

Ancient Toponyms as Historical Witness: Linguistic Reconstruction of Turkic Place Names in Eastern Chronicles

Bibatpa A. Koshimova

Department of Kazakh Language and Literature, Faculty of Philology, Yessenov University, Aktau, Kazakhstan

Bibaysha I. Nurdauletova

Department of Kazakh Philology, Faculty of Philology, Yessenov University, Aktau, Kazakhstan

Ainur Z. Kenbayeva*

Department of Foreign languages and Literature, Faculty of Philology, M. Utemisov West Kazakhstan University, Uralsk, Kazakhstan

Sharapat S. Abisheva

Department of Kazakh Language and Literature, Faculty of Philology, Yessenov University, Aktau, Kazakhstan

Akmaral N. Otarova

Department of Kazakh Language and Literature, Faculty of Philology, Yessenov University, Aktau, Kazakhstan

Abstract—Recent years have witnessed a surge of interest in toponyms as an important historical and cultural heritage. Analyzing their historical etymology provides a better understanding of the role place names perform in the culture and history of a people, as well as in the preservation and transmission of their value to future generations. The aim of this study is to reconstruct the etymology of place names associated with the territory of Central Asia. The qualitative analysis of historical etymology goes beyond the examination of national Turkic sources to include historical sources from other regions, such as East Asia and the Middle East, in which Central Asian geographical names are mentioned. As a result, the study identified Arabic-Persian and Turkic roots in a large number of toponyms, demonstrating close contacts and interconnections between various onomastic traditions. The qualitative research into the origins of ancient Central Asian toponyms strengthens a sense of solidarity and historical continuity among Turkic-speaking communities, highlighting the common linguistic and cultural characteristics of the region. The deep analysis of semantic patterns and historical language contacts contributes to the preservation and revival of the Turkic languages. The research results represent a valuable contribution to linguistics, ethnography, and history, helping to better understand the collective heritage of Turkic peoples and its impact on modernity.

Index Terms—toponym, Central Asia, East Asia, Middle East, historical etymology

I. INTRODUCTION

The study of toponyms is a fundamental task in modern historical linguistics, archaeology, and ethnography. The region's toponymy bears witness to the past, objectively recording centuries-long processes of ethnic migration, linguistic contacts, and geopolitical transformations (Assadorian, 2017). In the absence of written sources from the early Middle Ages, toponyms become primary historical documents, which permit reconstructing the chronology of settlement, trade routes, borders, and cultural interactions between peoples (Qian et al., 2016; Zhao et al., 2020). Recently, there has been growing interest in the study of toponyms as tools for the historical and archaeological study of different regions in the world, including West and East Asia. Thus, the analysis of toponyms in the County of Tripoli, which is currently Lebanon and Syria, enabled to a better understanding of Crusader settlement patterns between the 12th and 13th centuries (Chlela, 2024). The investigation into natural and cultural landscape toponyms in traditional villages in western Hunan, China, provided valuable insights into their spatial distribution and formation mechanism essential for the sustainment and preservation of the cultural heritage in the region (Liu et al., 2024). The study of water-related settlement names contributed to the current knowledge of regional development and demographic shifts in the Manas River Basin (Zhang et al., 2025).

Toponyms of Central Asia have also gained considerable attention in recent research. Central Asia has been demonstrated to be not only the historical and cultural center of the ancient Turkic civilization, but also the main area of ethnogenesis and civilizational development of the Turkic peoples (Yama & Wasino, 2025). Furthermore, ancient toponyms in this region provide clear evidence of the historical memory and cultural code of the nation (Pangereyev et

al., 2023), and the discovery of their origin and meaning plays an important role in studying the spiritual and historical foundations of modern Turkic integration (Kenbayeva et al., 2025), as well as preserving the nations' cultural heritage (Shakirova et al., 2026). In this respect, the place names left by the ancient Turks are not only geographical landmarks, but also linguistic and historical evidence of their ethnic history, social structure, migration routes, and connections with neighboring civilizations.

Since many toponyms associated with Central Asia have a rich history, their linguistic reconstruction presents significant challenges. The names of these places have undergone numerous changes, having been used differently in different eras and in different languages. To accurately reconstruct the etymology of toponyms, relying solely on the sources of a given people is insufficient. To overcome these challenges, the current study looks at broader historical sources that permit going beyond the Turkic peoples to trace the development of Central Asian place names. It emphasizes the key role of chronicles from the Middle East and East Asia in reconstructing the historical etymology of toponyms belonging to Central Asia, as they allow tracing the evolution of these names, their structure, and their variants in languages other than the Turkic ones. Studying the etymology of ancient toponyms preserved in Chinese and other Eastern chronicles helps restore lost cultural connections of the past. Of special interest is the identification of connections with Arabic-Persian and Turkic origins in the toponyms of Central Asia. Hence, this research aims to reconstruct the historical etymology of geographical names associated with Central Asia using Eastern historical documents, as well as to systematize them based on their Arabic-Persian and Turkic origins. This historical and linguistic analysis of toponyms is important not only for understanding the historical legacy but also for understanding the spiritual and cultural unity of the modern Turkic world.

