another. Scholars working with existing parallel corpora suggest that such tools allow for the identification of translation strategies, the evaluation of their effectiveness, and the comparative analysis of vocabulary and grammar between the original and translated texts. Furthermore, they help in recognizing stylistic features and determining how these elements are conveyed in translation [5; 22].

Parallel Kazakh-Russian sub-corpora are currently being developed as part of the broader effort to create a National Kazakh Language Corpus [6]. One such sub-corpus includes the original Kazakh text of M. Auezov's epic novel *Abai Zholy* and its Russian translation [7]. In addition, a parallel Kazakh-Russian corpus has been compiled using multilingual news sites in Kazakhstan, encompassing a total of 50,410 words, with 25,600 in Kazakh and 24,800 in Russian [8].

Research initiatives at Al-Farabi Kazakh National University have focused on constructing an English-Kazakh parallel corpus containing one million words from legal texts. This project also involves developing an English-Kazakh database of legal translations and establishing a statistical machine translation system [9].

Translation transformations are a central aspect of translation studies, and numerous academic works have explored this topic. Understanding translation transformations is crucial for any translator, as it forms the foundation of their professional knowledge and practice. Scholars have offered various perspectives on this issue, resulting in multiple definitions and classifications of translation transformations. Despite extensive research in this area, experts in the field, including Vinae et Darbelnet, P. Newmark, L.S. Barkhudarov, V.N. Komissarov, A.D. Schweitzer, R.K. Minyar-Beloruchev, and Y.I. Retsker, have yet to reach a consensus regarding the precise nature and interpretation of translation transformations.

Results

The development of a concordance and parallel corpus for Akhmet Baitursynuly is being undertaken for the first time in Kazakhstani lexicography. This project aims to provide a broad audience with a software and information environment for studying the literary and scholarly legacy, language, and biography of the prominent Kazakh educator. The Kazakh-Russian parallel corpus of Baitursynuly is currently in development, offering users an interface that allows for the retrieval of words, word forms, and sentences within their contextual framework (Fig. 1).
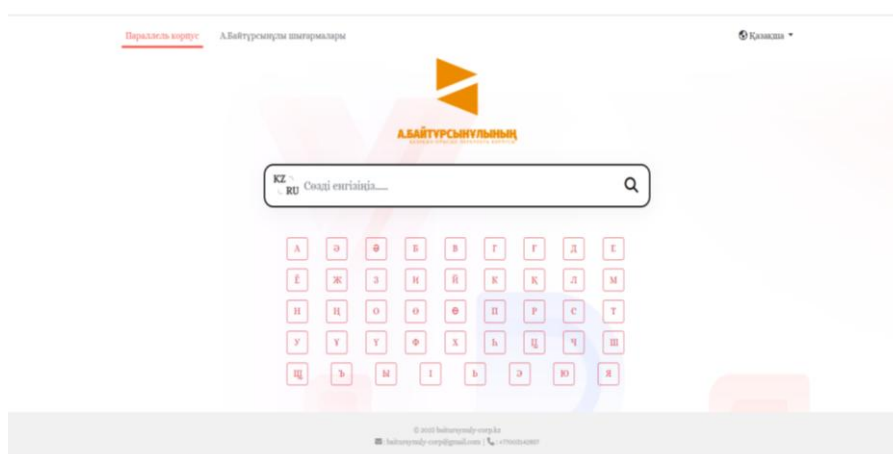
Figure 1. Home page of the site.

The interface of the parallel corpus allows you to find words, word forms and sentences with reference to the context (Fig. 2).
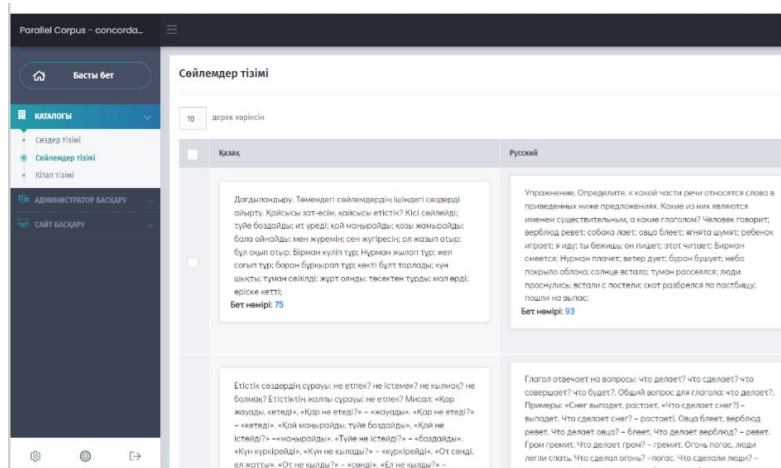

Figure 2. A. Baitursynuly Parallel corpus interface

Examining the features of the parallel corpus and its potential applications in translation work highlights the necessity of translation transformations. Achieving an adequate and equivalent translation often entails certain unavoidable losses. However, a range of linguistic transformations allows for the preservation of overall textual adequacy. These transformations serve to address structural and semantic differences between languages, making them an essential tool for translators.

L.S. Barkhudarov defines translation transformations as a diverse set of modifications aimed at achieving equivalence despite formal and semantic discrepancies between the source and target languages [10; 237].

The theoretical foundation for this study is based on V.N. Komissarov's classification of translation transformations, which includes lexical, grammatical, and complex (lexical-semantic) types. This study focuses specifically on lexical transformations, which involve replacing lexical units from the source language with non-dictionary equivalents in the target language, conveying meaning differently from the original units [11; 77].

Kazakhstani researchers have explored various aspects of translation transformations. For example, studies have examined the strategies used to translate culturally specific Kazakh vocabulary into Russian and English, particularly in works such as Ilyas Yesenberlin's novel-trilogy The Nomads [12]. The research highlights the frequent use of combined transformation methods, such as word order modification alongside lexical substitutions, or transliteration combined with descriptive explanations. Similar approaches have been identified in translations of Abai's Words of Edification into Russian and English, where translators employ different transformation techniques, including reordering sentence elements, substituting words through generalization or antonymic translation, adding explanatory content, and omitting inconsistencies: 1) order of

place (alternation of individual words and phrases in a sentence); 2) substitution (generalization, antonymic translation, compensation); 3) addition (inclusion of linguistic units by the translator from his side in conveying meaning in the text); 4) omission (exclusion from the text of some grammatical, lexical, semantic inconsistencies) [13], [14].

There are studies devoted to various approaches and tools in translation and computational linguistics. In particular, the use of corpora in teaching translation [15, 16], empirical and computational methods of processing parallel texts [17]), as well as the role of translation in intercultural communication [18]) are considered. In addition, modern machine translation technologies are being studied, including transformer models [19]) and their application to historical and classical languages [20, 21].

