

Open
AccessCheck for
updates

RESEARCH ARTICLE

Section(s): *Literature, Linguistics & Criticism***Linguistic reconstruction of Central Asian toponyms: Evidence from Eastern historical sources**Bibatpa A. Koshimova¹ Bibaisha Nurdauletova¹ Marzhan Zhylykybayeva¹ Sharapat Abisheva¹ Akmaral Otarova¹ Ainur Kenbayeva²¹Yessenov University, Kazakhstan;²M. Utemissov West Kazakhstan University, KazakhstanCorrespondence: bibaisha.nurdauletova@yu.edu.kz**ABSTRACT**

Central Asia is a key region for the ethnogenesis and civilizational development of the Turkic peoples, yet its toponyms have a complex history shaped by centuries of migration and multilingual interaction. Drawing on a historical-etymological approach and the concept of topoformants, this study linguistically reconstructs the historical etymology of approximately 1,000 Central Asian toponyms based on a wide range of Eastern historical documents. The corpus was compiled through purposeful sampling from Eastern chronicles and corroborated by works of renowned Orientalists and Turkologists. The analysis helped reconstruct original forms, identify probable root languages, and explore semantic shifts caused by social and environmental changes. Toponyms were classified by geographic features and linguistic origin, revealing Iranian and Turkic layers, as well as Arab-Islamic influences. The study demonstrates that Central Asian toponyms are multilayered constructs reflecting complex historical and cultural contacts. By integrating various Eastern sources, the study fills the gap in the comparative analysis of regional toponymy, contributing to Turkic onomastics, and highlighting the role of toponyms in preserving collective heritage and strengthening solidarity among Turkic-speaking communities. The study offers valuable insights into the shared historical heritage of Central Asia and its significance for contemporary identity and cultural memory.

KEYWORDS: Central Asia, cultural heritage, historical etymology, topoformant, toponym .

Research Journal in Advanced Humanities

Volume 7, Issue 2, 2026

ISSN: 2708-5945 (Print)

ISSN: 2708-5953 (Online)

ARTICLE HISTORY

Submitted: 11 March 2026

Accepted: 30 May 2026

Published: 25 June 2026

HOW TO CITE

Nurdauletova, B., & Kenbayeva, A. (n.d.). Linguistic reconstruction of Central Asian toponyms: Evidence from Eastern historical sources. *Research Journal in Advanced Humanities*, 7(4). <https://doi.org/10.58256/h6yq4r52>



Published in Nairobi, Kenya by Royallite Global, an imprint of Royallite Publishers Limited

© 2026 The Author(s). This is an open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

Place names, or toponyms, play an important role in preserving cultural and historical memory and maintaining the local identity of people (Abdrakhmanova, Kurmanbaeva, Ibzharova, & Yesbulatova, 2025; Shakirova et al., 2026). As “linguistic relics” (Perono Cacciafoco & Cavallaro, 2023), they are fundamental to reconstructing the diachronic development of peoples, especially in the absence of written sources (Hua & Perono Cacciafoco, 2026). When written sources are not available, toponyms become one of the main carriers of historical information, since they allow us to reconstruct the chronology of settlement, trade routes, boundaries, and the nature of cultural and interethnic interactions (Qian, Kang, & Weng, 2016; Zhao et al., 2020). Historical etymology provides access to place names that reflect historical events, migration dynamics, and language contacts (Assadorian, 2017) and that have subsequently been integrated into oral folklore, written texts, and contemporary digital forms of expression (Saltakova, Oralbek, Mursalimova, Kudaibergenova, & Kenbayeva, 2026). A thorough historical and etymological analysis of toponyms reveals patterns in the spatial distribution and formation of the natural and cultural landscape (Liu, Liu, Xu, Yi, & Qiu, 2024), population patterns (Chlela, 2024), and demographic development in the region (Zhang, Zhou, & Ren, 2025).

Recently, there has been a growing interest in the toponymy of the Central Asian region, as it is considered a key area for the ethnogenesis and civilizational development of the Turkic peoples (Yama & Wasino, 2025). Studying ancient sources from the Central Asian region for their linguistic reconstruction is particularly important and at the same time challenging, as many of the region’s toponyms have a long and complex history, entwined with complex migration routes and ties with neighboring civilizations. Over the centuries, these names have been transformed under the influence of various languages and cultures, and have been used in modified forms in different historical periods. Therefore, to accurately establish the etymology of toponyms, it is insufficient to limit ourselves to sources presented within a single ethnolinguistic complex. Therefore, this study draws on a wide range of Eastern sources, traditionally used only fragmentarily, to trace the transformation and preservation of Central Asian names. By recording the evolution of toponyms, their morphological and phonetic variations, and their reflection in languages other than Turkic, historical chronicles of the Middle East and East Asia offer opportunities for restoring lost cultural and linguistic connections of the past.