II. METHODOLOGY

A. Data Collection

This qualitative study is based on the linguistic and historical analysis of ancient toponyms of Central Asia in Chinese and Arabic-Persian sources, which contained descriptions of such medieval regions as Mawerennahr (Maverennahr), Khorezm, Khorasan, Semirechye, Turkestan. These regions relate to modern Central Asia and include the territories of modern Uzbekistan, Turkmenistan, Kyrgyzstan, Tajikistan, Kazakhstan, the northern part of Iran, and Afghanistan. Among Eastern chronicles containing such data are the following sources: 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-a'lak an-nafisa", Yakut al-Hamawi's "Mu'jam al-buldan". These sources should form the basis of the study, ensuring the reliability of the data, as they contain early references to Turkic territories in the context of conquests, as well as route descriptions with precise locations of names.

Due to the lack of publicly available medieval Eastern sources, we used secondary sources as material for our research, that is, works by Western and Russian scholars who studied Central Asia based on these Arabic-Persian and Chinese sources. Historical and linguistic analysis of toponyms from Arabic-Persian texts relating to Central Asia and Kazakhstan occupies an important place in the works of many orientologists and Turkologists. Such research not only explains the origins of names but also reveals the ethnocultural processes underlying them. Toponyms associated with Central Asia were collected by purposive sampling from secondary sources containing references to Eastern historical sources that belong to East Asia and the Middle East. Reconstructing the toponymy of Central Asia based on Eastern chronicles confirms the authenticity of the forms found.

The qualitative study identified lexical units denoting place names associated with Central Asia from Chinese and Arabic-Persian historical sources as well as the explanation of their origin and cultural background. Our research corpus comprised 1000 units that were used in Eastern chronicles naming cities, towns, villages, caravanserais, canals, fortresses, and wells. Except for large cities, most of the place names found in these chronicles are no longer in use today.

B. Data Analysis

In our etymological analysis of toponyms, we took into account not only linguistic analysis, but also the history of the people, local geography, as well as the peculiarities of the perception and description of nature by the people. Toponymic research should involve the analysis of different types of data from linguistics, history, ethnography, and geography altogether. We utilized the qualitative method of intensive topology developed by Tent (2015), which involves creating a "biography" of toponyms through answers to questions about their etymology and semantics including the meaning of a place name, the reasons for naming this place, its location, etc.

The corpus of 1000 toponyms was coded according to the following analytical categories: toponym type (city, town, village, caravanserai, canal, fortress, well, etc.); linguistic origin (Turkic, Iranian, Sogdian, Persian); chronological layer (early Islamic, Samanid, Seljuk, later medieval attestations); geographic location (modern states and historical regions: Mawerennahr, Khorezm, Semirechye, Turkestan, etc.); semantic field (water-related, relief-related, ethnic or tribal, administrative, religious names). These categories were systematically applied during the corpus annotation so that each toponym was assigned multiple intersecting codes, enabling both linguistic and historical-etymological analysis.

To perform a detailed analysis of the origins and historical significance of toponyms, as well as their similarities and differences with modern forms, the study combined historical-comparative, analytical-interpretive, historical-linguistic, functional-systemic, and etymological methods. The data obtained were systematized by toponym type, their geographic location, modern and original names, and cultural significance in the context of Central Asia. The historical-comparative and historical-linguistic methods helped trace the evolution of names over time and identify differences between historical and modern linguistic forms. The analytical-interpretive and functional-systemic approaches allowed the identification of common characteristics, as well as the cultural and semantic features of toponyms. The integration of these methods consisted in semantic and morphological analysis to classify the etymology of each toponym using standard Turkic and Iranian etymological references, as well as historical-comparative tracking across Chinese, Arabic–Persian, and secondary scholarly sources to reconstruct chronological variants and identify shifts in name usage, spelling, and cultural interpretation.

III. RESULTS AND DISCUSSION

A. Toponyms of Arabic-Persian Origin

Among the oldest toponyms associated with Central Asia are the names of Arabic or Persian origin. They are common both in the Turanian Plain and in the Pamir Mountains, but are much less common in the mountainous regions of the Inner Tien Shan, in Jetysu and the steppe regions of Kazakhstan (Karayev, 1991). It has been established that the majority of ancient toponyms related to Central Asia in Arab-Persian sources (about 70%) 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 and reflect the ethnic composition and historical ties of that era.

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). The emergence of toponyms in Central Asia comprising such Western-Iranian topoforms as *-an*, *-kird* (or *-gird*, *-jird*), *-diza* (or *-diz*), *-rud*, and *-stan* (or *-istan*) also dates back to this time (Khromov, 1974). However, elements of truly Arabic origin are not very numerous in the corpus, being mainly found in the south of Uzbekistan and in the Zarafshan valley (Karayev, 1991).