Despite extensive studies on translation transformations in Kazakh literature, there are no published works analyzing lexical transformations in the Russian translations of Baitursynuly's writings using a parallel corpus. This gap further underscores the originality and significance of the current research.

The effectiveness of a translation largely depends on the translator's theoretical training and professional expertise. A skilled translator must be able to identify optimal solutions to linguistic challenges and apply accumulated knowledge and experience effectively. A high-quality translation requires not only a faithful reproduction of the author's ideas but also adherence to the stylistic and linguistic norms of the target language. Achieving this level of accuracy involves a thorough and multi-stage translation process, during which various linguistic strategies are employed to ensure fidelity to both content and form. This rigorous approach is particularly evident in the translation of Baitursynuly's works, where careful selection of lexical transformations plays a crucial role in maintaining the integrity of the original text (Table 1).

T a b l e 1

Comparative study of translations

Text in Kazakh	Russian translation variant	The final version	Translation into English
Sauattanyu zhas balalardynn zhuriuge uirengeni **mysaldy: Bastapqy kezde** balalar **ayagyn bir basyp, bir turady, onan sonn apyl-tapyl basyp** zhuriuge zharaidy. **Onan keyin aqyryn** da bolsa kadimgidei zhuretin bolady. Onan ari zhugiriuge zharaidy, **zhugiriuge zharagannan ari zharysqa tusip,** zhuristi oner esebinde qoldanyu darazhaga **zhetedi**.	**Ovladenie gramotoi pohozhe** na to, kak rebyenok uchitsya hodit: **v pervyi period** rebyenok **delaet odin shag i ostanavlivaetsya, posle chego** nachinaet topat. Posle etogo on **ponemnogu** nachinaet hodit obychnym shagom. **Potom** on sposoben begat, **posle bega mozhet sorevnovatsya, dostigaya** vysokogo urovnya v masterstve hodby.	**Obuchenie gramote podobno** tomu, kak rebyenok uchitsya hodit: **snachala** rebyenok **uchitsya perestavlyat nozhki, a potom uzhe** nachinaet ponemnogu topat. Posle etogo on **potihonku** nachinaet hodit obychnym shagom. **Dalee** on **uzhe mozhet** begat, **zatem on sorevnuetsya v bege i dostigaet** vysokogo urovnya hodby	**Literacy training is similar** to how a child learns to walk: **first**, the child **learns to move his feet, and then** he begins to stomp a little. After that, he **slowly** begins to walk at his usual pace. **Then** he **can run, then he competes in running** and reaches a high level of walking.

This example clearly illustrates the distinctions between the initial and final versions, which are largely influenced by the choice of lexical units and the subsequent process of their repeated transformation.

The use of a parallel corpus significantly accelerates the identification of various types of lexical transformations. When translating between languages, it is essential to consider their typological characteristics. In general, lexical transformations can be categorized into several types, including transcription, transliteration, concretization, generalization, lexical addition, omission, semantic development, and holistic transformation.

Transcription involves reproducing the sound structure and phonemic composition of the original lexical unit using the letters of the target language. This process occurs not only in the translation of the author's text into Russian (Table 2), but is also present in the original Kazakh text (Table 3).

Translating ethno-cultural vocabulary presents challenges, as it requires the translator to accurately convey linguistic elements that are unique to a specific language and culture. Non-equivalent vocabulary typically includes realia or culturally specific terms such as names of traditional foods, beverages, tools, clothing, exotic expressions, ethnonyms, professional jargon, and authorial neologisms.

In these instances, historic realia serve as a lexical tool for national and historical stylization, helping to recreate the cultural context of the original text. Transliteration, in particular, enhances the stylistic impact of

realia, preserving the national and cultural essence of the text. Compared to descriptive explanations, transliteration retains more of the original's emotional and expressive nuances, maintaining its authenticity in translation.

T a b l e 2

Comparative study

Original text in Kazakh	Translation into Russian	Translation into English
Zhaqsy kelse qut, zhaman kelse zhut.	*Pridyet khoroshiy chelovek – blago, pridyet plohoi chelovek – jut.*	*If a good person comes, it's a blessing, if a bad person comes, it's zhut.*
Khan degendi qosyp aitqanda, qai Zhanibek ekeni ashylyp, anyq bolady.	*Pri upotreblenii slova khan stanovitsya yasno, o kakom Janibeke idet rech.*	*When using the word khan, it becomes clear which Zhanibek we are talking about.*
Zhigit qyz tanndau ushin zhuripti.	*Jigit hodil dlya vybora nevesty.*	*Zhigit went to choose a bride.*
Shapan zhanna.	*Chapan novyi.*	*Shapan is new*
Shai ishti, kitaptaryn aldy, oquga ketti.	*Popil chai, vzyal knigi, poshyel v shkolu.*	*I had tea, picked up books, and went to school.*
Qazaqtynn oz sozi bolyp ketken arab ya farsy sozderi bar.	*Est slova iz arabskogo yazyka ili farsi, kotorye stali sobstvenno kazakhskimi slovami.*	*There are words from Arabic or Farsi that have become Kazakh words.*

In the Russian translation, transcription is employed to retain the ethnic and cultural distinctiveness of lexical units. Traditionally, the Kazakh letter "zh" at the beginning of words is rendered as "j" in Russian. Additionally, common transcription patterns in the Russian version include replacing "sh" with "ch", "y" with "i", and "I (soft)" with "i", among others.