The purpose of this study is to linguistically reconstruct the historical etymology of geographical names associated with Central Asia, drawing on Eastern historical documents. This study also systematizes them according to typological features that reflect the specifics of their origins in relation to a particular historical era. This article examines toponyms as multilayered constructs that can be “decoded” by sequentially revealing their linguistic and cultural layers. This approach permits connecting individual names to broader processes of cultural continuity, transformation, and language contact, and presents toponymy as an important starting point for understanding the region’s shared historical heritage and collective identity. The comprehensive use of diverse Eastern historical chronicles helps clarify the historical forms of Central Asian toponyms and reveal their Turkic roots. The results of this study confirm the complex, multilayered nature of the formation of Central Asian toponyms and contribute to the development of Turkic onomastics in general. This study sheds light on the processes of Turkification in the Central Asian region, as well as the interaction of Iranian, Turkic, and Arabic linguistic elements.

2. Topoformants in Historical Etymology

Theoretically, this study is based on a historical-etymological approach to toponyms, also known as historical toponomastics, which seeks to reconstruct the original forms, meanings, and linguistic sources of toponyms (Perono Cacciafoco & Cavallaro, 2023). This approach involves tracing a name through its attested forms in historical sources, determining its probable root language, and explaining semantic shifts due to social and environmental changes. In the case of Central Asia, such analysis often reveals multiple layers of borrowing and adaptation, where Turkic, Persian, Arabic, and, sometimes, other languages intersect in place names.

Topoformants are recurring word-forming morphemes that systematically shape the form and meaning of toponyms in a given linguistic tradition. In historical etymology, they usually represent suffixes, prefixes, or word-forming elements that head a whole group of toponyms and carry a constant semantic meaning, for example: *-burg/-gor/-grad* often denoting “a fortified place or town” in East European toponymy, or *-abad/-tan/-ak/-kan*

denoting settlements or territories associated with a name or group in many Turkic and Iranian traditions. The identification of toponyms allows classifying toponyms by morphological type and semantic subclass; reconstructing historical naming systems by identifying changes in word-formation preferences and observing how language contacts and migrations modified the local toponymic system; revealing historical layers, economic and social organization, as well as implicit cultural meanings.

In recent years, a number of studies have dealt with historical toponymics based on the analysis of toponyms in Slavic, Finno-Ugric, Turkic, Germanic and Romance languages. For example, the identification of repeated segments, or toponyms, allowed Garcia-Quera (2025) to relate Pyrenean toponyms to old cognates based on a conceptual thinking. Based on her corpus of Chilean Hispanic geonames, Martynenko (2023) identified unique linguo-culturological, morphological, and semantic features in the toponymic system of Chile alongside with certain universal features typical of Latin American toponymy. In her analysis of Hungarian toponyms, Bába (2016) differentiated between primary and secondary lexical toponyms. The analysis of toponyms in Zangazur area enabled Murtuza (2017) to trace Turkish roots in Azerbaijani toponymy.

Despite the growing body of research within the historical-etymological approach, most existing studies are limited to European, South American, or isolated regional contexts, often relying on national or monolingual corpora. To the best of our knowledge, Central Asian toponyms, particularly those preserved in Arabic-Persian and Chinese chronicles, have not been systematically studied to reconstruct the historical etymology of toponyms in the region. Nor has it been considered how the toponyms in these names reflect the complex language contacts, ethnic shifts, and cultural continuity among Turkic-speaking communities. As a result, the ethnocultural codes embedded in Central Asian toponyms and their role in preserving collective memory remain understudied. This study fills these gaps by reconstructing the etymology of Central Asian toponyms through a comprehensive analysis of Eastern historical sources, including Chinese and Arabic-Persian chronicles.

3. Methodology

3.1. Research Design

This study utilizes a qualitative approach to reconstruct the etymology of toponyms associated with Central Asia. The approach is based on a comprehensive analysis of a wide range of Eastern historical sources, including Arabic-Persian chronicles, Chinese historical records, and Turkic textual traditions containing references to the region's geographical names. Rather than relying solely on national Turkic sources, the study employs a comparative, cross-cultural approach, integrating materials from diverse linguistic and cultural contexts. This broader perspective allows for the identification of interconnections between various onomastic traditions and the collective heritage of Turkic-speaking peoples.

3.2. Data Collection

The data collection is based on the systematic extraction of toponyms from published Eastern historical documents concerning East Asia and the Middle East. These sources contain descriptions of medieval regions such as Mawarannahr (Maverenahr), Khorezm, Khorasan, Semirechye, and Turkestan, which correspond to modern-day Central Asia and encompass the territories of modern-day Uzbekistan, Turkmenistan, Kyrgyzstan, Tajikistan, Kazakhstan, northern Iran, and Afghanistan. Among the sources used for collecting data are the following Eastern chronicles: al-Balazuri's "Kitab futuh al-buldan", al-Ia'qubi's "Kitab al-buldan", al-Istakhri's "Kitab mesalik al-memalik", al-Maqdisi's "Ahsan at-takasim fi ma'arif al-akalim", as-Samani's "Kitab al-ansab", Ibn Fadlan's "Ar-Risala", Ibn Haukalya's "Kitab al-mesalik wa-l-memalik", Ibn Ruste's "Kitab al-'lak an-nafisa", Yakut al-Hamawi's "Mu'jam al-buldan". These chronicles serve as the basis for the research and guarantee the reliability of the data, as they contain some of the earliest references to Turkic territories in the context of military conquests, as well as detailed descriptions of routes with precise indication of the location of toponyms. Toponyms associated with Central Asia were selected from Eastern chronicles using a targeted sampling method. To clarify and confirm the data obtained from Eastern chronicles, secondary sources were used – the works of Western and Russian orientalists and Turkologists who studied Central Asia based on these chronicles. These studies not only clarify the origins of names but also reveal the ethnocultural processes underlying them. Reconstructing Central Asian toponymy based on Eastern chronicles confirms the authenticity

of the identified forms.

