Exploration of Loan Words in the Kipchak Translation of Gülistān

Ayiguli Abuliemiti

Faculty of Chinese Language and Literature, Northwest Minzu University, Lanzhou, Gansu, China

Abstract: The Kazakh translation of the Gülistān represents a rare documentary witness to living Kipchak vernacular speech. Alongside indigenous Turkic vocabulary, the text preserves stratified layers of loanwords and neologisms constructed from these borrowings; a single lemma may surface in multiple morphological guises, each bearing subtle semantic nuances that attest to successive stages in the diachronic biography of the Kipchak lexion. Employing the historical-comparative method and intergrating both synchronic and diachronic perspectives, this study examines the 1391 Kipchak translation of the Gülistān to systematically analyze the structural patterns, morphological characteristics, and semantic evolutionary trajectories of foreign loanwords within its vocabulray. Within this lexical system, Arabic-Persion borrowings numerically predominate over native Kipchak stock. A salient feature this textual language is the prolofic derivation of neologisms through the affixation of both borrowed and indigenous morphemes. These lexical elements survive in various forms in modern Kazakh; certain borrowings have become fully naturalized components of the Kazakh lexion, having undergone diverse adaptive modifications through prolonged language contact and historical development phonologically via epenthesis, deletion, and systemic alternation; semantically through extension, narrowing, or categorical shift. Furthermore, specific loanwords have shed their autonomous semantic functions, now requiring recourse to auxiliary lexical strategies to fulfill particular grammatical or expressive roles.

Keywords: Gülistān; Kipchak Translation; Loan Words; Structural Characteristics; Semantic Features

1. Introduction

Gülistān also translated as The Garden of Truth or The Rose of Garden, is a masterpiece by the medieval Persion poet Sa'dī [1,2]. Composed in a dinstinctive folk idiom based on materials accumulated during the author's extensive travels, it enjoys worldwide renown as a classic of world literature. The manscript under examination was trasnlated in 1391, comprising 372 folios with thirteen lines per page, and is currently held at Leiden University Library. According to scholars abroad, three Turkic translations of Gülistān have been identified to date: a Kipchak translation, a Chaghatay translation, and an Oghuz translation [3], on the identification and classification of these versions, see Sayf Saray Gülistan Tercümesi, preserved respectively at Leiden University Library, the Süleymaniye Liberary in Turkey, and the Faculty of Letters, Istanbul University.

Gülistān was compused from materials accumulated throughout Sa'dī's lifetime, presenting a rich tapestry of content that crystallizes the author's lived experiences and sisdom. Its advocacy of thruth, goodness, and beauty, together with its profound concern for the suffering of common people, prominently articulates the aspirations and demands of the masses [4-6]. Consequently, the Kipchak translation achieved circulation during its time and earned profound esteem among local populations.

The Kipchak versions of Gülistān exhibits a significant predominance loanwords over native vocabulary, with the intergration of loanwords with native suffixes to form neologisms constituting one of its distinctive features. Kazakh likewise contains numerous cognate loanwords from Arabic and Persion, yet the chronology of their incorporation into Turkic-whether contemporaneous with the Kipchak stratum, or earlier or later and their semantics trajectories remain to be elucidated. Analysis of these phenomena will, on the one hand, illuminate the historical development of loanwords in both the document language and

Kazakh; on the other hand, it will contribute to clarifying the status and functional roles of loanwords across different diachronic stages of Turkic varieties.

According to our statistics, the literary language contains 1989 lexical items, of which 1108 are native words and 1881 are loans. In terms of origin, the borrowings are drawn primarily from Arabic and Persion, with a smaller stratum of Sogdian, Greek, Chinese, Mongolian, Sanskrit, and Khotanese items. Numerically, Arabic loans predominate, followed by Persian. Arabic terms cover kinship and social address, administrative and professional titles, plant names, color terms, animal names and associated vocabulary, food and clothing words and thier derivatives, human body parts and organs, and items relating to celestial bodies and natural phenomena. Persion loans, for their part, include plant and animal names, military terminology, body-part and organ terms, and everyday-life vocabulary.