Let us demonstrate different versions of toponyms based on some examples. One of the oldest names for the Central Asian region is *Turan*, based on the ethnonym *Tur*, which was the general tribal name of ancient nomadic and semi-nomadic peoples. During the reign of the Sassanid king Shapur I, who reigned between 242 and 272 A.D., the name *Turistan*, meaning “the land of the Turks”, became widely spread among the Persians, along with the name *Turan* (*Tugan*). On Sasanid seals of the 3rd-4th centuries, along with the names *Turan* and *Turistan*, the name *Turgistan* (*Turgistan*) was also found as their equivalent. Later, during the Hephthalites (5th century) and the Turks (6th–8th centuries), this name was changed to *Turkestan* and began to be used in the meaning of “the land of the Turks” (Kamoliddin, 2006, p. 10). In the understanding of the ancient Iranians, *Turkestan* geographically included the basins of the Amu Darya, Syr Darya and Tarim rivers. The Arab conquerors of the 7th-8th centuries introduced their own term, *Ma wara’ an-nahr* (a simplified form of *Mauerannahr*). *Mauerannahr* was the Arabic name for the Central Asian lands located north of the Amu Darya and literally meant “the land that lies beyond the river”. This geographical name was widely used in the Middle Ages, and in Persian there were also less well-known equivalents, such as *Vararud* and *Varajaihun*, which had a similar meaning. In the 19th century, a new term was introduced into the European scientific and historical tradition, *Transoxania* or *Transoxiana*, meaning “the land beyond the Oxus” or “the other side of the Oxus River”, which was a Latin equivalent of the Arabic name *Mauerannahr* (Kamoliddin, 2006, pp. 11-12).

Thus, the Central Asian region was repeatedly influenced by foreign conquerors throughout its history, each of which gave this place its own name. As for the Central Asian local historical and geographical tradition, the most widespread among these names among the local population were the names *Turkestan* and *Mauerannahr*. In Chinese sources, Central Asia was called *Siyu*, or “the Western Region”, standing for East Turkestan in a narrow sense and Central Asia in a broad sense. *Mawerennahr* was one of the largest civilization hubs of the Great Silk Road, covering many ancient cities from Samarkand and Bukhara to Otrar and Tashkent. This land, known as *Mawerennahr*, now occupies a large and central part of the Republic of Uzbekistan, also including the southern parts of the Kyzylorda and Turkestan regions of Kazakhstan and the western region of Tajikistan. Major cities such as Samarkand, Bukhara, Shash (Tashkent), Fergana, Khujand were important centers in this region. This historical region, which is home to cities of spiritual, political, and economic importance, has long been the cradle of civilization.

Fergana was one of the oldest cultural centers in Central Asia. Fergana is called the “Gateway to Turkestan” in sources. Chinese sources contain information about the city called *Hu-xun* (*Hu-sun*), which is located in the eastern part of the Fergana Valley. The Fergana Plain or Fergana is a commodity valley located in the upper reaches of the Syr Darya River and at its source, surrounded by the Pamir-Alai and Tien Shan mountain systems. According to written sources, in the early Middle Ages, Fergana occupied a special place as a historical and cultural region among the Central Asian states and played an important role in the formation of a unified civilization of the region. The Fergana Plain is surrounded by the Tien Shan mountain ranges in the north and the Hissar-Alai mountain ranges in the south,

and is located at the junction of important strategic routes, so it participated in all political, economic and ethnocultural processes in Central Asia. Early medieval Arabic-Persian authors such as Al-Istakhri left information about Fergana and its cities: *Akhsikat*, *Kubo* (*Kuwa*), *Hikat*, *Uzgand*, *Nisya*, *Asbara* (*Isfara*), *Nekad* (*Navkat*), *Miyonrud*, *Zhigil*, *Urest*, *Biskent*, *Salot*, *Vanket*, *Saj* (*Sukh*), *Khuwakand* (*Kokkan*), *Rishtan*, *Marginon*, *Zandaramin*, *Bereng*, *Ushtikan*, *Andukon* (*Andijan*) (Ayubov, 2016).

