

ARTICLE

Semantic Analysis of the Linguistic Heritage of the Golden Horde Period: Based on the Material of “Gulistan Bit-Turki” by Seif Sarai and “Khosrov and Shirin” by Qutb

Ainur Seitbekova *, *Gulfar Mamyrbek*, *Myrzabergen Malbakov*, *Assel Seidamat*, *Temirlan Zhorabay*

A. Baitursynyly Institute of Linguistics, Almaty 050010, Kazakhstan

ABSTRACT

The aim of this study is to examine the written monuments from the Golden Horde period and contemporary literary works, focusing on the frequency of use, functionality, and stylistic features of both indigenous and borrowed vocabulary within literary texts. The lexical units were analyzed using methods such as linguistic-statistical, contextual, historical-comparative, semantic, and contextual analysis. This study explored the literal and figurative meanings of words of Turkic, Arabic, and Persian origin used in the written monuments of the Golden Horde. The research provides insights into the primary lexical layers of the Turkic languages, the time of their formation, and their frequency of use in the literary tradition. Emphasis was placed on examples of contexts that confirm specific lexical meanings, as well as on metaphorical and figurative transformations aimed at achieving stylistic effects. The study also considered the history of the formation of Turkic languages, the adaptation of borrowed lexical layers, and the modification of primary meanings (metaphorical, figurative). Kipchak, Karluk, and Oghuz language elements were highlighted through the analysis of the written monuments, along with their frequency of integration. Contemporary literary texts were examined in terms of the use of lexical elements (both indigenous and borrowed) for artistic purposes. This research can be used for comparative analysis of closely related Turkic languages, the study of contemporary literary texts in terms of the use of indigenous and borrowed lexicon, and for forming an understanding of the ancient lexical layers of various Turkic languages.

Keywords: Middle Ages; Arabic and Persian Lexicon; Kazakh and Turkish Languages; Lexico-Semantic Adaptation;

*CORRESPONDING AUTHOR:

Ainur Seitbekova, A. Baitursynyly Institute of Linguistics, Almaty 050010, Kazakhstan; Email: ainurseit@mail.ru

ARTICLE INFO

Received: 30 August 2024 | Revised: 22 September 2024 | Accepted: 14 October 2024 | Published Online: 4 November 2024

DOI: <https://doi.org/10.30564/fls.v6i5.7175>

CITATION

Seitbekova, A., Mamyrbek, G., Malbakov, M., et al., 2024. Semantic Analysis of the Linguistic Heritage of the Golden Horde Period: Based on the Material of “Gulistan Bit-Turki” by Seif Sarai and “Khosrov and Shirin” by Qutb. *Forum for Linguistic Studies*. 6(5): 116–130. DOI: <https://doi.org/10.30564/fls.v6i5.7175>

COPYRIGHT

Copyright © 2024 by the author(s). Published by Bilingual Publishing Co. This is an open access article under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License (<https://creativecommons.org/licenses/by-nc/4.0/>).

1. Introduction

In the 13th to 15th centuries, the Golden Horde encompassed a vast territory from the Irtysh River to the Dnieper, and from the Volga Bulgaria to Khwarezm. The majority of the tribes inhabiting these areas were Turkic. During the 13th to 15th centuries, a Turkic written literary language developed based on the ancient Turkic language and the local Kipchak dialect within the territory of the Golden Horde^[1]. This language contributed to the emergence of numerous poetic and prose works until the early 20th century. One of the unique and valuable works that have survived to this day is Seyf Sarai's "*Gulistan bit-Turki*" and Qutb's "*Khusraw wa Shirin*".

This article will not discuss the discovery and publication of the aforementioned written monuments by Seyf Sarai and Qutb, as these details have been researched by many Turkologists, notably Nadzhip^[1, 2], Fazylov^[3], among others. Seyf Sarai's "*Gulistan bit-Turki*", which consists of eight chapters, does not follow a specific plot. The content of the text comprises various short stories, poems, and interesting narratives. The poet depicts individuals from different social strata, incorporating characters from all segments of society based on the events' content. He aims to convey his observations of goodness and evil, justice and tyranny, truth and falsehood, among other themes, using eloquent language, summarizing them with quatrains or proverbs, and delivering them with a deeper meaning to his people. In contrast, Qutb's "*Khusraw wa Shirin*" is focused on themes of love and affection.

In Turkology, there has not yet been a consensus regarding the language of medieval written monuments. This is due to the emergence of mixed languages such as Old Turkic, Oghuz-Kipchak, Kipchak-Oghuz, Oghuz-Turkmen, and Karluk-Oghuz from various Turkic tribes that lived in Central Asia, the Golden Horde, Syria, and Egypt, which laid the foundation for the development of Turkic literary language^[1]. For instance, among global Turkologists, there are differing opinions on the language of Seyf Sarai's "*Gulistan bit-Turki*". In a 1966 article published in the Kazan newspaper "*Qazan Utlari*", Tatar scholar H. Usmanov considered

"*Gulistan bit-Turki*" to be a written legacy of Old Tatar^[4]. In contrast, Uzbek scholar E. Fazilov compiled the lexical units of this written monument into his two-volume dictionary titled "Old Uzbek Language: Khwarezmian Monuments", categorizing the language of these monuments as Khwarezm Turkic, transcribing the text into Uzbek script and viewing it as a model of Uzbek literary language^[3].

Koprülü^[5] states that the language of this written monument is composed in a Kipchak-Oghuz mixed literary language. Turkish scholar Abdullah Battal-Taymas, in his extensive article written in 1955, emphasizes that the language of Seyf Sarai's written monument is related to the Kipchak language^[6]. Scholar Nadzhip^[2], further clarifying Abdullah Battal-Taymas's assertion, notes the influence of Oghuz-Turkmen language, adding that later, during the Mamluk period in the 14th century, a Kipchak-Oghuz literary language developed in Egypt.

Historically, after the fall of Genghis Khan's rule, the Chagatai Khanate and the Jochi Ulus, composed of various Turkic tribes, formed the Golden Horde state across the region from the Caspian Sea to Altai. Northern Khwarezm became part of the Golden Horde during this period. The flourishing of the Golden Horde, established during Mongol dominance, did not last long. Internal wars and conflicts hindered the development of science and culture, leading many prominent scholars and poets to migrate to Syria and Egypt. This migration somewhat elevated the status of the Mamluks. The cultural and economic ties between Egypt and the Golden Horde strengthened, resulting in the establishment of a new major center for Turkic literature based on the Kipchak language in Egypt by the second half of the 14th century. Consequently, during this period, a Turkic literary language based on the Kipchak language, influenced by Oghuz dialects, emerged in Egypt^[1]. For centuries, the close relations and shared struggles against common enemies between the Kipchak and Oghuz tribes inevitably influenced their languages. Thus, numerous Oghuz words and grammatical forms appear in the Kipchak language within the works of that time^[1].

Scholar Ibatov^[7] categorizes Seyf Sarai's "*Gulistan bit-Turki*" among written monuments and substantiates its

relation to the Kipchak language as follows: "The primary language of this monument is Kipchak. While it has been influenced by a few elements of Oghuz languages, it has had little exposure to the Karluk-Uighur literary tradition. The elements of Oghuz, Karluk, and Uighur languages are so thoroughly integrated that their foreignness is hardly detectable in the language of these monuments. Initially emerging in the lower region of the Syr Darya, this language spread beyond the Volga and reached Egypt during the Mamluk period. As it distanced from its original core, its distinctive features became more pronounced. In monuments like "*Gulistan bit-Turki*", the traces of Oghuz elements become barely perceptible. In such works, Oghuz and Karluk-Uighur elements blend seamlessly with general Turkic words. Thus, a Kipchak-Oghuz mixed language emerges, formed by the languages of diverse tribes living in Egypt. This language is referred to as the "Kipchak-Oghuz literary language"^[7].