We can observe this through the origins of loanwords in the literature. The majority of borrowings derive from Arabic and Persian, with a smaller number originating from other languages. The example sentences in the Table1 below serve as the best evidence. Citations follow the format 'page/line', wherein the numeral preceding the slash denotes the page number and that following it denotes the line number. It should be explicitly noted that the semantic definitions provided for each lexical item in the Table 1 are contextually anchored that is, they reflect the specific meaning the word carries within the document rather than its general or dictionary definition.

Table 1. Selected Loanwords in the GÜListĀN: Etymology, Word Class, and Meanings

	Page/Line	Etymology	Word Class	Meaning
alfāẓ	13-12	Ar.	n.	words, speech
alqïṣṣa	28-11	Ar.	adv.	finally, at last
anwā'	11-9	Ar.	adj.	variously
arzāq	309-6	Ar.	n.	food, provisions
aṭlas	5-11	Ar.	n.	damask, satin
āgāh	54-1	Far.	adj.	known, cognizant
ba'ẓī	70-6	Ar.	pron.	some, certain
bahāna	126-4	Far.	n.	excuse, pretext
baḫš	85-3	Far.	n.	portain
barbaṭ	126-5	Far.	n.	lyre
baṭn	304-13	Ar.	n.	stomach, belly
bezzāz	150-6	Ar.	n.	cotton merchant
birke	103-12	Ar.	n.	pond, pool
češme	278-8	Far.	n.	spring water
dāye	5-12	Far.	n.	wet nurse, nanny
čust	341-3	Far.	n.	beautiful, pretty
farāsat	23-3	Ar.	n.	insight, foresight
farrāš	94-6	Ar.	n.	servant, attendant
elḫān	220-13	Ar.	n.	sound, voice
zāġ	239-1	Far.	n.	crow
za'ferān	247-4	Ar.	n.	saffron
'unnāb	308-5	Ar.	n.	maroon, date red
tāza	12-7	Far.	adj.	fersh, novel
ṭawīle	20-7	Ar.	n.	stable
tarjuma	355-5	Ar.	n.	translation
śu'bān	242-6	Ar.	n.	dragon, loong
ṣaḥn	251-12	Ar.	n.	courtuard, venue
šängül	225-6	Far.	adj.	glad, joyful

2. Structrual Characteristics

This section examines the structural characteristics of loanwords, with particular attention to their morphological intergration and syntactic distribution. A subset of loanwords attested in the document exhibits phonological develpments in Kazakh characterized by epenthesis, as exemplified by farq (Ar.) > parïq (Kz) 'difference'[4]; mīrāś (Ar.) > mïyras(Kz)

'heritage'; mūsīqī (Ar.) > muwziyka (Kz) 'music'; qadr (Ar.) > qadïr(Kz) 'dignity'; rāžī (Ar.) > riyza (Kz) 'satisfaction'; deletion, illustrated by quwwat (Ar.) > quwat (Kz) 'strength', dost (Far.) [7] > dos (Kz) 'friend'; and consonant alternation, as in qïṣṣa (Ar.) 'story, tale' > qïysa (Kz) 'epic, poem', respectively.

A number of loanwords attested in the Gülistān manuscript have been retained in Kazakh through derivational and compounding processes. Among derivationally formed words, those suffixed with -lï, -sïz predominate and exhibit high frequency of usage. Within compounds, those formed through combination with auxiliary verbs are particulary numerous. Additionally, certain loanwords persist in Kazakh in binomial constructions, paired with native elements.

jarāḥat (Ar.) +-lï > jarāḥatlï / jaralï (Kz) 'injured'

bahā (Far.) + -lï > bahālï / baġalï (Kz) 'valuable'

fāyïda (Ar.) + -lï > fāyïdalï / paydalï (Kz) 'benefical'

ṭāqat (Ar.) + -sïz > ṭāqatsïz / taġatsïz (Kz) 'impatient'

gedāy (Far.) + -lik > gedāylik / kedeylik (Kz) 'poor'

zārī (Far.) + -lïq > zārīlïq / zarlaw (Kz) 'weeping, wailing'.

