BILINGUAL PUBLISHING GROUP

Forum for Linguistic Studies

https://journals.bilpubgroup.com/index.php/fls

ARTICLE

From Foreign to Familiar: Transliteration Challenges in the Latinization of Kazakh Exonyms

Kyzdarkhan Rysbergen [1,*] **, *Nursaule Rsaliyeva*** [1] **, *Dana Pashan*** [1] **, *Eugen Schochenmaier*** [2] **, *Seyitbatkal Arlen*** [1]

[1] *Department of Onomastics, Akhmet Baitursynov Institute of Linguistics, Almaty 050000, Kazakhstan*

[2] *International Council of Onomastic Studies, 77933 Lahr, Germany*

ABSTRACT

The transition of the Kazakh language from Cyrillic to the Latin alphabet, initiated as part of a broader cultural modernization strategy, has significant implications for the standardization of exonyms - foreign geographical names adapted for use in Kazakh. This article explores the phonological, orthographic, and sociolinguistic challenges that arise in this process, particularly in relation to exonyms inherited from Russian-language conventions or derived from typologically unrelated source languages. Building on transliteration principles established by the United Nations Group of Experts on Geographical Names (UNGEGN) and recent corpus-based research, the study develops a stratified model for exonym adaptation in Kazakhstan. Through the analysis of country names, urban toponyms, and natural landmarks, the article examines key linguistic factors such as phoneme inventory compatibility, semantic transparency, and morphological adaptability. A quantitative analysis of Kazakh-language corpora and map-based data reveals that Russian-influenced forms still dominate public discourse and cartographic materials, although there is an observable trend toward direct transliteration from English and other global languages. The study also considers the complexities involved in adapting exonyms from non-alphabetic writing systems, such as Chinese. Ultimately, the paper advocates for a nuanced standardization approach that balances Kazakh phonological integrity with international recognizability, recommending policy measures such as phonetic consistency, public-access digital databases, and educational support for wider implementation.

Keywords: Geographical Names; Exonym; Transliteration; Latin-Script Alphabet; Corpus Linguistics

*CORRESPONDING AUTHOR:

Kyzdarkhan Rysbergen, Department of Onomastics, Akhmet Baitursynov Institute of Linguistics, Almaty 050000, Kazakhstan;
Email: dr.ms.rysbergen@mail.r

ARTICLE INFO

Received: 20 April 2025 | Revised: 19 May 2025 | Accepted: 20 May 2025 | Published Online: 23 May 2025
DOI: https://doi.org/10.30564/fls.v7i6.9617

CITATION

Rysbergen, K., Rsaliyeva, N., Pashan, D., et al., 2025. From Foreign to Familiar: Transliteration Challenges in the Latinization of Kazakh Exonyms. Forum for Linguistic Studies. 7(6): 42–55. DOI: https://doi.org/10.30564/fls.v7i6.9617

1. Introduction

Exonyms, as an integral part of the toponymic system of any language, do not originate within that language. Although they function and accumulate within the boundaries of a given linguistic system, exonyms denote the names of physical and geographical objects located outside the territory and beyond the borders of the country. Hence, they are toponyms of foreign lands. An exonymis defined as "a toponym bestowed from outside and in a language from outside"[1]. Every world language has established traditional conventions for writing the geographical names of foreign countries. These conventions are typically recognizable to speakers of the language but may be unfamiliar or incomprehensible to representatives of other linguistic communities. This article explores the process of exonymization in the Kazakh language.

Historically, the onomastic system of the Kazakh language, including its exonyms, has evolved under the direct influence of the Russian language and its orthographic rules. Before gaining independence in 1991, Kazakhstan's toponymic system – especially its exonymic component – was heavily Russified, with nearly 90% of toponyms bearing Russian linguistic influence. The challenge lies in the linguistic disparity between Russian, an inflectional language, and Kazakh, an agglutinative language. These differences in phonetic systems complicate the transcription and transliteration of foreign names, as the Russian language lacks specific Kazakh sounds and letters. Consequently, the phonetics and orthography of the Kazakh language have been distorted, leading to a gradual and "painful" erosion of linguistic authenticity.

In this context, this article raises, for the first time, the critical issue of the necessity for transliteration and adaptation of exonyms according to the grammatical and orthographic norms of the Kazakh language. This effort aligns with the transition to the Latin script and the broader agenda of modernization, including the "Kazakhization" of foreign toponyms. The relevance of this research is underscored by the need to establish principles for translating and transliterating foreign geographical names into Kazakh, minimizing the influence and pressure of Russian orthographic conventions.

Kazakhization is the process of integrating an exonym (a foreign toponym) into the national linguistic system through phonetic adaptation, morphological modification, and orthographic standardization in accordance with the rules of the Kazakh language. This transformation ensures that the toponym aligns with Kazakh phonological norms, grammatical structures, and script conventions, thereby facilitating its usage within spoken and written discourse, official cartographic representation, and administrative nomenclature.

In the global information landscape, the Kazakh communicative space encounters an ever-growing influx of foreign geographical names (countries, capitals, cities, physical-geographical objects, etc.). It is crucial to accurately identify these names across various languages while preserving geoinformation integrity. For instance, the same name may be represented differently in different languages: English *Japan, Japanese* □□ (Nihon, Nippon), Russian *Япония,* Kazakh *Жапония*; or English *Algeria*, Russian *Алжир*, Arabic الجزائر (al-Jazair), Kazakh *Әл-Жазира*[2]. Ensuring the general "recognizability" of these names in the geoinformational space often requires preserving their phonetic forms from the original language.

In English-speaking academic circles, some debates exist regarding the designation of geographical names for foreign territories. However, experts generally support the term "exonym", relying on the definition provided by the United Nations Group of Experts on Geographical Names[3]. Within this framework, endonyms refer to indigenous geographical names within a native territory, while geographical names for external territories are termed exonyms[4].

The necessity for a standardized international approach to writing certain exonyms is evident[5]. Both transcription and transliteration aim to faithfully represent a "foreign name" – an exonym – in another language. This practice is closely associated with demonstrating respect for foreign toponymic traditions.