Despite the extensive historical and geographical data, the problem of the etymology of the names of cities in early medieval Fergana has not yet been fully resolved. For instance, the component *kat* in the toponym *Akhsikat*, which was the name of the main city of Fergana, comes from the ancient Sogdian language and could be found in the forms *kata*, *kand*, and *kant* meaning “city”, “settlement”, “house”. Parallels can be drawn with other languages: *kata* meaning “underground house, tomb” in Avestan; *kada* meaning “house” and *kanda* meaning “mound” in Tajik-Persian; *khandak* in Arabic; *kand* meaning “mound” in Afghan; *kantha* meaning “city”, “mound”, “wall” in ancient Indian languages; *kant* meaning “city”, “settlement” in Yagnob; *kaent* meaning “building” and *xatan* meaning “room” in Ossetian; *kandai* meaning “part of the city” or “quarter” in Afghan; *cīd*, *cod* meaning “house” in Pamir languages. Later, the component *kat* was transferred from Iranian languages to Turkic languages in the form of *kant/kent* retaining its original meaning. The emergence of settlements with the component *kat* or *kent* in their names dates back to the 10th century, of which the form *kat* is considered the most ancient. The term *kat* is interpreted differently in different locations: in documents related to the lands of Bukhara in combination with the word *khudo* meaning “chief”, “owner” it can denote being “strong” or “firm”; it is also found in the mountainous regions of Tajikistan in the meaning of “house”. Later, the meaning of the term *kat/kand/kent* expanded and became a general term denoting “city”. In documents found in the Mugh Mountains of Sogd, the name *Akhsikat* is found in the form *ysykandh/ysykt/xsykte*, which is formed from the Turkic words *ak* and Sogdian *kat* meaning “white city” (Ayubov, 2016). According to some researchers, the name of the city may be associated with the name of the Aksu River, the ancient name of the medieval Naryn River. The *kat* component is also found in a number of toponyms in the Fergana region: *Uzgand*, *Hikat*, *Nekad* (*Navkat*), *Biskent*, *Vanket*, *Huvakand* (*Kokkan*), *Vasket*, *Kashuket*.

Based on the data from Eastern chronicles, it can be stated that the majority of ancient toponyms found in Central Asia reflect the linguistic and cultural heritage of the ancient inhabitants of the region, as well as the ethnic composition and historical ties of that era. With the spread of Islam in the 8th-9th centuries, Arabic culture and language began to actively influence the 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, the addition of grammatical structures in the Arabic language, and the semantic shift of some words. This phenomenon did not consist only in linguistic adaptation, but it was also a manifestation of political and cultural dominance. Arabized toponyms determined the formation of new administrative structures and religious centers in the Islamic era, for example: *Ahsisak* meaning “a river, a ditch, a spring with little water”; *Arapkhana* meaning “the house of an Arab”; *Baban* meaning “gate”, “door”, “entrance”; *Babshir* meaning “the gate of the city”; *Nuwayze* meaning “new settlement”. These toponyms clearly indicate the point of intersection of the ancient Indo-Iranian culture of the region and the later Islamic Arab culture. Through their analysis, it is possible to gain a better comprehension of the continuity between linguistic layers, as well as the historical development and ethnic ties of the peoples of the region.

B. Toponyms of Turkic Origin

Names of Turkic origin, widespread in Central Asia even in ancient times, made up a significant part in the Middle Ages and still occupy a dominant place in modern toponymy. In the early Middle Ages, they constituted the second largest group in terms of number after toponyms of East Iranian origin, and in the Middle Ages they formed the most powerful layer in the toponymy of Central Asia. At that time, many cities and settlements had two versions of their names, Sogdian and Turkic (Kamaliddinov, 2013). Turkic toponymy is found in the northern, central, and southern regions of Central Asia. Names of Turkic origin reflect the multifaceted nature of the ethnic and linguistic composition of the region.

Of interest are ancient toponyms created by the method of calque, that is, by literally translating the meaning of a word or phrase from one language into another. Some geographical names in Chinese are semantic calques (direct semantic translations) of Turkic toponyms. For instance, Chinese sources mention a city called *Chi-gu-cheng*, located in the Semirechye region. Researchers usually associate it with the region on the southeastern shore of Issyk-Kul. This name in Chinese means “city of the red valley” and corresponds to the Turkic name *Kyzylsu*. In this region, the Chon-Kyzylsu and Kishi-Kyzylsu rivers flow into Issyk-Kul (Umurzakov, 1978). Some other geographical names found in Chinese sources can be considered calques of local names of Turkic origin, including, for example, *Zhe-hai* meaning “Warm Sea”, *Zhen-hai* meaning “Salt Lake”, *Ling-shan* meaning “Ice Mountains”, *Chi-shan* meaning “Red Mountains”, *Bin-yu* meaning “Thousand Springs”, *Qian-quan* meaning “Thousand Springs”.

Another crucial issue in the analysis of Central Asian toponyms in Eastern chronicles is related with topoformants, which not only reveal the meaning of place names, but also determine the time of their formation and linguistic basis. The linguistic influence of each era and ethnic group is clearly visible in the formation of toponyms in the regions of Central Asia. Thus, topoformants such as *-diz*, *-kat*, *-gird*, *-kent* are important indicators of the historical periods of the region, the continuity of cultural influences and linguistic layers. The majority of medieval and ancient toponyms in the

study area were formed on the basis of topoformants introduced from the Sogdian, Persian and Turkic languages. Their comparative-etymological analysis makes it possible to identify the features of the ancient toponymic system of the region.

