

CROSSBORDER DIALECT DIFFERENCES AND SPEAKERS
PERCEPTIONS: KAZAKH LANGUAGE IN CHINA AND
KAZAKHSTAN

by

Moldir Bizhanova

A thesis submitted in partial fulfilment of the requirements for the
degree of

Master of Arts in
Eurasian Studies

at

NAZARBAYEV UNIVERSITY - SCHOOL OF SCIENCES AND
HUMANITIES

2024

THESIS APPROVAL FORM
NAZARBAYEV UNIVERSITY
SCHOOL OF HUMANITIES AND SOCIAL SCIENCES

Crossborder dialect differences and speaker perceptions: Kazakh
language in China and Kazakhstan

BY

Moldir Bizhanova

NU Student Number: 201684948

APPROVED

BY

Dr. EVA-MARIE DUBUISSON ASSOCIATE PROFESSOR

ON

The 14th of May, 2024

E.M. Dubuisson

Signature of Principal Thesis Adviser

In Agreement with Thesis Advisory Committee
Second Adviser: Dr. Benjamin Brosig Assistant Professor
External Reader: Dr. Abish Aynur Assistant Professor

Abstract

The Kazakh language is a Turkic language of the Kipchak branch and an official language of Kazakhstan with more varieties spoken by minorities in other countries. My research focuses on one of such Kazakh varieties spoken in China, with the fieldwork done among the ethnic Kazakhs born in the Ili region of Xinjiang Uyghur Autonomous Region.

My thesis aims to identify the linguistic differences between the Kazakh spoken in China and Kazakhstan, and investigate the perceptions and attitudes of ethnic Kazakhs born in China towards the Kazakh language variations. For that, the study first contrasts phonetic, lexical, and semantic differences in the spoken Kazakh in China and Kazakhstan based on the audio files recorded during the fieldwork. Even though similarities exist between spoken Kazakh in China and regions of Kazakhstan, distinct lexical and dialectical differences are evident based on the results of my work.

Then based on the interviews conducted with the ethnic Kazakhs from China, my study reveals attitudes towards the Kazakh language varieties where language ideologies were detected regarding language purity. It was shown how semiotic processes of language ideologies further influence the perceptions of the dialect borders in the minds of the speakers.

Table of Contents

Table of Contents.....	2
Chapter 1. General Overview.....	3
1.1. Introduction	3
1.2. Literature Review.....	4
1.2.1. Kazakh language and its variations	4
1.2.1. Kazakhs in China.....	7
1.2.3. Research on Kazakh language variation in China	11
1.2.4. Dialects and perceptual dialectology	14
1.3. Research Questions.....	18
1.4. Methodology.....	20
1.4.1. Analysis of the linguistic differences	20
1.4.2. Investigation methods of perceptual dialectology.....	24
1.5. Conclusion for the chapter	26
Chapter 2. The results and discussion of the linguistic differences of the spoken Kazakh language in China.....	27
2.1. Phonetics.....	28
2.2. Lexical new words.....	35
2.2.1. Lexical words that existed earlier but got lost in modern life.....	35
2.2.2. Words specific to Ch-K.....	37
2.2.4. Derived dialect words.....	41
2.3. Words/phrases that have different meanings among the speakers of the Kazakh language in China in comparison to Kazakhstan.....	42
2.4. Overall discussion.....	48
2.5. Conclusion.....	50
Chapter 3: Language attitudes and ideologies of ethnic Kazakhs from China towards Kazakh linguistic variation.....	52
3.1. ‘Pure’ Kazakh language.....	56
3.2. The erasure of regional differences.....	59
3.3. Fractal recursivity.....	65
3.4. Erasure related to the ‘purenness’ of the Kazakh language.....	69
3.5. Conclusion.....	71
Chapter 4. Conclusion.....	76
5. References.....	79

Chapter 1. General Overview

1.1. Introduction

This research aims to identify the linguistic differences between the Kazakh spoken in China (C-K) and Kazakh spoken in Kazakhstan (K-K) and investigate the perceptions and attitudes of Kazakhs born in China and currently living in the south of Kazakhstan towards the Kazakh language variations. First, the Kazakh language and its varieties are presented, followed by the historical and linguistic analysis of one of such dialects, namely the spoken Kazakh language in China. There the historical and political overview is discussed, adding information on the educational and linguistic background of the situation in Xinjiang with a focus on the ethnic Kazakh people. By referring to the research done on spoken Kazakh in China, the linguistic characteristics and existing language attitudes are demonstrated. I then discuss my own project work and approaches used to accomplish the goals of the thesis. Specifically, first, the term "dialect" and the research areas of "language variation and attitude" and "perceptual dialectology" are introduced with further elaboration on the existing studies. The methodology of this study is explained by further clarifying the techniques and procedures used in previous studies in the field of language documentation and perceptual dialectology. By referring to the earlier research done on spoken Kazakh in China, the linguistic characteristics and existing language attitudes are demonstrated. Finally, the methodology of this study was explained by further clarifying the techniques and procedures used in previous studies in the field of language documentation and perceptual dialectology.

1.2. Literature Review

1.2.1. Kazakh language and its variations

The Kazakh language is a Turkic language of the Kipchak branch and an official language of Kazakhstan. The number of speakers of this language is estimated to be 13.4 million in Kazakhstan itself, while according to the latest statistics, the number of ethnic Kazakhs residing in Xinjiang totals 1.6 million (Qazaqstan Respwblıkası Strategiyalıq josparlaw jäne reformalar agenttigi, and Ulttıq statistika byurosı 2022: 23; National Bureau of Statistics of China 2022: 25). The grammar books of the Kazakh language provide descriptions of phonology, morphology and syntax with the examples of the modern usage of spoken and written Kazakh varieties (Muhamedowa 2015), while Dotton and Doyle also include examples and reference charts Arabic-script used in Xinjiang and the new 2017 Latin script (2008). Krippes's 1996 Grammar of Kazakh Language emphasizes affixes and their types, while shorter articles by Kara (2002) and Kirchner (1998a) aim to cover the major aspects of Kazakh grammar. Even though there is a disagreement in terms of depicting the sounds of the Kazakh language via IPA (International Phonetic Alphabet), which is a systematic representation of the language in a uniform phonetic notation description, all of the studies have tried to systematize the phonetic and phonological characteristics of the Kazakh language (McCollum & Chen 2021; Vajda 1994). For the purposes of my study, the works that provide detailed descriptions of Kazakh sounds using IPA prove to be invaluable and accessible such as the research conducted by Dotton and Doyle (2008) which includes possible variations and glossing with transcription. Conversely, there are works like that of Janpeyisov et al. (2002) which describe sounds, meanings, and pronunciations in Kazakh in detail, albeit without IPA. It's important to highlight that researchers

have utilized various versions of the Latin script in syntactic analyses of Kazakh, resulting in inconsistencies in grammar and sound descriptions.

The dialects or varieties of the spoken Kazakh language were first claimed by Amanzholov to be three in number and differ geographically (1997: 171–600). Before that, according to Sawranbaev and Bolatov, most specialists of the Kazakh language, relying on the opinion of scientists such as W. Radlov, P. Melioransky, N. Ilminsky, A. Pozdneev, came to the conclusion that there are no dialectal features in the Kazakh language (1957). Later, based on the expedition results, Amanzholov explained that the geographical dialectical continuum resulted from the Three Juz (Hordes) of the Kazakh steppe who had earlier been the speakers of different but related Turkic languages. He believed that even though throughout history these languages underwent various changes through areal convergence with neighboring Turkic and Kazakh varieties, the division of *Juz* (Hordes) should not be ignored. On the basis of materials collected during expeditions between 1937 and 1952, a list of lexical, phonological, and grammatical characteristics was compiled for each of three dialectical regions: the south, the northeastern, and the west (which correspond to Great, Middle, and Small Juz/Horde, respectively). According to this division, northeastern dialect speakers reside in the area starting from the Tarbagatai and Altai to the Ural region, while people who speak western dialects are mostly situated in the western region of Kazakhstan starting from the southwestern borders of Kazakhstan up to the Ural River (Amanzholov 1997: 169-232). The third dialect covers the southern area of Kazakhstan starting from Almaty city until the Kyzylorda region. While Amanzholov's work offers valuable insights for data comparison, it is outdated due to the fieldwork being conducted half a century ago. However, it still holds significance for dialectology studies and serves as a reference for comparison. For example, the geographical differentiations in dialects detected by

Amanzholov were later accepted by other researchers of the Kazakh language such as Muhamedowa (2016: 19), Sawranbaev & Bolatov (1957).

All three dialects have their own lexical, grammatical, morphological, and phonetic characteristics even though they do not differ hugely and can be understood by the speakers of other Kazakh dialects. For example, the northeastern regional dialect became the standard variety of the Kazakh language as most of the prominent Kazakh writers came from this region and contributed to the written literature (Amanzholov 1997: 245–251). Therefore, this variety was taken as a basis to compare with other dialects: the way of speaking in the south region included borrowings from the Uzbek language as they were in intense language contact, preserved archaic versions of verbs, using different affixes, phrases with different case suffixes, and others (Amanzholov 1997: 314–374); the westerners' speech was differentiated by having more [ʃ] sound instead of [s], use of plural marking, lexical variations similar to southern dialects, and other features (Amanzholov 1997: 375–447).

