

EDN: DOAOGQ
УДК 811.512.122–049.35(510)

Problems of Preserving the Kazakh Language in China: Threats of Extinction and Ways of Solution

Sholpan K. Zharkynbekova^{*a}, Saule Zh. Tazhibayeva^a
and Andrey U. Filchenko^b

^a*L.N. Gumilyov Eurasian National University
Astana, Republic of Kazakhstan*

^b*Nazarbayev University
Astana, Republic of Kazakhstan*

Received 18.01.2025, received in revised form 27.01.2025, accepted 07.02.2025

Abstract. This article examines the Kazakh language spoken in China, focusing on its unique characteristics, historical development, and challenges in the context of globalization and sociolinguistic transformations. The study highlights the preservation of archaic linguistic features among Kazakh-speakers in China while they have been lost in standard Kazakh. This exemplifies how the language evolves over time while retaining its fundamental identity. It also discusses how educational reforms have weakened the status of the Kazakh language in China, contributing to its decline in both daily communication and formal settings. The article addresses the risks of losing distinctive linguistic forms due to generational shifts, globalization, and assimilation processes. It underscores the importance of proactive measures for language documentation, revitalization and the preservation efforts as aspects of ethno-cultural identity to ensure the continuity of the Kazakh language in China for future generations. Kazakh variety spoken by repatriates from China is shown to retain some archaic grammatical and lexical elements, no longer found in modern Kazakh, which underscores the crucial need to document this linguistic and ethno-cultural heritage through the development of an electronic corpus of glossed and translated texts. Such efforts not only contribute to preserving the ethno-cultural legacy for Kazakhs in Kazakhstan and the diaspora, but also facilitate contemporary scientific research of the unique linguistics cultural diversity of Kazakhstan in the wider context of Central Asia.

Keywords: the Kazakh language spoken in China, the standard Kazakh language, language variation, language change, language contact, language documentation, anthropological linguistics, ethno-cultural heritage.

This research has been funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP19676809) and supported by Nazarbayev University grant “Multimedia Corpus of modern Spoken Kazakh Language (MULTICORSKL)” Collaborative Research Projects grant No 021220CRP1422.

Research area: Social Structure, Social Institutions and Processes; Languages of the Peoples of Foreign Countries (Kazakh); Theoretical, Applied and Comparative Linguistics.

Citation: Zharkynbekova Sh. K., Tazhibayeva S. Zh., Filchenko A. U. Problems of Preserving the Kazakh Language in China: Threats of Extinction and Ways of Solution. In: *J. Sib. Fed. Univ. Humanit. soc. sci.*, 2025, 18(3), 502–515. EDN: DOAOGQ



Проблемы сохранения казахского языка в Китае: угрозы исчезновения и пути решения

Ш.К. Жаркынбекова^а, С.Ж. Тажибаева^а, А.Ю. Фильченко^б

^аЕвразийский национальный университет им. Л. Н. Гумилева
Республика Казахстан, Астана

^бНазарбаев Университет
Республика Казахстан, Астана

Аннотация. Статья посвящена исследованию языка этнических казахов из Китая, выделены его уникальные особенности, обусловленные историческими, культурными и социальными факторами. Вариант казахского языка, используемый казахами-репатриантами из Китая, содержит архаичные грамматические и лексические формы, утерянные в современном казахском языке. Анализируется влияние образовательных реформ на ослабление позиций варианта казахского языка в Китае, что приводит к его вытеснению в повседневной жизни и официальной сфере. Рассматривается угроза утраты уникальных языковых форм вследствие смены поколений, глобализации и процессов ассимиляции. Подчеркивается важность фиксации языкового и этнокультурного наследия путем создания базы данных аудио- и видеоматериалов, а также разработки электронного корпуса текстов с глоссированием и переводами. Документирование уникального языкового материала казахов-репатриантов из Китая способствует не только научному изучению, но и сохранению этнокультурного наследия, необходимого для формирования чувства сопричастности у будущих поколений как в Казахстане, так и за его пределами.

Ключевые слова: вариант казахского языка в Китае, нормативный казахский язык, языковые изменения, языковые контакты, языковое документирование, антропологическая лингвистика, этнокультурное наследие.

Исследование проведено в рамках грантового проекта «Механизмы интеграции этнических казахов из Китая и Монголии в РК: социокультурный и лингвоэкологический мониторинг» на 2023–2025 гг. (ИРН AP19676809); и грантового проекта Назарбаев Университета «Мультимедийный корпус современного казахского языка (MULTICORSKL)» Грант на совместные исследовательские проекты № 021220CRP1422.

Научная специальность: 5.4.4. Социальная структура, социальные институты и процессы; 5.9.6. Языки народов зарубежных стран (казахский); 5.9.8. Теоретическая, прикладная и сравнительно-сопоставительная лингвистика.

Цитирование: Жаркынбекова Ш. К., Тажипбаева С. Ж., Фильченко А. Ю. Проблемы сохранения казахского языка в Китае: угрозы исчезновения и пути решения. *Журн. Сиб. федер. ун-та. Гуманитарные науки*, 2025, 18(3), 502–515. EDN: DOAOGQ

Introduction

In the era of globalisation and cultural assimilation, minority languages face the risk of reduced functional domain, loss of unique features and, in some cases, total extinction. These challenges are particularly pressing for minority languages in multinational states with strong centralised language policies, such as Kazakh in China.

The migration of Kazakhs to China in the first half of the 20th century occurred amidst significant socio-economic and political upheaval. The primary drivers included opposition to Soviet-era policies that discriminated against the Kazakh population, dismantled wealthy social groups, and imposed unfamiliar and undesirable norms of daily life. These factors compelled thousands of Kazakhs to seek refuge in China, where they now form the second-largest Turkic-speaking ethnic group after the Uyghurs, numbering between 1.5 and 2 million people (Turkic Languages and Ethnicities, 2015: 33–34; Abish, 2015).