2. Theoretical Framework

The standardization and transliteration of geographical names are foundational aspects of modern linguistic, cultural, and cartographic practices. These processes aim to establish consistency in the use and representation of geographical names across languages and scripts. They are particularly significant in an increasingly interconnected world, where geographic names must function effectively in global communication while preserving their linguistic and cultural integrity.

Standardization efforts are largely guided by frameworks established by organizations such as the United Nations Group of Experts on Geographical Names. Since its inception in the 1970s, UNGEGN has been central to defining the principles of standardization, focusing on differentiating between endonyms – localnames used within the area they denote – and exonyms – names used by external linguistic communities. This distinction is more than semantic; it encapsulates the interplay between local linguistic identity and global recognition. As noted by Jordan, the endonym-exonym divide represents a complex, and at times contested, aspect of linguistic and cultural geography [6].

Transliteration, as a component of standardization, involves converting names from one script to another while striving to retain their phonetic and semantic essence. The challenge of transliteration arises from the structural and phonetic differences between languages. For instance, some phonemes in the source language may lack direct equivalents in the target language, necessitating approximations that can distort the original pronunciation. Such distortions are particularly evident in cases where transliteration moves between fundamentally different script systems, such as from Cyrillic or Arabic to Latin [7]. These approximations highlight the inherent tension between maintaining phonetic accuracy and ensuring usability within the target linguistic system.

International frameworks, such as the GOST standard for Cyrillic transliteration or the Beirut System for Arabic, have provided critical guidelines for standardizing transliteration practices. These frameworks are designed to ensure consistency across languages while accommodating the diverse linguistic characteristics of each script. However, the application of these standards often requires adaptation to account for local linguistic and cultural nuances. For example, Zagórski emphasizes the role of such adaptations in aligning transliterated names with the phonetic and grammatical norms of the target language [8].

The practical implications of transliteration extend beyond linguistic accuracy. They influence how names are perceived and recognized across cultural and linguistic divides. This is particularly evident in the case of geographic names rendered in Latin script, where transliteration decisions must balance global accessibility with fidelity to the original name. Dhieb highlights how double distortions can arise when names pass through multiple scripts, leading to forms that are ambiguous or even unrecognizable to native speakers [7]. Such challenges underscore the need for robust transliteration frameworks that can mitigate these issues while enhancing the clarity and functionality of geographic names.

Theoretical studies have further illuminated the complexities of transliteration. Scholars such as Kudła and Ormeling have explored how different transliteration systems handle the phonetic and orthographic challenges inherent in rendering exonyms [9, 10]. Their work demonstrates that transliteration is not merely a technical process but also a cultural and political one, reflecting broader considerations of linguistic identity and historical context. Transliteration systems must therefore strike a balance between these competing demands, ensuring that the names they produce are both faithful to their origins and functional within their new linguistic environments.

Literature Review

2.1. International Perspectives on Exonymization and Transliteration

The study of exonyms and their transliteration occupies a significant place in modern onomastics, reflecting the broader linguistic, cultural, and geopolitical dynamics embedded in geographical names. These names serve as markers of identity and heritage while functioning as tools for cross-linguistic communication. The dual challenges of standardizing exonyms and ensuring their accurate transliteration have prompted extensive scholarly discussions and institutional initiatives. This review synthesizes key contributions to the field, focusing on the theoretical foundations, transliteration methodologies, and practical implementations that guide the treatment of exonyms in international contexts.

The foundational concepts of exonyms and endonyms are central to understanding their role in language and culture. As defined by UNGEGN, an endonym is the name used within a local linguistic community for a geographical feature, while an exonym refers to an externally applied name that differs from the local usage. This distinction, while conceptually clear, becomes complex in application due to linguistic, historical, and political factors. Kudła emphasizes that exonyms are not anomalies but natural linguistic phenomena, resulting from historical interactions and the

adaptation of foreign names into local linguistic systems [9]. His classification of exonyms into eleven types highlights the spectrum of their linguistic transformation, from minor phonological adaptations to completely unrelated forms, illustrating the diversity of exonymic processes across languages.

One of the earliest and most influential frameworks for addressing the standardization and transliteration of exonyms was established by the United Nations. In the 1970s and 1980s, UNGEGN initiated discussions aimed at reducing the use of exonyms, emphasizing the need for consistency and cultural sensitivity. This period saw resolutions discouraging exonyms as relics of colonialism and revisionism. However, recent decades have marked a shift in perspective. Contemporary resolutions recognize the cultural value of exonyms, positioning them as part of intangible heritage rather than mere linguistic deviations. This evolution reflects a broader understanding of the role of exonyms in maintaining cultural identities while fostering international communication [6] .

For instance, the Reference Dictionary of Ukrainian Names explores the etymology and transliteration challenges of Ukrainian names, offering a model for analyzing naming conventions in regions with complex linguistic landscapes [11]. Similarly, the Handbook of Top Thai Names presents a detailed account of the most frequently used Thai names, focusing on phonetic and cultural considerations [12]. These works demonstrate the importance of systematic approaches to studying names across diverse languages and contexts, which can be applied to the Kazakh exonymic system during its transition to the Latin script.

The transliteration of exonyms poses unique challenges. Ormeling identifies the difficulty of preserving phonetic fidelity during transliteration, especially when moving from non-Latin scripts to Latin [10]. He notes that the omission of diacritical marks often leads to phonetic distortions, reducing the comprehensibility and authenticity of transliterated names. The role of diacritics is pivotal in ensuring phonetic accuracy, yet their inconsistent use across transliteration systems highlights the ongoing need for standardized approaches [8] .

In Europe, projects like EuroGeoNames (EGN) have made significant strides in standardizing geographical names, including exonyms. Stani-Fertl outlines the EGN database, which integrates national names data with supplementary databases of exonyms and variant names. By linking exonyms to their standardized endonyms, the EGN platform facilitates access to consistent geographical information across linguistic and national boundaries. This initiative exemplifies how technological tools can support the harmonization of naming practices while accommodating linguistic diversity [13] .

Arab countries provide a case study in the complexities of transliteration. Dhieb explores the challenges of double distortion, a phenomenon where names are altered during their initial transcription into French or English and subsequently modified upon retranslation. He argues that such distortions compromise both phonetic fidelity and cultural representation. His proposed solutions emphasize the need for pragmatic transliteration systems that minimize deviations from original forms while adhering to international standards [7]. This approach aligns with UNGEGN recommendations, which advocate for the preservation of phonetic and orthographic integrity in transliterated names.