Indeed, since the primary language of communication and the dominant language in Egypt and Syria was Kipchak, it is quite possible that the works written in that region feature a predominance of Kipchak elements over Oghuz elements. Therefore, Seyf Sarai's "*Gulistan bit-Turki*" can be classified as one of the works written in a mixed Kipchak-Oghuz literary language.

Scholars have differing hypotheses regarding the language of Qutb's epic "Khusraw and Shirin". A. Samoylovich posits that the majority of Turkic words used in the monument are characteristic of the Oghuz, suggesting it was written based on the Oghuz language^[8]. In contrast, Nadzhip^[1] argues that Oghuz elements prevail over Kipchak elements and that the work has also been influenced by Uighur literary traditions. Baskakov^[9] acknowledges the presence of Uighur elements but maintains that the work is based on the Kipchak language. Scholar A. Ibatov compares affixes and concludes that the affixes characteristic of the Kipchak language are predominant^[7]. Polish scholars Zayackowski^[10] and Shcherbak^[11] also conclude that the work is written in Kipchak.

Turkish scholar A. F. Karamanlioğlu notes the close relationship between the languages of the written monuments "*Gulistan bit-Turki*" and "Khusraw and Shirin"^[12].

When examining the historical development of the Kazakh language and the formation of its literary language, we should primarily seek materials from these Kipchak lan-

guages. To determine whether a particular monument's language is dominated by Oghuz, Kipchak, Karluk elements, or by Arabic and Persian words, one must meticulously analyze its lexicon, phonetics, and grammatical forms. Without this, one cannot definitively conclude that it was written based on a specific language. In this regard, we can substantiate the relation of the languages of "*Gulistan bit-Turki*" and Qutb's "Khusraw and Shirin" to the Kipchak language, particularly the Kazakh language, through a detailed analysis of their lexical resources (Turkic, Arabic-Persian).

The identification of the lexical closeness of the written monuments from the time of the Golden Horde to Turkic languages (Kipchak, Karluk, Oghuz), as well as to Arabic and Persian languages, represents an important philological task that requires an analysis of the lexical, phonetic, and morphological characteristics of the lexical stock. A thorough analysis of the language of each written monument is relevant, as it allows for the determination of the history of the formation of Turkic languages and their lexical connections with modern Turkic languages.

For Kazakh linguistics, there is a significant need to investigate the lexical stock of the written monuments of the Golden Horde, as many linguistic phenomena in contemporary Kazakh are still difficult to explain. According to M. Sabyr, historical reconstruction helps approach the linguistic history of medieval written monuments, which can shed light on issues related to the history of literature and language^[13].

The aim of this work is to study the Turkic and borrowed lexicon in the written monuments of the Golden Horde period, as well as in the literary works of modern Turkic and Kazakh authors. It seeks to establish an understanding of the historical periods of development of Turkic languages and the lexical layers of varying chronology, and to conduct contextual and semantic analyses of texts of Turkic origin. The subject of this work comprises ancient written monuments of the Golden Horde and contemporary literary works.

To date, the following works have been undertaken by the authors regarding the written monuments of Seyf Sarai's "Gülistan bit-türki" and Qutb's "Khusraw wa Shirin":

- (1) For the first time, the texts of the works have been transcribed from Arabic script into Cyrillic script.
- (2) For the first time, all words from the texts of the works have been compiled into two comprehensive explanatory dictionaries, which are currently being prepared

for publication.

Based on the developed dictionaries, the following tasks have been set to analyze the lexical units of the works:

- (1) Conduct a linguostatistical study of the lexical units in both works.
- (2) Determine that the language of the works is written in the Kipchak dialect by analyzing the phonetic and morphological characteristics of Old Turkic lexical units.
- (3) Analyze the semantic features of Arabic and Persian words.
- (4) Identify the formation characteristics of lexical units in the old language by comparing them with contemporary literary works.

The innovative aspects of this research include the preparation of two explanatory dictionaries based on the written monuments “Gülistan bit-türki” by Seyf Sarai and “Khusraw wa Shirin” by Qutb, which will serve as the foundation for semantic markup of these works in the historical subcorpus of the National Kazakh Language.

Study of Language Evolution: The semantic analysis of ancient texts can reveal how the meanings of words have changed over time, providing a unique opportunity to explore the process of forming the modern Kazakh language. This can aid in understanding linguistic dynamics, social structures, and cultural changes.

Interdisciplinary Approach: The research can integrate philology with data from other disciplines, such as historical semantics, cultural studies, and even computer science. This will allow for a deeper understanding not only of the lexical meaning of words but also of their cultural context and changes in usage over the centuries.

Thus, the innovation in this research can be expressed through the application of new technologies and methods, as well as the creation of a knowledge base that will enhance the understanding of the development of the Kazakh language and culture during that period.

2. Materials and Methods

2.1. Materials Used

The theoretical foundation of this study comprises works by contemporary Kazakh, Turkish, British, and other scholars who focus on indigenous and borrowed lexicon

within Turkic languages. In particular, the works of the prominent scholar and Turkologist who studied the language and culture of the Kipchaks of the Golden Horde and Mamluk Egypt during the 13th–15th centuries^[1, 2] were used as key references. Regarding the semantics of ancient and loanwords in the Middle Ages, the research findings of domestic scholars (Sagyndykuly,^[14] (Sabyr,^[13]), (Ibatov,^[7]), as well as Uzbek and Tatar scholars (Fazylov,^[3]), (Usmanov,^[4]), Turkish scholars (Battaltaymas,^[6]), (Köprülü,^[5]), (Karamanlioğlu,^[12]), Russian scholars (Samoylovich,^[8]), (Shcherbak,^[11]), (Baskakov,^[9]), and the Polish scholar (Zajączkowski,^[10]) were employed in our research.

As a primary source for the study, copies of the manuscript of Seyf Sarai’s “*Gulistan bit-Turki*”, published in 1980 in Kazan by Khatib Usmanov and Zeynab Makhsudova (Usmanov; Makhsudova,^[15]), were used. Additionally, the manuscript of Kutb’s “*Khosrow and Shirin*” published in the 16th volume of the Literary Artifacts series (Literary Artifacts,^[16]) served as a source for our research.

To facilitate the semantic analysis of different lexical layers, the following dictionaries were utilized: A Comprehensive Persian-English Dictionary, Including Arabic Words and Phrases Occurring in Persian Literature by Steingass^[17], *Dīwān Lughāt al-Turk* (Dictionary of Turkic Dialects) by Kashgari^[18], Arabic-Kazakh Dictionary by Ondasynov^[19], Dictionary of the Kazakh Literary Language by Iskakov and Uali^[20], and the Ottoman Turkish Dictionary^[21].

For the contextual analysis of lexical units, the following written monuments from the Golden Horde period were selected: *Gulistan bi’t-Turki* by Saif-i Sarai (1968) and *Khosrow and Shirin* by Qutb (1958), as well as contemporary literary works: Sarsenbaykyzy’s *Baghalym*^[22] and Perker’s *Şinto—Şinto’nun Yolu Doğanın Kalbine Gider* (Shintoism—The Path of Shinto Leads to the Heart of Nature)^[23].

2.2. Description of Research Methods

Using linguistic-statistical, diachronic, and synchronic analysis methods, the correspondence of lexical units found in the written monuments to Kipchak-Oghuz language was studied, and the frequency of usage of indigenous (Turkic) and borrowed (Arabic, Persian) lexicon in ancient and modern literary texts was assessed.

Diachronic analysis provided insights into linguistic processes by examining the lexical layers of the written monuments from the Golden Horde, while synchronic analysis demonstrated the current stage of development and use of these languages for artistic purposes (Kazakh and Turkish).

To determine the semantic proximity between the lexical layers of Kazakh and Turkish, historical-comparative and lexico-semantic analysis methods were employed. Historical-comparative analysis was used to compare lexical layers and analyze semantic connections in the written monuments of the Golden Horde and in contemporary literary works. This analysis contributed to an understanding of ancient lexical layers, the nature of their interactions, and the active periods of lexical borrowing from various languages.