Today, ethnic Kazakhs live in the Ili Kazakh Autonomous Region, More-Kazakh Autonomous County, Barkol-Kazakh Autonomous County, Aksai-Kazakh Autonomous County, Gansu Province, the northern regions of the Xinjiang Uygur Autonomous Region, and the Haixi Mongolian-Tibetan Autonomous Region in Qinghai Province (Benson & Svanberg, 1988: 9), preserving their language and culture in compact living conditions.

The Kazakh language spoken by ethnic Kazakhs in China is a unique and little-studied phenomenon. It has retained elements characteristic of the Kazakh language before the 1938 orthographic reform, including archaic lexical and grammatical forms lost in Kazakhstan. At the same time, it has been significantly influenced by the languages of surrounding peoples, such as Chinese and Uighur. This variant of Kazakh contains borrowings and adaptations associated with long residence in a different cultural environment, making it a valuable object of study for linguists, ethnographers and cultural anthropologists.

Statement of the problem

The purpose of this study is to investigate the linguistic characteristics of ethnic Kazakhs in China, to analyze changes under the influence of social and cultural factors, and to document unique linguistic forms essential for preservation of ethnic identity.

To achieve this goal, the study addresses the following questions

- 1) What archaic lexical and phonetic elements of the Kazakh language have been preserved in the speech of ethnic Kazakhs in China?
- 2) What influence have Chinese and Uighur languages had on the Kazakh variety of China?
- 3) What factors contribute to the weakening of the position of the Kazakh language among ethnic Kazakhs in China?
- 4) What is the potential of language documentation for preserving the unique Kazakh linguistic heritage?

The study contributes to the study of the Kazakh language as a multifaceted cultural phenomenon and emphasizes its role as a key component of ethnic identity. It focuses on the need to maintain linguistic diversity and preserve cultural values in the face of globalization and changes in language policy.

Theoretical framework

The Sociolinguistic Situation of Kazakh in China: Status, Functions, Regional Features and Educational Policy

The Kazakh language is one of the recognized minority languages in China. Among the 55 officially acknowledged ethnic minorities in the People's Republic of China, five groups hold "state-level" status: Uyghurs, Mongolians, Tibetans, Kazakhs, and Koreans. By law, regulations in China must be translated into the languages of these five ethnic groups (https://english.www.gov.cn/archive/lawsregulations/201911/20/content_WS_5ed8856ec6d0b-3f0e9499913.html).

Regarding educational policy, it is worth mentioning that prior to 1935, there were no

general education schools in regions with significant Kazakh populations. Kazakh children were mainly religiously educated. The first religious school, a madrassa, was opened in Xinjiang in 1870. In the mosques, Kazakh boys learned the basics of religion, as well as Persian, Arabic, Chagatai and Turkic literary languages of Central Asia. And in 1949, the first general education schools were opened in the Xinjiang Uygur Autonomous Region. However, during the Cultural Revolution (1966–1976), Kazakh-language schools were closed, and the restoration of minority education in Xinjiang began after 1976.

In 1991, there were 588 Kazakh-language primary schools with 138,973 students, 89 secondary schools with 31,880 students, and 42 post-secondary schools with 16,067 students (Abish & Csató, 2011: 276). In Kazakh-speaking educational institutions, the teaching of Chinese began in the third grade. In the 2004 school year, all subjects in such schools were taught exclusively in Kazakh, and teaching materials were translated from Chinese into Kazakh. In 2005, a system of bilingual education was introduced, providing for the use of Chinese from the first year of elementary school. Subsequently, bilingual education was completely replaced by monolingual education with the exclusive use of Chinese.

Economic development of China changed the role of the language. Mandarin Chinese has become essential worldwide. Researchers have found that knowledge of Mandarin ensures not only the privilege of good job prospects but also “the option of entry into the identity of being Chinese” (Nelson, 2005: 26). Mandarin Chinese penetrates all minority areas and all aspects of community life. For examples, in Tibet, the widespread usage of Mandarin Chinese had destructive impact on the Tibetan language learning, official communication, administrative and educational institutions, which leads to a loss of the minority heritage language (Nima, 2001).

Over the past decade, the weakening of the Kazakh language in the People’s Republic of China has been particularly noticeable, mainly as a result of ongoing socio-economic and demographic processes. The number of Kazakh-language educational institutions is

decreasing year by year, which means that ethnic Kazakh children are increasingly forced to attend Chinese-language schools. As a result, the functional domain of the Kazakh language is decreasing significantly, and its status is weakening. Kazakh is gradually losing its position as a dominant language of the regional community and is being transformed into a predominantly a language of informal daily communication.

The current policy of the People’s Republic of China towards ethnic minorities is characterized by relative loyalty (Abish, 2016). Ethnic Kazakhs have the opportunity to develop their own media and cultural centers. However, the problem of preserving the Kazakh language, cultural values and traditions remains very relevant. According to K. Yesimova, a specialist on China, the functions of the Kazakh language are currently significantly limited. The researcher links this to the introduction of the bilingual education system, within the framework of which the volume of teaching of academic disciplines in the Kazakh language has been significantly reduced. At the same time, the production of educational literature and methodological manuals in Kazakh is decreasing. In addition, teachers who do not speak Chinese are leaving educational institutions, which leads to the dominance of teachers who are fluent in Chinese (Yesimova, 2016).

The policy of modernization and *Chinising* the western provinces of the country is gradually extending to all regions of the PRC and certainly affects Kazakhs living in compact and isolated regions. Kazakhs who are not fluent in Chinese face certain problems in finding work. The analysis of the data from the sociological surveys conducted by the Kazakh scientist B. Rakisheva in 2010, 2011, 2013 and 2014 in the areas of compact residence of ethnic Kazakhs in the XUAR of China demonstrate the attitude of ethnic Kazakhs to their mother tongue.