The corpus of collected data consists of approximately 1,000 toponyms extracted from Eastern historical sources, selected to represent a wide range of geographical features (hydronyms, oronyms, settlements, and other landscape elements) and reflect the linguistic and cultural diversity of Central Asia. Another criterion for selecting place names was that they appear in multiple sources or possess clear etymological markers (e.g., ethnonymic components or common Turkic, Arabic, or Persian morphemes). Each toponym was recorded in its original form, transliterated, and contextualized within the source text.

3.3. Data Analysis

The basic unit of analysis is an individual toponym, with an emphasis on its historical forms, semantic layers, and morphological composition. The analytical procedure includes three main stages: (1) identification of the toponym and its historical variants; (2) etymological decomposition into morphological components; and (3) interpretation of semantic and cultural meanings, taking into account the Arabic, Persian, and Turkic language systems. Comparative analysis examines patterns of semantic continuity and change, as well as evidence of language contacts and borrowings. The systematic description of Central Asian toponyms relied on two main features: the type of geographical feature and linguistic origin.

Based on the type of geographic feature, Central Asian toponyms, as in most other regions of the world, are divided into basic semantic categories corresponding to different types of natural territorial units such as oikonyms, hydronyms, and oronyms. Oikonyms, or names of populated areas, constitute the most significant portion of the corpus. These include major centers such as *Almaty, Bishkek, Tashkent, Samarkand, Ashgabat*, as well as historical towns and villages such as *Serketabad, Serdar, Keneurgenish, Kara-Kongur, Dargan-Ata*. These names often mark key stages in the political and cultural development of the region and reflect its ethnolinguistic dynamics. In addition to oikonyms, hydronyms, or names of water bodies such as *Murghab, Uzunsu, Sumbar, Geyman, Adir, Su-yatak, Chin, Gazma, Gui, Guruk* play an important role. Hydronyms highlight the economic specificity of the steppe and semi-desert zones of Central Asia, where the availability and accessibility of water determined the location of settlements, caravan routes, and the development of irrigated agriculture. Among oronyms, or names of landforms, in Central Asia, the names of large landforms are particularly prominent, such as *Karakum, Kum-Dag, Merv-Dag, Sumbar-Dag*, as well as individual names of mountains and dunes, reflecting ideas about the form and function of the area.

In addition to typological classification by type of geographic feature, Central Asian toponyms were also systematized by linguistic origin, making it possible to trace the multilayered history of toponymic background and identify the influence of various linguistic and cultural circles. Several main layers have been identified in the toponymy of Central Asia: Iranian, Turkic, and Arabic (Islamic-religious). The Iranian (ancient) layer is represented primarily by hydronyms, oronyms, and a number of oikonyms. It includes names that originate in ancient Iranian dialects, for example: *Murghab, Sumbar, Serakhs, Firuza, Dasht, Madau*, and many toponyms of the southern and foothill zones. These names are often associated with the natural and geographical characteristics of the area and reflect the early Iranian-language layer. The Turkic layer is the most widespread in the modern toponymic corpus and includes names such as *Kara-Kongur, Karakum, Kum-Dag, Kara-Kala, Buynnu-Uzun, Kyzyl-Ayak, Uzunsu, Dargan-Ata, Keneurgenish, Turkmenbashi*. The Arabic and religious-cultural layer includes toponyms associated with Islamic culture and Muslim shrines, being derived from the Arabic language or Islamic tradition, for instance: *Khodjambaz, Khodjagala, Imam-Baba, Guraish*. The typology of Central Asian toponyms by geographic type and linguistic origin allowed us not only to systematize the toponymic fund of Central Asia, but also to interpret it as a multi-layered socio-cultural text reflecting historical migrations, ethnic contacts and linguistic dynamics of the region.