T a b l e 3

Comparative study

Original text in Kazakh	Translation into Russian	Translation into English
*Ony **Istepany** syzyqpen bylaisha korsetedi.*	*Eto **Stephani** pokazal v vide chyertochek.*	***Stephanie** showed this in the form of dashes.*
***Ameriykan** adisi dep aitylyp zhurgen adistinn negizi Zhakotoniki.*	*V osnove tak nazyvaemogo **amerikanskogo** metoda lezhit metod Zhakoto.*	*The so-called **American** method is based on the Jacotot method.*
*Muny aueli qoldanyushylar: **Pogel, Tomas, Belme** degen nemis oqytyushylary, sonan basqalar alyp turlishe qoldanady.*	*Vpervye stali shiroko polzovatsya dannym metodom nemetskie uchitelya **Fogel, Tomas, Belme** i drugie.*	*For the first time, German teachers **Vogel, Thomas, Belme** and others began to widely use this method.*
***Orynborga** shaqyrgan qazaq qairatkerlerininn siezi adam az kelgendikten siez esebinde bolmai, kishkentai gana kennes turinde bolyp otkeni ozgan nomirlerinde zhazyldy.*	*V predydaschih vypuskah pisalos o tom, chto iz-za malochislennosti priehavshih v **Orenburg** sezd kazakhskih obschestvennyh deyatelei proshyel ne kak sezd, a v vide nebolshogo soveschaniya.*	*In previous issues, it was written that due to the small number of Kazakh public figures who came to **Orenburg**, the congress was held not as a congress, but as a small meeting.*
*Qazaq balasy ultym, zhurtym, bauyrym dep uirenip qalgan qazaqtynn bauyrmal qalam qairatkerleri **oktabir** ozgerisi bolganda birden iinternatsiyanal (beibauyrmal) bolyp ozgere almady, ozgelerdei «alymsaqtan beri» **kammunist, internatsianalist** edim dep aituga auzy barmady.*	*Schitauschie kazakhov svoim rodnym narodom, vernye svoei natsii, kazahskie pisateli ne mogli posle **oktyabrskoi** revolyutsii v odnochasie izmenitsya i stat **internatsionalistami** (vne natsionalnosti), po utverzhdeniyu nekotoryh, yakoby davno yavlyayuschihsya **kommunistami, internatsionalistam**.*	*Kazakh writers who consider Kazakhs to be their native people, loyal to their nation, could not change overnight after the **October** Revolution and become **internationalists** (beyond nationality), according to some who have allegedly been **communists** and **internationalists** for a long time.*

A. Baitursynuly adapted the pronunciation of foreign words to align with the phonetic system of the Kazakh language, while translators adjusted them to fit Russian pronunciation. For instance, "Istepany" became "Stephani", "Pogel" was rendered as "Fogel", "Belme" was softened to "Belme", "oktabir" was adapted as "oktyabr(skoi)", and "kammunist" was translated as "communist".

In certain cases, transcription and transliteration were combined, as seen in "ameriykan" transforming into "amerikanskiy" and "internatsiyanal/internatsianalist" becoming "internatsionalist". In these instances, both the phonetic and graphic forms of the words were adjusted using the alphabet of the target language.

The variability in spelling of foreign words can be attributed to historical and linguistic factors, particularly when standardized phonological norms had not yet been established for newly introduced terms.

Transliteration involves preserving the graphical representation and letter composition of the original lexical unit using the target language's script. Examples include "qurt" remaining "qurt", "quyrdaq" transliterated as "quyrdak", "bauyrsaq" as "bauyrsak", and names such as "Alzhan" and "Otegeni" appearing as "Alzhan" and "Utegen", respectively. As indicated by the analysis of proper names and parallel contexts, transliteration is most commonly applied to proper nouns (Fig. 3).

Figure 3. Example of searching for proper names by frequency

By analyzing the frequency of proper names, we can demonstrate the effectiveness of a parallel corpus as a valuable tool for both research and translation (Table 4). Within A. Baitursynuly's text corpus, 256 proper names were identified, including personal names, country names, and city names. Many of these were transcribed or transliterated into Russian, considering linguistic, historical, and cultural factors. A quantitative analysis of non-equivalent vocabulary distribution revealed the following proportions: 226 instances (88%) were personal names, while 30 (12%) were geographical names. Additionally, the texts contained 50 common nouns classified as non-equivalent vocabulary, including names of traditional dishes, objects related to nomadic life, clothing, social classes, and kinship terms. Overall, proper names constitute the majority of non-equivalent vocabulary at 88%, while common nouns account for the remaining 12%. In translation, proper names are predominantly rendered through transcription or transliteration.

T a b l e 4

Proper names in Parallel Corpus

In the text of A. Baitursynuly	Frequency	In translation
Abai	105	Abai
Alash	22	Alash
Alashorda	13	Alashorda
Orynbor	19	Orenburg
Europa	18	Evropa
Semei	9	Semipalatinsk
Usen	5	Usen
Alzhan	5	Alzhan
Nurman	4	Nurman
Qarqaraly	2	Karkalinskoi
Qunanbai	2	Kunanbai
Ulzhan	1	Ulzhan
Abdirei	1	Abdirei

Concretization involves substituting a word or phrase with a more general meaning for one with a more specific and precise meaning.

*Auzy zhaman eldi bylgar; ayagy zhaman tordi bylgar. — **Klyauznik** maraet narod, a tot, u kogo nogi gryaznye, maraet pochyetnoe mesto (A **slanderer** dirties the people, and the one with dirty feet dirties the place of honor).*

*El qulagy elu. — U **lyudei** pyatdesyat ushei (**People** have fifty ears).*

*Ozi qandai bolsa, sozi sondai. — Kakov **chelovek**, takova **ego** rech (What a **man** is, such is **his** speech).*

The phrase "auzy zhaman" (literally "one who has a bad mouth") is rendered more specifically in translation as "klyauznik", meaning a person who speaks ill of others, slanders, or files petty complaints that tarnish someone's reputation. The Kazakh word "el", which broadly refers to "people" or "population", is translated as "lyudi" (people) in Russian, aligning more naturally with the numeral "fifty". The pronoun "ozi", meaning "you, yourself", is translated as "chelovek" (man) and reinforced with the possessive pronoun "ego" (his).

Generalization, the opposite of concretization, involves replacing a specific term with a broader, more generic concept. Examples of this technique in Russian translation include condensing the sequential enumeration "ekinshi, ushinshi, tagysyn tagy" into the generalized term "drugih" (others), replacing the repeated phrase "bir zhagynan" (from one side) and "armen qarai" (farther) three times with the single word "gradually", and translating "ekeui Qatar" as "vmeste" (together).

*Sonan arpi **ekinshi, ushinshi, tagysyn tagy** buundardy alyp, isteitin isininn reti bari osylaisha bolady. — Posle etogo na primere slogov **iz drugih** bukv povtoryaet te zhe deistviya (After that, using the example of syllables **from other** letters, he repeats the same actions).*

*Suitip, on shaqty soilemdi oqyi bilgen kezde, balalar **bir zhagynan** aripterdi, **bir zhagynan** buuyn turlerin tanyp alyp, **bir zhagynan** oqyuga zhattygyp, **armen qarai** oz betterinen oqyuga zharagan. — K tomu vremeni, kogda oni mogut prochitat okolo desyati predlozheniy, deti, **vo-pervyh**, vyuchat bukvy i slogi, a **vo-vtoryh**, obretut navyki chteniya i **postepenno** smogut chitat samostoyatelno (By the time they can read about ten sentences, **firstly,** children will learn letters and syllables, and **secondly**, they will gain reading skills and will **gradually** be able to read on their own).*

*Suitip, bir zamanda dybysty zhalqylau adispen dybysty zhalpylau adis – **ekeui qatar** orkendep kusheigen kezi bolgan. — Takim obrazom, nastalo vremya, kogda metody sinteza i analiza stali sosuschestvovat **vmeste** (Thus, the time has come when the methods of synthesis and analysis began to coexist **together**).*

The technique of lexical addition involves inserting extra words into the translation or restoring words and expressions that are implied but not explicitly stated in the original text.