In response to the question “Are you concerned about the present and future of your native language?”, 61.5 % of the respondents show concern about the present and future of their native language, 32.7 % declare somewhat concerned, and only 5.8 % of the respon-

dents said that they were not concerned about the present and future of their native language (Rakisheva, 2015:143).

The effects of globalisation, as well as the influence of Chinese on Kazakh through media and education, are certainly evident in the unique linguistic and cultural characteristics of ethnic Kazakhs living in China. At the same time, the growing national consciousness of Kazakhs, their desire to preserve their language and culture for future generations testifies to the importance of Kazakh language and cultural heritage as key components of ethnic identity (Zharkynbekova, Tazhibayeva, 2024; Zharkynbekova, Tazhibayeva, 2024).

This aspiration is manifested in the increased activities of cultural organizations, in the creation of public initiatives to preserve the language, and the support of traditional rites and customs. Despite the challenges associated with education and language policies, ethnic Kazakhs are making efforts to ensure intergenerational transfer of language and cultural values, which contributes to strengthening of the national identity and preservation of the unique cultural code.

The return of ethnic Kazakhs within the modern borders of Kazakhstan contributes to the strengthening of national identity, the preservation of cultural heritage, the revival of the Kazakh language and the expansion of its functions. The Kazakh language has low dialectal variation, which ensures mutual understanding in oral speech among Kazakhs, regardless of their place of residence. However, written communication is difficult due to the use of different alphabetic systems. For example, Kazakhs in Kazakhstan, Russia and Mongolia use the Cyrillic alphabet; Kazakhs in China, Afghanistan and Iran use the Arabic alphabet; and in Turkey and European countries, the Latin alphabet prevails. This fragmentation of writing systems has two main consequences. First, there is a separation from the common written heritage, including scientific and artistic literature and historical documents, which become inaccessible to speakers of different alphabets. Second, the mutual understanding between the representatives of the Kazakh diaspora and the citizens of Kazakhstan in written communi-

cation decreases. Thus, the written language loses its unifying function, which weakens the position of the Kazakh language as a means of preserving cultural and historical community among Kazakhs living outside Kazakhstan.

With Kazakhstan's independence, the process of ethnic Kazakhs returning to their ethnic homeland from near and far abroad began. The first wave of migration began in 1991. According to the Committee on Migration of the Ministry of Labour and Social Protection of the Population of the Republic of Kazakhstan, more than 1.14 million ethnic Kazakhs arrived in the country during this period. In 2024, 12,325 ethnic Kazakhs were granted Kandas status, of whom 45.6 % were from the People's Republic of China (PRC) (<https://www.gov.kz/memleket/entities/kqon/press/news/details/851228?lang=ru>).

Repatriated Kazakhs (Kandas) from the People's Republic of China in Kazakhstan are concentrated in the regions of Eastern Kazakhstan and Almaty, as well as in the cities of Almaty and Astana. Such localization helps them to preserve their traditional way of life, original culture, customs and peculiarities of the Kazakh language variant typical of the Kazakh diaspora in China.

Methods

This article investigates the linguistic features of the Kazakh language as spoken by ethnic Kazakhs residing in China, with an emphasis on the retention of archaic forms, distinct phonetic characteristics, and the influence of cultural assimilation on the language. Through a multidisciplinary approach involving fieldwork, interviews, audio-visual recordings, and linguistic analysis, the research aims to document the unique features of variation of the Kazakh variety spoken in China. Additionally, it explores the challenges of language endangerment resulting from generational language shifts and the growing influence of dominant languages, such as Chinese and Uyghur.

Participants were selected based on their region of origin in China, notably the Ili Kazakh Autonomous area, in Urumqi as well as urban and rural areas in Kazakhstan. Different speech events were recorded in Kazakh-

stan and in the border regions with China. The study sample comprised 30 participants aged between 22 and 97, predominantly from East Kazakhstan oblast', Akmola oblast', and the city Astana. Efforts were made to ensure a diverse representation of individuals with firsthand experience of Kazakh as spoken in China. Interviews were conducted in Kazakh, utilizing questions designed to capture linguistic features, personal narratives, and cultural practices associated with language use. The collected audio and video recordings were transcribed using the International Phonetic Alphabet (IPA) to enable a comprehensive phonetic analysis. These transcriptions facilitated the identification of unique phonetic characteristics, including vowel shifts, consonantal variations, and the integration of loanwords from Chinese and Uyghur.

A comparative analysis was conducted against standard Kazakh, highlighting differences in vowel harmony, consonant patterns, and overall sound systems. Additionally, the study examined the contact influence in the lexicon from Chinese and Uyghur, particularly in vocabulary associated with daily life, cuisine, clothing, and ethno-cultural traditions. The comparative analysis of the linguistic data from ethnic Kazakhs in China with the standard Kazakh of Kazakhstan revealed both commonalities and distinctions, particularly in phonetics and vocabulary. The research placed a strong emphasis on lexical borrowings, especially terms related to food and everyday items influenced by Chinese and Uyghur. This comparative framework provided insights into the impact of language contact on the Kazakh spoken in China and its divergence from the standard Kazakh language. The findings emphasize the distinctive characteristics of the Kazakh variety spoken in China and highlight the urgency of documenting these features before they are diminished by language shift and cultural assimilation. Additionally, the research provides a basis for future investigations into language endangerment and revitalization efforts among Kazakh-speaking communities in China.

Ethical Considerations. Every interview started with filling out meta data with a promise

to respondents of keeping personal information confidentiality, and the usage of the recorded data only for academic purposes.