In addition to the aforementioned three main dialects of the Kazakh language spoken in Kazakhstan, other varieties of the Kazakh language exist which are spoken by minorities in other countries. This includes areas such as Xinjiang (China), Bayal-Ulgii (Mongolia), Astrakhan, Samara, Orenburg, Omsk oblasts (Russia), Golestan province (Iran), Tashkent and Karakalpkstan (Uzbekistan), etc. As a result of language contact with the languages of these nations, the spoken Kazakh language in these regions is anticipated to have undergone changes in its linguistic features. Similarly, the spoken Kazakh language in China can be categorized as one of these variations of the Kazakh language. Though the official orthography used in Kazakhstan today is Cyrillic, ethnic Kazakhs in China use the older Arabic-based script. To expand more on this topic, and why Kazakhs in China still use the Arabic script, the background information about

Xinjiang Kazakhs should be given with historical details.

1.2.1. Kazakhs in China

This section reviews the Kazakhs in China, their historical background, precarious political status since the twentieth century, and how these historical tensions resulted in linguistic landscape changes in the Kazakh language. Additionally, it discusses how language contact of Kazakh with other languages (e.g. Mandarin Chinese and Uyghur languages) influenced the spoken Kazakh language in China by giving an overview of the research done on them.

Xinjiang is an autonomous region of the People's Republic of China located in the northwest area where Han Chinese (10.9 million) and Uyghurs (11.8 million) make up most of the population. being with a higher population (National Bureau of Statistics of China 2022: 25). Similar to Uyghurs, Kazakhs of China are called an ethnic minority (*minzu*) in Xinjiang Uyghur Autonomous Region and are estimated to be 1.6 million in 2020 according to the Xinjiang Yearbook (National Bureau of Statistics of China 2022: 25). The Kazakhs have begun moving to the Xinjiang region starting from the eighteenth century because of the wide range of available pastures in these regions (Benson & Svanberg 2016: 11), or with the aim of returning back to the ancestral home of Ili Valley they had been pushed out of during 1716-1723 (Salimjan 2021: 113). Moreover, a significant number of Kazakhs migrated to China at the beginning of the twentieth century when the Russian Empire declared a law calling for the conscription of Central Asian men aged 19-43 years into labor battalions during World War I. This was the main reason for many Kazakh revolts, which ended unsuccessfully for local Kazakhs who were forced to move to the eastern side of Kazakhstan, i.e. to what is today parts of the Xinjiang region of China.

According to Benson and Svanberg, the number of Kazakhs in Xinjiang in 1953 was 594 500, while by 1990 it had doubled (1998: 16). Among the existing eight prefectures in the Xinjiang region, the Kazakhs are commonly concentrated in the Ili-Kazak, Barkol-Kazak, and Mori-Kazak Autonomous prefectures, which were founded in the middle of the twentieth century as administrative units (Benson & Svanberg 1998: 128). Looking at the social organization of the Kazakhs who migrated to China, several main tribes can be highlighted. Benson and Svanberg talked about the six tribes or *ru* (*uru* how they wrote it) from the Middle Horde or Jüz: Kerey, Naiman, Waq, Kongrat, Qipchaq, and Arghin (Benson & Svanberg 1998: 46–47). Salimjan reinforces this view as well by documenting the descendants of the Middle Horde living in Xinjiang. Additionally, she mentioned that the tribes such as Naiman (Qizay), Alban, and Suwan tribes of the Senior Jüz in China had started to compile their ancestral genealogical *shejire* books (Salimjan 2021: 108). Thus, we can see that a major part of the Kazakhs in China came from the Middle and Senior Horde.

Despite the fact that initially, Chinese authorities wanted to keep the migrated Kazakhs at their new locations in Xinjiang, later the political and governmental changes worsened their situation which then led to the uprisings by Ospan (Osman) Batyr in the middle of the twentieth century. The main grievances were the restrictions on Islam, the attacks on mosques by Chinese troops, and the wish to retain their rights to weapons and pastures (Benson & Svanberg 1998: 72). Such complications led to socio-political riots, and Kazakhs fought against the Chinese Army. It should also be pointed out how the Kazakhs who moved to China, as a result of strict governmental rules which did not allow the Kazakhs to live freely as nomads, had to migrate further to India and Turkey (Noda 2019: 30). In the 1950s, some Kazakhs returned to the Soviet Union, aided by the USSR in migration and conflicts against the Chinese government. Their

decisions were also significantly influenced by Ospan Batyr's rebellions. The inter-ethnic and nationalistic problems between Kazakhs and Uyghurs were also a factor for local Kazakhs of Xinjiang to migrate to other locations where special associations existed that accepted Kazakhs such as India, Pakistan, Iran and Turkey (Noda 2019: 38). Benson and Svanberg also addressed this issue, primarily drawing from archival documentary records of the Chinese government (1998: 38). These sources detail the tensions experienced during the twentieth century and provide insights into the political context of Kazakhs residing in Xinjiang.

Both Nationalist and Communist regimes of the Chinese Government led the ethnicities in Xinjiang to feel powerless but at the same time brought a sense of nationalistic identity and a strong urge to protect the freedom they had before. Initially, the Chinese government's policies aimed to bring different ethnic groups in Xinjiang together without making them lose their own identity which can be illustrated in Article 53 of the September 1949 Common Program saying "minorities should have freedom to develop their dialects and languages, and to preserve or reform their traditions, customs, and religious beliefs" (Dwyer 2005: 7) . However, after the 1980s the focus shifted from cultural accommodation towards an overt policy of assimilation (Dwyer 2005: xi). This reinforced ethnic nationalism and created tensions not only between the government and ethnicity but also between the ethnicities by leading to an even higher level of nationalistic identity in XUAR. Being the highest population of Xinjiang, the Uyghur nationalists openly expressed skepticism and concern over China's perceived aims for regional development, which are the promotion of monoculturalism through cultural assimilation towards Han Chinese (Dwyer 2005: xii). Thus, the issue of freedom and nationalistic identity was relevant not only to Kazakhs but all ethnicities in Xinjiang including Uyghurs, Kyrgyz, Tatars, and others.

Noticing the increase in nationalistic movements and instability, the Chinese government implemented new changes in language policy in the form of changing the writing orthography and later changing the language instruction to Putonghua (Standard Mandarin Chinese). According to Benson and Svanberg, the introduction of different versions of the writing system several times brought new problems, as every time literate people became illiterate, as they had to learn new scripts (1998: 174-175). As a result, each generation learned different scripts: traditional Arabic script was used before 1949 until the Chinese government officially announced that the Cyrillic should be used. This was made because the Kazakhs in the USSR moved to the Cyrillic version in 1940. However, a modified version of the alphabet was introduced for Kazakhs in China. In 1958, to facilitate the unification among the minorities, it was suggested by the Chinese authorities that Kazakhs and all other minorities use only Latin scripts in China. Various books, newspapers, and dictionaries were printed in this new Latin script until 1981 when Kazakhs had to return back to Arabic script. The intention was to foster cultural expression among minorities (Kim 2020: 790). The consequences of frequent changes resulted in a limited number of books in one standard script. Nevertheless, there has been a growing demand for literary works, leading to a new wave of books about oral folklore, folktales, and songs, that were preserved orally from generation to generation (Benson & Svanberg 1998: 174-175). In addition, it should be noted that this quick change in scripts and having different variations of writings could influence the language, and therefore this may have generated variations in the spoken Kazakh language of Chinese Kazakhs.

Other factors that could influence the language variation of the Kazakhs in China can be the increasing number of students getting an education in the Chinese language, and the language contact of Kazakhs with other nationalities, specifically Uyghur and Han Chinese speakers. A

new system of “bilingual education” starting from the end of the twentieth century that reduced Kazakh to native language classes while promoting Mandarin Chinese as the mode of instruction for all other subjects brought the students in close contact with the Chinese language. Statistics have shown that the number of students among minorities in the Xinjiang region pursuing education in Chinese has increased, largely due to the potential for better employment opportunities in the future (Zhang & Yang 2021: 53-55; Kim 2018). Furthermore, speaking in Mandarin Chinese for Xinjiang Kazakhs became a usual, "modern" and adapted norm showing that this language is constantly used through code-switching (Salimjan 2017: 268-269). For example, in the Kazakh traditional duel (*aitys*), the participant (*aqin*) tried to imitate Chinese Kazakhs by code-switching to Chinese phrases by hinting *at the* differences in speaking the Kazakh language by ethnic Kazakhs in China (Salimjan 2017: 268-269). Furthermore, the impact of the Uyghur language was also detected, since this Turkic language has a high level of contact with Kazakh speakers in Xinjiang, so Kazakhs understand the Uyghur speech though they may not fully be able to express themselves, and this may have significantly influenced the spoken Kazakh language. Because of that, Abish (2016: 1) has compared the Kazakh language as it is spoken in China with the Uyghur language by demonstrating the impact of the language contact. Therefore, the particularities of the spoken Kazakh speakers who came from Xinjiang may be influenced by Uyghur and Mandarin Chinese languages.

1.2.3. Research on Kazakh language variation in China

Following an examination of the political and linguistic circumstances of ethnic Kazakhs in China, it becomes relevant to delve into studies conducted on Kazakhs of this region and the

linguistic characteristics specific to their dialects. This section provides an overview of the situation of the ethnic Kazakhs in China and the studies done on them from a linguistic perspective. Furthermore, it pinpoints the lack of linguistic analyses of phonology and lexical differences in these studies and how my own research will try to contribute to the discussion.

Starting from 1997 the Kazakhstan Government helped ethnic Kazakhs to migrate to Kazakhstan by assisting with applying for citizenship and helping them financially. It was observed by Kazakhs in Kazakhstan how the Kazakhs from Mongolia and China use the Kazakh language and have preserved the cultural traditions to a greater extent which marks the difference in using the language among the local population and repatriates (Werner et al. 2017: 1566). In addition, previous research has shown some difficulty with learning a language and having differences in speaking upon returning to the original homeland for repatriates. This list of challenges included the hardships of speaking in a different way, which was the case for Pontain Greeks when they faced the Modern Greek language (Mariou 2020: 8). In another study concerning Chinese repatriates returning to Hong Kong from America, Australia, and Canada, local residents displayed no favorable attitudes towards these individuals. The underlying reason was attributed to the migrants' accents, which were perceived as being strongly influenced by languages originating from the countries they had returned from which shows the presence of the detectable changes in the speech of repatriates (Sussman 2010). Since in Kazakhstan the Russian language became a dominant language in communication along with Kazakh, while the Kazakhs in China had language contact with Uyghur and Chinese languages, there could be noticeable variation between spoken Kazakh speeches from these two regions.