The historical dimension of exonym studies further enriches our understanding of their evolution and usage. Kudła notes that the perception of exonyms has shifted over time, influenced by geopolitical changes and linguistic policies. For instance, in the post-World War II era, exonyms in German and other European languages were often stigmatized due to their association with nationalism. However, the revival of exonym usage in the post-Cold War period underscores their enduring functional and cultural relevance. This shift is evident in the increasing recognition of exonyms as symbols of historical interactions rather than instruments of appropriation [9].

Modern contributions to the field have expanded on these foundational ideas, incorporating new methodologies and technologies. For example, Watanabe highlights the sociolinguistic dimensions of exonyms, examining how they reflect power dynamics and cultural perceptions. His work underscores the importance of considering the social implications of exonym usage and transliteration, particularly in multilingual and multicultural settings [14]. Similarly, Kudła's classification system provides a nuanced framework for analyzing exonymic transformations, offering valuable tools for comparative studies across languages [9].

Cartographic practices also play a crucial role in shaping the representation of exonyms. Ormeling emphasizes

that exonyms often occupy prominent positions on maps, particularly in small-scale cartography where they designate key geographical features [10]. This prominence underscores the importance of standardizing exonyms to ensure their usability and consistency in international contexts. The integration of standardized exonyms into digital mapping platforms, as seen in the EGN project, represents a significant advancement in this area, bridging the gap between linguistic theory and practical application.

P. Jordan's research emphasizes the shared understanding of exonyms across languages and the process of de - exonymization [15]. While this principle may not fully apply to Kazakhstan, as its onomastic practices often reflect de- Russification (renaming Russian *Омск* to "Omby" in Kazakh), the insights from international studies offer valuable frameworks for the evolving field of Kazakh onomastics.

2.2. Kazakhstani and Kazakhstan-Related Studies on Exonymization

Exonyms in Kazakh onomastics have not been extensively addressed as scientific issues. While some academic works touch on the subject, Sh. Qurmanbayuly highlights the lack of adaptations for foreign country and nation names borrowed from Russian into Kazakh. These borrowed names often fail to capture the national essence and hinder ease of articulation for Kazakh speakers [16].

Recent efforts in Kazakhstan aim to standardize geographical names in the Latin script and align them with international norms. A pilot project for the Regulation for Standardization of Foreign Geographical Names was developed by Rysbergen and Rsaliyeva [17]. Scholars Q. Rysbergen and I. Şahin were among the first to analyze issues surrounding the standardization of exonyms in Kazakh. They examined the treatment of suffixes like *-ия (-ia)* and *-стан (-stan)* in exonyms, the transliteration of Slavic toponyms, and the adaptation of Turkic-language toponyms into Kazakh [18].

In Russian linguistic studies, exonyms are primarily addressed in the context of transliteration and transcription. Between 1955 and 1999, Russian authorities issued 72 instructions on the romanization of geographical names from various languages [17]. These included toponyms from continents such as Asia, Europe, and Africa, as well as underwater features and Antarctic locations. The *Dictionary of Geographical Names of Foreign Countries* in 1986 catalogs around 40,000 foreign names in Cyrillic, including countries, capitals, islands, rivers, and mountains, reflecting a primarily practical approach [19].

Aleksandra Superanskaya's seminal work *Theoretical Foundations of Practical Transcription* written in 1978 explored the theoretical underpinnings of transliteration and transcription. She emphasized the variability in how names are adapted globally, distinguishing between fully localized names, such as Dead Sea, and names transliterated with cultural adjustments, as seen in Russian contexts like Игольный 'Cape of Needles' and Желания 'Cape Zhelaniya' [20]. On the contrary, there are names translated in many languages yet given by their pronunciation in Russian: *Шварцвальд*, German *Schwarzwald*, English *Black Forest*, French *Forêt-Noire*. These principles also resonate with Kazakh toponymy practices. If written in the new Latin-based Kazakh alphabet, it would likely appear as *Şvartsvald*. The transliteration aims to approximate the German pronunciation while adapting it to Kazakh phonetic and orthographic norms.

Dmitry Yermolovich's *Names at the Intersection of Languages and Cultures* further enriches the field by focusing on the transliteration and translation of names in cross-cultural contexts. His guidelines for transliteration span 23 languages, including the Kazakh language, and area valuable resource for translators and educators [21].

3. Methodology and Data Sources

3.1. Methodological Framework

When acquiring onomastic terms from other languages, the phonetic structure and the initial phonetic configuration play significant roles in their adoption into the target language. Research methods such as transcription, transliteration, transformation, and calque are commonly employed to designate exonyms.

In this study, the methods of preserving the official English transliteration of foreign country names and maintaining the traditional transliteration established in Kazakh are used. However, many exonyms found on World Atlas have entered the Kazakh language through Russian and have been adapted according to Kazakh linguistic conventions. This process often leads to orthographic inconsistencies. For example, in Russian, *Тегеран* is transliterated as *Tegeran*, whereas in the new Latin alphabet of the Kazakh language,

it is transcribed as Tehran, which aligns more closely with the phonetic rules of Kazakh. Moreover, even in the current Cyrillic script, the latter has become the established transcription in Kazakh.

The method of phonetic transformation is also noteworthy. This occurs when a geographical name is adapted to fit the phonological rules of a different language. Such transformations are frequently observed in related Turkic languages. For instance, the Kyrgyz name *Алатоо* becomes *Alatau* in Kazakh, and the Uzbek name *Учкудук* transforms into *Üşqūdyq* in Kazakh. The phenomenon of phonetic transformation is a recurring topic of discussion within the UNGEGN, particularly concerning the standardization of toponyms.

In our study, a corpus-based analysis method was utilized, allowing for the identification of the usage frequency of major exonyms. Corpus linguistics has emerged as one of the most effective yet underexplored methods for researching onomastic names globally. German researcher Heiko Motschenbacher has conducted substantial work in this area, emphasizing the benefits of corpus linguistics in onomastic studies. He notes that corpus linguistics provides powerful empirical methods for studying names in actual language use through frequency-based evidence. It can, therefore, serve to check and refine more traditional, normative descriptions of name usage as found in grammar and other reference works[22]. Additionally, descriptive methods, along with statistical analyses, were employed in the research, providing a comprehensive framework for examining exonyms and their transliteration patterns.