One of such topoformants is *-diz* (*-diza*). *Di* means “settlement” in the Tatar language and Tajik dialect; the forms *de/di* are derived from the Iranian word *diza* meaning “fortified place”, “fortress, castle” in some Pamir languages; the Avestan version of this word means “to pile (earth, stone)”. In Turkic, this word is preserved in the Azerbaijani language in the form of *diza* in the meaning of “fortress”, “fortress on the top of the mountain”, which is easily comparable with the Persian words *diz*, *daz*, *duz* denoting “fortress on the top of the mountain, fortress, castle-palace”. The Sogdian topoform *diz/diza* is explained to mean “mound”, “fortress” (Khromov, 1983). Such names date back to the pre-Islamic era and were associated with small early medieval fortresses such as *Suvadiza*, *Farkhuzdiza*, *Ustugdadaiza*, *Khashyandiza* in the Nasaf region. The Sogdian formant *-diz/-diza* was shortened in later times to *-di* and is also found in modern toponyms, for example: *Pakandi* (*pai* + *kan* + *diz*) meaning “a fortress on the lower reaches of the river”; *Lagandi* (*lakka* + *diz*) meaning “a fortress on the open ground”; *Kamandi* (*kam* + *diz*) meaning “a fortress on the river”; *Kakaidi* (*kuh* + *kad* + *diz*) meaning “fortress in a mountain settlement” (Nafasov, 1988). Thus, being the names of early medieval castles and fortresses, these toponyms have survived since the pre-Islamic era and have been transformed into villages and settlements in later centuries. Such fortresses were the centers of small feudal estates and are considered an important source for studying the history of early medieval Central Asia.

In medieval sources, toponyms derived from the Sogdian and Old Turkic languages, associated with sacred places and temples of the pre-Islamic era, were often found. Such names contain the Sogdian components *bagn*, *vagn*, *fagn*, which mean “temple”, “house of God”. The occurrence of these elements indicates the presence of pre-Islamic temples there (Smirnova, 1971). Thus, the name of the city *Iskifagn* in the Nasaf region of the 10th century comes from the Sogdian words *isku* and *bagn* (“high temple”) or from the Old Turkic words *isku* and *pagn* (“old temple”). Documents from the 15th-17th centuries mention the settlement of *Khurbagn* coming from the Sogdian words *khur* (“sun”) and *bagn* (“temple”). Another widespread topoform of Sogd is *-misan/-mitan*, which means “temple” and is found in the toponymy of such regions as Bukhara, Samarkand, Khorezm, Merv, Fergana (Khromov, 1974). This element is preserved in the tribal name of the Karakalpaks and Uzbeks in the form *muyten* (Tolstova, 1978). Such names in the Kashkadarya valley include *Anshamisan*, *Farimisan*, *Kabamisan*, *Misan-i Kalan*, *Misan-i Khurdak*. Thus, toponyms with the components *bagn* (*wagn*) and *misan* have preserved traces of pre-Islamic religious places of worship, as they arose near religious centers of the pre-Islamic era. Such names are an important source of information for determining the history of ancient cultural and spiritual centers in Central Asia.

Another Sogdian topoform, *-kas/-kam* was used to create medieval toponyms, in which these components denoted “fortress”, “fortified settlement” or “small town”. For instance, the toponym *Shirkas* (*Sirkas*) meaning “fortress located on a hill” was widely used in all regions where the Sogdian language was widespread in the Middle Ages (Nafasov, 1988). This toponym has survived to the present day in the Turkic form *Shirkent* naming a small town located in the upper reaches of the Surkhandarya River and a settlement in the Kasbi district of the Kashkadarya region. In the 12th century, a settlement called *Fijkas* was named in the Nasaf region, and it can be found in the 16th-century documents as *Bijkat* or *Wijkat*. According to the documents, it was located to the west of the lands of the Gubdin settlement, on the banks of the canal of the same name (Mukminova, 1966). The toponym *Fijkas* or *Bijkat* has survived to the present day in the Turkic form *Beshkent*. Today, this name is given to the district center, located 20 km southeast of the city of Karshi. This toponym means “lower fortress” in Sogdian (Nafasov, 1988). In the Kesh region, al-Sam’ani also mentions a settlement called *Surkhakas*, which is located in the administrative region Khuzar rustag and denotes “red fortress” in Sogdian.

The Sogdian topoform *-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 such a *-kent* or *-kand* component, the origin of which is associated with the early medieval period. For example, the village of 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 (Nafasov, 1988). The modern village of Pakhtaobod in the Kitab district was previously called *Parchakent*. This name translated 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 the Sogdian language. Thus, as a result of the study of the toponymy of medieval Central Asia, a number of groups of toponyms containing the Sogdian topoforms *-kas*, *-kam* or *-kad* can be distinguished. All of these names belong to the pre-Islamic toponymic layer. This topoform has been preserved to this day in the Turkicized form *kent* and is widely used in the modern toponymy of Central Asia (Kamaliddinov, 2013). This is evidenced by many place names in the territory of modern Central Asia and Kazakhstan, for example: *Beshkent*, *Yangikent*, *Shurkent*, *Shymkent*, *Manket*, etc., all of which have a clear Turkic etymology.