4. Results

4.1. Toponyms of Iranian Origin

Toponyms of Iranian origin are among the oldest place names in the region. They are widespread in Central Asia, being widespread in the Turanian Plain and the Pamir Mountains, but are much less common in the mountainous regions of the Inner Tien Shan, in Semirechye, and the steppe regions of Kazakhstan. The beginning of the active formation of Western-Iranian (Persian) toponymy in Central Asia corresponds to the era of the Arab conquests

(Khromov, 1980). It was during this period that, along with the spread of the new Persian language (Farsi) of Western-Iranian origin in Central Asia, new Persian topoforms *-deh* and *-abad*, meaning “settlement”, began to come into widespread use (Barthold, 1963). For instance, *Zurabad* is a place (or region) that unites several settlements in the Serakhs region. The name is formed from the combination of the words *Zur* and *Abad*. In Persian, *zur / sur / sor / sorkh* stands for “red”, while *abad* is a very common topoformant in Persian toponyms meaning “settlement, village, city”. Or the topoform *-gird* (variants: *-gard*, *-kird*, *-kirt*, and *-jird*) belonged to the toponymic layer in pre-Islamic Western Iranian toponymy and was widespread in ancient and early medieval times in the territories of Iran and the South Caucasus (Transcaucasia). It originates from the Middle Persian word *kard* meaning “to create, build”. For example, *Sanuvajird*, the name of a village in the Serakhs region, is formed from the combination of the words *Sanuvad* and *jird* altogether meaning “a village built by Sanuvad”. In the historical geography of Central Asia, in the register of settlements of oases engaged in agriculture, toponyms with the suffixes *-kon*, *-gon*, *-jon* or *-zhan* are often found in toponyms like *Konimeh*, *Kumishkon*, *Khaydarkon*, *Mchgkon*, *Vobkan*, *Barkan*, *Savakon*, *Gondara*, *Andagon*, *Anjirfagon*, *Buzgon*, *Sagon*, *Jondar*, *Samjon*, *Arbinjon*, *Robinzhan*, *Obijaya*, *Khuzhan*. Most of the names of these settlements have a broad meaning and were formed on the basis of hydronyms associated with water sources and local hydrography. Therefore, the topoformants *-kon*, *-gon*, *-jon* or *-zhan* primarily denoted various water resources, such as rivers, canals, lakes.

Of particular interest is the topoformant *-zam*, which was found in *Zamma* (Kerki), one of the most important crossings on the Amu Darya in the Middle Ages. The city of *Zamm* or *Zam* is mentioned in Firdawsi’s “Shahnameh” in the era of the mythical Afrasiab. According to Yakut, the local pronunciation of its name is *Zam*, and in the original *Zam* form, this city is mentioned in the work of Ibn Haukal. According to Bal’ami, *Zamm*, along with Termez and Qubadian, was a “royal city” founded by the Iranian shah Qubad. The *Zamm* crossing is first mentioned in al-Tabari’s work, and in al-Muqaddasi’s list of the Amu Darya crossings it is called *Karkuh an-nahr*. The name *Karkuh* has been preserved in a distorted form as *Kerki* and is the name of a small town located on the right bank of the Amu Darya in the Chardzhou region of Turkmenistan. The name *Kerki/ Karkuh* means “city on the mountain”. Old Iranian **zam-* came into common use in many Iranian languages with the meaning of “land, region, country, nation”.

It has been established that the majority of ancient toponyms found in Central Asia in Arab-Persian sources were originally formed in Indo-Iranian, Persian, and Middle Persian languages. These names reflect the linguistic and cultural heritage of the ancient inhabitants of the region, ethnic composition and historical ties of that era.

4.2. Toponyms of Turkic Origin

Names of Turkic origin are also common in sources. These names reflect the multifaceted nature of the ethnic and linguistic composition of the region. For example, medieval sources mention the toponym *Su-kara* (*Karasu*) meaning “black water”, or *Sairam* (*Saryam*) in the Syrdarya basin meaning “shallow land”. *Küzeligyr* is an ancient Turkic city that was the first capital of ancient Khorezm in the 6th-5th centuries BC. *Küzel/Güzel* means “beautiful, wonderful” in the Turkic languages. The component *Igyr* may come from the ancient Turkic concepts of “plain, lowland”. Hence, the meaning of the name can be “Beautiful plain”, “Beautiful field”, or “Picturesque place”.

Medieval sources often contain toponyms derived from the Sogdian and Old Turkic languages, associated with sacred places and temples of the pre-Islamic era. Such names contain the Sogdian components *-bagn*, *-vagn*, *-fagn* meaning “temple”, “house of god”. The occurrence of these elements indicates the presence of pre-Islamic temples there. Thus, the name of the city *Iskifagn* in the Nasaf region may come from the Sogdian words *isku* and *bagn* (“high temple”) or from the Old Turkic words *isku* and *pagn* (“old temple”).

There are also hybrid toponyms or mixed component names. This category includes toponyms formed by combining two or more words of different origin: Turkic-Sogdian, Turkic-Persian, Arab-Turkic. For example, the name *Turkestan* used by the ancient Iranians was formed by combining the Pahlavi topoformant *-stan* (“country”, “region”). A number of Turkic toponyms formed using the Sogdian topoformant *-kat* (*-kath*) are known. For example, the names *Jabgukat* and *Khatunkat* in the Shash region; *Akhsikat* (*Akhshikat*), *Ardlankat*, and *Kukat* in the Fergana region belong to this group. Other examples include *Hirkhizkat* in the Zhetysu region, and *Jumulkat* and *Irguzgukat* in the Togyz-Oguz region. The name of the city of *Turar* (*Otrar*), located in the