The lexico-semantic method involved an epigrammatic analysis of the structure of unambiguous and polysemous words, revealing the internal semantic structure of lexemes, along with elements of paradigmatic analysis focusing on external semantic relations, and syntagmatic analysis reflecting syntactic connections. Attention was given to the modifications of lexical meanings within one or more languages (Turkic, Arabic, Persian).

Through contextual analysis, the meanings of polysemous Arabic and Persian words, used in both literal and figurative senses, were examined. Metaphorical and figurative nominations related to the transformations of the core meanings of indigenous and borrowed lexicon were explored. Examples of contexts were provided to confirm the semantic connections between words in Turkic written monuments and literary works.

3. Results

During the XII–XV centuries in the territory of the Golden Horde, a written form of the Turkic literary language was developed based on the Old Turkic language and the local Kipchak dialect. This language influenced the creation of numerous poetic and prose works up until the 20th century^[24]. Among these unique literary works, which have been preserved to this day, are *Gulistan bi't-Turki* by Saif-i Sarai^[25] and *Khosrow and Shirin* by Qutb^[26].

3.1. Usage of Turkic Words in the Written Monuments of the Golden Horde

During this period, the core of the lexical layer in Turkic languages was considered to be Common Turkic and

Old Turkic words related to social, domestic, genealogical, and onomastic layers^[27]. These ancient Turkic words made up an average of 50–60% of the lexical corpus of medieval written monuments^[14]. Many of these words are still found in modern Turkic languages such as Turkish, Kazakh, Uzbek, and Kyrgyz.

Some Old Turkic words used in medieval written monuments (Yenisei-Orkhon, Uighur) have become archaic. This lexical layer can be divided into three groups: 1) archaic words that have undergone phonetic changes: “qanda” (where), “adaq” (foot), “egu” (virtuous), “bılık” (knowledge), “chechek” (flower); 2) archaic words that have undergone grammatical changes: “erse”, “erdi”, “aima” (to narrate); 3) obsolete words that are no longer used: “ajun” (world), “üküş” (many), “telim” (many), “künäş” (sun), “jimgen” (lawn)^[18]. Archaic lexemes that have fallen out of use today are more common in *Khosrow and Shirin* (18%) than in *Gulistan bi't-Turki* (9.6%)^[18].

Various dialectal forms of Oghuz, Kipchak, and Karluk languages are also used. Karluk words in *Gulistan bi't-Turki* are few: “andağ” (here), “belgür” (defining), “jümğaq” (parsley), “ian” (return), “asiğ” (profit), while Oghuz elements are more widespread: “qavra” (embracing), “iaryn” (tomorrow), “iüksäk” (high), “elik” (fifty)^[25]. In *Khosrow and Shirin*, Karluk elements are more numerous than in *Gulistan bi't-Turki*: “kurag” (collection), “sakysh” (thought), “sayu” (always), “sychan” (mouse), while Oghuz elements are also present: “karuncha” (ant), “surmak” (pursuit), “kapsamak” (fence), “yildirim” (lightning), “yanak”, “dudag”(lip), “elik” (fifty), “kapug” (door), “chunki” (so), ect.

According to scholars, the main phonetic differences between the Kipchak language and the Oghuz and Karluk languages are as follows:

- 1) The presence of 8–9 vowels, with the sounds *y, i, u, ı* being pronounced in various manners;
- 2) The absence of rounded vowels in affixes;
- 3) Weak labial harmony in word structure;
- 4) The absence of long vowels;
- 5) The use of new combinations of vowels and consonants, replacing the ancient Turkic combinations of *ğ* and *g*, such as *au, oy, iu, ey, iy, üy, iuu, uu*;
- 6) Frequent use of the voiceless consonants *q, k, t* at the beginning of words;
- 7) Interchangeable use of the consonants *y, j*, and *dj* at

the beginning of words;

- 8) Softening of the voiceless consonants *p*, *q*, *k* between two vowels^[13].

A key feature of the Oghuz language is the use of the sound *b* instead of *m* at the beginning of words, as in *baña* (instead of *maña*) and *ben* (instead of *men*). However, this characteristic of the Oghuz language does not appear in either of the two manuscripts, where only the pure Kipchak forms (*maña*, *men*) are used. For example:

“Shirin sözün maña kim aytur erse. Tanimdin sürgemin öz canım erse.”

(“If someone tells me Shirin’s words, I will drive out my soul from my body.”)

“Shirin sözün kim mağan aytar bolsa, Tanimnen öshiremin öz janım bolsa”^[16].

“Äsar qıldı maña gülniñ jamalı.”

(“The beauty of the flower affected me.”)

“Äser etti mağan gülniñ jamalı”^[15].

Another feature typical of the Oghuz group is the omission of the initial *b* sound in words, resulting in *oldı* instead of *boldı*. This Oghuz form is absent in both manuscripts, where the pure Kipchak form *boldı* is used consistently:

“Aññ teğ körki boldı elde mäshür. Yüzin körgen bu Yusuf teyü säqnr.”

(“His beauty became famous in the land. Those who saw his face would think of him as Yusuf.”)

“Körki elge qanday boldı mäshür. Yüzin körgen bul Yusuf dep oylayr” (Literary artifacts, 2012).

“Atası ayttı ey uğluñ, bu yol’da falak saña yär qıldı dağy dawlät дәlil boldı kim könilini tikendän tikkeni ayağın’dan çığardı.”

(“His father said: ‘O son, fate has helped you in this journey, and fortune was the cause that removed the thorn from your heart and from your foot.’)

“Äkesi ayttı: ‘Ey ulım, bul saparda sağan tağdır jәrdem berdi, baqyt sebep bolıp, könilin tikennen, tikendi ayağınnan şığardı”^[15].

The only phonetic feature of Oghuz languages reflected in these works is the parallel use of the sounds “*i*”, “*j*”, and “*dj*”. In *Gulistan bi’t-Turki*, only words beginning with “*i*” are found: “*Rauza ianlyğ hus bağynda ianğy nar erür*” (“Her face was a pomegranate in the garden of paradise beauty”)^[25]. In *Khosrow and Shirin*, there are 3011 words with “*i*” and 595 words with “*j*”: “*Bü kün ianymda bolsa erdi ol iar. Bü*

ianlyğ bolmas erdim ğaşyq ta zar” (“If she were with me today, I would not be so in love”), “*Janynda qalmady aram taqat. Täinde qalmady sabyr etgu quat*” (“There is no peace in the soul, no strength to wait in the body”)^[26].

Kipchak elements may be monosyllabic: “*öz*” (self), “*men*” (I), “*sen*” (you), “*sız*” (you), “*bü*” (this), “*köz*” (eye), “*baş*” (head), “*iüz*” (face), “*mal*” (property), “*qül*” (slave), “*üch*” (three); disyllabic: “*ağach*” (tree), “*dağy*” (again), “*qarsy*” (against), “*sary*” (yellow), “*qyzyl*” (red), “*beşik*” (cradle), “*iağmur*” (rain), “*büryn*” (before), “*temür*” (iron); trisyllabic: “*belgılı*” (determined), “*tırlık*” (life), “*tobraq*” (soil) (Kashgari, 2005). Specifically, there are many verbs that have undergone phonetic changes and are used in modern Kazakh: “*ait*” (say), “*ata*” (name), “*al*” (take), “*aş*” (open), “*bağla*” (bind), “*bas*” (press), “*başla*” (begin), “*kel*” (come), “*keltir*” (bring), “*ketür*” (go away), “*közle*” (see), “*körün*” (appear), “*küze*” (not touch) (Iskakov, 2011).

In the grammatical and lexical forms of the written monuments, both Kipchak and Oghuz and Karluk signs are used. The Oghuz suffix “*-myş (-miş)*”, indicating the past tense, is used 35 times in *Gulistan bi’t-Turki*, while the suffix “*-ğan (-gen)*” is used 45 times. In *Khosrow and Shirin*, this suffix is used 215 times, while the Kipchak variant “*-ğan (-gen)*” is used 128 times (parallel forms are sometimes used)^[25, 26].