The Language Documentation of Kazakh spoken in China

Language documentation or documentary linguistics is a relatively new direction of empirical linguistics. Language documentation is the creation of corpora of annotated and translated speech events in audio and video format. (Himmelman 2006, Woodbury 2003). Language documentation is also significant for comparative analysis of Kazakh spoken in China and standard Kazakh. Such kind of research helps to focus on linguistic changes in the main layers of the language, problems of terminology, and standardization of transcription.

Kazakh spoken in China belongs to a group of languages known as the primary minority language, which has historical and modern prestige, a large population, and moderate political clout (Dwyer 1998:71). Kazakh as one of the significant minority languages in China plays an important role in the areas where Kazakhs dominate. In some regions of Ili Kazakh Autonomous Prefecture, Kazakh functions as a lingua franca; i.e. it is used as a common language between speakers whose native languages are of Turkic origin as Uyghur, Tuvanian, Kyrgyz and non-Turkic origin as Mongol, and Xibe, etc..

Kazakh is a language of communication among Kazakhs in other Kazakh autonomous counties. Ili Kazakh Autonomous Prefecture is a home for the majority of Kazakhs. In this Prefecture the government authorities use both Kazakh and Chinese as official languages. The official documents issued by the authorities to the township-level administrations are mostly written in Kazakh. The Congress of the Party in this prefecture employs a translation agency for Kazakh. Public signs including names of places, streets, etc., and official stamps are in Kazakh and Chinese. Kazakh is also used in the courts when they deal with a case concerning a Kazakh person. Kazakh is a language of education, is an object of research, and has its own print and broadcast media in China (Abish, 2016).

According to Abish and Csato (2011), referencing the Ethnologue, the use of Kazakh was described as vigorous, with an estimated 85 % of Kazakhs being monolingual and only 15 % fluent in Chinese. The strong proficiency in the native language among Kazakh communities was attributed to their socio-cultural organization, rural living environments, preservation of ethno-cultural traditions, and involvement in traditional livestock keeping and herding practices (Abish & Csato 2011). This fact was also mentioned in the interviews of the respondents. The interview was conducted among the elder generation of repatriates who migrated to Kazakhstan in the period 1992–1995, and who currently live with their relatives in the village Qoyandy not far from Astana. In China the families of our respondents used to live in the remote area near the Altai mountains, leading the traditional pastoral way of lifestyle. They mentioned that they had never had contact with Hans (Chinese people) in their herder's households and had never communicated in the Chinese language. The only language they spoke was their native Kazakh:

'yeşkaşan qandardı körgen žokpız, qanzu-ša söylegen žokpız, tek özimizdiñ qazaq tilde söyledik' (NQ, 88 years old, Qojandy village, Astana, 2023).

'We have never met Chinese (in our region), we speak only Kazakh'. As for their children, they graduated from schools with a Kazakh program of education, starting study Chinese from the third grade of the elementary school, and continued mastering the advanced Chinese at university foundation courses (Qoyandy region, Astana, 2003). However, most native speakers of Kazakh from the urban regions have acquired native-like competence in Chinese and those in contact with the largest Turkic minority language speak also Uyghur.

The language situation of ethnic Kazakhs in China has significantly changed after 2005. The main reason was connected with the active introduction of bilingual education. The initial effort to establish bilingual education for minority groups began in various regions in 1978, aiming to introduce Chinese and "gradually achieve the transition from the speaker's native language to Chinese" (Dwyer, 1998: 81).

The modern generation of ethnic Kazakhs in China is multilingual. The linguistic situation of Kazakhs born in the late 1990s and living in the major mega-cities of China differs radically from the linguistic situation of the older generation. They get their education today dominantly in Chinese, and use Kazakh only in a restricted linguistic domain, mostly as a language of informal communication within the family and with Kazakh friends. The number of schools with the Kazakh language of education is declining from year to year. Children of ethnic Kazakhs have to study in Chinese schools (Abish, 2016). Thus, the functions of the language of ethnic Kazakhs are narrowing and weakening. Kazakh from the dominant language turns into a household language. So, Kazakh, as a dominant minority language, has a serious problem of becoming endangered (Abish, 2016).

The ongoing processes of globalization, the influence of the Chinese language through the mass media, and education contribute to the destruction of the unique culture and language of ethnic Kazakhs living in China. A new generation of Kazakhs from China are losing their native language proficiency. They understand the mother tongue but cannot speak or write. Johanson outlines that language endangerment starts when "younger generations begin to pay less attention to their ancestral code and switch over to a dominant code because they find it more useful, attractive and prestigious" (2024: 11). In this respect, it is necessary to document as many speech events as possible to be able to register important changes in language practices. The situation is exacerbated by the generation shift. The older generation of Kazakhs, who have preserved archaic forms of the language, customs, national traditions, ethno-cultural peculiarities and who remember the history of their ancestors' migration to China, are passing away.

The study of the history and spiritual heritage of the Kazakhs from China is conducted in Kazakhstan by documenting and researching the language peculiarities of Kazakh repatriates. The historical heritage, language, and culture retained by the older generation of repatriates from China are recorded and docu-

mented by our research team. Documentation of different variants of Kazakh is important for linguists to get more knowledge about the diachronic and synchronic development of the language. Besides, linguistic documentation provides a stimulating effect for the repatriates from China. In the interviews, they proudly emphasized that they speak an old version of Kazakh and have successfully preserved it outside of Kazakhstan. In this regard, it is essential to conduct interviews with audio and video recordings of older generation representatives who hold valuable memories of captivating stories, significant events, and various traditions and customs that are handed down through generations and preserved as part of historical memory. The study, recording, and description of linguistic peculiarities will not only enrich scientific ideas about the development of the Kazakh language but also create a basis for further research.

Language documentation is also significant for comparative analysis of Kazakh spoken in China and the standard Kazakh, it helps focusing on contact-induced language change at the main levels of the language system, as well as more specifically – on the problems of terminology, and standardization of transcription.