lower reaches of the Syr Darya, is also given in the form *Turarband* with the Persian topoformant *-band*. The name *Mangyshlak* is also formed by the union of the Indo-Iranian formant *man* and the Turkic form *kishlak*. Of significance is the Sogdian topoformant *-misan* (*-mitan*) meaning “temple”, being found in the toponymy of such regions as Bukhara, Samarkand, Khorezm, Merv, Fergana. This element is preserved in the tribal name of the Karakalpaks and Uzbeks in the form *muyten*. Such names in the Kashkadarya valley include *Anshamisan*, *Farimisan*, *Kabamisan*, *Misan-i Kalan*, *Misan-i Khurdak*. They arose near religious centers of the pre-Islamic era. *Farimisan* is preserved in the modern name *Parmitan*; *Misan-i Kalan* means “big temple” and *Misan-i Khurdak* stands for “small temple”; *Kabamisan* means “mountain temple” or “cave temple”. The Sogdian formant *-misan* has been preserved to this day in the forms *-mitan*, *-matan*. For example, the settlements of *Namatan* (“new temple”) and *Shamatan* (“black temple”) are located near Shakhrisabz. Toponyms with the component *-misan* have preserved traces of pre-Islamic religious places of worship and thus are an important source of information for determining the history of ancient cultural and spiritual centers in Central Asia.

Another widespread topoformant of Sogdian is *-kas* or *-kam* meaning “fortress”, “fortified settlement” or “small town”. For example, in the 12th century, al-Sam’ani mentions a settlement called *Shirakas* (Κρήθη) in the medieval Nasaf region. The toponym *Shirkas* (*Sirkas*) translated from Sogdian means “fortress located on a hill”. This name was widely used in all regions where the Sogdian language was widespread in the Middle Ages. According to the data, in the 12th century, at least three settlements in the territory of Sogd were known by the name *Shirkas* or *Sirkas*: one was located in the Nasaf region, the second in the Kesh region, and the third near Samarkand. This toponym has survived to the present day in the Turkic form *Shirkent*. Today, this name is given to a small town located in the upper reaches of the Surkhandarya River and a settlement in the Kasbi district of the Kashkadarya region.

The Sogdian topoformant *-kas* (*-kat*) is also found in medieval toponyms in the forms *-kad* and *-qad*, which have the same meaning “fortress”, “fortress”, “fortified settlement”. In the 10th-century Arab sources, the city (medina) of *Naukad Quraish* is mentioned in the Kesh region. And in the Persian work “Hudud al-Alam” it is found in the form *Naukat-i Quraish*. The word *Quraish* in the name indicates that this place was inhabited by the Arab-Quraish tribe in the early centuries of Islam. Ibn al-Athir mentions a fortress called *Naukat* in the Kesh region during the al-Mukanna revolt in 159/775–776, which can be equated with *Naukad Quraish*. There were even several settlements with this name in the Kashkadarya valley. In the 12th century, al-Sam’ani mentions two settlements in the Nasaf region: *Naukad Sawa* and *Naukad Khurdahur*. Nowadays, villages named *Naukat* or *Naukat* are found in the Shahrissabz and Kasbi districts of the Kashkadarya region, as well as in the Fergana and Namangan regions. This toponym in Sogdian means “new fortress” or “new settlement”. Al-Sam’ani writes about another settlement in the Nasaf region, *Fankad*, which is formed from the Sogdian words *fan* and *kad* meaning “upper settlement” or “upper fortress”.

The Sogdian topoformant *-kas*, *-kat* or *-kad* has survived to the present day in the Turkicized form *-kent* and is widespread in Central Asia. In the Kashkadarya valley, there are three settlements with the *-kent* or *-kand* component, the origin of which is associated with the early medieval period, for example, the village *Tashkorgan*, located in the Yakkabag district, is called *Ishkent* in ancient sources. According to researchers, this name is associated with the name of al-Ishkand or al-Ashkand, who was the ruler of Nasaf at the beginning of the 8th century and participated in the struggle against the Arab invaders. The modern village of *Pakhtaobod* in the Kitab district was previously called *Parchakent* and in translation from the Sogdian language means “outskirts” or “outer fortress”. The village of Chorshanba, located in the Shakhrisabz district, is called *Shovkand* in ancient sources, meaning “black settlement” in translation from the Sogdian language. All of these names belong to the pre-Islamic toponymic layer, being evidenced by many place names in the territory of modern Central Asia and Kazakhstan, for example: *Beshkent*, *Yangikent*, *Shurkent*, *Shymkent*, *Manket*, all of which have a clear Turkic etymology.