3.2. Corpus and Data Sources

The necessary linguistic materials for the research were sourced from the World Atlas. This large-scale map, with a scale of 1:50,000,000, includes only the names of major geographic features worldwide. The statistical composition of the geographic names on this map is as follows: country names – 203, city and settlement names – 2,172, hydrographic feature names – 392, and orographic feature names – 705, totalling 3,472 names. These were categorized according to the type of geographic object. The preparation of this map was based on a Cyrillic-script map in the Kazakh language created by the National Cartographic and Geodetic Foundation of the Republic of Kazakhstan in 2014. This institution also provided the transliteration of the map's names into the Latin script of the Kazakh language[23].

Additionally, the research utilized the National Corpus of the Kazakh Language database. Exonyms and other general linguistic materials were collected from various sources, including atlases, educational maps, dictionaries, and multiple online resources. Among these were websites related to cartography, such as Earth 3D Map[24].

4. Results and Frequency-Based Statistical Analysis

The results of this study are presented as quantitative analysis of exonym usage frequency. These findings highlight the challenges and opportunities in standardizing exonyms within the context of the Kazakh language's transition to the Latin script.

The frequency of usage for the limited number of unique exonyms in the *National Corpus of the Kazakh Language database*[25] was also examined, providing additional clarity regarding their global distribution and usage patterns:

- Resei – 30
- Qytai – 8
- Özbekstan – 7
- Moñğolia – 6
- Kanada – 6
- Norvegia – 5
- Türkımenstan – 4
- Qyrğyzstan – 3
- Auğanstan – 2
- İran – 2
- Türkia – 1

Exonyms in Kazakh related to natural hydrographic and orographic features frequently include terms that denote attributes such as "higher", "lower", "large", or "small". The presence of exonyms, such as Canada and the Netherlands – geographically distant from the region – can be attributed to the translation of geographic terms and proper nouns into Kazakh through physical-geographical nomenclature (e.g., *Úlken Eriksizder kóli – Great Slave Lake*). Based on the data from the World Atlas, a brief statistical analysis of Kazakh exonyms is presented, highlighting the extent of their semantic adaptation to "complete Kazakhization" and "exonymization", as shown in **Table 1** below.

Table 1. Exonyms Semantically Adapted into Kazakh (World Atlas).

Countries	Total Number of Names on the Map	Names Translated into Kazakh
United States of America	250	6
Afghanistan	6	2
Vanuatu	12	1
Germany	12	1
New Zealand	22	3
Jordan	4	1
Canada	185	6
Kyrgyzstan	6	3
China	138	8
Mongolia	21	5
Egypt	16	1
Norway	24	5
South Africa	22	2
Uzbekistan	7	7
Papua New Guinea	22	2
Russian Federation	350	30
Saudi Arabia	16	1
Sudan	17	1
Tunisia	7	1
Turkey	17	1
Turkmenistan	4	4
United Kingdom	16	1
Fiji	8	1
Philippines	26	1
France	41	1
Sweden	14	1
Ethiopia	9	1
Total number of names:	**1269**	**97**

The number of names adapted into Kazakh accounts for 7.6% of all names. The ratio of digital data in **Table 1** indicates a predominant dominance of Russian orthography and a low degree of adaptation of exonyms in the Kazakh language. There remains a strong dependence on the Russian language as an intermediary in the translation and transliteration of exonyms from English. In turn, foreign geographical names, as well as names of countries and capitals, are assimilated from English in accordance with the phonetic and orthographic norms of the Russian language. Thus, 92.4% of exonyms found on maps, in textbooks, dictionaries, and other printed and electronic sources in Kazakhstan remain in Russian spelling: *Shvetsiya* (Sweden), *Finlyandiya* (Finland), *Frantsiya* (France), *Italiya* (Italy), *Rim* (Rome), *Parizh* (Paris), *Gamburg* (Hamburg), *Kaliforniya* (California), *Edinburg* (Edinburgh), etc.

We deliberately do not touch upon exonyms written in hieroglyphic scripts – Chinese, Japanese, Korean – as they are fully adapted to the norms of the Russian language and have been assimilated into Kazakh in an unchanged form (*Huanghe, Guangzhou, Changsha, Kyoto, Pyongyang*, etc.). Therefore, the issue of exonymization of hieroglyphic space remains unresolved. The quantitative coverage of exonyms from different countries on the Political World Map depends on the area occupied. Large territories such as the USA, Canada, and Russia have the highest coverage – 250, 185, and 350 exonyms respectively. The relatively small number of exonyms for countries such as Jordan, Turkmenistan, Tunisia, and others is directly related to their level of representation on the map.

In **Figure 1**, ten of the most well-known exonyms worldwide are selected based on the principle of proximity: closely situated → moderately distant → far away. The frequency of usage does not always correlate with the geographical proximity or remoteness of an object; instead, it may depend on the political, cultural, or economic sig-

nificance of the state or country in the global geopolitical landscape. For instance, the high frequency of exonyms such as Moscow, Russia, and China in the National Corpus of the Kazakh Language can be attributed to objective factors such as shared borders, a long- standing common history, and close political, cultural, and economic ties.