Field research and data collection can be conducted not only in China but in Kazakhstan as well. The speakers of unique linguistic variants are residents of the regions where repatriates are compactly inhabited after migration from China. Representatives of the generation who is under 50 years old graduated from Kazakh middle schools and studied advanced Chinese at universities.

Members of the research team utilized the International Phonetic Alphabet (IPA) for phonetic notation (<http://ipa.typeit.org/full/>). The transcription system used in the paper is proposed by Johanson (2021), Johanson & Csato (2022), and adopted by Abish (2016) and Jumabay (2022). The analyzed audio and video recordings had been collected during fieldwork in 2023, 2024 from Kazakh repatriates from China, aged 22 to 97, currently residing in East Kazakhstan, Akmola oblast', and the city of Astana. The analysis of the audio and video recordings revealed distinctive features

in the sound system of the language in the spoken speech of older generations of repatriates. It is well-known that the sound systems of territorial dialects within specific languages may differ in both qualitative and quantitative characteristics of their vocalic and consonantal subsystems. To study the phonetic features of the Kazakh language variants, word forms, sentences, and texts recorded under fieldwork conditions were utilized.

Discussion

We posit that the language variant of Kazakh repatriates keeps archaic elements in the vocabulary and grammar. This study primarily focused on phonetic and lexical features of Kazakh as spoken in China, with less emphasis on other linguistic aspects as morphology and syntax. Future research could expand on these areas to provide a more comprehensive understanding of the linguistic situation.

We are focusing our interest on peculiarities of sound system and two groups of lexemes of Kazakh as spoken in China.

The sound system of the Kazakh variant in China exhibits the following distinctive features:

In the vocalic system:

(A) The vowels *i* [i], *ы* [ɨ], *у* [u], and *ү* [ʊ] are regularly used before sonorant consonants *l* and *r*, which is not characteristic of the standard modern Kazakh language.

No	Kazakh language variant spoken in China	Kazakh language spoken in Kazakhstan
1	ілегі [i'legi]	легі [legi]
2	ілезде [i'lezde]	лезде [lezde]
3	ірет [i'ret]	рет [ret]
4	ылақ [ɨ'lak]	лақ [lak]
5	ыриза [ɨ'riza]	риза [riza]

(1) Oralman-niŋ oqwşı-lar-dıŋ aldınyı i'legi boı-dı-қ.

Repatriate-GEN student-PL-GEN first wave BOL.COP-PAST-1PL

'We were the first wave of students among the repatriates.'

(U., Astana, Qoyandy region, 29.10.2023)

(2) Soñ kişi ayt-baq-tay, ĭras kelĭn-ge
that person say-**DER-EQUA** true daughter-in-
law-**DAT**
kelĭn boĭ-a-dĭ eken-siñ.
daughter-in-law **BOL.COP-PRES-3 E.COP.IN-**
DIR-2SG

‘As that person says, it is true that you
(a mother-in-law) behave like a daughter-in-
law.’ (in the meaning ‘a mother-in-law wants to
help a daughter-in-law’)

(O. K., Astana, 08.02.2024)

(B) The use of the vowel sound a in the Ka-
zakh language variant instead of the soft vowel
sound found in standard Kazakh. For example,
ары [ari] instead of *ери* [äri], and *ажым* [ažim]
instead of *эжим* [äžim]. Conversely, there is
also the use of the soft front vowel ä in spoken
language instead of the back vowel a found in
standard Kazakh.

No	Kazakh language variant spoken in China	Kazakh language spoken in Kazakhstan
1	кiтeп [kĭtär]	кiтeп [kĭtär]
2	жеңе, жеңе ғана [žäŋe; žäŋe ғана]	жаңа, жаңа ғана [žäŋa; žäŋa ғана]
3	жеңеғы [žäŋeγĭ]	жаңағы [žäŋaγĭ]
4	бeγäγi [bäyγi]	баяғы [baγγi]

(C) The preservation of labial vowel har-
mony in the Kazakh language variant of China,
for instance, *turmustuқ žayday* is pronounced
turmĭstĭq žayday in standard Kazakh; *kötoredĭ*
is pronounced *köteredi* in standard Kazakh.

1) So-niñ bala-sĭ-n bay-ĭp kempir-šal
Ūrümžĭ-de žatĭr.

that-GEN child-poss3-ACC take care-CONV
grandparents Urumqi-LOC lie-3

‘The grandparents take care of her child
in Urumqi.’ (P.K., Astana, 08.02.2024)

As a remark, in both standard Kazakh and
the Kazakh language variety spoken in China,
labial vowel harmony, is consistently main-
tained in spoken language, as exemplified by
bügĭn. However, in written, labial vowel har-
mony is not reflected, resulting in forms such
as *bügĭn*, where harmony is disrupted.

A peculiarity in the consonant sound sys-
tem is the use of the consonant ‘č’ at the begin-
ning and in the middle of words:

Kazakh language variant spoken in China	Standard Kazakh language
Чырағым / čĭraγĭm	Шырағым / šĭraγĭm
Чимайлайды / čĭmaylaydĭ	Шимайлайды / šĭmaylaydĭ
Чарчағанын / čarčaγanĭn	Шаршағаның / šaršaγanĭn
патЧасы / patčasĭ	патШасы / patšasĭ
аманЧылық / amančĭlĭq	АманШылық / amanšĭlĭq

(1) Češe-m qarapayĭm žumĭsčĭ boĭ-dĭ.
mother-POSS 1SG ordinary worker **BOL.COP-**
PAST3

‘My mother was a worker.’ (Ж.К., 50 years old,
Aqmola oblast’)

(2) Balalĭq čay-ĭm-da üš tiĭ-dĭ қатар қолдан-
dĭ-қ.