Both works contain affixes such as “*-ar*”, “*-er*”, “*-r*”, “*-maq*”, “*-mek*”, “*-la*”, “*-le*”, “*-da*”, “*-de*”, “*-n*”, “*-in*”, “*-yn*”, as well as verbs formed with their help, proving their kinship with modern Kazakh. Based on the statistical data obtained from the research, it can be concluded that Kipchak elements predominate in the written monuments.

3.2. Analysis of the Use of Turkic and Borrowed Lexicon in the Monuments of the Golden Horde

In the XII–XVI centuries, spiritual and cultural ties were established between the Turkic, Arabic, and Persian peoples through Islam. Persian became the lingua franca, used as a means of interethnic communication in all spheres of everyday life for the Turks, while Arabic became the language of science, education, and religion in the Muslim world^[28].

In the Middle Ages, as a result of the economic, spiritual, and cultural ties between the Mamluk Empire and the

Golden Horde, bilingualism (Turkish and Arabic) and trilingualism (Turkish, Arabic, and Persian) emerged among the intellectuals of these regions. Among these intellectuals were the Turkic poets Seyf Sarai and Kutb, who are the focus of this study. Their works reveal that they were well-versed not only in the Turkic language but also in Arabic and Persian, and they freely employed these languages in their writings. Therefore, it is natural that a significant portion of the vocabulary in these works consists of borrowed elements. Secondly, as medieval Turkic poets often translated from Eastern languages, Arabic and Persian words were used in place of Turkic equivalents where such terms did not exist. Thirdly, they skillfully utilized Arabic and Persian words as stylistic devices in their compositions.

In 2023–2024, the authors of this study worked with the original manuscripts of Seyf Sarai’s *Gulistan bit-Turki* and Kutb’s *Khosrow and Shirin*. The texts of these written monuments were fully transcribed, and two extensive explanatory dictionaries were prepared: the explanatory dictionary of Seyf Sarai’s *Gulistan bit-Turki* and the explanatory dictionary of Kutb’s *Khosrow and Shirin* (currently in manuscript form). During the lexicographical transcription of the texts, each lexical unit was recorded in a register, meaning that every word was included in the dictionary. The origin of each lexical unit was determined, with markers identifying whether the word was of Turkic, Arabic, or Persian origin. Through this method, it was possible to determine which words belonged to the Turkic layer and which to the Arabic and Persian layers. Notably, a large number of Arabic and Persian words are found in the language of these works.

A linguistic statistical analysis of all words in the texts of the written monuments, excluding repetitions, revealed the following results (**Tables 1 and 2**).

Table 1. Analysis of the written monument “Gulistan bi’t-Turki” by Saif-i Sarai.

Turkic Words	Arabic Words	Persian Words
1282	1539	408

Source: Prepared by *Gulistan bit-Turki* Sayf-i-Sarai^[25].

Table 2. Analysis of the written monument “Khosrow and Shirin” by Kutb.

Turkic Words	Arabic Words	Persian Words
2380	852	569

Source: Prepared based on the results of *Khosrow and Shirin* by Kutb^[26].

Arabic words in *Gulistan bi’t-Turki* occur half as often as in *Khosrow and Shirin*. This difference can be attributed to the fact that *Gulistan bi’t-Turki* was written in Egypt, where Arabic was widely used among the local population. There are significantly fewer Persian borrowings, as the Persian lexicon is largely composed of Arabic loanwords^[25, 26].

In this study, the main semantic groups of Arabic and Persian words used most frequently in the monuments were identified:

- 1) Common words without semantic changes;
- 2) Lexemes with semantic changes;
- 3) Turkic words that have fallen out of use in modern Kazakh;
- 4) Compound Arabic and Persian words.

The first group includes Arabic and Persian lexical units whose meanings are widely understood throughout the Turkic community and which are used in relation to the religious, everyday, and socio-political life of the Golden Horde period. An analysis of the introductory units in the two written monuments reveals many such lexemes: aḥṣam (evening), bağ (garden), bülbül (nightingale), kırıp̄tar (Christians), gauhar (jewel), ġazal (gazelle or ode), kümān (doubt), kör (blind), hākım (wise man), qairat (courage), fārman (order), zeināt (decoration), farasat (wisdom), mamlakat (kingdom), ṣahar (city)^[17, 19].

The second group includes words with lexical transformations. Since Arabic and Persian languages are polysemous, not all meanings of words are used in one written monument due to the limitations of the genre. The same words may be used in different works with different meanings.

For example, the Persian word andiṣe (اندیشه) is used in Persian with the following meanings: 1) thought; 2) doubt, anxiety. Examples:

- 1) “Aitty bır zaman andiṣege baryb qoilar chuban qatynnda iaraq tūrūr qūranny men oqiğaiym” (“He thought for a while and said: ‘The sheep are far away with the pastor, so I will read the Qur’an’”^[25]).
- 2) “Bū andiṣe birlā könlü bolüb tar” (“His heart ached with this anxiety”)^[26].

In Persian, the Arabic word aiam (إيام) is used with its primary and secondary Persian meanings:

- 1) rest (primary);
 - 2) time, period (secondary). Examples:
- 1) “Zamanadan ṣekaiat qylyb ġabatanndan elin eli üstine

qoiüb aitür edi büný tale dun ua aiam bũ malauan durur” (“With a complaint against his fate, he put one hand on the other and said: ‘Why such a terrible fate, dreary days and nights?’”)^[25].

2) “Aitqannyñ menim halime muafeq tūrūr ııgtlık

aiamynda bır kım erse bılān mūhabbat bađlab edım” (“What you described is very close to my experience; in my youth, I also fell in love”)^[25].

The meanings of some polysemous words used in the monuments are presented in **Table 3**.

Table 3. Multivalued Arabic and Persian borrowings.

Transliteration of the Word in Latin	Original	Translation from Arabic	Translation from Persian
ali	ا	big above tall	
andam	اندام		whole body, silhouette flesh, body body structure
gustah	گستاخ		brave rude
ahlaq	اھلاق		symbol virtuous a set of humane traits

Source: Prepared by “*Gulistan bi-turki*” Sayf-i-Sarai^[25] and “*Khosrow and Shirin*” Qutb^[26].

The third group consists of Arabic and Persian lexical elements that are found only in written monuments and are not used in modern Kazakh. For instance, instead of Turkic words denoting cardinal directions, Arabic lexemes are used: janub (south), şamal (north), mağrib (west), maşraq (east), atraf (around), taraf (side). Similarly, medieval Arabic words are used instead of words for “poem”, “prose”, “poet”, and “language”: nazym (poem), nasyr (prose), lāfz (word), şaiyr (poet)^[19]. At the current stage of Kazakh lexical development, these words have become obsolete in journalism and literary texts and have been replaced by Kazakh words^[20].

Words not used in modern Kazakh but common to both written monuments include: auaz (melody, sound, voice), asir (to imprison), nasim (wind), farauan (numerous), falak (sky), ğişrat (celebration). Examples of lexemes found in one written monument but absent in the other include: baqıue (remainder), batyl (brave), bahyl (stingy), bādraqe (protector), ğurab (crow)^[25], barid (lamp player), başarat (good news), gilä (resentment), gülāb (rose water), ğalat (to err)^[26].

The fourth group includes words that represent the worldview of Eastern peoples through Arabic and Persian lexical-semantic fields in the monuments of the Golden Horde: 1) Arabic and Persian words reflecting the mystic-

ascetic thought in Islam—Sufism; 2) lexemes representing complex structures composed of Arabic and Persian words.