The topoform *kan*, derived from the Sogdian word *kanta* or *kanda*, was one of the most common concepts in ancient times, meaning “digging”, “digging a pit”, “dug land”, as well as “water”. These meanings are especially closely related to the life of peoples engaged in agriculture, since digging water channels and ditches for agriculture, and cultivating the land were one of the main conditions. In folk linguistic usage, the initial sound “k” of this word underwent phonetic

changes over time, and in some regions it was replaced by the sounds “g” or “zh”. For example, the word *kan* has been preserved in certain place names, changing into *kon*, *gon*, *jon* or *zhan*. This phenomenon is considered a natural manifestation of the linguistic evolution of the peoples of Central Asia. For this reason, these forms are often found in the historical geography of Central Asia, especially in the names of settlements located in oases where agriculture was practiced. Such names testify to the economic specificity of the region and the existence of ancient irrigation systems, for example: Andukan, Ankulkan, Vivakan, Balujuzjan, Bugavizjan, Buzmakan, Kutankan, Surfukan, Gulkan, Darakan, Konimeh, Kumishkon, Khaydarkon, Mchgkon, Vobkan, Barkan, Savakon, Gondara, Andagon, Anjirfagon, Buzgon, Sagon, Jondar, Samjon, Arbinjon, Robinzhan, Obijaya, Khuzhan, etc. Most of the names of these settlements are based on hydronyms associated with water sources and local hydrography. Therefore, topoforms such as *kon*, *gon*, *jon* or *zhan* were originally used as general names for various water bodies. They meant “river”, “channel”, “lake”, “main canal” and similar water systems.

All these names show traces of ancient agricultural culture. They reflect not only the development of land cultivation and irrigation, but also the socio-economic life of the peoples of the region. In addition, these toponyms are clear evidence of cultural and linguistic ties of that period. For example, most of the names with the *kan* component are located along the Silk Road, which indicates that they were centers of not only agriculture, but also trade and cultural exchange. In addition to the above-mentioned topoforms, toponyms were also created using the following geographical terms: Sogdian *barz* (hill) – Barzan, Khalbarzan; *kar*, *gar* (mountain) – Batkar; *behar* (church) – Behar; *kaah* (tower, palace) – Bulkah Abu Vahb; *dar*, *dara* – (pass, “mountain pass”, “ravine, river) – Darakan; *djay* (in the toponymy of Upper Zeravshan this word means “reservoir” or “pond”) – Jayasar; *ab* (water) – Divab, Suyab, Dulab; *huk*, *kuh* (mountain) – Khukandja, Siyaku; *khan* (well, pool) – Zandakhan, *kut* (fortress) – Kutankan, *khan* (house) – Maikhana, Arabkhana; *van* (house) – Vahuvan; *zhar* (mountain) – Najjar; *var*, *vara* (mound) – Sadivar, Huvaran; *sang* (stone) – Sang; *shah* (river tributary, river, branch) – Shahmin, Shakalan; *an* (land, country, settlement) – Shaukan, Turan, Barzan, etc.

A number of names of settlements found in medieval sources are associated with ethnonyms, that is, the names of certain peoples and tribes. Among them, along with Turkic tribes, there are names of Arab and Iranian-speaking peoples. For example, the name of the Qajar settlement in the Nasaf region comes from the Turkic ethnonym *qajar* (Nafasov, 1988). The name of the Chagyroguz or Jafragun mountain and the Chagirchik (Chirchik) river on the coast of the Khorezm (Aral) Sea is associated with the Turkic ethnonym *chagirak* or *chagirat* (Barthold, 1963). The name of the Khuttalan region is associated with one of the most powerful tribes among the Turks, the Khuttalakhs, the people who used mule meat as food (Kamoliddin, 2006). The emergence of such names is often the result of the migration of ethnic groups, settlement areas and historical and cultural influence. Therefore, ethnotoponyms are an important historical source that reflects not only a geographical object, but also the ethnic history of the region, migration routes, and cultural ties.

IV. CONCLUSIONS

The toponyms of Central Asia, recorded in Eastern sources, represent unique historical and linguistic material. They not only designate geographical features but also reflect the ethnic, linguistic, and cultural history of ancient peoples in the region. The study revealed the multilayered nature of ancient names preserved in Arabic-Persian and Chinese chronicles. They intertwine Iranian, Sogdian, Persian, and Turkic elements, attesting to complex historical and cultural contacts, as well as changes in ethnic composition. The historical and linguistic approach has proven effective in studying the toponymy of the region, permitting tracing not only the evolution of languages but also the cultural continuity of peoples. The results demonstrate close connections between the onomastic systems of various cultures. Analysis of Eastern sources allowed reconstructing the origins of toponyms, determining their ethnoterritorial significance, and identifying the influence of linguistic contacts. These data have practical implications for Turkic onomastics, as they can be utilized in the development of interactive migration maps, corpora of parallel texts, and interdisciplinary models of ethnogenesis using GIS technologies and corpus linguistics.