For example, the words "zhamannan" (literally "from the bad") and "zhaqsydan" (literally "from the good") are supplemented in translation with the word "chelovek" (person) to clarify the intended meaning. Similarly, the implicit intensity conveyed by the word "qarqyndary" (literally "paces") is rendered in Russian as "tvorcheskaya aktivnost" (creative activity), making the meaning more explicit.

*Zhamannan zhaqsy tuady, adam aitsa nangysyz, **zhaqsydan** zhaman tuady, bir ayaq asqa alqysyz. — Ot **plohogo cheloveka** rozhdaetsya horoshiy, ne poverish, esli skazhet chelovek, ot **horoshego cheloveka** rozhdaetsya plohoi, ne blagodaryaschiy za chashku edy (A good person is born from a **bad person**, you won't believe it, if a person says, a bad person is born from a **good person**, who doesn't thank you for a cup of food).*

*Soitip, qazaq qalam qairatkerlerininn khalinde dagdarys bolyp, ozgeristen buryngy **qarqyndary** qaityp, tartynnqyrap qaldy. — V itoge, v mirovozzrenii kazakhskih pisatelei proizoshyel krizis, po sravneniyu s dorevolyutsionnoi deyatelnostiyu oni snizili svoyu **tvorcheskuyu aktivnost** (As a result, there was a crisis in the worldview of Kazakh writers, compared with their pre-revolutionary activities, they reduced their **creative activity**).*

Omission, in contrast to addition, involves leaving out words that are semantically redundant in translation. For instance, in Russian, the word "pora" (time) would be unnecessary in the literal translation of "balalyq kui" (literally "children's time"), so the more natural term "detstvo" (childhood) is used instead. Similarly, the phrase "qalam qairatkerleri" (literally "figures of the pen") is translated as "pisateli" (writers), while the words "arirek, terennireken" (literally "farther, deeper") are conveyed simply as "glubzhe" (deeper) to maintain fluency and clarity.

*«Balalyq kui» — **Detstvo (Childhood)**.*

*Qazaq **qalam qairatkerleri** qaidan tugan, qashan shyqqan dep surau qoiylsa, zhauap qiyn emes. — Kogda i kak poyavilis kazakhskie **pisateli**, otvetit na etot vopros neslozhno (When and how Kazakh **writers** appeared, it is not difficult to answer this question).*

*Munynn manisin **arirek, terennirekten** izdeu kerek. — Prichiny kroyutsya **glubzhe** (The reasons go deeper).*

Semantic development (modulation) involves substituting a dictionary equivalent with a contextual translation that is logically connected to the original meaning. This process replaces a lexical unit in the source language with a corresponding unit in the target language whose meaning is inferred from the original through logical association (causality). For example, the Kazakh word "zhygar" literally means "to knock down", which implies "to be exhausted". However, in context, it conveys the opposite meaning — "odolet siloi" (to defeat by force).

*Bilegi tolyq birdi **zhygar**; bilimi tolyq mynndy **zhygar**. — Tot, u kogo silnye zapyastiya, **odoleet** odnogo, tot, kto silyen svoimi znaniyami, **odoleet** tysyachu (The one with strong wrists will **defeat** one, the one who is strong with his knowledge will **defeat** a thousand).*

From the meaning of the phrase "dauys turleri" (lit. types of voice), the meaning of "ottenok" (tone) is contextually and lexically derived:

*Aripter — sol dauys **turlerine** arnalgan belgiler. — Bukvy yavlyayutsya znakami dlya etih **ottenkov** golosa (The letters are the signs for these **tones** of voice).*

By analyzing the semantic connections within the context, the phrase "ony bylai istegeni" (literally "this is done because of...") is translated as "sleduya etomu priemu obucheniya" (following this method of teaching), effectively preserving the causal relationship in the translation.

***Ony bylai istegeni** mugalim birinshi sozdi alyp, «a-r» dep ekeuin qospai, zhikterin ashyp aitady da, neshe bolinip, aitylyp tur dep suraidy, sonan sonn ar bolimin bolek aitqyzady. — **Sleduya etomu priyomu obucheniya**, uchitel snachala dayot zadanie proiznesti slovo po slogam, ne soedinyaya dve bukvy «a-r», potom sprashivaet, iz skolki chastei sostoit slovo, kak oni proiznosyatsya, i tolko potom dayot proiznesti kazhduyu chast otdelno (**Following this method of teaching**, the teacher first gives the task to pronounce the word syllabically, without connecting the two letters "a-r", then asks how many parts the word consists of, how they are pronounced, and only then gives each part to be pronounced separately).*

One type of semantic development is **holistic transformation**, which involves altering a word or even an entire sentence as a whole rather than modifying its individual components. For instance, the Kazakh proverb «Zhoqtyq uyat emes, bailyq murat emes» literally translates to "Poverty is not a disgrace, wealth is not a goal". However, since Russian already has a proverb with a similar meaning — "Bednost ne porok" (Poverty is not a vice) — the Kazakh proverb undergoes a complete transformation in translation.

***Zhoqtyq uyat emes**, bailyq murat emes. — **Bednost ne porok**, bogatstvo ne predel mechty (**Poverty is not a vice**, wealth is not the ultimate dream).*

The sentence «Qazaq airyqsha soz bolatyn qali aiyryqsha bolgandyqtan» literally translates into Russian as "Kazakh was a special word, as his position is special". However, a holistic transformation is applied to emphasize causality, clarify contextual meanings, and ensure that the phrasing is more comprehensible in Russian.

*Qazaq airyqsha **soz** bolatyn **qali** aiyryqsha bolgandyqtan. — «Kazakh» — eto obosoblennoe **ponyatie**, potomu chto kazakhi — eto nezavisimyi **narod** ("Kazakh" is a separate **concept**, because Kazakhs are an independent **people**).*

Discussion

One of the key challenges in translation is preserving an author's unique style, as translation involves not just comparing individual words and forms but conveying coherent semantic units that form a unified conceptual whole. When translating culturally marked vocabulary, specialized knowledge and skills are required not only to reflect the author's stylistic and linguistic nuances but also to capture the historical and cultural essence of the era.

A. Baitursynuly's writing style is distinguished by its vivid and metaphorical narrative, characterized by frequent reiterations of ideas in various forms. This approach stems from his dedication as a scholar and educator to making complex scientific concepts and the fundamentals of Kazakh spelling accessible to illiterate students through clear, expressive, yet simple and illustrative language. Here are some specific examples.