Studies about spoken Kazakh in China do exist but in quite a small number (Abish 2016; Kim 2018; Abish and Csato 2011; Mustafaulı 2003; Mäsımıxanuli 2004; Naqısbekov 1982). One

study looked at expressions of modality and compared them with Uyghur and Chinese languages as the relevant contact languages (Abish 2016). In a similar vein, Kim (2018) has investigated the discourse field of linguistics in relation to the Kazakh language spoken by Chinese Kazakhs by indicating the strong need to know, understand, and speak the Chinese language as it is the language of the job market. In his fieldwork study in the Xinjiang region during 2012-2014, he examined the ethical management among the Altai Kazakhs by concluding that honorific speech is used to display a certain image of modesty rather than a sociological obligation to speak in this way (Kim 2018). Also, he noticed ideological and perceptual comments by ethnic Kazakhs towards Mandarin seeing it as merely a language of mundane concern (Kim 2022: 7), and sounding loud and too simple (Kim 2018: 88-126). This point was also echoed in a study by Abish and Csato (2011: 278-280) who showed how the older generations of Kazakhs in China are not happy with the new tendencies in the language. Namely, the changes are including phrases and words from Chinese languages even though there are alternative words in Kazakh language and how this kind of speech can be called a "contaminated variety" and a sign of a lack of respect, which clearly can be identified as a negative attitude. According to the same study, the linguistic practices adopted by young Kazakh speakers illustrate the insertion of Chinese words and phrases and incorporate Uyghur words into the Kazakh language. Thus, it can be seen that spoken Kazakh in China is mostly examined from a sociological or discourse-analytic point of view with the presence of certain language attitudes. My own research tries to further elaborate on this topic by adding more discussion on how the Kazakh language is perceived by the ethnic Kazakhs in China.

1.2.4. Dialects and perceptual dialectology

The term "dialect" is distinguished as being a variety of the same language where the speakers of different varieties can understand each other (Payne 1997: 18). Though there are nuances surrounding the precise usage of the terms "language" and "dialect," the latter has been observed to be usually connected with low status and being substandard of a certain variety which creates an ideological perception of being "not correct" (Chambers & Trudgill 1998: 3). The differences between dialects may happen at every level of linguistic characteristics – phonology, morphology, syntax, phonetics, discourse (Myers-Scotton 2006: 17-18; Chambers & Trudgill 1998: 18–19), and generally, two main division groups exist in the dialect continua – geographical and social. In the former, the dialects are separated based on regional variations, and the latter can be an example of the socially structured phenomenon where the community has divided the language variations based on social status. A geographical dialect continuum concerns the variations of a language that differ based on location (Payne 1997: 19) and region (Myers-Scotton 2006: 16-34) and depends on how far the groups of people are located from each other, and how isolated these communities are (Chambers & Trudgill 1998: 5). In a social dialect continuum, the people of different classes, for example, higher and lower class, speak different variations of the language which can lead others to have some specific ideological beliefs (Labov 1972: 51-69). Social factors that influence the dialects to be defined as languages can be national, religious, and cultural borders through which the same language can be differentiated from each other (Myers-Scotton 2006: 20-23). Even in the same geographical location, people may try to imitate certain variations of a language due to the socially existing language varieties (Payne 1997: 19). Thus, it can be seen how the dialect is a variety of the language that is spoken

in a different region or by a certain class of people in comparison to a standard version of the language.

Since the main term ‘dialect’ was clarified, the focus can now shift to introducing the specific area under investigation in this research, namely perceptual dialectology, also referred to as folk dialectology. It is a branch of social dialectology, and it investigates practical (social) problems by focusing on conscious evaluations of language varieties (Inoue 1999a: 147-148). The variations of the language can be perceived differently by people from certain geographical locations or of different social statuses. Early research has investigated topics like the differences in perceiving the language varieties by focusing on the social stratification of English in New York (Labov 1972), the linguistic structures of grammar, pronunciation, and words as in the Linguistic Atlas of New England (“LANE”, Kurath 1939), or, still in this line of research, the perception of English in England and in the USA (Wikle & Bailey 2019). In later works, new methods were applied due to the advancement of technologies that allowed better dialect images and maps through the addition of multivariate analysis to the data generated through computer technologies (e.g. Inoue 1999a). Other works tried to mix the results of folk linguistic investigations with the research models on late classical language attitude studies by adding “cognitively real” mental maps (e.g. Preston 1999: 131). Earlier studies in perceptual dialectology primarily utilized Preston's methodology, which he initially applied to examine Hawaiian perceptions of United States dialects. Subsequent investigations explored the perceptions of Southern Indiana locals regarding "correct" English language usage, as well as attitudes toward language varieties in Brazil and other regions (Preston (ed.) 1999). Then following his 5-step principle in studying the dialect perceptions, other researchers started applying his method to their own works. For example, Benson (2003) uses Preston’s method to

look at the people of the Ohio region and how they see speech in relation to other communities, while Chen (2018) constructed his experiment in a way to see the linguistic differences between standard Mandarin and Taiwan Mandarin based on the comments of the respondents. Buchholtz et al. (2007) followed the general methods established by Preston and investigated the perceptual dialectology within California in detail to see the salient features between the regions of the state. Similarly, my own research follows Preston's methodology and adapts to investigate the attitudes towards the linguistic differences between the spoken Kazakh language varieties in China and Kazakhstan.

Dialect attitude and perceptions are closely related to language ideologies which can be defined as “sets of beliefs about language articulated by users as a rationalization or justification of perceived language structure and use” (Silverstein 1979: 57). The language ideologies, as was indicated by Woolard and Sheefiln (1994), can be analyzed through different frameworks such as ethnography of speaking, multilingualism and literacy studies, linguistics and public discourse on language, even metapragmatics. In their study about language ideologies, Judith Irvine and Susan Gal (2000) show three semiotic processes that are central to the production of language ideologies and explain each of them based on the examples. *ICONIZATION* is used to describe a process when the linguistic features are the iconic representations of the indexed social groups, while *FRACTAL RECURSIVITY* is used to describe the process when a salient feature from a certain level is projected onto another level through a recursive structure. *ERASURE* happens when the features and indexical associations are eliminated intentionally by the social group to justify certain ideological beliefs (Irvine & Gal 2000: 38-39). These are the main approaches to how language ideologies may occur or be constructed based on indexicality, and in each case, it is

required to understand which specific semiotic process is happening in the society to properly understand the indexical associations.

Studies on language ideologies also include the semiotic processes and consider their appearances in languages around the world by providing examples of indexicality. For example, Blommaert (2008) discusses the Artefactualisation of African languages, and how research and textual production of these languages were bound to ideological understandings of ‘form’ and ‘language’. Another example was discussed by Andronis (2004), where Quechua language has been associated by Ecuador people with being ‘rural’ and ‘backward’ due to iconization. A study conducted by Chen (2018) contrasted the standard Mandarin Chinese language with Taiwan Mandarin where the author relied on the information from dialectal maps, self-reports, and metalinguistic comments by the participants. As a result, it was presented how Taiwanese people were rationalizing the language structure based on the ideologies that would be indexical of Taiwan Mandarin in contrast to the Mandarin of the PRC. Moreover, the author presented the distinctions between non-retroflex and retroflex being used within Taiwan and how they are used to highlight the linguistic boundaries between the Mandarin varieties across the Taiwan Strait (Chen 2018: 10-21). My study follows the same methodology as it tries to differentiate the linguistic features of the Kazakh dialects spoken in China and Kazakhstan by focusing on the meta-analysis and comments made by the speakers towards the spoken Kazakh varieties.

Previous studies in perceptual dialectology considered language attitudes, mapping of the dialect borders, investigation of specific linguistic features that influence dialect perception, consideration of social status, and regional division. Two factors were highlighted to impact the formation and creation of subjective dialect division: individual (the first-hand experience of people) and collective (the stereotypical beliefs coming from others) (Inoue 1999b: 161-176). In

other words, the respondents may just assume and think that the dialect division exists based on rumors and stereotypes or can actually divide the dialect borders based on their experience, a distinction which is important to consider. Moreover, the method of mapping language variation and changes was used a lot in studies of the English language in the USA, Canada, and England, and the Japanese language in Japanese prefectures (Inoue 1999b: 161-176; Preston 1999 (ed.); Wikle & Bailey 2019). Thus, by applying this method of mapping, my research provides information on the dialect borders in Kazakh language variations and dialects both in Kazakhstan and in China.

1.3. Research Questions

My study tries to further contribute to this discussion and see what language differences and beliefs exist regarding the Kazakh language spoken in China. The research starts by identifying and comparing the linguistic differences on the phonetic, lexical, and meaning levels by contrasting the spoken Kazakh language varieties of Kazakhstan and China. Thus, the research questions of the first part of the project are:

- What are the linguistic differences between the spoken Kazakh language in China in comparison to the spoken Kazakh language in Kazakhstan?
- Are there distinct phonetic, lexical, and meaning differences?
 - If yes, what are they?