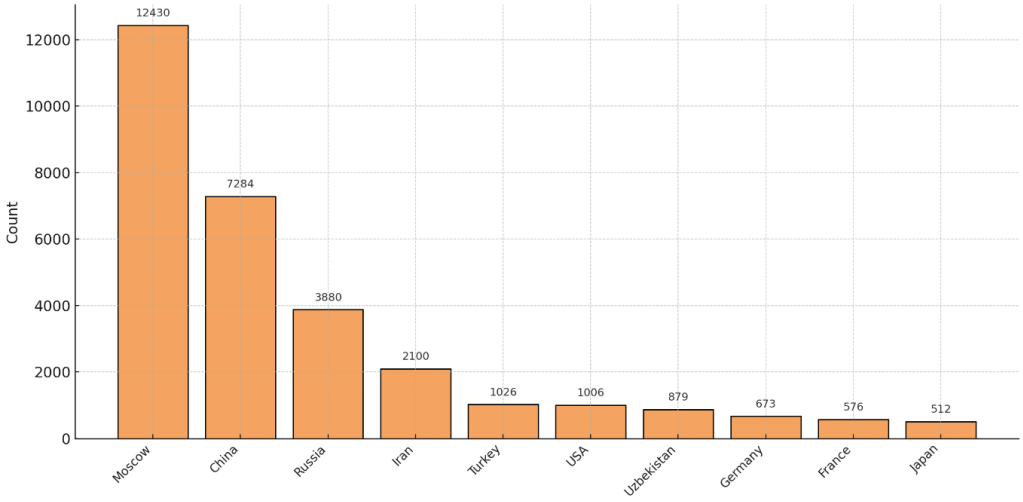

Figure 1. Frequency of Selected Exonyms in the National Corpus of the Kazakh Language.

However, the frequency of an exonym's usage is not always determined by the geographical distance of the state. In the Kazakh language, the exonym *Mysyr* (709 occurrences) is used instead of the Russian *Egipet* (224 occurrences), as reflected by their respective frequencies in the National Corpus. The parallel names *Iran* (2,100 occurrences) and *Persia* (19 occurrences) for the same country highlight a significant issue in exonymization: the simultaneous use of multiple variants, which necessitates standardization. The high frequency of Mysyr and Iran is associated with shared ancient history, culture, traditions, and religion with the Arabic- and Persian-speaking East. The exonyms *Mysyr* and *Iran* appear in historical texts, journalistic works, and the poetic compositions of medieval poets. For example, in Kazakh mythology, the phraseological expression *İran bağy* symbolizes the "Garden of Eden," serving as a metaphor for an idyllic paradise.

5. Discussion

The phenomenon of exonymization – the practice of assigning a geographical name from one language into another – remains a complex intersection of phonological accommodation, orthographic tradition, and sociopolitical influence. As observed in the case of the Kazakh language, the variation between related (e.g., Turkic) and unrelated (e.g., Slavic, Romance, Arabic, or Sino-Tibetan) languages demands a nuanced understanding of both inherited linguistic convergence and externally imposed naming conventions.

The Kazakh language, with its agglutinative structure and rich phonemic inventory, presents a unique case where the exonymic system has evolved under multiple external pressures – most notably Russian, but increasingly also English and Turkish. The degree to which a language adapts or resists these influences often depends on the perceived cultural proximity, linguistic compatibility, and geopolitical relationships with the source language[14]. This section analyses the patterns and principles of exonym adaptation in the Kazakh context through the lens of linguistic relatedness, historical contact, and modern standardization efforts.

5.1. Phonological Affinity and Turkic Exonyms

Languages within the same family – particularly those of the Turkic group – demonstrate a higher rate of phonological and morphological accommodation. In the Kazakh language, Uzbek, Kyrgyz, Turkmen, and Turkish toponyms are often adapted in ways that preserve both their semantic load and phonetic characteristics. Examples such as Uzbek *Qoraqalpog'iston* becoming Kazakh *Qaraqalpaqstan*, or

Kyrgyz *Ala-Too* becoming *Alatau*, demonstrate how Turkic linguistic kinship enables a smoother phonological transformation with minimal semantic distortion [19].

It is obvious that the letters C, X, W increase the capabilities of the new alphabet. In this regard, the question arises, should we adapt exonyms such as Washington, Holliwood, Zurich to the phonetic features of the receptor language – Kazakh: *Vaşiñton, Gollivud, Tsürih*, do we preserve the original English text by transliteration? This issue will hopefully be determined by the discussion of experts as well.

Yet even within the Turkic group, divergences arise. Turkish placenames such as *Çanakkale, Eskişehir*, and *Ölüdeniz* are adapted in Kazakh as "Şanaqqala", "Eskişahar", and "Öliteñiz", respectively. These Kazakh renderings demonstrate phonotactic regularization, especially with consonant-vowel alignment and the elimination of phonemes absent in Kazakh (e.g., Turkish "e" in word-final positions often becomes Kazakh "a"). However, this adaptation may result in semantic drift or orthographic distance from internationally recognizable forms.

This tension is visible in the case of İstanbul, which appears in the Kazakh corpus both as "Stambul" (following Russian conventions) and "Ystambul" (closer to Kazakh morphology), though neither form reflects the Turkish preference for the initial "İ" (Dotless-I). This exemplifies the challenge of balancing phonetic faithfulness, typographic compatibility, and international legibility.

5.2. Russian and European Exonyms: Legacy and Resistance

In contrast to Turkic kinship, Slavic languages – especially Russian – exerted a dominant influence on Kazakh exonym formation throughout the Soviet period. The prevalence of Russified forms such as Frantsiya (France), Gamburg (Hamburg), and Parizh (Paris) stems from centralized cartographic and administrative practices wherein Russian served as the mediating language for foreign names [23, 24].

The persistence of such exonyms, even in the Latinization process, demonstrates the depth of historical inertia. However, recent corpus-based evidence shows a gradual shift. For example, while "Orenburg" remains frequent (409 tokens), "Orynbor" has overtaken it (590 tokens) in Kazakh corpus usage. This transition represents not merely phonetic correction but also symbolic de-Russification.

In European names with strong international salience – such as Germany, Austria, or Australia – Kazakh often uses intermediary forms via Russian: "Germaniya", "Avstriya", etc. Yet alternative variants like "Almania" (from Arabic or Turkish influence) are also found, indicating the coexistence of multiple exonymic traditions. Here, international coordination, especially under UNGEGN guidance, plays a role in recommending standard forms, though compliance is uneven across regions [26].

5.3. Arabic and Persian Exonyms: Historical Depth and Religious Resonance

Unlike the Russian- or English-derived exonyms that often arrive through geopolitical channels, Arabic and Persian exonyms occupy a more venerable position in Kazakh due to historical, cultural, and religious proximity. The spread of Islam in Central Asia during the medieval period facilitated the early integration of Arabic and Persian toponyms into Kazakh, often through religious texts, poetic heritage, and oral tradition [17].