*Bul kungi musylmansha oqyp zhurgen zholdynn **bulgalaq-zhyrgalagy kop,** onymenen zhurip khat zhaza biluge zhetkenshe russha biluden uaqyt kem ketpeidi. — V sovremennom musulmanskom obuchenii **mnogo podvodnyh kamnei,** i na to, chtoby nauchitsya pisat pri takom obuchenii, uidyet ne menshe vremeni, chem na to, chtoby vyuchit russkiy yazyk. (There are many **pitfalls** in modern Muslim education, and learning to write with such training will take no less time than learning Russian).*

In this example, A. Baitursynuly's occasionalism *«bulgalak-zhyrgalak»* (literally "throwing-staggering") is translated as the phraseological expression *«podvodnye kamni»* (pitfalls), conveying the idea of unforeseen difficulties and hidden obstacles.

Eski ukimet tusynda, qazaqtynn basyna qaptagan **qara bulttai qalynn pale torlagan zamanda**, *zhurt ushin, ult ushin qyzmet qylgan, qinalgan adamdarga «zhurtshyl», «ultshyl» degen tannba basylyp, olar qoraga zholatpaityn qotyr mal syqyldy shetke shygarylyp, aleumet zhumysynan aulaq ustaldy. — Pri staroi vlasti,* **v samye mrachnye dni** *kazakhskogo naroda, na lyudei, predanno sluzhivshih svoemu narodu, svoei natsii, naveshivali kleimo «narodnika», «natsionalista», ih pritesnyali i kak prokazhyennyh ne dopuskali k obshchestvennoi rabote. (Under the old government,* **in the darkest days** *of the Kazakh people, people who faithfully served their folk, their nation, were branded as "populist", "nationalist", they were oppressed and like lepers were not allowed to public works).*

Bul aitylgan sebeptinn ustine **zhamalgan** *ekinshi sebep boldy. — Eto vtoraya prichina, tesno* **svyazannaya** *s pervoi (This is the second reason, closely* **related** *to the first).*

The metaphors used by the author, such as *«qara bulttai qalynn pale torlagan zamanda»* (literally "in an era thickly shrouded in disaster, like a black cloud") and *«zhamalgan»* (literally "patched"), are rendered in translation using words with similar semantic components. In the first case, it is translated as *«v samye mrachnye dni»* ("in the darkest days"), while in the second case, it is conveyed as *«sshityi, soedinennyi, svyazannyi»* ("stitched, connected, related"). The translators successfully preserved A. Baitursynuly's distinct figurative and metaphorical style by incorporating phraseological units, similes, and metaphors. For example, the phrase *«qaita kelip zhetken zhoq»* (literally "has not returned yet") is replaced by the Russian idiom *«vzyat sebya v ruki»*, meaning "to pull oneself together and act calmly". Similarly, *«isten shygara bastaidy»* (literally "withdraw from business") is translated as *«othodit na vtoroi plan»*, meaning "to lose importance".

Sol qaitqan qarqyn ali kunge **qaita kelip zhetken zhoq**. *— Do sikh por oni ne mogut* **vzyat sebya v ruki** *(They still can't* **pull themselves together**.)*

Alipshil adisti syryp birte-birte **isten shygara** *bastaidy. — Pobukvennyi metod chteniya postepenno nachinaet* **othodit na vtoroi plan** *(The letter-by-letter method of reading is gradually starting to* **fade into the background**).*

In A. Baitursynuly's characteristic style, the writer is vividly compared to a stray dog, effectively capturing the essence of the artist's mission.

Qalam qairatkeri — zholyn tastap koringen zholaushyga ere beretin **buralqy it** *emes. — Pisatel — ne* **brodyachaya sobaka**, *kotoraya shodit so svoego puti i sleduet za lyubym prohozhim (A writer is not a* **stray dog** *that goes out of its way and follows any passerby).*

A crucial aspect of translating A. Baitursynuly's works involved handling terms that served as key elements of his grammatical and spelling theories. It is important to note that the terminological framework of Kazakh linguistics, literary studies, and teaching methods emerged and developed through the scholar's research. Given this, it is natural that certain terms were transformed to align with modern concepts and terminology: «aqyn tili» (literally "the poet's speech") was rendered as «poeticheskaya rech» ("poetic speech"), «ansheyin til» (literally "ordinary speech") became «razgovornaya rech» ("colloquial speech"), «tannbalar» (literally "signs") was translated as «ieroglif» ("hieroglyphs"), «tubirshik» (literally "root") was replaced with «amorphnyi» ("amorphous"), «synalu» (literally "to be tested") was adapted as «aprobatsiya» ("approbation"), and «qosu» (literally "connect") was rendered as «kombinirovat» ("combine").

Shygarma tili eki turli bolady: 1) Aqyn tili. 2) **Ansheyin til.** *— Khudozhestvennaya rech delitsya na dva vida: 1) Poeticheskaya rech. 2)* **Razgovornaya rech** *(Artistic speech is divided into two types: 1) Poetic speech. 2)* **Colloquial speech**).*

Ar sozge arnalgan tannbany tanytsa, sol **tannbalardy** *qoidyra bilse, qytaisha oqyudy da, zhazyudy bilip sauattangany. – Ego tsel – obuchit pravilnomu napisaniyu i zapominaniyu* **ieroglifov**, *chto i oznachaet umenie chitat i pisat po-kitaiski (His purpose is to teach the correct spelling and memorization of* **hieroglyphs**, *which means the ability to read and write Chinese).*

Tubirshik *til tupki qalpynan ozgerilmei zhumsalady, masalan, qytai kham zhapon tilderi. —* **Amorfnyi** *yazyk ispolzuetsya bez izmeneniya osnovnoi formy, naprimer, kitaiskiy i yaponskiy yazyki (An* **amorphous** *language is used without changing the basic form, for example, Chinese and Japanese).*

Semeneriyelerde turli adister qaralyp, is zhuzinde **synalyp**, *tazhiriybeler zhasalgan. — V seminariyah raznye metody podvergalis tschatelnomu analizu i* **aprobatsii** *na praktike (In seminaries, various methods were thoroughly analyzed and* **tested** *in practice).*

*Sozdinn dybystaryn aiyra bilgennen keyin turli dybystardy **qosyp** soz qyla biletin bolgannan keyin arippen tanystyryuga tusedi. — Posle togo, kak deti smogut opredelyat zvuki slova, **kombinirovat** raznye zvuki, chtoby sostavit slovo, oni poznakomyatsya s bukvami (After the children are able to identify the sounds of the word, **combine** different sounds to make up a word, they will become familiar with the letters).*

We have determined the average statistical data for all lexical transformations identified in the analyzed texts (Table 5).