Next, the research has a goal of discussing the language ideologies and comments based on the recordings and statements given by the speakers during the fieldwork. By applying the methods used in perceptual dialectology, this research tries to conduct a first-hand investigation

among the ethnic Kazakhs addressing language use, perceptions, and ideologies. The research questions for the second part are as following:

- What is the language attitude of ethnic Kazakhs towards the spoken variations of the Kazakh language?
 - What are the perceptions of ethnic Kazakhs to different variations of the Kazakh language?
 - What are the variations based on their perceptions of the Kazakh varieties?

In other words, the objective of the project can be described as registering the divergence of the Kazakh language variety in China in comparison to other variations and detecting the language attitudes to the Kazakh language variations from Kazakhs in Kazakhstan and China.

Initially, the project also had the aim of constructing a dialect map or dialect borders based on the comments provided by the participants of the research regarding the language perceptions and attitudes towards the Kazakh language varieties with the help of mapping techniques. It was expected that this would lead to a dialect map that was not considered before in Kazakh dialectology - creating a perceptual dialect image generated by the Kazakh speakers using the classical methods of perceptual dialectology introduced by Preston (1999). For that, the following questions were asked during the fieldwork:

- How do they visualize the dialect map with the geographical borders in their minds?
- How can the dialect map be constructed based on the generated data and given comments by the participants?

Unfortunately, during the fieldwork, the results for these research questions were not attained during the fieldwork. The participants of the interviews encountered difficulties in

reading and understanding the map, making it unsuitable for the purposes of the thesis, as detailed later in Chapter 3.

1.4. Methodology

As the border between Kazakhstan and Xinjiang was locked between 2019 and 2023 and due to the difficulty of visiting this region and recording the spoken Kazakh language by these people, migrants from Xinjiang are relied on as the main source of data. The fieldwork took place in villages such as Shelek, Jana Turmys, Bahkar, Qara Jota in the Almaty region as the ethnic Kazakhs from China, specifically Tekes county of Ili region are concentrated there. In addition, a list of frequently used words among the ethnic Kazakhs of China was provided by Abish Aynur (which is used for documenting Kazakh spoken language in the project “the Resource and Protection of Languages in China” during 2018 that has not been published) that was compiled based on the speech of Altay and Tarbagatay people.

1.4.1. Analysis of the linguistic differences

The first goal of this thesis is to study the linguistic (phonetic, lexical, meaning) differences of the spoken Kazakh language in China and Kazakhstan. This first requires a standardized writing system that would allow us to register and illustrate the sound system and the lexical and morphological structures. The standard Kazakh language is depicted through the Cyrillic alphabet, but Kazakhs in China use the Arabic script. Also, as mentioned, previous

studies used different versions of the Latin script, which even further complicates the situation as no unified writing system exists for now.

Looking at the techniques implemented by researchers who examined the Kazakh language variation in China, methodological differences can be found, too. Specifically, Kagan claims that he wrote quotes orthographically how they were pronounced by the consultant, not how they should be written based on the standard version of the language during his fieldwork (1999: 4-7). To be specific, his work uses transliterations using a modified Latin script based on the Common Turkic Latin Alphabet (CTLA) (Kagan 1999: 3). Thus, the differentiated number of vowels (11 vowels) exceeds the standard Kazakh alphabetic script (9 vowels) which adds clarity and better understanding of the pronunciation, and full representation of the vowel harmony and how it is pronounced by the speakers.

The use of certain systems may also depend on the aim of the research, namely how precise the investigation should be. For example, Kim (2018) has created his own writing system using the Latin alphabet with the aim of being able to avoid special symbols such as umlauts (using diacritics) and using the phonemic system to grasp all the phonological expressions and sounds of the spoken Kazakh language by Xinjiang Kazakhs. This system and transliteration are much simpler and pronounced exactly how they are written with examples of approximate pronunciation from English. In comparison to works by Kagan (1999) and Abish (2014), this system does not include many varieties of sounds which may decrease the quality of being able to detect the differences between allophones but it allows the researcher to represent the spoken Altai Kazakh phonemically, and ultimately be able to write frequent Chinese terms that are present in the speech of Kazakhs based on the Pinyin system as well. A similar study done by Kirchner includes the text written based on the spoken language and how it was articulated

during the pronunciation of the speaker, not how it should be pronounced, and relies on the *Philologiae Turcicae Fundamenta* writing system (1998b: 189). The author himself warned that the linguistic inaccuracies and unusual sentence structures were preserved as original to depict the language differences which indicates the uniqueness of the spoken Kazakh language in China. Regarding the used methods, the author mentions that different notations of the same lexemes mean allophonic variations (Kirchner 1998b: 189), and the number of vowels (20) is higher than in other kinds of methods that were applied before in existing research which again shows that the method for sound registers depend on the focus of the research and how precise one wants to be.

Considering Abish's work which examined the linguistic feature of modality (2019), the author relies on the working framework of Lars Johanson and his model for describing the modal expressions in Turkic languages. For that, the transcription system developed by Johanson was used with Turcological transcription and transliteration. It is similar to the standard orthographic writing system of the Kazakh language and has some additional symbols such as the differentiation of [l] into voiced lateral approximant and voiced lateral velarized approximant ([l] vs. [ɭ]), and [y]/[w] sounds. The author decided to rely on Johanson's transcription system and also included the Arabic transcription as it is the alphabet used by the Kazakhs in China due to the ability to represent the Uyghur language, too, to highlight their similarities and differences with the Kazakh language.

Comparing all of the used writing system methods, in this study the IPA system was chosen as it is the common practice for describing languages and can be available to a larger number of researchers. In case of need, the relevant sources can be transferred to the IPA sounds as well, and respective phonemes can be indicated. In this way, my research with this writing

system will advance the studies in this field by questioning the standardization and better use of writing systems in depicting the sounds of the Kazakh language, especially for the spoken Kazakh language varieties. Furthermore, this study contributes to Kazakh linguistics by broadening the field of Kazakh language dialects, especially by adding more data about its spoken variations with the described linguistic features through IPA.

Moving to the etymological analysis of the detected differences to track back to the original version of the word or phrase, the Old Turkic dictionaries were consulted. Clauson's etymological dictionary of the pre-13th century can be used to understand the earlier etymological versions of the words or phrases (1972). Due to the ease of navigating it and listing based on the morphemes, Radlov's volumes can be handy (Radlov 1893a; Radlov 1893b;). It is not just big in terms of volume and amount but also easy to find the words based on the order of sounds. On the other hand, Dybo's list comprises 110 English words with simple meanings to assist in refining methods for comparative historical reconstruction of vocabulary (2013). Additionally, the etymological dictionary of Turkic languages with 7 volumes that lists and provides information on the common Turkic and inter-Turkic stems are relevant for this study as Servortyan, Dybo, and other prominent Turkologists have worked for more than 30 years to construct this dictionary (Sevortyan 1974; Levitskaya et al. 2003). It stands as the first work in the history of Oriental studies, in which the origin of common and inter-Turkic words is examined on the basis of materials from all modern Türkic languages (more than 25), their dialects, and monuments. In the case of loanwords, books that list the words in the Kazakh language that originated from Arabic and Persian languages can exactly direct to the source of the borrowed phrase (Mamyrbekova 2017). For the Kazakh dictionaries, the biggest one which contains 15 volumes was used to reference the words, phrases and their meanings (Januzaqov et

al. 2006; Qaliev et al. 2006; Cüyerqulova et al. 2006; Äşimbaeva et al. 2007; Janabekov et al. 2013). Additionally, a further volume was released subsequent to the completion of these extensive dictionaries of the Kazakh language. Since some words were missing from the previous volumes, Qaliyev decided to publish one more book which I also used to gain an understanding of some phrases (Qaliev 2014). Also, Kazakh-Russian, Uyghur-English, and Uzbek-English dictionaries were consulted to understand the language contact effects too (Awde et al. 2002; Schwartz 1992; Abdrakhmanov et al. 2002).

1.4.2. Investigation methods of perceptual dialectology

Secondly, this study has the goal of detecting the language attitudes and language ideologies of the ethnic Kazakhs born in China towards Kazakh language variations of Kazakhstan and China. For this, the Xinjiang Kazakhs' natural conversations were recorded where they were asked about their attitudes about the Kazakh language in China and Kazakhstan to understand their language ideology and were asked to draw dialect maps. This is based on Preston's five-point method where the first step in investigating perceptual dialectology studies is to ask the respondents to draw a map (1999). Thus, a simplified map of Kazakhstan with the borders of the Xinjiang region was given to participants with pencils. After getting opinions and perception comments from the participants with the indication of the degree of difference in general, which was the second step in the five-point method, the third phase was accomplished by allowing the speakers to express their opinions regarding the "politeness" and "correctness" of different Kazakh dialects spoken both in Kazakhstan and China. Then, to understand the dialect division, the speakers listened to recordings derived from the Multi Corpus Spoken

Kazakh Language Project (Nazarbayev University) that consisted of the three Kazakh language dialects (western, southern, and northeastern) to see if there are any kinds of comments regarding the language attitude and descriptions from the Kazakhs in China (Filchenko et al. 2023). That is a “Dialect Identification” step which is designed to understand more about the perceptions of the respondents about the samples from each dialectical region (Preston 1999: xxxiii-xxxvii). Finally, the last step “Qualitative data” consists of questions that were designed to open up more discussion and better understand the perceptual comments. For that, semi-structured interviews were conducted with the following questions:

1. What do you think of the speech played in the recording?
2. Do you think it is different from your way of speaking?
3. Do you think that there are certain regions in Kazakhstan where the way of speaking the Kazakh language differs from other regions?
4. If yes, which ones? (Asking for descriptions)
5. What do you think of them?
6. Can you draw these regions?