Certain lesemes in the manuscript exhibit suffixation with native Turkic morphemes, as exemplified by ḥurmet (Ar.) + -siz + -lik > ḥurmetsizlik / qurmetsizdik (Kz) 'disrespect'; jomard (Far.) + -lïq > jomardlïq / jomarttïq (Kz) 'generosity'; ẕärrä (GT) + -čä > ẕärräčä / zere-dey (Kz) 'a tiny bit'; qaṣṣāb (Ar.) + -čï > qaṣṣābčï / qasapčï (Kz) 'buther'.

Compounde words formed throgh compounding in Kazakh fall into three categories. The first type involves the compounding of loanwords with native Turkic vocabulary, such as Persion-Turkic combinations or Arabic-Turkic combinations: ḥurmā aġačï 'date palm', ḫoš iyi 'fragrance', 'aqrab ini 'scorpion nest', tāzī aṭ 'Arabian horse', teŋiz mawjï 'sea wave', dawūdī naġme 'brisk melody / lively tune', 'ahl-e makān 'owner / master', 'aqd nekāḥ 'marriage contract', jädd wa jehd 'effort / hard work', sar-gerdān 'be at a loss', i'tidālï mizāj 'health / normal state', qāyïm maqām 'heir / successor'. The second type consists of loanwords combined with Turkic Suffixes. For example, nawḥa qïl- 'to cry one's heart out', 'ilāj qïl- 'treat / cure / heal', ġāyib bol- 'disappear', paydā bol- 'appear', safar et- 'travel / journey', jawr kör- 'suffer', jawāb ayt- 'answer', ḫejālat keltür-

'lose face', ġam ye- 'worry / be anxious', sākit bol- 'stop / halt'. A notable characteristic of this type is that the same word is followed by different native Turkic auxiliary verbs in the document language and Kazakh. The use of these compound forms demonstratesm on the one hand, that they are ancient formations that have persisted to the present day with high frequency in modern usage. The variation in auxiliary verbs between the two periods is closely related to lexical semantics. As meanings shift, the attached auxiliary verbs also change, as in tamāšā qïl- (GT) / tamaša -la- (Kz) 'to view, appreciate'; ḫabar qïl-/ḫabar ber- (GT) / ḫabar -la- (Kz) 'to inform, to convey news'; 'afw et- (GT) / 'apïw et- (Kz) 'to forgive'; naẕar qïl- (GT) / nazar sal- (GT) 'to pay ettention to'. The third type involves the compouding of loanword with loanword. Certain words in the document do not occur independently in Kazakh but appear only in binomial constructions. This indicates that these loanwords have lost their original function and now require support from other words to realize thier meanings, such as Persian-Persian combinations:šād šādmān > šat-šaẟïman 'joyful, delighted'; and Arabic-Arabic combinations: 'ömr bāqī > ömir baqïy 'forever, lifelong'. In the latter combination, the loanword 'ömr is of high frequency, whereas bāqī does not occur independently; this word is generally used in this fixed combination in Kazakh. The high frequency of this combination in Kazakh is largely attributable to the role of 'ömr. Māl denotes 'livestock, property, goods' in both the document language and Kazakh with essentially identical meaning. Mülk signifies 'property, state, territory' in the document language, but in Kazakh its meaning has narrowed to 'property' only. Māl and mülik frequently appear as a binomial pair meaning 'property' (māl-mülik) in Kazakh [8].

3. Semantic Features

Although the evolution of vocabulary in any language of the world can be summarized under three headings the admission of neologisms, the elimination of obsolescent forms, and the expansion or contraction of meaning it is semantic change that function as the true pivot. By assigning new contect to existing forms, it steadily enriches and refines the expressive network; in so doing it may also precipitate the birth or death of lexical entries and alter their

internal relations, thereby forcing the entire system to re-equilibrate. An examination of the loanwords preserved in the literary language shows that a considerable number of them have survived in Kazakh, and having undergone precisely such semantic adjustment, have become firmly naturalized and now rank among the core items of the lexicon.