Childhood time-POSS 1SG-LOC three
language-ACC together use-PAST-1PL

‘In my childhood, we spoke three languages.’
(Aqmola oblast’)

(1) Moĭda azan čakĭr-a-dĭ...

Mullah call to prayer recite-PRES-3SG

‘The mullah recites the azan (the prayer).’
(E.A., Astana, 28.11.2023)

(d) The use of the voiced consonant sound
‘d’ in the Kazakh language variant instead
of the voiceless ‘t’ in the standard language:
*düzüw – tüzüw; duz – tuz; düzetüw –
tüzetüw*, etc.

(7) Žaksĭlap duzda-p, duz-ĭ siñ-gen soñ
properly season-CONV salt-POSS3 absorb-PTCP
after

düzüw aγaš-қа қатарlas-tĭr-ĭp il-ĭp қой-atĭn.
straight wood-DAT line.up-CAUS-CONV hang-
CONV PUT.POSTV-INTRA3

‘After properly seasoning and salting it,
and once the salt is absorbed, they line it up on
a straight piece of wood and hang it up.’ (HA,
Astana, 18.02.2024)

(D) The use of the voiced consonant *d* in-
stead of the sonorant *l* in standard Kazakh.

Kazakh language variant spoken in China	Standard Kazakh language
қытайдармен [kĭtaydarmen]	қытайлармен [kĭtaylarmen]
балалар [baladar]	балалар [balalar]
одар [odar]	олар [olar]

(2) *Sonan keyin bala-dar-ım osi qazaqstan-dar-dı öte unat-tı.*

then after child-PL-POSS1SG this Kazakhstan-PL-ACC very like-PAST3

Odar talap et-ti, äke-čeşe osi qazaqstan-ya bar-ıp oқи-sa-k

They request-PAST3, father-mother this Kazakhstan-DAT go-CONV study-HYP-1PL

‘After that, my children liked Kazakhstan very much. They requested, “Father and mother, if only we go and study in Kazakhstan.”’ (A.A., Astana, 27.02.2024)

(E) The use of the sound d instead of ž:

(1) 80-ge kel-eyik, 90-ya keleyik, düz-ge kelip soł

80-DAT reach-VOL1PL 90-DAT reach-VOL1PL hundred-DAT reach-CONV that

Čöbere odan arı kör-ıp dekeñde-p žür-eyik...

great-grandchild beyond that see-CONV enjoy-CONV MOVE.POSTV-VOL1PL

‘Let us reach 80, 90, even 100 years old, and wish to see great-grandchildren, enjoy them, and live happily.’

(F) In copied words from other languages, sounds that are absent in the Kazakh language are adapted to the phonetic system: *ministırlık* ‘ministry’, *samawrın* ‘samovar’, *wagon* ‘wagon’, *woleibol* ‘volleyball’, *pakıt* ‘packet’, *pak-wltet* ‘faculty’, *perma* ‘farm’, *äktip* ‘active’, *ap-tonomiya* ‘autonomy’, which contributes to the enrichment and expansion of the vocabulary of the Kazakh language variant in China.

In standard Kazakh, the consonants *f, v, ts, š, č* are used only in words borrowed from Russian or through Russian. In standard Kazakh, the spelling and pronunciation of copied words are fully preserved: *ministrlık, vagon, samovar, voleybol, fakt, fakultet, ferma, aktiv, aftonomiya*. It should be noted that the older generation of Kazakhs in Kazakhstan adapts sounds in copied words in the same way as ethnic Kazakhs in China.

The analysis of the audio-video recordings collected during field research shows the presence of some differences in the sound systems of Kazakh and the Kazakh language variant in China. However, these differences are minor and do not alter the original meaning of the words.

Lexical features of the Kazakh language variant in China

A significant part of the vocabulary of the language of ethnic Kazakhs from China, who now live in Kazakhstan, consists of everyday vocabulary. This includes terms related to material culture, closely associated with daily life and household activities. In the article we are focusing only two groups of words of everyday usage: lexical items denoting food and clothing

Field research data from various speech events reveal that Kazakhs from China use Chinese-derived terms for certain vegetables and fruits, such as *лобы* [lobı] for “carrot” instead of *сәбіз* [sebız], *қызыл лазы* [kıził łazi] for “red pepper” instead of *қызыл бұрыш* [kıził bwrıš], *джизы* [žizı] for “mandarin,” and *чыңзы* [čıñzi] for “orange.” Additionally, Uyghur borrowings include terms like *сүйқаш* [süykaš] for “soup” and *зәш* [göš], *менті* [mantı] for “meat dumplings”.

The names of vegetables, fruits, and food products in the Kazakh language variant of China are influenced by regional languages, including Chinese, and Uyghur. Due to the introduction of new, non-traditional foods and agricultural products in China, it became necessary for ethnic Kazakhs to adopt foreign words to name these items. The borrowed terms were copied from the languages with which the Kazakhs had contacts: Chinese, Uyghur, and Russian. These adaptations reflect the changes in their lifestyle and the integration of new elements into their daily lives, which were influenced by the cultures and languages surrounding them in China.

№	Spoken Kazakh variant in China	Spoken Kazakh variant in Kazakhstan	Standard Kazakh	Translation
1	ішпек-жемек <i>išpek žemek</i>	продукты	азық-түлік	food
2	чезы (qiezi)	баклажан/баялды	баялды/baıaldı	eggplant

№	Spoken Kazakh variant in China	Spoken Kazakh variant in Kazakhstan	Standard Kazakh	Translation
3	чинцай (qíncài)	сельдерей	балдыркөк/ baldırkök	celary
4	боцай (bōcài)	шпинат	саумалдық/sawmałdıq	spinach
5	шиаң жау (xiāngjiāo)	банан	banan	banana
6	чыңзы (chéngzǐ)	апельсин	apelsin	orange
7	жұзы (júzi)	мандарин	mandarin'	mandarin
8	жиуцай (jǐucài)	жусай	жусай žwsay	chives
9	шекер/ құм шекер	қант/ шекер/ құм шекер қант/	қант қант	sugar
10	пірендік (pīrǎndīk)	пряник prjanik	пряник/ prjanik	gingerbread

Repatriates use familiar terms for vegetables and fruits that have been copied from the Chinese language. In standard Kazakh, however, such terms are typically borrowed from Russian or through Russian influence. Although new Kazakh equivalents for food products exist in standard Kazakh, they are rarely used in everyday conversation, as Kazakhs tend to favor Russian-derived words.