4.3. Toponyms of Arabic Origin

With the spread of Islam in the 8th-9th centuries, Arabic culture and language began to actively influence this region. As a result, many ancient names underwent phonetic and morphological changes and began to be used in an Arabized version. Such changes were most often carried out through the adaptation of the original sounds to the sound system of the Arabic language, addition of grammatical structures in the Arabic language, and

semantic shift of some words. This phenomenon is not only a linguistic adaptation, but also a manifestation of political and cultural dominance. Arabized toponyms determine the formation of new administrative structures and religious centers in the Islamic era, for example: *Absisak* meaning “a river, a ditch, a spring with little water”; *Arapkhana* meaning “the house of an Arab”; *Baban* meaning “gate”, “door”, “entrance”. The toponym *Babshir* is formed by adding *bab* meaning “door” to *shar* (*shir, sar*) meaning “the gate of the city”. These toponyms clearly indicate the point of intersection of the ancient Indo-Iranian culture of the region and the later Islamic Arab culture. By studying them, it is possible to better understand the continuity between linguistic layers, as well as the historical development and ethnic ties of the peoples in the region.

Central Asia is one of the important junctions of the Great Silk Road and a major center of culture and civilization in the Islamic era. As a result of the settlement of Arabs in this region in the 8th-9th centuries, many new settlements appeared. The names of these new settlements were often formed as anthrotoponyms, that is, the tradition of naming them mainly by the names of the first Arab settlers or influential rulers and dynasties took hold. A significant layer of the historical anthrotoponymy of Central Asia is formed by names derived from the names of historical figures associated with the early Islamic era and the period of the Arab conquest. The names of some settlements are associated with the names of the companions of the Prophet Muhammad, people who died in the cause of religion and were later included in the ranks of saints. Thus, the name of the village of Safid Bulan, located in the Ala-Bukin district of the Jalal-Abad region of Kyrgyzstan, is associated with the name of a girl named Bulan, who was a servant of Shah Jarir, one of the companions of the Prophet, who came to Fergana from Medina to spread Islam. The name of the settlement of Sahrabad near Merv is associated with the name of Sakhra ibn Abdullah al-Aslami, the grandson of the companion Buraydah ibn al-Khusayb al-Aslami.

In the Middle Ages, about four farsakhs (about 20–25 km) from Merv, there was a settlement called Bal’aman. This name is associated with a man named Bal’aman who came to Merv with the army of Qutayba ibn Muslim. He was the great-grandfather of the famous vizier of the Samanid era, Abu-l-Fadl al-Bal’ami.

There was a settlement called *Shahr-i Saman* in the Termez region, from where the founders of the Samanid dynasty came. However, in the Middle Ages, several settlements with this name were known in Central Asia. For example, settlements called *Saman* were also found in Balkh, Samarkand, Bukhara, Khuttalan, Tokharistan and Fergana. Recent studies allow us to conclude that some of these settlements were named in honor of the ancestor of the Samanid dynasty - Saman-khudat. According to legend, Saman-khudat allied with the Arab commander Asad ibn Abdullah, and then received land from the Arabs in various regions of Maurennakhr. In those places, he founded settlements, and later these settlements were named after him.

In the 10th century, among the cities of Tokharistan, the city of *Sarai ‘Asim* is mentioned. The name of the city often contains the Arabic name *Asim*, which was a common name among Arab warriors who came to Central Asia. Perhaps the name of the city is associated with the name of one of the Arab commanders or settlers of that name. In the Guzgan region, on the road from Andhud to the banks of the Amu Darya, medieval sources mention a settlement called *Kanni Ghiyas*. This name is also associated with the Arabic name *Ghiyas* (*Ghiyas*). One of the settlements near the city of Valwalij in Tokharistan is called Bistama ibn Suri ibn ‘Amira ibn Musavira, and there is also a city called Karid ‘Amir [29, 34]. The name ‘*Amir*, which appears in these toponyms, indicates that it belonged to one of their Arab settlers. In the Merv region, several settlement names associated with the names of the first Arab settlers have also been preserved. Among them are the following toponyms: *Asadabad* associated with the name *Asad*; *Khalidabad* from the name *Khalid* (as-Sam’ani); *Hafsabad* associated with the name *Hafs* (as-Sam’ani); *Ma’marar* from the name *Ma’mar* (as-Sam’ani); *Sinj al-‘Abbadi* associated with a representative of the Abbad dynasty (as-Sam’ani); *Dizak Hafs* from the name *Hafs* (as-Sam’ani); *Rikanj ‘Abdan* associated with the name *Abdan* (as-Sam’ani); *Mahdiabad* from the name *Mahdi* widespread in Islamic tradition; also *Muhammadabad* located in the Dehistan region and *Musaabad* in the Jurjan region.

These toponyms indicate that during the era of the spread of Islam in Central Asia, Arabic names were widely used and began to be included in the names of places and water bodies. Such toponyms are associated with the establishment of a new Islamic culture and Arab political dominance among the local population and are of great historical value. They immortalized not only the name of the settlement, but also the names of the clans, dynasties, and even famous historical figures who settled in the region.