In *Gulistan bi-turki*, the following combinations are used: ab e haiat (water of life), ab e haiuany (water of eternity), adib e ğarib (foreign poet), huş habib u muntazir u uşşaq dılbari (beloved lover), aşaar e ajib (wonderful poems)^[25]. In *Khosrow and Shirin* by Qutb: bağyñ mivesını (fruit of your garden), gulchehra saqy (like a tulip), dilbar fyraqy (loss of a lover), auaz e halhal (various melodies), bağ e chimgan (green garden)^[26]. Context examples include: bādr e dāulāt (like a full moon): “Bādr e dāulāt alāmā qutlu iüziñ” (“The full moon of your face is a joy for the world”)^[25], bad e hazan (autumn wind): “Chechek kım üzmasān uaqtynda any esib eltür kezın bad e hazani” (“A flower not plucked in time will be blown away by the autumn wind”)^[26].

It is evident that such ornamental words and phrases, found in classical Eastern poetry, influenced Kazakh poetry as well. In the 11th–12th centuries, the Persian language played an active role in the cultural life of Turkic tribes living in northeastern Central Asia. Variants of Turkic poetry marked by Persian influence are found in ancient Kazakh poems and tales.

3.3. The Use of Turkic, Arabic, and Persian Elements in Modern Kazakh and Turkish Literature

Contemporary literature of Turkic peoples is united by a large layer of common Turkic lexicon and Arabic-Persian borrowings. Additionally, European influences (words of Greek, Italian, German, and English origin) and Eastern influences (Chinese and Japanese lexemes) are noteworthy.

Among the common Turkic lexicon used in the works under consideration, the following groups can be highlighted: names of household items, parts of the human body, times of day, plants, processes, and others. Arabic and Persian borrowings in the texts are generally represented by lexemes related to religious concepts and a group of abstract concepts.

In the poetry collection *Bağalım* (“My Dear”) by the Kazakh poetess Sarsenbaykyzy^[22], a whole layer of common Turkic lexicon is used. Since the Kazakh language is aggluti-

native, various forms of root words of Turkic origin are found: zaman (era)—zamandaşım (contemporary), kün (day)—kün-degen (daily)^[20]. Examples: “Zaman mynaü, alayaqtar keriskän, Aldındağı asyñdy aldap jep-işkän” (“These are the times when the swindlers are here; they deceived and ate your food right before your eyes”), “Biriñ—dosym, biriñ—jaqyn, zamadaşym, qırdaşym!!!” (“One of you is my friend, one of you is close, my contemporary, my peer!!!”)^[22].

It should be noted that Arabic borrowings have been fully adapted to the Kazakh language and, in general, to the Turkic agglutinative structure, while they can be used for word formation purposes with standard suffixes: mahabat (love)—mahabattıñ (possessive suffix —ıñ), as well as compound words for artistic effect: mahabattıñ isi (“the smell of love”)^[20]. Examples: “Mahabattıñ isi anqığan bar aimağı” (“A bar area filled with the smell of love”), “Mahabattıñ tım joğarı, bağası, ä?” (“The price of love is too high, right?”)^[22]. Statistical data on the use of lexemes is presented in **Table 4**.

Table 4. Statistical data on the use of Turkic words and loanwords in “Bagalym” (“My dear”) by E. Sarsenbaykyzy.

Turkic Origin	Quantity	Arab Origin	Quantity	Persian Origin	Quantity
«kun» (day)	115	«nur» (light)	22	«aspan» (sky)	7
«basyn» (head, to lead)	12	«mahabbat» (love)	17	«dari» (medicine)	2
«zaman» (time, epoch)	9	«sabyr» (patience)	11	«zor» (big)	4
«job» (way)	9	«ajyp» (fault)	5	«tamasha» (beautiful, wonderful)	2
«tur» (view, form)	1	«tylsym» (mysterious)	1	«dostyq» (friendship)	1

Borrowed words are often used as expressive or stylistic devices. For example, in the context, “Köshe tolı dorba üstağan qalıñ jurt, Azıq-tülik bağasy tur aspandaq...” (“The street is full of people with bags, food prices are skyrocketing...”)^[22], the Persian lexeme *aspandaq* (literally meaning “heavenly”) is used to hyperbolize and metaphorize the meaning of “high (price).” The lexeme *aspan* (sky) also transforms into the adjective “eternal”^[17, 19], which allows for metaphorization, as seen in: “Jüziñ jarıq bolyp, däyım aşıq bolsyn aspañ!” (“May your face be bright and your sky eternal!”).

In the novel “Şinto – Şinto’nun Yolu Doğanın Kalbine Gider” (“Shintoism—The Path of Shinto Leads to the Heart of Nature”) by A. Perker, there is a significant number of Old Turkic lexemes whose meanings transform depending on the suffixes used:

- “-lar (a-i-o-u)”, “-ler (e-i-ö-ü)” (plural suffixes): “Denizler var ya, okyanuslar...” (“There are seas,

oceans...”);

- “-daki”, “-deki”, “-taki”, “-teki” (the suffix “-ki” used with the locative case to denote the location of the defined word): “Öyle bir dünyada yaşıyoruz ki bir kuş denizdeki bir canlıyı yiyor ve pembeye boyanıyor” (“We live in a world where a bird eats a living creature from the sea and turns pink”);
- “-dan”, “-den”, “-tan”, “-ten” (suffixes used to mean “from whom”, “from what”, “due to what”, “because of whom”): “Japon mutfağına topraktan ve denizden gelmeyen ürünler böyle eklendi” (“This is how products not from land and sea were added to Japanese cuisine”)^[21, 23].

Perker’s work^[23] demonstrates a higher frequency of Persian words than Arabic ones, unlike the poems of Sarsenbaykyzy^[22]. These lexical layers constantly intertwine. For example, in the given context, the word *din* (religion) of Arabic origin and *günah* (sin) of Persian origin (both belong-

ing to the religious group) are used^[19]: "İnsan düşünmeden edemiyor, büyük dinler doğaya dua etmeyi, taş toprağa tapmayı günah gibi gördüklerinden acaba doğa ile olan bağımızı da zayıflatmış olabilirler mi?" ("One can't help but think, could they have weakened our connection with nature, as the great religions consider praying to nature and worshiping stones and soil as sins?")^[23].

The Arabic layer in "Şinto—Şinto'nun Yolu Doğanın Kalbine Gider" ("Shintoism—The Path of Shinto Leads to the Heart of Nature") is minimal, usually appearing only once. At the same time, it is worth noting the aphoristic na-

ture of the expressions: "Bir rahip ya da imam gibi aracı yok" ("There is no intermediary like a priest or imam"). The combination of different lexical layers makes the text stylistically richer and more interesting: "Belki o gün rüzgâr kamisine dua edecek. Kami tanrılardan sadece biri" ("Perhaps on that day the wind will pray to its kami. Kami is just one of the gods")^[23]. In this context, the Persian word rüzgâr (wind) and the Japanese word kami (kami) are introduced to anthropomorphize the wind and endow it with the traits of a Japanese deity. The frequency distribution of different lexical layers is presented in **Table 5**.

Table 5. Statistical data on the use of Turkic words and loanwords by Perker in "Şinto – Şinto'nun Yolu Doğanın Kalbine Gider" ("Shintoism—The path of Shintoism Leads to the Heart of Nature").

Turkic Origin	Quantity	Arab Origin	Quantity	Persian Origin	Quantity
«bir» (one)	213	«din» (religion)	7	«rüzgâr» (wind)	9
«deniz» (sea)	10	«sebep» (reason)	4	«günah» (sin)	2
«balık» (fish)	2	«hamur» (dough)	1	«pilav» (pilaf-food)	1
«ekin» (sowing)	2	«imam» (imam- religion's term)	1	«çorba» (soup)	1
«söz» (word)	1	«sabah» (morning)	1	«pembe» (pink)	1

Source: Prepared by Perker in "Şinto-Şinto'nun Yolu Doğanın Kalbine Gider"^[23].

The study of ancient written monuments of the Golden Horde and contemporary literary works has confirmed that Arabic and Persian borrowings constitute a significant lexical layer in the literary Turkic languages. Three main lexical layers correspond to different historical periods: the oldest is Persian, followed by Arabic, and then Old Turkic.