In the case of clothing, ethnic Kazakhs in China use mostly Kazakh traditional names for clothing. However, in standard Kazakh clothing items are copied from Russian culture or copied from other cultures via the Russian language. These borrowed terms reflect the clothing items that were either introduced or became more common due to interactions with Russian culture. In everyday speech, some traditional Kazakh clothing items might have different names when spoken by ethnic Kazakhs in China compared to those living in Kazakhstan, because of regional influences and the adoption of new items of clothing.

Repatriates employ traditional Kazakh terms to refer to basic clothing items. Terms such as баупақ (socks), шақая (shoes), yetek köylek (skirt), and күрме / күрмесhe (vest) belong to the archaic layer of vocabulary and have fallen out of use in modern Kazakh.

The names for clothing in the Kazakh language represent one of the most interesting and rich layers of the lexicon. Many words that have disappeared from everyday use in modern Kazakh and have transitioned into the category of archaic vocabulary are actively used in the variant of Kazakh spoken in China. The discussion above illustrated this with two exemplary categories of everyday vocabulary: words related to food and clothing. Terms describing traditional Kazakh headdresses, attire, and jewelry represent a distinctive lexicon that reflects the ethno-cultural characteristics of the Kazakh language. Furthermore, this ethno-cultural vocabulary serves as a valuable resource for studying the lexicon of the endangered variants of Kazakh spoken in China.

№	Kazakh in China	Spoken Kazakh in Kazakhstan	Standard Kazakh	Translation
1	етек көйлек yetek köylek	юбка jwbka	белдемше beldemše	skirt
2	күрме/күрме күрмесhe/күрмесhe	желет/кеудеше želet/kewdeše	кеудеше/kewdeše	velt
3	жейде žeyde	футболка/ жейде	футболка/fwtbořka жейде/žeyde	T shirt
4	қылта көйлек qılta köylek	мини юбка/қысқа юбка	мини юбка /қысқа белдемше/ mini jwb- ka/ qıřqa jwbka	mini skirt

№	Kazakh in China	Spoken Kazakh in Kazakhstan	Standard Kazakh	Translation
5	ұйқы киім uıqı kiim	пижама pižama	пижама / pižama	pijamas
6	сым/шалбар sım/šalbar	сым/шалбар sım/šalbar	шалбар / šalbar	trousers
7	дамбал dambal	дамбал dambal	ішкиім/iškiim	underwear
8	ұзын байпақ uzın baupraқ чаң тоң уа (liánkùwà)	колготки	колготки/ kołgotki	tights
9	ұйық/шұлық uıyıq/šulıq	шұлық/šulıq	шұлық/ šulıq	stokings
10	байпақ baupraқ	нәски/носки näski	шұлық šulıq	socks

Documentation and analysis of these and other aspects of local language contact, variation and dynamic change may be performed most effectively and with empirical rigor within the contemporary corpus linguistics methods. An important initiative in this domain has been the development of the Multimedia Corpus of modern Spoken Kazakh Language (MCSKL (Filchenko, Troiani, Du Bois, Sarseke, Akanov, Bizhanova, Mikhailov, Temirbekova, Seitak, Turaliyeva, 2023)) at Nazarbayev University, Astana. This project employing discourse-centered approach to corpus development, inspired by best practice of the Santa Barbara Corpus of Spoken American English (Du Bois et al., 2000), prioritizes naturally occurring speech events, integral to participants' lives (Troiani, Du Bois, Filchenko, 2024). Documenting diverse speech events across Kazakhstan and in the border regions of China, including informal and formal interactions, traditional celebrations, online communications in their natural contexts, the corpora will include deep annotation schemas encompassing orthographic, phonetic, and morphemic information, alongside morphemic glossing, and translations. Such corpora and their associated methodology provide rich resource for diverse research needs, from formal linguistic analysis and lexicology, to sociocultural, dialectological, and historical investigations of ethnolinguistic communities. The corpora projects such as MCSKL exemplify successful interdisciplinary collaboration, promoting the

study and preservation of linguistic and cultural diversity of Kazakhstan, and serves as a benchmark and framework for other corpus development projects in the larger Central Asian region (Troiani & DuBois & Filchenko, 2022) that will help to preserve unique linguistic heritage for future generations of Kazakhs, both in Kazakhstan and abroad.

Conclusion

The study of the Kazakh language spoken by ethnic Kazakhs in China is pivotal for understanding the processes of linguistic and cultural preservation in a rapidly evolving global context. Despite the challenges posed by cultural assimilation, globalization, and language shift, Kazakh language remains a vital symbol of identity for ethnic Kazakhs, particularly in rural areas of China. This variant of Kazakh spoken in China, which retains archaic features and exhibits unique lexical influences from

Chinese and Uyghur, provides valuable insights into the linguistic history and evolution of Kazakh prior to significant reforms in Kazakhstan. Documenting and analyzing the linguistic features of the Kazakh spoken in China by comparing its peculiarities with the standard Kazakh variant shed light onto a range of phonetic, lexical, and syntactic aspects, that are critical not only for advancing linguistic scholarship but also for safeguarding the cultural heritage of Kazakh communities in China. Fieldwork, including interviews, elicitation and audio-visual documentation of nat-

urally occurring interactional speech events, is essential for capturing and preserving the spoken varieties in their natural contexts, especially those from the older generation. However, generational language shifts and the rise of multilingualism in urban areas pose serious challenge to the continued use of Kazakh, underscoring the urgent need for ongoing documentation and revitalization efforts.