This study has some limitations. The selection of sources reflects the availability and accessibility of extant works in Western and Russian-language scholarship, which may underrepresent certain regions or periods. Furthermore, the quantitative-qualitative balance of the 1000-unit corpus, though substantial, might not fully capture the geographic and temporal diversity of Central Asia, for some remote or less-documented locations may be underrepresented in the corpus. Despite these limitations, the study lays a methodological groundwork for future, digitized, and GIS-supported toponymic corpora that can refine and expand the proposed reconstruction model. Overall, the study of ancient toponyms contributes to the preservation of cultural heritage, strengthening historical memory, and fostering a sense of unity among Turkic-speaking peoples. This work contributes to the understanding of the collective heritage of Central Asia and its significance for the present day.

DECLARATION OF THE USE OF AI

Perplexity.ai was used to translate fragments from Kazakh into English. These translations were afterwards carefully checked by the authors of the manuscript for clarity and precision.

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’.

ACKNOWLEDGEMENTS

The authors express their deep gratitude to Michael Erdman, Head of the Middle East and Central Asia Collections at the British Library, for his assistance in providing access to unique and rare sources in Arabic and Chinese, as well as manuscript materials from the collections of Central Asian Turkic languages.

REFERENCES

- [1] Assadorian, A. (2017). On the toponymy of the Iranian Azerbaijan. *Advances in Anthropology*, 7, 146–153. <https://doi.org/10.4236/aa.2017.73010>
- [2] Ayubov, A. R. (2016). History and etymology of toponyms of Fergana according to written sources. *Scientific Notes*, 1, 42–49.
- [3] Barthold, V. V. (1963). *Turkestan during the Mongol Invasion*. Nauka.
- [4] 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>
- [5] Kamaliddinov, Sh. S. (2013). *Research on the historical toponyms of Central Asia. Toponyms, topoformats, etymology*. Sharq.
- [6] Kamoliddin, Sh. S. (2006). *Ancient Turkic toponymy of Central Asia*. Sharq.
- [7] Karayev, S. K. (1991). *Toponymy of Uzbekistan (Sociolinguistic aspect)*. Fan.
- [8] Kenbayeva, A., Abisheva, S., & Amangalieva, A. (2025). Representando um conceito nacional de espaço na literatura épica. *Bakhtiniana: Revista De Estudos Do Discurso*, 20(4), e67597e. <http://dx.doi.org/10.1590/2176-4573e67597>
- [9] Khromov, A. L. (1974). *On the structural features of the Iranian toponymy of Maverannahr in the period of the 9th-12th centuries*. Tajik State university.
- [10] Khromov, A. L. (1980). *Some features of the formation of toponymy of Central Asia in the 9th – 13th centuries*. Nauka.
- [11] Khromov A. L. (1983). *The term “diza” – “fortress” in the toponymy of Khorasan, Maverannahr and Tokharistan in the early Middle Ages: Bactria-Tokharistan in the ancient and medieval East*. Nauka.
- [12] 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>
- [13] Mukminova, R. G. (1966). *On the history of agrarian relations in Uzbekistan in the 16th century: Based on materials from “Vakfname”*. Fan.
- [14] Nafasov, T. (1988). *Complete list of toponyms of Uzbekistan (Southern regions of Uzbekistan)*. Fan.
- [15] Pangereyev, A. S., Umatova, Z. M., Azamatova, A. K., Ibrayeva, Z. K., Karagoishiyeva, D. A., Umarova, G. S., & Kenbayeva, A. Z. (2023). Ethnocultural originality of color toponyms in Turkic folklore. *Eurasian Journal of Applied Linguistics*, 9(3), 216–225. <http://doi.org/10.32601/ejal.903019>
- [16] 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>
- [17] Shakirova, K. M., Kenbayeva, A. Z., Suleimenova, Z. Ye., Myrzagaliyeva, A. S., Obayeva, G. S., Amangalieva, 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>
- [18] Tent, J. (2015). Approaches to research in toponymy. *Names: A Journal of Onomastics*, 63(2), 65–74. <https://doi.org/10.1179/0027773814Z.000000000103>
- [19] Tolstova, L. Kh. (1978). *Echoes of the early stages of the ethnogenesis of the peoples of Central Asia in its historical onomastics*. Nauka.
- [20] Umurzakov, S. U. (1978). *On the historical toponymy of Kyrgyzstan*. Nauka.
- [21] 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>
- [22] 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, 337. <https://doi.org/10.1038/s40494-025-01912-7>
- [23] 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), 393. <https://doi.org/10.3390/e22040393>