After conducting this perceptual investigation, in this study I try to draw the dialect maps and images based on comments made by the participants during the fieldwork. It allows further understanding of specific language descriptions assigned by the locals to certain language variations. For example, in the study by Preston, the possible explanations and suggestions were provided in regard to “pleasantness”, “standard”, “correct”, “good” or “normal” English characteristics (1999). Thus, this mapping technique is used in this study as it constitutes a valuable method to form an understanding of the Kazakh language variation perception from the respondents, and visually be able to see the borders of the dialects.

1.5. Conclusion for the chapter

This study has the goal of identifying the linguistic differences in the spoken Kazakh language in China and investigating the language attitudes and perceptions of ethnic Kazakhs from China towards the Kazakh language variations. In general, the paper has looked at the literature review, first, by focusing on the Kazakh language, and its varieties in Kazakhstan followed by the relevant studies of the spoken Kazakh language. Next, analyses of the historical, political, and sociolinguistic situation of the ethnic Kazakhs in China were provided to familiarize the audience with the background context of the study. Additionally, the research done on spoken Kazakh in China was surveyed to understand the gaps in the existing studies. After that, the research questions and methodologies of this study were explained with a discussion of the methods used in linguistic and perceptual dialectology and how I adapted them to my study.

The next chapters introduce the results and discuss the obtained data from the fieldwork. The second chapter is a linguistic analysis and comparison of the features of the spoken Kazakh language in China as compared to Kazakhstan obtained from the fieldwork recordings of migrants from Xinjiang to Kazakhstan. The third chapter contains the discussion of the language attitudes and perceptual dialectology resulting from the semi-structured interviews with the ethnic Kazakhs born in China.

Chapter 2. The results and discussion of the linguistic differences of the spoken Kazakh language in China

This chapter presents the results of my study on the linguistic differences between the Kazakh language as spoken in China and the one spoken in Kazakhstan. The study was done based on the recordings gathered during my fieldwork among the ethnic Kazakhs of China, where their everyday natural speech and conversation were recorded, and then analyzed to investigate the differences of language varieties at different analytic levels such as phonology, lexical structures, and derived/compound dialect words.

The participants of the project were speakers of the Kazakh language who came from the Xinjiang region and were older than 20 years. The total length of the recordings is 10 hours 37 minutes in which overall more than 10 speakers have participated. This material was used to detect linguistic differences such as phonetic distinctions, lexical, compound words, and phrases from the speech of the speakers of the Kazakh language in China by contrasting it with the one in Kazakhstan. The latter one comes from the standard Kazakh language which is a dialect spoken in the Northeastern region of Kazakhstan (Amanzholov 1997: 245–251). Moreover, a list of 1000 common words in the speech of Kazakh as spoken in China was provided by Professor Aynur Abish and was also incorporated and analyzed as a data source. What should be noted is that the Xinjiang region itself has different variations in the spoken Kazakh language of the dialectal regions such as Altay, Tarbagatay and Ili areas. The participants of my recordings primarily came from Tekes County in the Ili region of the Xinjiang Uyghur Autonomous Region. The list provided by Abish Aynur, by contrast, is based on the spoken Kazakh from the Tarbagatay and Altai regions. Thus, the results of this thesis represent the spoken Kazakh language of

Kazakh-speaking regions in Xinjiang even though it should be noted that the differences in the spoken Kazakh languages between these regions exist as well.

This chapter first discusses phonetic differences that were identified in the recordings. Secondly, it discusses lexical words that are not used by the Kazakhs in Kazakhstan such as dialect phrases and compound words. The next part is a discussion of words that have a slightly or totally different meaning among the speakers of Chinese Kazakh in comparison to the spoken Kazakh language in Kazakhstan. Then loanwords and calque phrases derived from Mandarin Chinese, Uyghur, Russian and other languages are presented with their respective original form.

2.1. Phonetics

The spoken Kazakh language in China (henceforth abbreviated as “C-K”) exhibits certain phonological differences when compared to its counterpart in Kazakhstan (henceforth “K-K”). For this work, the latter was based on the spoken Kazakh in the Northeastern dialect as it was considered the standard one (Amanzholov 1997: 245–251).

Table 1 illustrates a limited number of examples that illustrate the phenomenon, while the whole data set can be found in the appendix of this work. It's important to highlight that the spoken Kazakh in China encompasses three regional dialects (Ili, Altai and Tarbagatay), and the listed words may not fully capture the nuances of the varieties spoken in Altai and Tarbagatay since my participants are originally from the Ili area. Similarly, the data graciously provided by Professor Abish Aynur may not accurately represent the spoken Kazakh in Ili.

	C-K (IPA)	K-K (Cyrillic)	K-K (IPA)	Meaning	Phenomenon
1	oqʃaw	ұқсау	uqʃaw	to be similar	[s] > [ʃ]
2	aʃxana	асхана	asxana	kitchen, canteen	[s] > [ʃ]
3	jertaraq ¹	ертрек	jertjerek	earlier	Breaking vowel harmony
4	keramat ²	керемет	kjerjemjet	great	Breaking vowel harmony
5	masalan	мәселен	mæseljen	for example	Back vowel ⇔ front vowel
6	tøk	тоқ	toq	electricity	Back vowel ⇔ front vowel
7	ʃaptoʃ	шабдалы	ʃabdaʎ	peach	Shortening
8	kæresin	керасин	kjerasin	kerosene	Adjusting to Kazakh language phonology
9	pi:redzektir	пражектыр	pi:radzektyr	projector	Adjusting to Kazakh language phonology
10	mænti	манты	mantɯ	dumpling	Adjusting to Kazakh language phonology

¹ This way of pronunciation was found among the speakers of Ili region but might not represent how people from other regions such as Altai or Tarbagatay speak.

² This way of pronunciation was found among the speakers of Ili region but might not represent how people from other regions such as Altai or Tarbagatay speak.

11	birden <u>ke</u>	бірдеңе	birdenje	something	One sound difference
12	by <u>tkyl</u>	бүкіл	bykyl	whole	One sound difference
13	tɛi <u>q</u>	шық	ʃɤq	go out	Use of older version of [tɛ]
14	tɛ'æq <u>paq</u>	шакпак	ʃaqpɑq	flintstone	Use of older version of [tɛ]

Table 1. Examples of the phonetic differences between Kazakh in Kazakhstan and China

The first phonetic difference to be shown is the change of voiceless alveolar fricative [s] into voiceless postalveolar fricative [ʃ]. The example [1] in the table illustrates such case, and by comparing it to other instances from the appendix, it becomes apparent that the root term ‘*ұқса-*’ has become ‘*ұқша-*’. Looking at the Old Turkic form for this word ‘*ука*’, it can be detected that the K-K version is older, while in the C-K version, the voiceless alveolar became voiceless postalveolar (Radlov 1893b: 1614). The same happens with the root of the [2] – ‘*ауу-*’ where more words are made with this term. The Old Turkic version is ‘*ac*’ similar to the K-K pattern (Radlov 1983a: 532). Amanzholov (1997: 320-321) noted the sound change [s] > [ʃ] including [1] as a possible feature of the Kazakhs living on the south side of Kazakhstan which also includes the Kazakhs from China. This feature was also commented on by Sawranbaev and Bolatov where they illustrated many examples of the alveolar becoming postalveolar (1955: 48) based on the speech of Kazakh speakers from South and Eastern regions. Thus, it can be safe to say that there is a presence of change in the direction from [s] to [ʃ].

The Kazakh language is a Turkic language that follows the vowel harmony rule meaning that all the syllables should follow the same vowel system – back or front (Dotton & Doyle 2008: 11; Kara 2002: 12; Vajda 1994: 633). It is also called the palatal harmony where the following suffixes in a word depend on the preceding syllable nature: if the stem is back vs. front, then the syllables are back vs. front respectively (Johanson & Csato 2022:28). There are sometimes exceptions for loanwords (Johanson & Csato 2022:27). However, for some words it was noticed that Kazakhs from China tend to break this rule by pronouncing both front and back vowels in one word. [3] and [4] are examples of it where the first syllable contains a front vowel but the following syllable contains a back vowel, and then the next syllable is also back vowel following the last syllable rule. However, this does not happen in the spoken Kazakh language in Kazakhstan for the same words: all syllables contain front vowels.

What was also apparent is that there are some words that use opposite (front vs. back) vowels that are normally used in Kazakhstan. To be specific, a word with a front vowel is pronounced with a back vowel in China, and exactly the opposite happens in some words with a back vowel in K-K. For example, [5] illustrates the case when the word [mæselʃen] contains only front vowels, while speakers from China pronounce the same word with front vowels only. The word was borrowed from Arabic مَسَلَّةٌ (мәсәлә) where the K-K followed the vowel harmony based on the first syllable structure (front vowel), while C-K is pronounced with back vowels only (Mamyrbekova 2017: 658). The same process happened in [6], not only the vowels changed but the consonants also were influenced by the vowel nature: the front vowel is followed by the velar sound, while the back vowel changed the consonant into the uvular sound. That is natural to the Turkic languages to signal the frontness and backness not only with vowels but also with consonants (Johanson & Csato 2022: 27). The change in the vowels' nature was

also discussed and illustrated with examples by Amanzholov where he has shown that [æ] changes into [ɑ] and [e] in some words (1997: 332). Though the clear pattern of when it happens was not detected, the author claims that it might be related to the Uzbek language and its influence on the speakers of the Southern region in Kazakhstan (Amanzholov 1997: 332). As Naqısbekov has shown the same phenomenon was detected among the speakers of the Kazakh language in Shymkent which are also located in the South region of Kazakhstan and include a high number of Uzbeks in comparison to other areas of Kazakhstan (1982: 32). That might be relevant, and the Uyghur language could also affect spoken Kazakh in that region, as the word in [4] is pronounced as [karamæt] (*карамәт*) in Uyghur (Kibirova & Cunvazo 1961: 101). The Uyghur version contains back vowels that are followed by the front vowel in the last syllable. The last two syllables are strongly similar to how it is pronounced by the Kazakhs in China and also breaks the vowel harmony. This may suggest that the word was assimilated to and was influenced by the Uyghur language by altering its phonetic features.