T a b l e 5

Average statistical data on lexical transformations in the analyzed texts

Types of lexical transformations	Relative frequency
Transcription	0.18
Transliteration	0.09
Concretization	0.08
Generalization	0.03
Lexical addition	0.05
Omission	0.02
Semantic development (modulation)	0.31
Holistic transformation	0.24

After analyzing the translational transformations in the studied texts, we concluded that the most characteristic types are semantic development, holistic transformation, and transcription. These lexical transformations are influenced by the structural differences between the source and target languages, as well as the priority of accurately conveying contextual semantic nuances.

The analysis of lexical transformations in the Russian translations of A. Baitursynuly's texts highlights that achieving an accurate, precise, and thorough translation requires the translator not only to have specialized knowledge and skills but also to possess a broad perspective, research abilities, and competencies in linguoculture, ethnolinguistics, intercultural communication, and other related fields.

Conclusion

The Kazakh-Russian parallel corpus of A. Baitursynuly serves as an effective automated tool for facilitating intercultural communication and translation practices. Parallel corpora enable quick retrieval of relevant contexts, selection of synonyms, phraseological units, and stylistic devices, as well as in-depth analysis of both the original and translated texts, while also assisting in the application of translation transformations.

This study examined lexical transformations in the works of A. Baitursynuly and their Russian translations. The findings suggest that various types and methods of lexical transformations play a crucial role in ensuring high-quality translation. The combination of different lexical transformations significantly enhances the accuracy and adequacy of both individual text fragments and the overall work. Future research directions include an in-depth examination of grammatical transformations, as well as the collection of statistical and linguistic data, which can expedite analytical work with both source and target texts, while also deepening the understanding of linguistic phenomena and their representation in translation. The variability in the spelling of foreign words can be attributed to historical and linguistic factors, as phonological norms for newly borrowed words were not yet firmly established.

One of the main challenges in translation is preserving the author's unique style, as it requires consideration not only of individual words and grammatical structures but also of entire semantic units forming a unified conceptual framework. Translating culturally marked vocabulary demands specialized knowledge and skills to convey not only the author's linguistic style but also the nuances, cultural depth, and distinctive features of the historical period.

A. Baitursynuly's writing style is characterized by its figurative and metaphorical narration, along with frequent repetition of key ideas in varied formulations. This stylistic choice stems from his intent to explain complex scientific concepts and the fundamentals of Kazakh spelling in an accessible, expressive, and visually clear manner for an illiterate audience.

A significant aspect of translating A. Baitursynuly's works is working with terminology, as these terms are fundamental to his theories on grammar and orthography. His research played a vital role in the formation and development of the terminological system in Kazakh linguistics, literary studies, and teaching methodologies. Consequently, adapting and modernizing these terms in translation appears logical and justi-

fied. Based on the analyzed translations, the most frequently observed transformations include concretization, equivalent substitution, and logical development of concepts.

The ongoing development of A. Baitursynuly's parallel corpus continues to expand, incorporating additional works from his complete collection, refining the search algorithm for grammatical queries, and enhancing metadata processing. Additionally, an alphabetical frequency concordance has been developed, with plans for further lexicographic data processing. Thus, A. Baitursynuly's parallel corpus functions as a modern searchable reference database and an electronic platform supporting diverse types of scientific research.

Funding source

References

1 Захаров В.П. Корпусная лингвистика: учебник / В.П. Захаров, С.Ю. Богданова. — 3-е изд., перераб. — Санкт-Петербург: Издательство Санкт-Петербургского университета, 2020. — 234 с.

2 Байтұрсынұлы А. Таңдамалы шығармалары / А. Байтұрсынұлы; құраст. Е. Тлешов, Н. Аитова (жау. ред.), О. Жұбай, А. Қадырхан. — Астана, 2022. — 240 б.

3 Байтұрсынұлы А. Избранные труды / А. Байтұрсынұлы; сост.: Е. Тлешов, Н. Аитова (отв. ред.), О. Жубай, А. Кадырхан; пер. с каз. Г.Ж. Байшукуровой, А.Б. Иргибаевой. — Астана, 2022. — 240 с. — [Электронный ресурс]. — Режим доступа: https://tilalemi.kz/viewer/viewer.php? file=/books/8175.pdf.

4 Казахско-русский параллельный корпус А. Байтурсынова. — [Электронный ресурс]. — Режим доступа: http://baitursynuly-corp.kz/. (Дата обращения: 17.03.2024).

5 Тао Ю. Разработка и использование параллельного корпуса русского и китайского языков / Ю. Тао, В.П. Захаров // Вестник НТИ. Серия Информационные процессы и системы, 2015. — № 4. — С. 18–29.

6 Национальный корпус казахского языка. — [Электронный ресурс]. — Режим доступа: https://qazcorpus.kz.

7 Ашимбаева Н.М. Параллельный подкорпус казахского и русского языков: разработка, функционирование и проблемы / Н.М. Ашимбаева, А.З. Бисенгали, С.К. Кулманов, Ғ.М. Аязбаев, М. Нурлан // Tiltanym. — 2023. — № 2. — С. 49–61.

8 Хаирова Н. Выровненный казахско-русский параллельный корпус, ориентированный на криминальную тематику / Н. Хаирова, А. Колесник, О. Мамырбаев, К Мухсина // Вестник Алматинского университета энергетики и связи. — 2020. — № 1 (48). — С. 84–92.

9 Калдыбеков Т.Е. Англо-казахский параллельный корпус для статистического машинного перевода / Т.Е. Калдыбеков // Молодой ученый. — 2014. — № 6 (65). — С. 92–95.

10 Бархударов Л.С. Язык и перевод (Вопросы общей и частной теории перевода) / Л.С. Бархударов. — М.: «Международные отношения», 1975. — 240 с.

11 Комиссаров В.Н. Современное переводоведение / В.Н. Комиссров. — М.: ЭТС, 2001. — 424 с.

12 Karagulova B.S. Transformation issues in translation / B.S. Karagulova, S. Mukhtarova, R. Koilybayeva // Tiltanym. — 2024. — 2. — P. 140–149. https://doi.org/10.55491/2411-6076-2024-2-140-149.

13 Дадебаев Ж.Д. Абай қарасөздерінің аударма үдерісіндегі трансформациялану типтері / Ж.Д. Дадебаев // Молодой ученый. — 2022. — № 12 (407). — С. 321–325. — [Электрондық ресурс]. — Қолжеткізу режимі: https://moluch.ru/archive/407/89660/ (дата обращения: 23.01.2025).

14 Мухамедина А.А. Абайдың «қара сөздерін» аудару барысында лексикалық трансформациялардың қолданылуына талдау жасау / А.А. Мухамедина, А.Е. Кішібаева // Наука о человеке: гуманитарные исследования. — 2011. — № 2 (8). — С. 101–108.