4. Discussion

In Turkic society, there is an idea of creating a unified literary language for all Turkic peoples. As U. Bingol notes, there are various movements today advocating for secularization and nationalization in the linguistic environment^[29]. The idea of a common Turkic language is intriguing in contemporary linguistic discourse but contradicts national views on language and cultural development in Turkic states.

Akalın's article^[30] examines the specifics of replacing borrowings with lexemes from Turkic dialects during Atatürk's language reform, which aimed to eliminate foreign elements (Arabic and Persian words) and popularize Turkish words. Research results confirmed that despite the eradication of Arabic and Persian words at different historical stages, around 70–80% of the modern Turkish lexicon consists of

borrowed words.

In the context of receptive multilingualism, Turkish and Azerbaijani are closely related as both belong to the Oghuz branch of Turkic languages. Sagin-Simsek^[31] focuses on the mutual understanding of speakers of these languages, revealing that despite linguistic similarity, the risk of misunderstanding, especially by Turkish speakers, is quite high. Thus, while languages can borrow lexical elements, each retains unique linguistic features at all levels, differentiating languages within the same group or family.

Pourhossein's research^[32] investigates the status of Persian in modern Turkish through the familiarity of Turkish university students with Persian loanwords. The study showed that students are poorly acquainted with this lexical layer. The research confirms that Persian and Arabic lexical layers in Turkic languages are perceived as native vocabulary by speakers, given their long usage and historical circumstances.

The phonological and morphological adaptation of Arabic borrowings in Turkic languages helps reveal historical processes of modification undergone by borrowed lexicon. W. F. Alshammari^[33] analyzed a corpus of 250 borrowings in Turkish, comparing them with their Arabic counterparts to identify phonological changes such as substitution, deletion,

vowel harmony, and epenthesis. The study compared Arabic, Persian, and Turkic loanwords by frequency of use, semantic proximity, and contextual function.

The study of borrowed lexemes is a significant linguistic contribution, as this lexical layer represents prehistoric cultural interfaces. Researcher Bjorn^[34] argues that successful identification of borrowed words can resolve long-standing debates about the function of different linguistic elements through comparative analysis. The study confirmed that comparing the lexical layers in Golden Horde written monuments demonstrated how Arabic, Persian, and Turkic lexemes interacted in the 13th–14th centuries.

Arabic borrowings in Turkic languages are well-known in linguistics. These lexemes underwent significant structural, grammatical, and phonological changes to integrate into Turkic languages. Rahimzade^[35] notes that Arabic-origin lexemes have a fusional structure and thus lose certain characteristics when transferred to agglutinative Turkic languages. A large number of lexical-semantic changes and transformations of Arabic words occurred during the Golden Horde period, making them today appear as native Turkic lexemes, indistinguishable from them phonologically, morphologically, or structurally.

Can's article^[36] presents a quantitative analysis of Turkish literature over a century, including a linguistic examination of 40 novels, and investigates gender differences in determining the most frequent words. Contextual analysis of Turkic, Arabic, and Persian lexicon in Turkish and Kazakh literature examined the frequency of lexical integration and the methods of embellishing and enriching the lexicon within literary genres.

As science and technological progress influence language transformations, continuous lexical updates reflect the world of science and new terminologies. Anggraeni^[37] studied the process of English lexical borrowing in Indonesian engineering articles related to informatics. Results showed that in one journal, 20% and 80% of the lexicon was borrowed, while in another, 35% and 65%. Statistical analysis of Golden Horde texts revealed that texts contained 30–35% Turkic, 50% Arabic, and 15–20% Persian words.

Nikeghbal's research^[38] explores Turkish-Persian language interaction in recipes and cookbooks. The analysis revealed only 148 common lexemes and evaluated etymological and historical factors affecting the Turkic language.

Results showed that Arabic lexicon is the most frequent, followed by Turkic and Persian lexicon.

The introduction of borrowed words enriches the recipient language's lexicon, but borrowed words must undergo certain transformation processes to fit the donor language's forms. Jdetawy^[39] identifies processes such as adaptation to the recipient language's sound structure and accommodation of phonological models. This work highlights borrowed lexicon that has fallen out of use in modern Turkic languages and that which remains in use.

Hoque's research^[40] illustrates four main types of borrowings: "with native language equivalents", "without native language equivalents", "with close equivalents", and "hybrid". The study focuses on obsolete lexemes no longer used in modern literary languages and the reasons for their archaism. It also addresses words that have survived in Turkic languages and undergone adaptation processes at various linguistic levels: phonetic, word-formative, morphological, and syntactic^[41, 42].

Given that Turkic languages have actively interacted with various South Asian languages, they have introduced many new lexical elements into Urdu, Punjabi, Pashto, Sindhi, and Balochi. Yucel's study^[43] provides tables of borrowed words and lexical elements, considering their frequency of use. Turkic lexicon may also be borrowed relative to other languages, indicating a high level of Turkic language involvement in forming the global linguistic fund.

Sylvester's study^[44] examines patterns of lexical retention, as well as semantic replacement, transformation, narrowing, and extension of meanings using English as an example. Despite the significant lexical-semantic, morphological, and phonological similarities between Turkish and Kazakh, they have experienced considerable linguistic influences from other languages in recent centuries, as demonstrated in literary works.

Comparing the results of this research with other studies leads to the conclusion that Arabic and Persian borrowings in modern Turkic languages are perceived as native Turkic lexicon, due to the adaptation of words at various linguistic levels. It is also noted that these lexical layers are interconnected not only linguistically but also historically, socio-culturally, and politically.

5. Practical Application

The practical application of the results lies in studying the relationship between different lexical layers and their frequency in ancient and modern contexts. This research can be used to enrich theoretical databases, address gaps in linguistic and literary studies, and contribute to diachronic, synchronic, and historical-comparative studies of Turkic languages. The findings can be applied in various humanitarian fields: semiotics, cultural studies, history, archaeology, conceptology, and literary studies.

6. Conclusions

Turkologists have expressed a variety of opinions regarding the language of Seyf Sarai's *Gulistan bit-Turki* and Kutb's *Khosrow and Shirin* manuscripts. Domestic scholars have refuted the claims that these works were written in "Old Uzbek" or "Tatar" languages, asserting that such views are unfounded. After analyzing the lexical composition of both manuscripts, scholars concluded that the language of these works belongs to the Kipchak-Oghuz language group. To support this conclusion, the entire vocabulary of the two manuscripts was collected and analyzed. Based on the objectives outlined in the article's introduction, the following conclusions were drawn:

First, using a linguostatistical method, it was determined that the text of *Gulistan bit-Turki*^[15] contains 1,151 Turkic words, 1,466 Arabic words, and 610 Persian words; *Khosrow and Shirin*^[16] contains 2,380 Turkic words, 852 Arabic words, and 569 Persian words.

Second, although elements of Karluk and Oghuz languages are found in the texts of both manuscripts, lexical units belonging to the Kipchak language group dominate in terms of frequency of use and form the core vocabulary. This supports scholars' conclusions that the primary language of both works is the Kipchak-Oghuz literary language. Furthermore, the frequent use of auxiliary verbs such as *bol*, *qil*, and *et*—as well as verbs derived through suffixes—demonstrates the Kipchak character of the language, highlighting its connection to the Kazakh language.

Third, the borrowed lexical layer—specifically the Arabic and Persian words—was semantically classified and analyzed in both manuscripts. It was found that many Arabic and Persian words in these texts are polysemous and that

their meanings often correspond to those in modern Kazakh. Additionally, some Arabic and Persian words that are no longer used today were also found in the manuscripts. The reasons for using such words include: (1) the absence of Turkic equivalents, (2) the use of synonyms despite the availability of Turkic alternatives, and (3) the use of these words to enhance stylistic sophistication and express nuanced meanings. In some contemporary Turkic languages, these words have become archaisms or have completely fallen out of use.

The Arabic and Persian words in these manuscripts reflect the artistic language tools and complex linguistic units characteristic of the intellectual and cultural world of the Eastern peoples during the Golden Horde period.