The words can change their pronunciation and structure due to the influence of the language's contact with other languages such as Uyghur, Mandarin, Uzbek, etc. The vowel and consonant systems may get altered similarly as in [7] C-K [tʃaptoʃ] vs. K-K [ʃabdaɫɣ] ‘peach’, where the word was shortened and the last syllable changed. The word is borrowed from Persian *Šāftālu* [ʃæftɒlu] which means the same fruit as in the Kazakh language (Doerfer & Tezcan 1980: 194). What can be said is that the K-K version is more similar to the C-K version of the borrowed word. To be specific, an alveolar plosive was added to the beginning, and the word-final sound was deleted in the C-K version. Such changes can be suggested to be affected by the Uyghur language where the ‘peach’ is given in the dictionaries as ‘шаптула’ [ʃaptula] (Kibirova & Cunvazo 1961: 223) or ‘шаптул’ [ʃaptul] (Kambarov 2005: 5). In this case, the

similarity can be detected between the Uyghur language and the spoken Kazakh language in China rather than the latter with the spoken Kazakh language in Kazakhstan.

The next phenomenon to be displayed is the phonetic comparisons of the words pronounced among the Kazakhs in Kazakhstan and China where the former one was assimilated and influenced by the Russian language. For example, looking at the case [10], the word for ‘dumplings’ is pronounced as [mantʏɪ] in the Russian language which is similar to how it is done by the speakers of Kazakh language in Kazakhstan. However, all the back vowels were changed to front vowels by the Kazakhs in China. This word is present in different dialects of the Uyghur language in the forms of *манты*, *манту*, *манта*, *манто* and can be traced back to the Chinese *mantou* ‘steamed bread’ (Levitskaja et al. 1974: 32-33). Similarly, the words that came from the Russian language, as in [8] and [9], were pronounced without any changes in Kazakhstan, while these exact words were assimilated to the easier pronunciation and adaptation to the Kazakh language in China as they do not speak Russian.

Other characteristics of the spoken Kazakh language in China include one sound difference in comparison to how these words are pronounced among the Kazakhs in Kazakhstan. It can be seen in [11] and [12] where the consonants [k] and [t] are present respectively. [11] is not listed in the Old Turkic dictionaries (Sevortyan 1978; Räsänen 1969; Clauson 1972), and the Kazakh language dictionary list both versions of the words as existing and used in various spoken dialects of the Kazakh language (Abdrakhmanov et al. 2002: 176-178; Äşimbaeva et al. 2007: 42&71). The word consists of two parts *бip* ‘one’ and *деңе* which has its origin in the Persian word *تانه* *танә* meaning ‘body’ (Mamyrbekova 2017: 658). This word changed into *дене* in the Kazakh language while retaining the original meaning of the word. Thus, the C-K version consists of a new sound [k]. For [12], though similar to [11] both versions can be found in the

Kazakh dictionaries in spoken Kazakh (Äşimbaeva et al. 2007: 163), Dybo lists *büt-kül* as a general Turk word used today by Tuvans, Tatars, Bashkirs and Kyrgyz (2013: 23). Hence, the C-K version is similar to the older Turkic word which includes the sound [t] in contrast to how it is spoken in K-K.

Another resemblance of the Kazakh language as spoken in China with the Uzbek and Uyghur languages is the change of the postalveolar fricative to the affricate when it comes at the word-initial position. This can be seen in cases [13] and [14] where [ʃ] was pronounced as [tʃ] by also changing the vowels of the initial syllables of the words by fronting them. Such changes leading to the affricates were noticed before by Amanzholov (1997: 322) and Naqısbekov (1982: 23) where they listed the words that have pertained such differences in the initial syllables. Sawranbaev and Bolatov considered it to be the main phonetic phenomenon characteristic of the south-eastern dialects covering the region from Altai and Tarbagatay to the Syr River, distinguishing them from the group of northwestern dialects (1955: 47). According to the Old Turkic form, *çık* has the same meaning of ‘going out’, while the root of [14] *çak-* means ‘to strike fire (with a flint and steel)’ (Clauson 1972: 405-406). This suggests that the use of affricates among Kazakhs in China (C-K) represents an older version compared to the one used in Kazakhstan (K-K). Therefore, Kazakhs in China, including Uzbeks and Uyghurs, adhere to the older version of affricates.

In general, the phonetic differences between the spoken Kazakh language in Kazakhstan and China suggest that the C-K has pertained to the older version of the form, while in K-K affricates at the beginning of the words were changed from the voiceless alveolar fricative. Also, the words tended to be altered from back vowels to front vowels or in the opposite order. Despite

the fact that the Kazakh language is a Turkic language that follows the vowel harmony rule, it was detected that this phenomenon is not consistent among Kazakhs from China.

2.2. Lexical new words

One of the fundamental characteristics of the spoken language of a dialect is the presence of lexical words that may have a new meaning. These words may even be used only by the speakers of a certain region or social class. Each of the dialects of the spoken Kazakh language in Kazakhstan has such differences, and the ethnic Kazakhs of China have their own set of words that are understood only by them but not by speakers from other regions. This paper looks at such words that were detected based on the recordings and categorize them into different subgroups.

2.2.1. Lexical words that existed earlier but got lost in modern life

Among the speeches of the Kazakhs in China, some ancient words can be noticed that are no longer used in modern Kazakhstan. For the person who speaks Kazakhstani Kazakh, these words may not be familiar and sometimes are difficult to understand. Such words are presented in Table 2.

	Word	C-K (IPA)	Meaning	K-K equivalent
15	шөшіре	ʃøʃɪre	dumpling	тұшпара
16	қарбалас	qarbaɫas	busy	бос болмау
17	әсілінде	æsiɫinde	actually	негізінде

18	шаптау	ʃaptaw	to stick	жабыстыру
----	--------	--------	----------	-----------

Table 2. The words that can be found in the speech of Kazakhs from China but not among the Kazakhs in Kazakhstan

The answer to the question of why lexical words that lost their use in Kazakhstan still exist among the speakers of Kazakh language in China might have several reasons. However, the main explanation could be the fact that these people have been living mostly in isolation in Xinjiang without interacting with the speakers in Kazakhstan that leads to some words and phrases to be still saved and intensely used.

To understand the use and presence of the phrases now, first, it should be checked with the speakers of the dialect regions of Kazakhstan. The words in Table 2 were not described as existing in the Kazakh dictionaries, which means that they can be searched for their meaning from the earlier sources. The word for dumpling used in Kazakhstan today ‘тұшпара’ [tʃpɑrɑ] is a loanword from Uzbek *chuchvara* [tʉtʉfɑrɛ] (Awde et al. 2002: 133), while [ʃɔʃɪrɛ] like in [15] is only found in one dictionary of the Kazakh language without giving the proper explanation. In other words, the dictionary indicated that this word exists based on the piece of writing that was published in 1955-1957 but has not provided any meaning (Qaliev 2014: 673). What it means is that [15] exists, and they were used earlier, but now their use has diminished or almost gone extinct in Kazakhstan, and hence the author was not able to provide the meaning. Radlov’s dictionary lists the word *uöu-böpä* as dumplings (1911: 1039), and it can be suggested that the C-K version came from the Uyghur *çöçürä* ‘چۆچۈرە’ which took its form from Uzbek *chuchvara* (Schwarz 1992: 350). Similarly, [16] and [18] can be found in the dictionaries of the bigger volume with the same descriptions though [18] is indicated to be a dialect word (Kaliyev 2006: 497; Abdrakhmanov et al. 2002: 486; Abdrakhmanov et al. 2002: 940; Zhanabekova 2013: 258).

The case with [17] is that the back vowels of [ɑsɪɫɪndɑ] got reversed to the front vowels phonetically [æsiɪnde] similar to the phenomenon like in examples [5] and [6] (Abdrakhmanov et al. 2002: 107). If [17] is used in everyday talk and speech, [ɑsɪɫɪndɑ] is only used in speeches connected to the religious texts or explanations around Islamic stories in Kazakhstan. This can be supported by the dictionaries as well where the explanation was provided based on religious books (Abdrakhmanov et al. 2002: 77). Thus, we can see that the speakers of the Kazakh language in China still use some archaic words, while the Kazakhs in Kazakhstan have abandoned them by replacing them with other versions of the phrases.

2.2.2. Words specific to Ch-K

The spoken Kazakh language in China is a dialect of the Kazakh language spoken outside of Kazakhstan, and it has its dialect words that might not be understood by the speakers of other dialects. They may not be even mutually intelligible; thus, the speakers of the dialect should explain the meaning of the words or phrases. Table 3 presents such a list where some of the examples were extracted and gathered there from the appendix of this work.

	Word	C-K (IPA)	Meaning	K-K equivalent
19	абақай	abaqaj	snake	жылан
20	дәнеке етпеу	dæneke etpew	not to pay attention	мән бермеу
21	борлау	borlaw	fattening of the stock	семірту
22	басы көзін жуу	basɪ kɵzɪn dʒuw	to take a bath	суға түсу

23	суретке тартып алу	suretke tartyp aluw	to take a picture	суретке түсіріп алу
24	сөйлеп қою	söjljep qojuw	send an audio	аудио жазу

Table 3. The dialect words based on the speech recorded from the ethnic Kazakhs from
China

All these are words that I could not find in dictionaries of 15 volumes of the Kazakh spoken dictionary, and I needed assistance during fieldwork from the speaker of Kazakh in China to understand them. [20] and [21] can be examples of such instances when no reference can be found.