Bibatpa A. Koshimova was born on July 20, 1970, in Mangistau Region, Kazakhstan. She is an Associate Professor at the Department of Kazakh Language and Literature of Yessenov University, Kazakhstan. She completed her undergraduate at the Faculty of Philology of Al-Farabi Kazakh National University. She received her Candidate of Sciences degree in Philology in 2001 from the A. Baitursynuly Institute of Linguistics (Almaty, Kazakhstan), defending a dissertation entitled ‘The Process of Deonymization in the Kazakh Lexicon’. Her research focuses on the selection and critical review of domestic and international scholarly works devoted to the lexico-semantic analysis of toponyms. She examines toponymic studies developed within the Turkic, Russian, Western European, and Orientalist scholarly traditions and conducts a comparative analysis of the methodological approaches used for the interpretation of place names. The aim of

her researches is to harmonise the semantic typologies proposed by various scholars and to develop a unified universal model for the semantic analysis and classification of toponyms. She teaches a course entitled 'The Historical and Linguistic Foundations of Place Names in the Mangistau Region'. <https://orcid.org/0000-0001-5544-3712>



Bibaysha I. Nurdauletova was born on July 24, 1964, in Kyzylorda Region, Kazakhstan. She is a Professor at the Department of Kazakh Philology of Yessenov University, Kazakhstan. She completed her undergraduate at the Faculty of Philology of Al-Farabi Kazakh National University. In 2008, she was awarded the degree of Doctor of Philological Sciences by the A. Baitursynuly Institute of Linguistics under the Ministry of Education and Science of the Republic of Kazakhstan. Her research interests include ethnography, linguistics, history, and the toponymy of the peoples of Kazakhstan and Central Asia. She teaches courses for master and doctoral students in 'General linguistics', 'The Kazakh literary language' and 'Cognitive linguistics'. Her online course "Sacred Mangyshlak" has been presented on the international educational platform COURSERA. She serves as the principal investigator of three research projects funded by grants from the Ministry of Science and Higher Education of the Republic of Kazakhstan. She is the author of more than 200 scholarly publications, including six monographs, as well as textbooks and methodological manuals. <https://orcid.org/0000-0002-6911-6509>



Ainur Z. Kenbayeva was born on February 24, 1984, in Uzbekistan. She is a Senior Lecturer at the Department of Foreign Languages and Literature at M. Utemisov West Kazakhstan University, Kazakhstan. She received her doctoral degree after successfully defending her dissertation entitled 'The Typology and Poetics of Toponyms in the Folklore of Turkic Peoples' at the Kazakh National Women's Teacher Training University. Her research focuses on aspects of cognitive linguistics, intercultural communication, methods for teaching non-native languages. She conducts extensive research and publishes widely on the linguistics and literary traditions of Turkic peoples, with a particular focus on toponymy and the historical linguistics of place names. Her research examines the linguistic, historical, and cultural layers of Turkic toponyms and their representation in early written sources, including ancient Chinese chronicles and other Eastern historical records. Through the analysis of these sources, she investigates the evolution, semantic transformation, and cultural significance of Turkic place names. She teaches courses in 'Cognitive linguistics', 'Linguistic pragmatics, and 'Introduction to linguistics'. <https://orcid.org/0000-0003-3536-8713>



Sharapat S. Abisheva was born on October 10, 1965, in Mangistau Region, Kazakhstan. She is Professor at the Department of Kazakh Philology of Yessenov University, Kazakhstan. She completed her undergraduate at the Faculty of Philology of Al-Farabi Kazakh National University. She received her Candidate of Sciences degree in Philology in 2005 from Al-Farabi Kazakh National University. In 2025, she was awarded the degree of Professor by the Ministry of Education and Science of the Republic of Kazakhstan. Her scientific research interests include toponymics. The scholar studies the origins of geographical names, their historical development, and their historical-ethnolinguistic and cognitive characteristics, namely how a particular people's worldview, culture and cognitive experience are reflected in geographical names. She teaches courses '19th-century Kazakh literature' and 'Folk literature'. <https://orcid.org/0000-0003-4856-9936>



Akmaral N. Otarova was born on May 24, 1963, in Mangistau Region, Kazakhstan. She is an Associate Professor at the Department of Kazakh Language and Literature of Yessenov University, Kazakhstan. She completed her undergraduate at the Faculty of Philology of Karaganda State University. In 2010, she received her Candidate of Sciences degree in Philology from Al-Farabi Kazakh National University. She teaches courses 'Early 20th-century Kazakh literature' and 'Regional literature'. Her research interests also include early twentieth-century Kazakh literature and regional literary traditions, particularly in relation to linguistic, cultural, and historical representations of place and regional identity. <https://orcid.org/0000-0002-9793-6256>