Moreover, Arabic and Persian exonyms often present fewer phonotactic conflicts with Kazakh due to shared phonemes and syllabic structures. However, controversies occasionally arise in Islamic contexts. For instance, the English exonym Mecca has been criticized in Arab and UNGEGN circles for being semantically trivialized in commercial contexts (e.g., "a mecca for tourists") and is replaced by Makkah in official transliteration [18]. In Kazakh, the use of "Mekke" rather than "Makkah" remains prevalent, but future reforms may revise this to align with international naming respect guidelines.

5.4. Sino-Tibetan, Japonic, and Korean Names: Transliteration without Adaptation

The adaptation of place names from hieroglyphic languages such as Chinese, Japanese, and Korean presents a unique challenge for Kazakh exonymization. These languages differ radically from Kazakh both phonologically and orthographically, with no historical phonetic channel through which the names were naturalized.

Unlike Turkic or Arabic exonyms, which may undergo

phonological adaptation, names like *Хуанхэ* (Huanghe), *Гуанчжоу* (Guangzhou), *Киото* (Kyoto), and *Пхеньян* (Pyongyang) are typically borrowed into Kazakh unchanged from Russian Cyrillic transliterations. These forms remain unadapted not only because of linguistic difficulty but also because Russian long served as the only mediating source for East Asian toponyms in Central Asian education and media [23].

For example, Хуанхэ retains its Russian-influenced syllabification, even though the Kazakh Latin script could theoretically accommodate closer phonetic renderings such as *Hwangkhe* or *Hwankhe*. However, such transformations would risk both orthographic inconsistency and user confusion, as the names are not phonetically transparent for most Kazakh speakers.

In the Latinization process, no standardized framework has yet emerged in Kazakhstan for adapting East Asian names. The principle followed in international cartography – namely preserving the Pinyin transliteration for Chinese (e.g., *Beijing* instead of *Peking*) – could offer a model, but this would require direct engagement with Chinese naming standards and a departure from Russian-based forms. UNGEGN suggests that if no culturally or linguistically suitable adaptation exists, it is preferable to use the internationally recognized Romanized endonym [3].

This leads to the suggestion that a separate protocol be developed for hieroglyphic names, particularly in countries like Kazakhstan where the geopolitical importance of East Asia is rising but linguistic affinity remains low.

5.5. Semantic Categories in Exonyms: Colours, Directions, and Universal Classifiers

An especially illustrative category of exonymization lies in semantically transparent names – those based on physical descriptors such as colour, size, direction, or geographic feature. These include globally attested terms like "White Sea", "Black Forest", "Upper Egypt", "Great Lakes", or "Cape of Good Hope". Such names often undergo partial translation, creating hybrid forms that combine original names with Kazakh or Russian descriptive elements.

In Kazakh, names such as *Ulken Eriksizder kóli* (Great Slave Lake) or *Qara teniz* (Black Sea) show how semantic elements are translated and merged with native classifiers. The semantic components "Ulken" (great), "Qara" (black), and "kóli" (lake) are readily intelligible to Kazakh speakers and follow local syntax yet preserve the original meaning. Internationally, this method is referred to as semantic exonymization, where the structure of the name is maintained while its lexical units are localized [16]. This method contrasts with phonological transliteration, which retains sound rather than meaning.

UNGEGN's Technical Manual proposes that semantic equivalence be prioritized for generic components (e.g., "Sea", "River", "Lake", "Island") while retaining phonetic or standard transliteration for the specific name. For instance, "Volga River" should become *Volga ózeni* in Kazakh – not Volga, nor a full calque like *Būlqyt ózeni*. The Kazakh corpus shows partial compliance with this guideline.

Colour terms also pose interesting challenges. For instance, *White Sea* (Белое море) could become *Aq teñiz*, but is often left unchanged or rendered via Russian *Beloye teñiz*. Standardizing such forms requires corpus evidence, public consultation, and linguistic modelling. According to best practice, if the translation of a geographical object is not available or unknown to the translator, the best solution is to keep the name of the settlement in the original language [17].

5.6. Semantic Categories in Exonyms: Colours, Directions, and Universal Classifiers

One of the major tools for evaluating the functional status of exonyms is corpus linguistics, which provides empirical evidence about how toponyms are used in real-world Kazakh discourse. As Motschenbacher asserts, corpus-based onomastics enables researchers to evaluate name frequency, contextual patterns, and co-occurrence with grammatical constructions. He points out the advantages of corpus-based research: "Such an analysis provides information on the commonness of individual names, onymic affixes and name-incorporating grammatical constructions in language use. This basic quantitative procedure can be usefully complemented by a qualitative analysis of concordance lines" [20].

The National Corpus of the Kazakh Language offers insight into usage trends and competing variants. For example, the Russian-influenced exonym *Orenburg* occurs 409 times, whereas the localized form *Orynbor* occurs 590 times. This suggests a significant trend toward phonological and orthographic alignment with Kazakh norms, particularly in

publicist and literary texts. As for *Germania* versus *Almania,* the corpus reveals the latter gaining frequency in religious and historical writing. The co-existence of these forms raises the question of standardization versus diversity. Should the Kazakh state promote a single preferred form in official media and education, or allow multiple historically grounded variants?

A significant pattern that emerges from corpus data is geopolitical and cultural proximity as a predictor of exonym preference. Countries with which Kazakhstan has strong ties – China, Russia, Turkey, and Iran – exhibit higher frequencies of exonym usage and greater variation. For distant countries like Tunisia, Vanuatu, or Fiji, Kazakh users tend to default to Russian- or English-derived forms with minimal localization. Moreover, the principle of transliteration via Russian as an intermediary remains statistically dominant, with 92.4% of exonyms on major Kazakh maps and educational resources using Russian orthography as the basis for rendering names from English, Arabic, French, and other languages. This heavy reliance has led to inconsistency and phonological distortions.

To address this, direct transliteration from English into Kazakh Latin script is increasingly recommended. For instance, *Tehran* (in Latin-based Kazakh) is now preferred over *Tegeran* (via Russian), aligning better with both international pronunciation and Kazakh phonotactics. This direct approach reduces cumulative distortions resulting from multi-stage transliterations – a problem documented by Dhieb [8] in Arab onomastics and applicable to Kazakh as well.