Fourth, modern texts were analyzed using semantic and contextual methods, focusing on key lexemes found in literary works. The role of metaphorical and figurative modifications of basic meanings was assessed, and the contexts in which the vocabulary interacts with other lexemes were examined.

It should be noted that the goals and objectives of this study were successfully achieved, as the key ancient layers of modern Turkic languages were examined, and crucial lexical processes from the Golden Horde period to the present day were traced. Additionally, both written monuments and modern literary texts were studied from a linguostatistical perspective.

As a novelty of this research, the statistical data obtained from the dictionaries compiled for Seyf Sarai's *Gulistan bit-Turki* and Kutb's *Khosrow and Shirin* manuscripts provide linguistic insights into the lexical layers of the Golden Horde period. These newly prepared dictionaries and semantic analyses will contribute to the inclusion of these works in the historical text database of the Kazakh National Corpus.

In this regard, the study of historical written sources, which reflect the social development, culture, and literary history of the past, is an essential step in shaping historical thought. The detailed study of the language of the written monuments from the Golden Horde period—such as Seyf Sarai's *Gulistan* and Kutb's *Khosrow and Shirin*—sheds light on the historical and cultural past of Kazakhstan and the neighboring countries and peoples of Central Asia within the chronological framework of the 13th–14th centuries.

Priority directions for future research include identifying semantic relationships and forming lexical-semantic

fields in closely related languages, studying the functionality, frequency, and stylistic features of different lexical layers, understanding the role of borrowings and their contribution to enriching the lexical layer, and automating text processing to identify the percentage ratio of native and borrowed vocabulary in literary works.

Author Contributions

A.S. (Ainur Seitbekova) considered lexical units in the text of Saif Sarai “*Gulistan bit-Turki*” on the basis of the method of contextual analysis and determined their semantic features. The inscription determined the stylistic function of Turkic and borrowed words found in the text of the monument. He analyzed the continuity with the modern Kazakh language. G.M. considered the lexical units in the text of Qutb “*Khosrow and Shirin*” on the basis of the method of contextual analysis and determined their semantic features. The inscription determined the stylistic function of Turkic and borrowed words found in the text of the monument. He analyzed the continuity with the modern Kazakh language. M.M. reviewed the research work of domestic, Turkish and English Turkic scientists regarding the language of the written monuments of the Golden Horde and conducted an analysis of scientific opinions. He wrote the scientific theoretical base of the article. A.S. (Assel Seidamat) collected ancient Turkic, Arabic and Persian words in the text of Saif Sarai “*Gulistan bit-Turki*”. The inscription published statistics of ancient Turkic, Arabic and Persian words in the text of the monument and determined the frequency. T.Z. collected ancient Turkic, Arabic and Persian words in the text of Qutb “*Khosrow and Shirin*”. The inscription published statistics of ancient Turkic, Arabic and Persian words in the text of the monument and determined the frequency.

Funding

The work was carried out within the framework of the grant project “Theory and practice of lexicography of Golden Horde era’s written artifacts language” (grant No. AP14869737) of the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

The theoretical foundation of this study comprises works by contemporary Kazakh, Turkish, British, and other scholars who focus on indigenous and borrowed lexicon within Turkic languages. These works explore several aspects: the frequency of lexeme usage, semantic proximity in closely related languages, assessment of the stylistic coloration of borrowings, examination of the lexical fund in percentage terms, and the processes of archaisation and adaptation of lexicon across various linguistic levels. To facilitate the semantic analysis of different lexical layers, the following dictionaries were utilized: A Comprehensive Persian-English Dictionary, Including Arabic Words and Phrases Occurring in Persian Literature by Steingass (1892), *Dīwān Lughāt al-Turk* (Dictionary of Turkic Dialects) by Kashgari (2005), Arabic-Kazakh Dictionary by Ondasynov (2010), Dictionary of the Kazakh Literary Language by Isakov (2011), and the Ottoman Turkish Dictionary (2002–2023). For the contextual analysis of lexical units, the following written monuments from the Golden Horde period were selected: *Gulistan bi’t-Turki* by Saif-i Sarai (1968) and *Khosrow and Shirin* by Qutb (1958), as well as contemporary literary works: Sarsenbaykyzy’s *Baghalym*^[22] and Perker’s *Şinto—Şinto’nun Yolu Doğanın Kalbine Gider* (Shintoism—The Path of Shinto Leads to the Heart of Nature)^[23].

Acknowledgments

The authors express their gratitude to the Ministry of Science and Higher Education of the Republic of Kazakhstan for financial support.

Conflict of Interest

The authors stated that there are no conflicts of interest.