5. Discussion

The results of a linguistic reconstruction of Central Asian toponyms based on Eastern historical sources confirm that the region's toponymic record is a complex, multi-component layer reflecting a long process of language contact and shifts in dominant linguistic superstructures. Reconstructed toponymic forms and basic roots allow us not only to clarify the origins of individual names but also to interpret them as “geographical documents” of ethnic movements, political realities, and cultural integrations observed over many centuries. Eastern sources record names in forms close to their historical pronunciation, enabling an accurate reconstruction of toponymic forms, as confirmed by the study of Usmonova and Xujboyeva (2025). Unlike later cartographic and administrative materials, Eastern historical sources record toponyms close to their phonetic and morphological originals, uninfluenced by modern standard languages and orthographic norms. Therefore, according to Tulibayev, Karimova, & Tulibayev (2025), reconstructing their original form and linguistic affiliation requires relying specifically on early Eastern sources.

The results of the linguistic reconstruction of Central Asian toponyms can also be linked to the broader issue of language contact and interlingual influence. Toponyms, like other components of the onomastic corpus, are a particularly sensitive indicator of language contact, as they are fixed in geographic space and codified in both written and oral traditions. In the Central Asian context, toponyms demonstrate how, under conditions of prolonged contact between Iranian-speaking, Turkic, and Arabic populations, not only the borrowing of names but also their morphological adaptation occurred: Turkic speakers, interacting with Iranian-language toponyms, transformed them according to their own morphology, creating stable hybrid constructions. Akmatov, Abdullaev, and Mushaev (2019) also emphasize that reconstructing topoformants becomes a tool for typological and diachronic comparison, allowing us to identify the spatial and temporal parameters of language contact. As a result, in Central Asian toponyms, we find both complete and partial preservation of Iranian roots and formants, as well as their transformation according to Turkic morphology, indicating a complex mechanism of language contact, including borrowing, calques, and phonetic simplification.

The results of this study can be useful for contemporary research in typology and comparative linguistics. The reconstructed topoformants allowed identifying common and distinctive features of linguistic layers and also demonstrated how Iranian, Turkic, and Arabic place names interact and transform under conditions of long-term contact. This enables a clearer understanding of the diachronic and territorial parameters of language contacts in Central Asia and contributes to the development of the theory of interlingual influence and the typology of toponymic systems. Thus, toponyms act as means of transmitting cultural memory between different linguistic communities and, in this respect, can be considered within the broader line of research devoted to language contacts under conditions of interdiscursive transfer (Liu, 2026; Aleshinskaya, 2020). The linguistic reconstruction of Central Asian toponyms based on Eastern historical sources is an essential step in understanding the complex history and language contacts of this unique territory.

6. Conclusions

The aim of this study was to linguistically reconstruct the historical etymology of Central Asian toponyms based on Eastern historical documents. By analyzing Eastern sources, we reconstructed the origins of toponyms associated with the region and revealed their ethnoterritorial significance. This study confirmed that the historical-etymological approach is effective in studying the toponymy of the region: it can be used to trace not only the development of languages but also cultural continuity between generations. The study demonstrated that the toponyms of Central Asia recorded in Eastern sources constitute unique historical-linguistic material because, in addition to designating geographical features, they preserve the ethnic, linguistic, and cultural history of the region's ancient peoples. It is demonstrated that ancient toponyms preserved in Arabic-Persian and Chinese chronicles have a multilayered structure, including Iranian, Sogdian, Persian, and Turkic elements. These results demonstrate complex historical and cultural ties, close links between onomastic systems across cultures, and the transformation of the ethnic composition of the Central Asian population. The study of ancient toponyms contributes to the preservation of cultural heritage, strengthening historical memory, and fostering a sense of community among Turkic-speaking peoples. This research has practical implications for Turkic onomastics, deepening understanding of the collective heritage of Central Asia and its significance for modern society. The data obtained can be used in the creation of interactive migration maps, corpora of parallel texts, and interdisciplinary models of ethnogenesis using GIS technologies.

Funding

This research was funded by the Ministry of Science and Higher Education of the Republic of Kazakhstan under the program-targeted financing project BR24993043 “Ancient Turkic Toponyms in Chinese and Other Eastern Sources – A Foundation for Modern Turkic Integration: Historical-Etymological and Linguistic Perspective”.

Acknowledgments

The authors express their deep gratitude to Aliya Zhanissenova of Yessenov University (Kazakhstan) for her invaluable assistance and exceptional efforts in organizing the Central Asian field expedition. They also extend their sincere appreciation to Michael Erdman, Head of the Middle East and Central Asia Collections at the British Library, for facilitating access to unique and rare Arabic and Chinese sources, as well as manuscript materials from the Central Asian Turkic collections which greatly contributed to this research successful completion.

Conflicts of Interest

The authors declare no conflict of interest.

Disclaimer Statement

The views, opinions, findings, and conclusions expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated institutions, funders, or publishers.