5.7. Towards a Functional Standardization Model for Kazakhstan

Given the tensions between historical usage, phonological fit, and international recognition, Kazakhstan requires a functional, context-aware standardization model for exonyms that is grounded in both linguistic evidence and cultural values. This model should be informed by best practices from UNGEGN, international cartography, and comparative onomastics. A five-level model of exonym standardization is proposed in **Table 2**.

Table 2. Five-level model of exonym standardization in Kazakhstan.

Level	Type of Adaptation	Example (English)	Kazakh Equivalent	Source Strategy
1	Full phonetic and semantic adaptat.	India	Ündistan	Turkic tradition
2	Phonetic transliteration	Tehran	Tehrān	English direct
3	Semantic translation + local classifier	Great Slave Lake	Ülken Eriksizderkóli	Semantic exonymization
4	Russian-based form retained	France	Frantsiya	Historical orthographic use
5	No adaptation (global form used)	Pyongyang	Pyongyang	International standard

This model allows for flexible decision-making depending on the function of the name (e.g., education, cartography, literary use), its degree of integration, and its symbolic or semantic value. For example, in poetic or religious contexts, *Mysyr* maybe preferred, whereas *Egipet* might serve as a formal term in diplomatic texts. The model also accounts for:

- Corpus frequency: Prioritize forms with higher public usage.
- Cultural resonance: Protect historically embedded exonyms.
- Ease of articulation: Favor forms that conform to Kazakh phonotactics.
- Recognition across borders: Avoid forms that obscure international identification (e.g., Orynbor vs. Orenburg in tourism literature).

A clear institutional mandate is needed for this standardization process. As Rysbergen and Rsaliyeva and Rysbergen & Şahin argue, the lack of coherent policies hinders linguistic modernization. They propose guidelines for suffix standardization (e.g., replacing "-ia" with "-stan" or "-ia" with Kazakh noun endings), which should be implemented by the Terminological Commission of Kazakhstan in collaboration with international bodies like UNGEGN and ICOS [17, 18].

Importantly, the process should avoid hyper-Kazakhization that renders names unintelligible. For example, changing "New York" to Jaña Jūrik may be phonologically logical but semantically alienating. A pragmatic

middle path is needed that respects both linguistic autonomy and global legibility.

5.8. Synthesis and Recommendations for Future Policy

The comparative and corpus-based analysis presented in this discussion confirms that Kazakh exonymization is a multidimensional process, shaped by phonological compatibility, historical legacy, political orientation, and cultural resonance. The interplay between related and unrelated languages – Turkic, Slavic, Arabic, and East Asian – demands a dynamic and stratified approach to toponymic standardization in the Kazakh language. To ensure coherent and culturally informed exonymization, we propose the following recommendations for Kazakh onomastic reform:

1. Adopt a Dual-Tier Standardization System: official exonyms for government, cartography, education and accepted variants for media, literature, and historical references.
2. Establish an Onomastic Standards Authority: A dedicated body, possibly under the aegis of the National Terminological Commission, should be charged with reviewing, approving, and updating exonym lists based on linguistic principles, corpus data, and international practices (cf. EGN and UNGEGN models).
3. Create an Open-Access Digital Exonym Database: Inspired by EuroGeoNames, Kazakhstan should maintain a national exonym database mapping each foreign toponym to its Kazakh equivalent(s), with annotations on etymology, source language, and corpus frequency.
4. Engage with International Naming Authorities: Coordination with UNGEGN, ICOS, and other global onomastic organizations will ensure that Kazakh exonyms are internationally intelligible and diplomatically neutral, while still reflecting local linguistic values.
5. Develop Educational Resources and Style Guides: Textbooks, media guides, and journalistic manuals should be updated to reflect standardized exonyms, along with pronunciation and orthographic guidance in the Latin script.
6. Train Translators and Educators: Professional development programs should include training on toponymic standards, pronunciation rules, and the sociolinguistic rationale behind Kazakh exonyms.

6. Conclusions

The study highlights the critical importance of adapting exonyms within the Kazakh linguistic context, particularly during the ongoing transition to the Latin script. By analyzing the phonetic, orthographic, and cultural complexities of transliterating and transcribing foreign geographical names, the research provides a comprehensive foundation for future standardization efforts.

The results underscore the challenges posed by the historical influence of Russian orthography, which often distorts the phonetic integrity of Kazakh exonyms. Through corpus-based analysis, the study identifies patterns of exonym usage and variability, illustrating the need for a balanced approach that respects both phonetic accuracy and cultural identity. Notably, the examples of toponyms with affricates ([dʒ], [ʒ]), the consonant [h], and the diphthong "ng" reveals areas where targeted adjustments can harmonize Kazakh exonyms with global linguistic standards.

Based on the results of the study, we conclude that it is advisable to transliterate exonyms into the Kazakh language directly from English-language cartographic sources, in a way that closely approximates the original, bypassing the Russian language.

This approach can help minimize orthographic deviations in the exonymization process in Kazakh.

The study also highlights the necessity for consistent and culturally sensitive transliteration frameworks to address discrepancies in parallel exonym forms. The findings have practical implications for cartography, language policy, and digital tools, supporting the usability of Kazakh in international contexts while maintaining its linguistic heritage.

Future research should expand on these findings by exploring the socio-political and technological dimensions of exonym standardization. By leveraging advancements in corpus linguistics and fostering collaboration with international naming authorities, Kazakhstan can enhance the effectiveness of its onomastic modernization efforts. The ultimate goal is to achieve a robust, culturally informed, and globally recognized exonymic system that strengthens the Kazakh language's role in an interconnected world.

Author Contributions

Conceptualization, K.R.; methodology, K.R., N.R. and D.P.; software, S.A.; validation, K.R., E.S. and D.P.; formal analysis, D.P.; investigation, N.R.; resources, E.S.; data curation, S.A.; writing—original draft preparation, K.R. and E.S.; writing—review and editing, E.S.; visualization, S.A.; supervision, K.R.; project administration, K.R.; funding acquisition, K.R. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan under Grant BR24993111 “DEVELOPING THE SPEECH SYNTHESIS SYSTEM BASED ON THE ORTHOEPIC NORM OF THE KAZAKH LANGUAGE”.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Woodman, P. (Ed.), 2012. The Great Toponymic Divide: Reflections on the Definitions and Usage of Endonyms and Exonyms. Head Office of Geodesy and Cartography: Warszawa, Poland.