Another example of dialect phrases are taboo words. In every language, such restrictions can be found, but there are some taboo words that are used only by the Kazakhs in China. [19] means ‘snake’, while the term that is used among the Kazakhs in Kazakhstan *жылан* [dʒɯlan] is not allowed to pronounce. This implies the existing taboo, which was also discussed by Talaspaeva and Tursıngalı, where they translated this word as ‘grandfather/big snake’ (2021: 249). They have also highlighted that this phrase is used by the Ili Kazakhs, and provided other taboo words used for this term in other regions of Kazakhstan and Xinjiang: *түйме* ‘button/snap’, *ұзын аю* ‘long/tall bear’, *омарта* ‘hive’, *ұзын кісі* ‘tall human’ (Talaspaeva & Tursıngalı 2021: 249). On the other hand, the taboo phrase used for taking a bath [22] cannot be found in any of the dialects in Kazakhstan, and it literally means ‘to wash your head and eyes’.

The dialect phrases can be also illustrated based on the terms used for the processes related to new technological advancements. To be specific, taking photos with a phone or sending audio messages via social messengers were found to be different as illustrated in [23] and [24] that could be shown in (1) and (2):

(1) сурет-ке	тарт-ып	ал-у	сурет-ке	түсір-іп	ал-у
picture-DAT	pull-CVB	take-INF	picture-DAT	drop-CVB	take-INF
C-K: ‘take a picture’			K-K: ‘take a picture’		
(2) сөйл-еп	қо-ю		аудио	жаз-ып	жібер-у
talk-CVB	put-INF		audio	write-CVB	send-INF
C-K: ‘to send an audio’			K-K: ‘to send an audio’		

The main difference in the phrases of taking a picture is in the verbs used for describing the action. If in Kazakhstan it is said as ‘dropping’, in China the verb of ‘pulling’ is used. Otherwise, the grammar structure is without any changes. However, for sending audio messages in Kazakhstan the term audio message is used, while in China it is directly translated as ‘to talk’ which was transformed into a meaning of ‘sending an audio message’. What should be highlighted is that the verb phrase of verb and auxiliary is also not similar in both of them. The Russian *фотографировать* (*photo-graph-ivaj-ova-INF*) did not motivate K-K speakers to use construction as in [23], but *скинуть аудиозапись* (*send-INF audio recording*) could have influenced to use the particular pattern as in [24]. The main verb in the verb phrase is ‘to send’, while ‘to record’ is presented with converb and the audio is presented as a noun, the meaning of which in total indicates a strong similarity of K-K [24] with the Russian version of the phrase. On the other hand, Mandarin Chinese for [23] is 拍照 Pāizhào where each character represents ‘to stick’ and ‘photo’, for [24] 发语音 Fā yǔyīn ‘to send’ and ‘sound/audio’ showing that C-K version was not influenced by the Mandarin language in forming these expressions in Kazakh language.

2.2.3. Compounded words

The speakers of the language may create new words by compounding 2 or even more words and construct new lexical compound words with the meaning. This merging technique was also practiced by the speakers of the Kazakh language in China where new words were made by combining two separate words as in Table 4.

	Word	C-K (IPA)	Meaning	K-K equivalent
25	қанбасым	qanbasym	high blood pressure	жоғары қан қысымы
26	майқұрам	majquram	blood fat composition	қандағы жоғары май үлесі
27	желпарақ	dzelparaq	fan	желдеткіш
28	ауасорғыш	awasorgysj	hood	вытяжка

Table 4. Compound words recorded in the speech of ethnic Kazakhs from China

As shown in Table 4, [27] has a meaning of ‘fan’ and was formed by compounding the words ‘wind’ and ‘paper’, while [28] stands for a ‘kitchen hood’ and made of the separate words ‘air’ and ‘suction’. It's important to note that in C-K all of these words are formed through blending techniques, while in K-K predominantly Russian phrases are used for these phrases. However, people usually use the Russian equivalent phrases instead of Kazakh translations such as ‘давление’ for [25] or ‘вентилятор’ for [27]. Even for the term in [26], neither a Russian alternative nor a unified term for this phrase exists in K-K.

2.2.4. Derived dialect words

In the recordings, there were some words that were created via derivational morphology. To put it another way, the words are grammatically correct and carry the meaning according to the morphemes, but such words cannot be found in the Kazakh dictionaries indicating that they are only used by the Kazakhs in China.

	Word	C-K (IPA)	Glossing	K-K equivalent
29	мұғалімдік	muɣalimdik	teacher.LIK	teachership
30	сабақтас	sabaqtas	class.LAs	classmate
31	масалық	masaɫɣɯ	mosquito.LIK	mosquito net

Table 5. The examples of dialect words that have grammatically correct structures recorded from the speech of ethnic Kazakhs from China

In the example [30], the word is made of the root 'class' and the suffix *-LAs* that denotes common attachment based on the meaning of the core noun or belonging to a group, and hence becomes 'classmate' (Muhamedowa 2015: 265; Kara 2002: 32). Nevertheless, this word is uncommon to use in Kazakhstan where another form of this word is used. To be specific, the Russian equivalent of 'class' ('класс') is used instead of 'сабақ' by forming *классмат*. In [31], the suffix *-LIK* is used which has many usages in Kazakh such as denoting objects with specific purposes, forming adjectives from nouns, or creating general and abstract noun from nouns and adjective (-ness) (Muhamedowa 2015: 266; Kara 2002: 32). In this case to denote objects designed for specific purposes, thanks to the suffix the word became a 'mosquito net' from 'mosquito', whereas in [29] a general noun was formed from noun.

2.3. Words/phrases that have different meanings among the speakers of the Kazakh language in China in comparison to Kazakhstan

The spoken Kazakh language in China included the differences in the meanings of the words and phrases in comparison to how it is used in Kazakhstan. This section presents such differences and discusses how it might be changed or used in various contexts as shown in Table 6.

6.

	Phrase	IPA	Glossing	Meaning in C-K	Meaning in K-K	Comment
32	күріш <i>басу</i>	kyryʃ basuw	rice cook.INF	<i>to cook</i> rice	<i>to press</i>	additional meaning
33	контейнерге <i>басу</i>	kanteinergje basu	container.DAT put	<i>to put</i> in a container	<i>to press</i>	additional meaning
34	қияр <i>айдау</i>	qijar ajdaw	cucumber plant	<i>to plant</i> cucumber	to drive	additional meaning
35	бата <i>қылу</i> ³	bata qɯluw	blessing make	to give a blessing	K-K uses <i>беру</i> instead	auxiliary verbs use differences
36	ЫҚЫЛАС ҚЫЛУ ⁴	ɣqɯlas qɯluw	intention make	to give a blessing	lit. to make an intention	new construction

³ This form was found among the speakers of Ili region, while for other regions such as Altai and Tarbagatay K-K version is used

⁴ This form was found among the speakers of Ili region, while for other regions such as Altai and Tarbagatay [iqilas qoyuw] is used

37	беріп <i>ету</i>	berip etuw	give.CVB AUX	to send	К-К uses <i>жіберу</i> instead	auxiliary verbs use differences
38	көмектесіп <i>ету</i>	kømjektjesip etuw	help.CVB AUX	to help	К-К uses <i>жіберу</i> instead	auxiliary verbs use differences
39	тұмау	tumaw	-	sickness	flu	generic name for common illnesses
40	бүлдірген	byldyrgen	-	berries	wild strawberry	generic name for berries (strawberry, raspberry, mulberry)
41	далаға шығу	dalawa ʃyɣaw	field.DAT go	to go to the toilet	to go outside	taboo word for going toilet
42	байып жатыр	baɣɣp dzatɣr	get rich.CVB AUX.3	is ending	is getting wealthier	taboo word for saying that something is finishing up
43	аяқ	ayaq	-	shoes	leg	change in use

44	үй	уј	-	room	house	change in use

Table 6. The list of the words and phrases that have different meanings among the speakers of the Kazakh language in China in comparison to Kazakhstan

The first group of words that can be used differently includes the verbs. They might have additional meaning to the one how it is used in Kazakhstan, or a different kind of verb can be used instead. According to the dictionary, *басы* can have multiple meanings based on the context such as ‘to press, to walk, to cover’, but in Ch-K it also means ‘to cook’ or ‘to put into something’ (Cüyerqulova 2006: 32-38). In [32] this verb is used to indicate the cooking process which carries a new meaning not indicated in the dictionary. Moreover, even though in [33] it is about putting something inside of an object, according to the dictionary this word is usually used to put something inside a car specifically, not in other objects like in a container as in [33] (Cüyerqulova 2006: 34). Unlike in K-K the agent of the verb *басы* does not necessarily need to be an inanimate object but also may include people. For example, similar to [33] this verb can be used to construct a phrase like in (3).

- (3) Адам-дар-ды машина-ға бас-ып ал-ды-м
 people-PL-ACC car-DAT take-CVB get-PST-1SG
 ‘(I) got people into the car.’

Thus, what can be seen is that this verb has an additional meaning of cooking and putting people, items and things into an object such as a car, container, etc.

Another verb that acquired a new additional meaning is *aı̋day* which means ‘driving’ in K-K. Nevertheless, this verb is used in the agricultural field to mean planting, and it is not included in the description of this word in the dictionary (Januzaqov et al. 2006: 146). The example in [34] illustrates a case when this verb was used to refer to the planting process of cucumber. This could be suggested by the fact that earlier the sowing process happened using the cows or bulls that are driven by humans in a plantation field. Thus, the verb underwent a generalization phenomenon, so that after some time it also meant the sowing.