References

- Abdrakhmanova, B., Kurmanbaeva, L., Izbharova, S., & Yesbulatova, Z. (2025). Kazakh folklore and the worldview: traditions, symbols, and identity. *Al-Farabi*, 89(3), 40–51. <https://doi.org/10.48010/2025.3/1999-5911.04>
- Akmatov, N. R., Abdullaev, S. N., & Mushaev V. N. (2019). Central Asian geolocation: Linguistic reception of toponyms (Turko-Mongolic area). *Oriental Studies*, 5, 945–956. <https://doi.org/10.22162/2619-0990-2019-45-5-945-956>
- Aleshinskaya, E. (2020). Translation in forming musical discourse: A case study of English-language song lyrics in Russia. *Research in Language*, 18(4), 395–405. <https://doi.org/10.18778/1731-7533.18.4.03>
- Assadorian, A. (2017). On the toponymy of the Iranian Azerbaijan. *Advances in Anthropology*, 7, 146–153. <https://doi.org/10.4236/aa.2017.73010>
- Bába, B. (2016). Lexical toponyms in toponyms. *Acta Onomastica*, 55, 17–24.
- Barthold, V. V. (1963). *Turkestan during the Mongol Invasion*. Nauka.
- Chlela, C. (2024). Some observations on the medieval rural settlements in northern Lebanon twelfth – thirteenth centuries. *Social Sciences & Humanities Open*, 9, 100780. <https://doi.org/10.1016/j.ssaho.2023.100780>
- Garcia-Quera, N. (2025). The etymology of opaque place names based on a cognitive and interdisciplinary method. *Language Sciences*, 107, Article 101688. <https://doi.org/10.1016/j.langsci.2024.101688>
- Hua, H., & Perono Cacciafoco, F. (2026). Kula toponyms: Preserving the cultural–linguistic landscape of Eastern Alor. *Data*, 11(3), Article 61. <https://doi.org/10.3390/data11030061>
- Khromov, A. L. (1980). *Some features of the formation of toponymy of Central Asia in the 9th – 13th centuries*. Nauka.
- Liu, Y. (2026). Translation as language contact: A multidimensional perspective of syntactic variation. *Frontiers in Language Sciences*, 4, Article 1712405. <https://doi.org/10.3389/flang.2025.1712405>
- Liu, Y., Liu, L., Xu, R., Yi, X., & Qiu, H. (2024). Spatial distribution of toponyms and formation mechanism in traditional villages in Western Hunan, China. *Heritage Science*, 12, 171. <https://doi.org/10.1186/s40494-024-01297-z>
- Martynenko, I. A. (2023). Hispanic toponymy of Chile: Linguocultural and linguopragmatic approach. *Linha D'Água: São Paulo*, 36(1), 160–182. <https://doi.org/10.11606/issn.2236-4242.v36i1p160-182>
- Murtuza, A. R. (2017). The Origin of toponym words (nomenclature) and the geographical terms in toponymy of Zangazur area. *International Journal of Science and Research*, 6(4), 1615–1619.
- Perono Cacciafoco, F., & Cavallaro, F. (2023). *Place names: Approaches and perspectives in toponymy and toponomastics*. Cambridge University Press.
- Qian, S., Kang, M., & Weng, M. (2016). Toponym mapping: A case for distribution of ethnic groups and landscape features in Guangdong, China. *Journal of Maps*, 12(sup1), 546–550. <https://doi.org/10.1080/17445647.2016.1201017>
- Saltakova, Z. T., Oralbek, A., Mursalimova, N., Kudaibergenova, K. T., & Kenbayeva, A. Z. (2026). Transforming national folklore in digital space: A case study of Internet memes in modern Kazakhstan. *Research Journal in Advanced Humanities*, 7(2). <https://doi.org/10.58256/txw8s555>
- Shakirova, K. M., Kenbayeva, A. Z., Suleimenova, Z. Ye., Myrzagaliyeva, A. S., Obayeva, G. S., Amangaliyeva, A. Ye., & Mambetova, G. Zh. (2026). Turkic language speakers' perception of toponyms as cultural heritage in the national epos “Edige”. *Theory and Practice in Language Studies*, 16(1), 215–223. <https://doi.org/10.17507/tpls.1601.23>
- Tulibayev, Zh., Karimova, N., & Tulibayev, T. (2025). Historical geography of Central Asia in oriental narrative sources of the 17th to the beginning of the 20th century. *Kazakhstan Oriental Studies*, 16(4), 247–259. <https://doi.org/10.63051/kos.2025.4.247>
- Usmonova, M., & Xujboyeva, M. (2025). Ancient Central Asia and Its Toponyms. *Development and Innovations in Science*, 4(6), 80–83.
- Yama, A., & Wasino, W. (2025). Kazakh ethnogenesis and the formation of Turkic identity in Central Asia: A historical and cultural analysis. *Paramita: Historical Studies Journal*, 35(2), 258–275. <https://doi.org/10.15294/paramita.v35i2.24979>
- Zhang, L., Zhou, S. & Ren, J. (2025). Spatiotemporal distribution of toponymic cultural heritage of water-

related settlement place names in Manas River Basin. *Heritage Science*, 13, Article 337. <https://doi.org/10.1038/s40494-025-01912-7>

Zhao, F., Fu, Y., Luan, G., Zhang, S., Cai, J., Ding, J., Qian, J., & Xie, Z. (2020). Spatial-temporal characteristic analysis of ethnic toponyms based on spatial information entropy at the rural level in Northeast China. *Entropy*, 22(4), Article 393. <https://doi.org/10.3390/e22040393>