Analysis of loanwords in the document language reveals that a portion have been retained in Kazakh and constitute significant members of its lexicon. These include Persion loans such as zīrek > zerek (Kz) 'clever, witty', hönär > öner (Kz) 'skill, technique, craft', ziyān > ziyan (Kz) 'harm, hazard / loss, damdage', and ḫurmā > qurma (Kz) 'date palm'; a Greek loan nāmūs > namïs (Kz) 'self-respect, honor'; a Chinese loan tūmān > tümen (Kz) 'ten thousand'; a Mongolian loan nöker > nöker (Kz) 'guard, attendant'; a Khotanese loan ton > ton (Kz) 'robe / long gown, clothing'; and Arabic loans such as ṣandūq > sandïq (Kz) 'box'; šekker > šeker (Kz) 'lump sugar', ṭāwūs > tawïs (Kz) 'peacock', and ya'nī > yaġnïy (Kz) 'that is, namely', qalam > qalam (Kz) 'pen', sïr > sïr (Kz) 'secret', ṣuḥbat > suḥbat (Kz) 'conversation / talk', kitāb > kitap (Kz) 'book', mu'allïm > muġalïm (Kz) 'teacher', waqt > waqït (Kz) 'time', ṭa'ām > taġam (Kz) 'food / foodstuff', ḫabar > ḫabar 'message / tidings', zamān > zaman (Kz) 'era / age / time', leẕẕat > läzzat (Kz) 'enjoyment / pleasure / delight'. While loanwords entered the Kazakh language from literature, they also underwent corresponding sematic changes, such as semantic narrowing and semantic shift.

3.1 Semantic Narrowing

The loan word ṣayyād 'hunter, fisherman' from the written sources is preserved in Kazakh as ṣayyād-čï, meaning 'hunter who catches gamen with birds of prey', no longer denoting 'fisherman'. Its semantic scope has narrowed considerably compared to the original. Although the term remains in Kazak, its semantic range fails to match that of the native Turkic word aŋšï [9].

The word pahlawān originally meant 'strong, wrestler', but in Kazakh it appears only as palwan 'wrestler'. The term 'imārat carried two distinct meanings in the written sources: 'bulding, mansion' and 'property, flourishing', whereas Kazakh preserves only the first sense 'bulding, mansion'. The word 'illat encompassed

multiple meanings including 'defect, fault', 'ilness, disease', and 'pretext, excuse', yet Kazak retains only the first. The term däwlät possessed meanings such as 'state, country', 'wealth', and 'happines', with the sense of 'wealth' still surviving in Kazak. The word ġaybat signified 'backbiting, absence' in the source language, corresponding to 'gossip, rumor' in Kazakh. The term ḥakīm denoted both 'physician' and 'learned person', but Kazakh preserves only the latter meanings as ḥakim 'learned person'. Finally, hewā expressed 'wish', 'air', and 'climate' in the literary language, yet survives in Kazak solely as awa 'air'.

3.2 Semantic Shift

Semantic change occurs alongside social development. Although this process unfolds gradually, a comparative analysis of the literary language and contemporary Kazakh reveals that numerous loanwords, despite initially preserving their full semantic range from the source language, have undergone various semantic transformations through prolongged contact with the Kazakh linguistic system and historical evolution. Notably, certain terms have develpoed meanings entirely distinct from their original literary sensesm representing cases of complete semantic shift. Without consulting etymological dictionaries to trace their origins and original significations, such semantic developments would remain largely opaque to modern speakers. For instance, ẕālim meant 'cruel, severe' in the literary sources but signifies 'cunning, sly' (ẕālim > zalïm) in Kazakh. Ṣaḥrā denoted 'desert, wilderness' in the original texts, yet Kazakh has lost the meaning 'desert' while acquiring the new sense 'grassland, steppe' (ṣaḥrā > saḥara); maṣlaḥat conveyed 'good strategy, affair, matter' in the sources, whereas in Kazakh it means 'consultation, discussion' (maṣlaḥat > mäsliyḥat); and ġayr signified 'belonging to others' in the literary language, but has shifted to 'other, different, special' in modern usage (ġayr > ġayry).