15 Aston G. Corpus use and learning to translate / G. Aston // Textus. — 1999. — No. 12. — P. 289–314.

16 Barlow M. ParaConc: Concordance Software for Multilingual Parallel Corpora / M. Barlow // Proceedings of the Third International Conference on Language Resources and Evaluation. LREC Workshop No. 8: Workshop on Language Resources in Translation Work and Research, 2002. — P. 20–24.

17 Dan Melamed I. Empirical Methods for Exploiting Parallel Texts. MIT Press, Cambridge (MA), 2001.

18 Liddicoat A.J. Translation as intercultural mediation: setting the scene / A.J. Liddicoat // Perspectives: Studies in Translation Theory and Practice, 2016. — No. 3.

19 Zhao Y. Transformer: A General Framework from Machine Translation to Others [Electronic resource] / Y. Zhao, J. Zhang, C. Zong // Mach. Intell. Res. — 2023. — 20. — 514–538. — Access mode: https://doi.org/10.1007/s11633-022-1393-5.

20 Palladino Chiara. Translation Alignment for Historical Language Learning: a Case Study [Electronic resource] / Chiara Palladino, Foradi Maryam, Yousef Tariq // In: Digital Humanities Quarterly. — 2021. — 15 (3). — Access mode: http://digitalhumanities.org/dhq/vol/15/3/000563/000563.html.

21 Maofu Liu. A cross-guidance cross-lingual model on generated parallel corpus for classical Chinese machine reading comprehension [Electronic resource] / Liu Maofu, Li Qiyuan, Qiu Chen, Hu Huijun // Information Processing & Management. — 2024. — Vol. 61. — Iss. 2. — Access mode: https://doi.org/10.1016/j.ipm.2023.103607.

Г.Ж. Байшукурова, А.Б. Иргебаева, Н.Н. Аитова

Ахмет Байтұрсынұлының параллель корпусы: аудармадағы лексикалық трансформациялар

Мақалада Ахмет Байтұрсынұлы таңдамалы шығармаларының орыс тіліне аудармасындағы лексикалық трансформация түрлері қарастырылған. Автор тілінің лексикасын ғылыми тұрғыдан талдау әзірленіп жатқан А. Байтұрсынұлының қазақ-орыс параллель корпусының негізінде жасалды. Мақала авторлары қазақ жазуының реформаторы еңбектерінің аудармашылары, сондай-ақ ғалымның қазақ-орыс параллель корпусының әзірлеушілері болғандықтан, зерттеу тақырыбы да осылай таңдалып, ұсынылды. Параллель корпус аударма және лингвистика саласындағы ғылыми және қолданбалы талдау мүмкіндіктерін едәуір кеңейтетін инновациялық құрал. Мақаланың мақсаты — аудару кезіндегі лексикалық өзгерістердің ерекшеліктерін талдау, А. Байтұрсынұлының қазақ-орыс параллель корпусымен аударма жұмыс істеу принциптерін сипаттау. Параллель корпусты аударма тәжірибесінде қолданудың артықшылықтары қарастырылады. Зерттеу барысында лексикалық трансформациялардың әртүрлі түрлері анықталды: транскрипция, транслитерация, нақтылау, жалпылау, лексикалық қосу және түсіру, семантикалық даму, тұтас түрлендіру тәсілдері анықталып табылған. Аударма кезінде жеке сөздер мен формалар емес, біртұтас ұғымдардың семантикалық сегменттері салыстырылады. Алғаш рет әзірленген бағдарламалық қамтамасыз ету орыс және қазақ тілдеріндегі мәтіндерді автоматты түрде салыстыруға мүмкіндік береді, баламаны іздеу уақытын қысқартады, аударманы талдауға, белгілі бір аударма стратегиясын қолданудың барабарлығын бағалауға мүмкіндік береді. Алынған нәтижелер аудармашылар, аударма теориясы мен практикасы, корпус лингвистикасы саласында жұмыс істейтін лингвистер үшін пайдалы болмақ.

Кілт сөздер: А. Байтұрсынұлы, корпустық лингвистика, параллель корпус, конкорданс, аударма, лексикалық трансформация.

Г.Ж. Байшукурова, А.Б. Иргебаева, Н.Н. Аитова

Параллельный корпус Ахмета Байтурсынова: лексические трансформации при переводе

В данной статье рассматриваются различные виды лексических трансформаций, применяемых при переводе избранных произведений Ахмета Байтурсынова на русский язык. Исследование основывается на анализе лексики оригинальных текстов автора с использованием разрабатываемого казахско-русского параллельного корпуса А. Байтурсынова. Выбор темы не случаен: авторы статьи непосредственно участвуют в переводе трудов выдающегося казахского просветителя, реформатора письменности и лингвиста, а также в разработке корпуса, позволяющего сопоставлять оригиналы и переводы для проведения лингвистического анализа. Параллельный корпус представляет собой инновационный инструмент, значительно расширяющий возможности научного и прикладного анализа в области перевода и лингвистики. Целью статьи является всестороннее исследование особенностей лексических трансформаций, осуществляемых в процессе перевода, а также описание принципов и методик работы с параллельным корпусом в переводческой практике. Авторы выделяют и классифицируют такие виды лексических трансформаций, как транскрибирование, транслитерация, конкретизация, генерализация, лексическое добавление, опущение, смысловое развитие и целостное преобразование. Перевод анализируется не на уровне отдельных слов, а с учетом смысловых сегментов, образующих единую понятийную структуру. Особое внимание уделяется программному обеспечению, разработанному для автоматического сопоставления казахских и русских текстов. Этот инструмент значительно ускоряет процесс поиска эквивалентов, повышает точность анализа перевода и позволяет оценить эффективность применения конкретной переводческой стратегии. Представленные в статье результаты вносят значительный вклад в развитие современных направлений переводоведения и корпусной лингвистики, открывая новые перспективы в изучении наследия А. Байтурсынова. Актуальность исследования определяется современными тенденциями использования корпусной лингвистики в переводоведении. Параллельный корпус, применённый в данной работе, не выполняет функции автоматического перевода, а служит инструментом глубокого анализа, помогающего повысить качество перевода. Впервые подробно описываются этапы формирования казахско-русского параллельного корпуса А. Байтурсынова и принципы его

применения. Эмпирические и статистические данные, полученные в результате анализа, нашли применение как в теоретических, так и в практических аспектах перевода. В статье рассматриваются примеры поэтапной и скрупулёзной работы переводчиков по выбору наиболее адекватных переводческих решений, направленных на обеспечение максимально точной передачи как содержания, так и формы оригинального текста на материале произведений Ахмета Байтурсынова. Даются среднестатистические данные по всем встретившимся в анализируемых текстах лексическим трансформациям.