[2] Languages by Countries, 2025. Language Distribution by Country. Available from: https://www.inyourlanguage.org/LanguageByCountry#A (cited 29 January 2025).

[3] United Nations Group of Experts on Geographical Names (UNGEGN), 2007. Technical Reference Manual for the Standardization of Geographical Names. United Nations: New York, NY, USA.

[4] Kadmon, N., 2006. Exonyms, Also Called Conventional Names. In: Manual for the National Standardization of Geographical Names. United Nations Group of Experts on Geographical Names: New York, NY, USA. pp. 129–134.

[5] Gábor, G., Mikesy, G., 2017. Does Google Serve as a Model for Using Place Names? Acta Geographica Slovenica. 57(1), 153–159.

[6] Jordan, P., 2021. Exonyms as Part of the Cultural Heritage. In Proceedings of the 2nd Session of the United Nations Group of Experts on Geographical Names (UNGEGN); New York, NY, USA, 3–7 May 2021.

[7] Dhieb, M., 2018. When Places Change Their Names on Maps: Case Studies from the Arab World. Proceedings of the ICA. 1, 1–5.

[8] Zagórski, B.R., 2010. Endonyms versus Exonyms: A Case Study in Standardization. With a List of Names of Arab Countries and Their Major Cities. In: Jordan, P., Bergmann, H., Burgess, C., et al. (eds.). Trends in Exonym Use. In Proceedings of the 10th UNGEGN Working Group on Exonyms Meeting; Tainach, Austria, 28–30 April 2010; Dr. Kovač: Hamburg, Germany. pp. 73–84.

[9] Kudła, M., 2022. Towards a Practical Classification of Exonyms. Lege Artis: Language Yesterday, Today, Tomorrow. The Journal of University of SS Cyril and Methodius in Trnava. 7(2), 127–144.

[10] Ormeling, F., 2012. Exonyms in Cartography. UNGEGN–ICA Webcourse on Toponymy. Available from: https://unstats.un.org/unsd/geoinfo/UNGEGN/docs/ (cited 2 April 2024).

[11] Chernovatyi, L., Schochenmaier, E., 2023. Reference Dictionary of Ukrainian Names. Mondonomo: Wilmington, DE, USA.

[12] Rutherford, A., 2023. The Handbook of Top Thai Names. Mondonomo: Wilmington, DE, USA.

[13] Stani-Fertl, R., 2011. Database of Exonyms and Other Variant Names (EVN-DB) within the EuroGeoNames Service. In Els noms en la vida quotidiana: Actes del XXIV Congrés Internacional d’ICOS sobre Ciències Onomàstiques, Annex, Secció 11. Generalitat de Catalunya: Barcelona, Spain. pp. 2692–2694.

[14] Watanabe, K., 2016. Whose Toponym? A Note for Deliberation on Place Names. In: Society for East Sea (Ed.). Seas and Islands: Connecting People, Culture, History, and the Future. Teikyo University: Tokyo, Japan. pp. 175–180.

[15] Jordan, P., 2023. Die Verwendung von Exonymen in mittel- und südosteuropäischen Schulatlanten als Ausdruck des politischen Wandels. In: Dahmen, W., Himstedt-Vaid, P., Schubert, G. (Eds.). Nomen est

omen: Die Bedeutung von Namen in südosteuropäischen Kulturen. Harrassowitz Verlag: Wiesbaden, Germany. pp. 3–18.

[16] Kurmanbayuly, S., 2018. Терминологиялық сөздік жасауға арналған әдістемелік нұсқаулық [Methodological Guide for Compiling a Terminological Dictionary]. Committee on Language Development and Socio-Political Work, Ministry of Culture and Sports of the Republic of Kazakhstan: Astana, Kazakhstan. Available from: https://termincom.kz/books/?id=82

[17] Rysbergen, K., Rsaliyeva, N., 2016. Onomastic Space of Kazakhstan: Current Condition and Problems. In: Hough, C., Izdebska, D. (Eds.). Names and Their Environment: Proceedings of the 25th International Congress of Onomastic Sciences; Glasgow, UK, 25–29 August 2014. pp. 153–165.

[18] Rysbergen, K., Şahin, İ., 2021. The Transfer of Foreign-Language Geographical Names into the Kazakh Language Based on the Latin Script as One of the Important Tasks of the Language Reform. Bulletin of the Karaganda University: Philology Series. 4, 7–14.

[19] Komkov, A.M. (Ed.), 1986. Словарь географических названий зарубежных стран [Dictionary of Geographical Names of Foreign Countries]. Nedra: Moscow, Russia.

[20] Superanskaya, A., 2018. Теоретические основы практической транскрипции [Theoretical Foundations of Practical Transcription]. 2nd ed. LENAND: Moscow, Russia.

[21] Yermolovich, D., 2005. Имена собственные: теория и практика межъязыковой передачи на стыке языков и культур [Personal Names: Theory and Practice of Interlanguage Transmission at the Junction of Languages and Cultures]. Valent: Moscow, Russia.

[22] Motschenbacher, H., 2020. Corpus Linguistic Onomastics: A Plea for a Corpus-Based Investigation of Names. Names: A Journal of Onomastics. 68(2), 88–103.

[23] National Cartographic and Geodetic Foundation of the Republic of Kazakhstan, 2014. World Atlas.

[24] Earth 3D Map, 2025. Online Map Service. Available from: https://earth3dmap.com (cited 29 January 2025).

[25] National Corpus of the Kazakh Language, 2025. Kazakh Language Digital Corpus. Available from: http://qazcorpus.kz (cited 29 January 2025).

[26] Woodman, P., 2007. Exonyms and UNGEGN: An Unhappy History. In: Jordan, P., Adamič, M.O., Woodman, P. (Eds.). Exonyms and the International Standardisation of Geographical Names: Approaches towards the Resolution of an Apparent Contradiction. LIT Verlag: Münster, Germany. pp. 7–10.