Another notable feature of loanwords in the literary sources is the use of both singular and plural forms. However, since Kazakh lacks grammatical number distinctions, these borrowings were preserved either in their singular or plural forms, with only a handful retainging both. Fro example, wäzīr (singular) / wüzirā (plural) appears as wäzir in Kazakh, preserving only the singular form, with the

meaning 'minister, vizier'; faqīr (singular) / fuqarā (plural) appears as fuqarā in Kazakh, preserving only the plulral form, with the meaning 'the poor, the needy'; 'ajāyip (singular) / 'ajīb (plural) appears as 'ajīb in Kazakh, preserving only the plulral form, with the meaning 'novel, strange, curious'; ḥakīm (singular) / ḥukamā (plural) appears as ḥakim in Kazakh, preserving only the singular form, with the meaning 'learned person, scholar'; ḥaywān (singular) 'animal, livestock, life' / ḥaywan (Kazakh) 'birds and beasts, animals'. Both the literary language and Kazkah also preserve its plural form ḥaywānāt. In Kzakh, both singular and plural forms are used concurrently with the meaning 'birds and beasts, animals' [10].

4. Conclusion

The lexical stratification of the Gülistān exhibits a markes exogenous profile: Arabic-Persion loans now outnumber native Kipchak items and constitute the primary stock of the literary language. The author habitually recombines borroued stems with original Turkic affixes, thereby coining new senses in real time and, through high-frequency reiteration, rapidly conventionalising them. This productive mechanism is both a signature strategy of the text and a replicable template for later Kazakh.

After entering Kazakh, these derived units undergo differential phonological reduction: initial glottal or pharyngeal fricatives are zeroed shen stress advances, but are sporadically retained in unstressed syllables, yielding an optional 'preserve-delete' continuum that mirrors adaptive phonotactic constraints.

Semantically, the source text is characterised by polysemy; yet in the Kazakh reception layer roughly 80% of loans retain only one core meaning, 7% keep two senses, and the remaining polysemic configurations are largely truncated. Such semantic contraction is not random: meanings salient to nomadic economy, kinship organisation and daily ritual are preferentially preserved, indicating that lexical filtering is governed by cultural ecology.

At the pragmatic level, a subset of loans has broken free of its literary context and joined the high-frequency stratum of moderen Kazakh, becoming indispensable systemic elements. Their adaptation trajectories display multidimensional adjustments: phonetically, prosthetic vowels, consonant deletion and alternations coexist; semantically, the usual typological pathways - expansion (mesk 'musk' → 'fragrance in general'), narrowing (ṣayyād 'hunter, fisherman' → fisherman) and shift (ṣaḥrā 'desert' → 'fragrance') are all attested; syntactically, some items have lost independent predicative force and must enlist auxiliaries or postpositions to realise clausal value, hinting at a drift from lexical to grammatical status. These layered data provide manipulable micro-evidence for tracing Turkic lexical history, gauging the depth of language contact and reconstructing cultural miggration routes.

References

[1] Bodrogligeti, András J.E., A Grammar of Chagatay, München:LINCOM EUROPA, 2001.

[2] Duysenäli Äbdiläšim, A Study of Written Kazakh "Document" Language (Beijing: Zhongyang Minzu Chubanshe, 2015), 17.

[3] Ali Fehmi Karamanlıoġlu, Sayf Saray Gülistan Tercümesi, Ankara: Türk Tarih Kurumu Basımevi, 1989.

[4] Sa'dī. 2000. *Gülistān.* Trans. Hongnian Zhang. Cahngsha: Hunan Literature and Art Publishing House.

[5] Sa'dī. 2003. *The Rose Garden.* Trans. Wanbao Yang. Yinchuan: Ningxia People's Publishing House.

[6] Sa'dī. 2008. *Gülistān.* Trans. Janatqan. Kuitun: Yili People's Publishing House.

[7] Arabic Teaching and Reserach Section, Department of Oriental Languages and Literatures, Peking University, ed. 1986. *Arabic-Chinese Dictionary.* Beijing: The Commercial Press.

[8] Persian Teaching and Reserach Section, Department of Oriental Languages and Literatures, Peking University, ed. 1981. *Persian-Chinese Dictionary.* Beijing: The Commercial Press.

[9] Serekenova Qarelğaš, Modern Kazakh Lecxicology, Astana: L.N. Gumilyov Eurasian National University Press, 2010.

[10] Nurbrk Abiken. 2005. Kazakh-Chinese Dictionary. Beijing: Minzu University of China Press.