Ключевые слова: А. Байтурсынов, корпусная лингвистика, параллельный корпус, конкорданс, перевод, лексические трансформации.

References

1 Zaharov, V.P., & Bogdanova, S.Yu. (2020). *Korpusnaia lingvistika [Corpus linguistics]*. Saint-Petersburg: Izdatelstvo Sankt-Peterburgskogo universiteta [in Russian].

2 Baitursynuly, A. (2022). *Tandamaly shygarmalary [Selected works]*. Tleshov, E., Aitova, N., Zhubay, O., & Kadyrhan, A. (Eds.). Astana [in Kazakh].

3 Baitursynuly, A. (2022). *Izbrannye sochineniia [Selected Works]*. Tleshov, E., Aitova, N., Zhubay, O., & Kadyrhan, A. (Eds.).; G.Zh. Baishukurova, A.B. Irgibayeva (Trans.) Astana. Retrieved from //https://tilalemi.kz/viewer/viewer.php?file=/books/8175.pdf [in Russian].

4 Kazakhsko-russkii parallelnyi korpus A. Baitursynova [Kazakh-Russian parallel corpus of A. Baitursynuly]. Retrieved from //http://baitursynuly-corp.kz [in Russian].

5 Tao, Yu., & Zaharov, V.P. (2015). Razrabotka i ispolzovanie parallelnogo korpusa russkogo i kitaiskogo yazykov [Development and use of a parallel corpus of Russian and Chinese languages]. *Vestnik Nauchno-technicheskoi Informatsii. Seriia Informatsionnye protsessy i sistemy — Bulletin of Scientific and Technical Information. Series Information Processes and Systems,* 4, 18–29 [in Russian].

6 Natsionalnyi korpus kazahskogo yazyka [National Corpus of the Kazakh Language]. Retrieved from /https://qazcorpus.kz [in Russian].

7 Ashimbaeva, N.M., Bisengali, A.Z., Kulmanov, S.K., Ayazbaev, G.M., & Nurlan, M. (2023). Parallelnyi podkorpus kazakhskogo i russkogo yazykov: razrabotka, funktsionirovanie i problemy [Parallel corpus of Kazakh and Russian languages: development, functioning and problems]. *Tiltanym — Linguistics,* 2, 49–61 [in Russian].

8 Hairova, N., Kolesnik, A., Mamyrbaev, O., & Muhsina, K. (2020). Vyrovnennyi kazakhsko-russkii parallelnyi korpus, orientirovannyi na kriminalnuiu tematiku [Aligned Kazakh-Russian parallel case, focused on criminal themes]. *Vestnik Almatinskogo universiteta energetiki i sviazi — Bulletin of the Almaty University of Power Engineering and Communications*, 1 (48), 84–92 [in Russian].

9 Kaldybekov, T.E. (2014). Anglo-Kazahskii parallelnyi korpus dlia statisticheskogo mashinnogo perevoda [English-Kazakh parallel corpus for statistical machine translation]. *Molodoi uchenyi — Young scientist*, 6 (65), 92–95 [in Russian].

10 Barkhudarov, L.S. (1975). *Yazyk i perevod (Voprosy obshchei i chastnoi teorii perevoda) [Language and translation (Issues of general and particular theory of translation)]*. Moscow: «Mezhdunarodnye otnosheniia» [in Russian].

11 Komissarov, V.N. (2001). *Sovremennoe perevodovedenie [Modern translation studies]*. Moscow: ETS [in Russian].

12 Karagulova, B.S., Mukhtarova, S., & Koilybayeva, R. (2024). Transformation issues in translation. *Tiltanym — Linguistics,* 2, 140–149. https://doi.org/10.55491/2411-6076-2024-2-140-149.

13 Dadebaev, Zh.D., & Atshybaeva, A.A. (2022). Abai karasozderinin audarma uderisindegi transformatsiialanu tipteri [Types of transformation of Abay's words in the translation process]. *Molodoi uchenyi — Young scientist*, 12 (407), 321–325. Retrieved from https://moluch.ru/archive/407/89660/ [in Kazakh].

14 Muhamedina, A.A., & Kishibayeva, A.E. (2011). Abaidyn «qara sozderın» audaru barysynda leksikalyq transformasialardyn qoldanyluyna taldau zhasau [Analysis of lexical transformations in the process of translating of Abai`s «qara sozder» Abai]. *Nauka o cheloveke: gumanitarnye issledovania — Science of Man: Humanitarian Studies*, 2 (8), 101–108 [in Kazakh].

15 Aston, G. (1999). Corpus use and learning to translate. *Textus*, 12, 289–314.

16 Barlow, M. (2002). ParaConc: Concordance Software for Multilingual Parallel Corpora. *Proceedings of the Third International Conference on Language Resources and Evaluation,* 20–24. LRE Workshop No. 8: Workshop on Language Resources in Translation Work and Research.

17 Dan Melamed, I. (2001). Empirical Methods for Exploiting Parallel Texts. MIT Press, Cambridge (MA).

18 Liddicoat, A.J. (2016). Translation as intercultural mediation: setting the scene. *Perspectives: Studies in Translation Theory and Practice*, 3. Retrieved from https://www.tandfonline.com/doi/full/10.1080/0907676X.2015.1125934.

19 Zhao, Y., Zhang, J., & Zong, C. (2023). Transformer: A General Framework from Machine Translation to Others. *Mach. Intell. Res.,* 20, 514–538. Retrieved from // https://doi.org/10.1007/s11633-022-1393-5

20 Palladino, Chiara, Maryam, Foradi, & Tariq, Yousef (2021). Translation Alignment for Historical Language Learning: a Case Study. *Digital Humanities Quarterly,* 15 (3). Retrieved from //http://digitalhumanities.org/dhq/vol/15/3/000563/000563.html.

21 Maofu, Liu, Qiyuan, Li, Chen Qiu, Huijun, Hu. (2024). A cross-guidance cross-lingual model on generated parallel corpus for classical Chinese machine reading comprehension. *Information Processing & Management, 61,* 2. Retrieved from https://doi.org/10.1016/j.ipm.2023.103607.

Information about the authors

Baishukurova Gulnur Zholaushybayevna — Candidate of Philological Sciences, Associate professor, Abai Kazakh National Pedagogical University, Almaty, Kazakhstan; e-mail: baishukurova@mail.ru

Irgebayeva Akerke Bayanovna — Candidate of Philological Sciences, Senior lecturer, Abai Kazakh National Pedagogical University, Almaty, Kazakhstan; e-mail: kaldanov70@mail.ru

Aitova Nurlykhan Nurullaevna — Candidate of Philological Sciences, Associate professor, Maksut Narikbayev University, Astana, Kazakhstan; e-mail: nurlykhan.an@gmail.com