The findings of this study demonstrate that while the Kazakh language in China is transforming, it still retains many distinctive features that reflect its historical and ethno-cultural origins. The contact influence of the neighboring Chinese and Uyghur languages, is evident in the vocabulary and phonetics, illustrating the interplay between Kazakh and its sociolinguistic environment. Although such

linguistic contact has enriched the language, it has also introduced changes that endanger traditional linguistic practices.

In conclusion, preserving the Kazakh language in China is an urgent issue that demands a multifaceted strategy involving linguistic research, comprehensive documentation, educational initiatives, and cultural programs. Raising awareness about linguistic diversity in China and promoting efforts to protect minority languages like Kazakh are essential. By continued study and documentation of Kazakh language varieties using contemporary research methods, we contribute to a deeper understanding of language evolution, cultural preservation, and the resilience of ethnic communities in the context of globalization.

References

- Abish A. Modality of Kazakh as Spoken in China. In: *Turcologica Series* (107). Harrassowitz Verlag Press, 2016, 259.
- Abish A., Csató É. Á. Recent developments in Kazakh as spoken in the People's Republic of China. In: *Turkic Languages*, 2011, 15, 275–290.
- Benson L., Svanberg I. The Kazakhs in Xinjiang. In: Benson L., Svanberg I. (Eds.) *The Kazakhs of China. Essays on Ethnic Minority*. Uppsala, 1988, 1–106.
- Constitution of the People's Republic of China. Ch. 1, Article 4. Available at: https://english.www.gov.cn/archive/lawsregulations/201911/20/content_WS5ed8856ec6d0b3f0e9499913.html (accessed 15 December 2024).
- Dwyer A. M. The texture of tongues: languages and power in China. In: Safran W. (Ed.) *Nationalism and Ethnoregional Identities in China*. Frank Cass, 1998, 68–85.
- Filchenko, A., Troiani, G., Du Bois, J. W., Sarseke, G., Akanov, A., Bizhanova, M., Mikhailov N., Temirbekova T., Seitak B., Turaliyeva Z. Multimedia corpus of spoken Kazakh language. 2023.
- Himmelman N. Language documentation. What it is and what it is good for? In: Jost Gippert, Nicolas P. Himmelman, Ulrike Mosel (Eds.) *Essentials of Language Documentation (Trends in Linguistics: Studies and Monographs 178)*. Berlin: Mouton de Gruyter, 2006, 1–30.
- Johanson L. Documentation of Turkic in comparative linguistics. In: Csató É. Á., Karakoç B. (Eds.) *Language and Documentation in Comparative Turkic Linguistics*. Ankara, 2024, 9–19.
- Nelson K. Language policies and minority resistance in China. Available at: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=3dc4e1a4cce628576b7eb01bb1958dc31d172436> (accessed 5 January 2024).
- Nima B. Problems related to bilingual education in Tibet. In: *Chinese Education & Society*, 2001, 34(2), 91–102. DOI: 10.2753/CED 1061–1932340291.
- Johanson L. *Turkic*. Cambridge: Cambridge University Press, 2021, 1084 p.
- Johanson L., Csató É. Á. (Eds.) *The Turkic Languages*. London & New York: Routledge, 2022, 512.
- Jumabay U. Clause combining strategies in Kazakh as spoken in China in comparative perspectives. PhD dissertation. Frankfurt: Goethe-Universität Frankfurt, 2023.
- Rakisheva B. Nekotorye aspekty etnichnosti kazakhov Kitaia [Some aspects of the ethnicity of Kazakhs in China]. In: *Tcentralnaia Aziia i Kavkaz [Central Asia and the Caucasus]*, 2015, 18(2), 128–146.

Yesimova K. Why Kazakhs in China Preserved Their Native Kazakh Language but Are Losing It Today. In: *Kazakh-Chinese Journal Esquire*, May 5, 2016. Available at: <https://esquire.kz/kazahskiy-kitayskiy/>.

Zharkynbekova Sh., Aimoldina A., Akynova D. Cultural and language self-identification of ethnic minority groups in Kazakhstan. In: *Sociolinguistic Studies*, 2015, 9(2–3), 289–314. DOI: 10.1558/sols.v9i2.26587.

Zharkynbekova S., Tazhibayeva S., Shakhputova Z., Azamatova A., Kuzar Z. Transnational practices of Kazakh repatriates: the role of family in the adaptation of ethnic Kazakh students from Mongolia and China. In: *Frontiers in Sociology*, 2024, 9. DOI: 10.3389/fsoc.2024.1452785.

Zharkynbekova S.K., Tazhibayeva S. Zh., Shakhputova Z. Kh., Kuzar Z.N. Repatriates in media: image formation and its transformation. In: *Journal of Siberian Federal University – Humanities and Social Sciences*, 2024, 17(10), 2036–2049.

Woodbury A. Defining documentary linguistics. In: Peter K. Austin (Ed.) *Language Documentation and Description*, Vol. 1. London: SOAS, 2003, 35–51.

Abbreviations

- 1 first person
- 2 second person
- 3 third person
- ACC accusative
- BOL.COP the copula bol – ‘to become, be’
- CAUS causative
- COP copula
- CONV converb
- DAT dative
- DER derivational suffix
- E.COP the copula e ‘to be’
- EQU equative
- GEN genitive
- HYP hypothetical/conditional mood
- INDIR indirective
- PAST past tense
- PART particle
- PRES present tense
- PRCTP participle
- POSTV–VOL post verb volative mood
- POSTV post verb
- POSS possessive
- PL plural
- SG singular