References

- [1] Nadzhip, E.N., 1975. The Turkic-speaking monument of the XIV century “Gulistan” of the Sarai Safe and its language: In 2 parts. Nauka: Alma-Ata, Kazakhstan. pp. 15–27.
- [2] Nadzhip, E.N., 1979. Historical and comparative dictionary of Turkic languages based on the material of Khosrov and Shirin by Qutb. XIV century. The Main Editorial Office of Oriental Literature: Moscow, Russia. pp. 18–27.
- [3] Fazylov, E.I., 1971. Khorezm monuments of the XV century. The old Uzbek language. Fan: Tashkent, Uzbekistan. pp. 115–132.
- [4] Usmanov, H., 1966. Turkic “Gulistan” translations, nazys, poems. Kazan Utlar, 12, 131–138.
- [5] Koprülü, F., 1926. History of Turkish literature. Istanbul: Publisher Location, Country. Istanbul, Turkey. pp. 18–34.
- [6] Battal-taymas, A., 1955. Seyf Sarayı'nın Gülistan Tercümesini gözden geçiriş [Review of the Gulistan Translation of the Seif Palace]. Türk Dili Araştırmaları Yıllığı. Ankara. 3, 73–98. [in Turkish].
- [7] Ibatov, A., 1974. Dictionary of Qutb's poem “Khusrau U Shyryn”: XIV century. Science: Almaty, Kazakhstan. pp. 151–157. (in Kazakh).
- [8] Samoilovich, A.N., 1928. On the history of the Central Asian Turkish language. Publishing House of the Ussr Academy of Sciences: Leningrad, Soviet Union. pp. 15–36.
- [9] Baskakov, N.A., 1960. Turkic languages. Publishing House of the East: Moscow, Russia. pp. 8–13.
- [10] Zajackowski, A., 1964. Language monument from the Golden Horde “Husrev-u-Slrln” Qutba. Oriental Yearbook. 1(4), 36–45.
- [11] Shcherbak, A.M., 1961. Grammatical outline of the language of Turkic texts of the X–XIII centuries from East Turkestan. Nauka: Moscow, Russia. pp. 18–21.
- [12] Karamanlîğlü, A.F., 2019. Seyf-i Saraî Cülistan tercümesi: (Kitab Cülistan bi't-Türki) [Saif-i Sarai Culistan translation: (Kitab Culistan bi't-Turki)]. Türk Dil Kürümü: Ankara, Turkey. pp. 68–69. [in Turkish].
- [13] Sabyr, M., 2022. Linguistic representation of the world in the language of the heritage of the Golden Horde. Tiltanym, 2(86), 77–78. Available from: <https://www.tiltanym.kz/jour/article/view/145/119>
- [14] Sagyndykuly, B., 2006. The written monuments of the Golden Horde (XIV century.) comparative vocabulary. Kazakh university: Almaty, Kazakhstan. pp. 55–59.
- [15] Usmanov, H., Makhstudova, Z., 1980. Safe Es-Sarai. Kitabe Gulistan bit-Turks [The Secure Palace. Epitaph Gulistan bit-Turks]. Saifa Sarai (2 Parts). KSU: Kazan, Russia. pp. 47–56.
- [16] Qorabay, A. (Ed.), 2012. Literary artifacts, Volume 16. Almaty: Times, 480. Available from: <https://adebipor.tal.kz/web/viewer.php?file=/storage/upload/iblock/6b0/6b0cc32618d88527752cedc2f2674028.pdf&ln=kz>
- [17] Steingass, F.J., 1892. A Comprehensive Persian-English dictionary, including the Arabic words and phrases to be met with in Persian literature. Routledge & K. Paul: London, UK. pp. 65–93. Available from: <https://dsal.uchicago.edu/dictionaries/steingass/>
- [18] Kashgari, M., 2005. Dictionary of Turkic dialects. Dike-Press: Almaty, Kazakhstan. pp. 15–30.
- [19] Ondasynov, N.D., 2010. Arabic-Kazakh explanatory dictionary. Bilim: Almaty, Kazakhstan. pp. page 14–23. (in Kazakh). Available from: <https://sozdikqor.kz/sozdik/?id=205>
- [20] Iskakov, A., Uali, N., 2011. Dictionary of the Kazakh literary language. Available from: <https://tilmedia.kz/js/pdfjs/web/viewer.html?file=/storage/2018/04/11/5acdb14e26c71.pdf>
- [21] Ottoman Turkish dictionary, 2002–2023. Available from: https://www.lexilogos.com/english/ottoman_turkish_dictionary.htm
- [22] Sarsenbaykyzy, E., 2021. Bagalym [Precious]. Zazhusy: Almaty, Kazakhstan. pp. 19-25. (in Kazakh). Available from: <https://www.google.com.ua/books/edition/%D0%91%D0%B0%D2%93%D0%B0%D0%BB%D1%8B%D0%BC/gGoWEAAQBAJ?hl=uk&gbpv=1>
- [23] Perker, A., 2021. The Shinto-Shinto Nun path leads to the heart of nature. Destek Yayinlari: Istanbul, Turkey. pp. 18–25.
- [24] Wilson, J.R., 2022. The role of Nogai in the Golden Horde: A reassessment. Oriental Journal of Hung. 75(4), 609–637. Available from: <https://akjournals.com/view/journals/062/75/4/article-p609.xml>
- [25] Sarai, S., 1968. Poetry. Guliston. Ghafur Ghulam Literary Publishing House. Tashkent, Uzbekistan. pp. 45-54. (in Uzbek). Available from: https://n.ziyouz.com/books/uzbek_mumtoz_adabiyoti/Sayfi%20Saroyi.%20She'rlar.%20Guliston.pdf
- [26] Qutb, 1958. The oldest Turkish version is Husräv u Sirin Qutba (translated by A. Zajackowski). Polska Akademia Nauk: Warsaw, Poland. pp. 15-28.
- [27] Oksal, C., Oguz, H.N., Catal, M., et al., 2022. Time travel in Turkish: WordNets for modern Turkish. In the proceedings of the Globalex Workshop on Linked Lexicography; 20 June 2022; Marseille, France. pp. 75–84. Available from: <https://aclanthology.org/2022.gwll-1.11.pdf>
- [28] Tat, D., 2022. Lexical borrowing targets spans. Languages. 7(4), 289. Available from: <https://www.mdpi.com/2226-471X/7/4/289/htm>
- [29] Bingol, Y., 2009. Language, identity and politics in Turkey: Nationalist discourse on creating a common Turkic language. Alternatives: Turkish Journal of International Relations. 8(2), 40–52. Available from: https://ciaotest.cc.columbia.edu/journals/tjir/v8i2/f_

- 0019739_16816.pdf
- [30] Akalın, S.H., Kumanlı, M.S., 2021. Turkic borrowings in the Turkish language reform: Past and today. *Bilig.* 98, 165–180. Available from: <https://avesis.hacettepe.edu.tr/yayin/b5fcef61-c62c-494b-bada-6b53cc371ec9/turkic-borrowings-in-the-turkish-language-reform-past-and-today>
- [31] Sagın-Simsek, C., Konig, W., 2012. Receptive multilingualism and language understanding: Intelligibility of Azerbaijani to Turkish speakers. *International Journal of Bilingualism.* 16(3), 315–331.
- [32] Pourhossein, S., 2012. Present status of Persian language in modern Turkish: The case of personal names. *International Journal of Linguistics.* 4(4), 157–170. Available from: https://www.researchgate.net/publication/235780094_Present_Status_of_Persian_Language_in_Modern_Turkish_the_Case_of_Personal_Names
- [33] Alshammari, W.F., Alshammari, A.R., 2020. Adaptation of Turkish loanwords originating from Arabic. *International Journal of English Linguistics.* 10(5), 388–398. Available from: <https://pdfs.semanticscholar.org/7828/d523094ebdae42e8fe85815795b41625064b.pdf>
- [34] Bjorn, R.G., 2022. Indo-European loanwords and exchange in Bronze Age Central and East Asia. *Evolutionary Human Sciences.* 4, e23. Available from: <https://www.cambridge.org/core/journals/evolutionary-human-sciences/article/indoeuropean-loanwords-and-exchange-in-bronze-age-central-and-east-asia/91132871EAF737CD98BDFC43C16C761A>
- [35] Rahimzade, D., 2021. Semantic adaptation of Arabic loanwords in Turkic languages (based on Oghuz languages). *Journal for Educators, Teachers and Trainers.* 12(4), 30–37. Available from: <https://digibug.ugr.es/handle/10481/72024>
- [36] Can, F., Patton, J.M., 2010. Change of word characteristics in 20th-Century Turkish literature: A statistical analysis. *Journal of Quantitative Linguistics.* 17(3), 167–190; Available from: https://www.researchgate.net/publication/220469204_Change_of_Word_Characteristics_in_20th-Century_Turkish_Literature_A_Statistical_Analysis
- [37] Anggraeni, D., 2021. English borrowing words in Indonesian informatics engineering articles: Types and meanings. *Deiksis.* 13(2), 134–144. Available from: <https://journal.lppmunindra.ac.id/index.php/Deiksis/article/view/7629/3982>
- [38] Nikeghbal, N., Yunc, R.H., 2022. Reflection of Turkish-Persian linguistic interaction on Turkish cuisine. *Journal of Tourism and Gastronomy Studies.* 10(3), 1908–1923. Available from: https://www.researchgate.net/publication/364116180_Reflection_of_Turkish-Persian_Linguistic_Interaction_on_Turkish_Cuisine
- [39] Jdetawy, L.F., Hamzah, M.H., 2020. Loanwords in the taxonomy of borrowing: A sociolinguistic analysis. *Language in India.* 20(12), 98–135. Available from: https://www.researchgate.net/publication/362903398_Loanwords_in_the_Taxonomy_of_Borrowing_A_Sociolinguistic_Analysis
- [40] Hoque, M.A., Ali, M.M., Puteh-Behak, F., et al., 2021. Lexical borrowings from the English language into Bangla short stories. *Journal of Language and Linguistic Studies.* 17(1), 158–172. Available from: <https://files.eric.ed.gov/fulltext/EJ1294604.pdf>
- [41] Khassenov, B.R., Adilova, A.S., Tuite, E.E., et al., 2022. A new approach in linguistics. *Bulletin of Karaganda University. Philology Series.* 106(2), 34–42.
- [42] Khassenov, B.R., 2024. Mytholinguistic interpretation of the 12+1 model. *Bulletin of Karaganda University. Philology Series.* 113(1), 6–15.
- [43] Yucel, D., 2021. Turkish/Mongolian lexical borrowings in South Asian languages. *Journal of Research in Social Sciences and Language.* 1(2), 105–127. Available from: <https://d-nb.info/1247366049/34>
- [44] Sylvester, L., Tiddeman, M., Ingham, R., 2022. Lexical borrowing in the Middle English period: A multi-domain analysis of semantic outcomes. *Information English Language & Linguistics.* 26(2), 237–261. Available from: <https://www.cambridge.org/core/journals/english-language-and-linguistics/article/abs/lexical-borrowing-in-the-middle-english-period-a-multidomain-analysis-of-semantic-outcomes/B83A68089915746EC1B613C20761E985>