The differences were also detected not just in the verb but also in verb auxiliaries. In [35] the meaning of the phrase is to say blessings which is a common religious practice done before or after the meal depending on the region. In Kazakhstan, this blessing phrase is constructed with auxiliaries such as *беру* ‘give’ and *emy* ‘do’, while in C-K the auxiliary *қылу* ‘make’ is used instead. Similarly, in [37] and [38] the auxiliaries *emy* ‘do’ are used instead of *жиберу* ‘send’. It should be mentioned that although the auxiliaries are used differently, there is no change in the meanings of the phrases.

A combination of the word and a different auxiliary can create a new phrase with a new lexical meaning in C-K that is not associated with the meaning used in K-K. One may not be able to explain the meaning of it without actually asking for some help to explain that. An example of it can be [36] which has the same meaning as in [35]. In this phrase, *ықылас* means ‘attention, wish, sincere feeling, zeal, desire’ and derives from Classical Arabic سورة الإخلاص (Qur’an Surah 112), plus the auxiliary *қылу* ‘make’. This transformation highlights how the phrase evolved from its religious origins to convey a sense of making a wish or offering a blessing.

What was interesting is that some words are used as a generic term to the types of the whole class. To be specific, a type of class is referred to represent the entire group. One of such examples is [39] which literally means ‘flu’, but nonetheless can be used to refer to a ‘sickness’ in general. It might be due to the taboo reasons of not saying the sickness or disease directly, or not giving much attention to the sickness in order to avoid the topic in conversation. In the fieldwork, what was noticed is that ‘flu’ was referred to when the person had any of such symptoms: fever, cough, sore throat, runny nose, body aches, headaches, vomiting or diarrhea. It might not be necessarily flu, but the sickness would be described as having flu without getting into the details with someone with whom the casual conversation is going on. Similarly, [40] is a wild strawberry that can be found in forests. Nevertheless, it is used to refer to all kinds of berries in C-K, including strawberries, raspberries, and mulberries, all of which have translations in Kazakhstan. Nevertheless, it should be noted that both Russian and Kazakh versions are used in Kazakhstan.

Following the topic of generalization, the shortening of certain phrases can be illustrated which might seem different and unfamiliar for the first time. The same happened to me when [43] was used in a conversation such as:

(4)	АЯҒ-ЫҢ-ДЫ ки-Ø leg-2SG-ACC wear Kazakhstan: Wear your legs! China: Wear your shoes!	(5)	АЯҒ-ЫҢ-ДЫ жин-ап қой leg-2SG-ACC collect-CVB put Clear up your legs. Clear up your shoes.
-----	---	-----	---

What (4) and (5) illustrate is the use of [43] in a conversation where it is used as the meaning of ‘shoes’ by ethnic Kazakhs from China, while it means ‘leg’ in Kazakhstan. It also

means ‘leg’ in China but can be used as shoes as well by broadening the spectrum of its use. To indicate the shoes in Kazakhstan, one should use *аяқ киім* which is a combination of leg and cloth. Such semantic meaning differences can show how certain words can get additional meanings besides the one used in another region.

The topic of taboo, which is a social custom of avoiding pronouncing certain words because of embarrassment or offensiveness, can be also discussed in lexical terms, differences based on the practice of going to the toilet. In a conversation, people usually may want to inform others about their intention of going to the toilets but might find it embarrassing. That is why another phrase is used in this case [41], *далаға шығу* which literally means ‘to go outside’. It can be easily explained by the fact that during the nomadic lifestyle, there were no toilets. Thus, ravines far from the house (yurt) were used for such places. Later, even now living in brick houses, people usually construct the toilets not in the house but in the back side of the house separately. Hence, going outside implies going to the toilet without explicitly saying so.

Another example of a taboo can be illustrated in case [42], where using a construction with the opposite meaning is required. In this case, it's considered taboo to say "something is finishing up" as it might imply being poor. Instead, an expression with the opposite meaning should be used, such as "getting wealthier," as demonstrated in [42]. For instance, if the milk is finished, one should express the opposite sentiment: "We are rich with milk."

The visual and structural construction of the houses also make a difference which led to the linguistic and lexical differences in naming and calling the rooms. The rooms by the Kazakhs in China can be labeled as ‘houses’ like in [44] even though there are many rooms in the house. This difference was confusing during the fieldwork as *үй* ‘house’ refers to a separate building in

the understanding of the Kazakhs in Kazakhstan, while the same term can mean ‘room’ for the Kazakhs in China. The explanation for that can be the nomadic lifestyle where one yurt or house was itself a room. Also, one of my participants said that the design of the houses in China was constructed in a way that each room had only one door which leads to the outside only. In comparison to that, in Kazakhstan one room may have two doors leading either to outside or to another room. Based on this explanation, it can be argued that the phrases developed in different ways that reflect differences in lifestyle.

In general, what can be concluded about the words and phrases that have different meanings among the speakers of Kazakh language in China in comparison to Kazakhstan can be the additional meanings to the existing ones, use of the different auxiliary verbs, some words being the generic terms and existent of taboo words or phrases.

2.4. Overall discussion

The linguistic differences of the spoken Kazakh language in Kazakhstan and in China include variations in the phonological pronunciation of the words, the meanings of the lexical items, the presence of new lexical words. The list contains the words and phrases that were identified previously by other researchers to be phrases used by the Kazakhs in China. For example, some of the lexical phrases that were highlighted to be different from the spoken Kazakh language in Kazakhstan based on my recordings and by the list provided by Professor Abish Aynur were also noted previously by Amanzholov about the Kegen-Narynqol region (where ethnic Kazakhs from China migrate). His expedition members presented 8-word-list from their expedition from 1937 (1997: 213-214) and some of them, namely *сақаю*, *қошина*, *жозы*, *қыжалат*, were found

during my fieldwork too (see the appendix for the full list). These loanwords came from Uyghur and Mandarin Chinese languages that were integrated into the speech of Kazakhs in Xinjiang. The meanings are also the same as were indicated by Amanzholov, and what makes it even more consistent and valid is that the speakers in the expedition were also people from the tribe of Alban similar to the most participants of my recordings. Amanzholov also cited the works of Katanov who conducted fieldwork at Durbuljin in 1890-1891 (in today's Emin County in Tarbagatay, Xinjiang area). Katanov was commenting that in comparison to the Russian Kyrgyz people (today's Kazakhstan's Kazakhs) the local speech included the pronunciation of affricates at the word-initial positions and the change from back vowels to front vowels or opposite in commonly used words (Amanzholov 1997: 143).

The spoken Kazakh language in China has been identified by previous studies as belonging to the dialects of the South regions which was also partially supported by the data elicited from the fieldwork recordings. This might be evident from the phonological and lexical similarities. For example, how it was shown in phonetic section, the assimilation of [s] to [ʃ], the shift from back vowels to front vowels or opposite were also described as the features of the south region (1997: 320-321). Similarly, the presence of [k] like was also shown to be ascribed to the same region as in the example of [11]. Additionally, the pronunciation of affricates [tɕ] instead of [ʃ] at the beginning of the words was detected among the speakers of Kazakh language in the southern dialects by Amanzholov and Naqisbekov where they have provided a list of such examples. Adding the lexical words that can be present in the speech of the Southern dialect, all of these distinctions indicate the partial alignment of the spoken Kazakh language in China with the dialect of the South region of Kazakhstan.

2.5. Conclusion

The chapter has illustrated the differences of the spoken Kazakh language in China in contrast to how it is spoken in Kazakhstan by indicating phonetic differences, new lexemes, compounded words, phrases that have different meanings. The data was derived from the fieldwork done in Almaty region based on the ethnic Kazakhs from China (Ili region) by me and the list of the commonly used words based on the fieldwork done in Altai and Tarbagatay regions by Aynur Abish.

The phonetic differences of the speech from the Kazakhs China were found to be similar to the spoken Kazakh language of the South regions. The main distinctions were the change from [s] to [ʃ]; use of affricate instead of the postalveolar fricative at word-initial positions by keeping the older version; presence or drop of a sound ([t], [k]); not following vowel harmony; shortening and change of back vowels to front vowels (or opposite) in some words. Researchers had previously observed such distinctions too, and the findings from my recordings have served to both reinforce and offer further insights into this observation.

The next section discussed the lexical words and phrases in the speech of Kazakhs from China. First, the words that do not exist among the Kazakhs in Kazakhstan were considered. Such words and phrases were found to be used earlier but have stopped their function lately. Second, the dialect lexemes were illustrated by providing the meaning and structural explanation for each of them. Among such words there were grammatically correct compound words and phrases which are not used in Kazakhstan but have linguistic meaning and speech function in China. It was followed by the list of the examples for the words that have different meanings among the speakers of Kazakh language in China in comparison in Kazakhstan. The main

findings were that some words have attained additional meaning; the auxiliary verbs were used differently; generic names existed for illnesses and berries; taboo words and phrases were present which cannot be found in Kazakhstan.

In general, the spoken Kazakh language in China has shown similarities to the spoken Kazakh language in both south and northeast regions of Kazakhstan and also had its lexical and dialectal distinctions. It should be highlighted that the Kazakhs language as spoken in China is also divided into smaller dialects and they may have linguistic differences among themselves as well.

Chapter 3: Language attitudes and ideologies of ethnic Kazakhs from China towards Kazakh linguistic variation

This chapter presents and discusses the results concerning the language attitudes and perceptions of Chinese-born ethnic Kazakhs towards different Kazakh language varieties. These insights were gleaned from interviews conducted with speakers during fieldwork, where they shared their views on various dialects spoken in both Kazakhstan and China. The process included asking questions regarding their language ideologies and dialect regions on a map, commenting and providing feedback on the recordings which featured speech in various dialects of Kazakhstan. This chapter first presents the information about the participants followed by the demonstration of the results by interpreting the answers based on the language ideologies. To be specific, the ideology of purism is defined first which emerged after the Russification of the Kazakh language during Soviet time. Then it was shown how Xinjiang and Kazakhstan are considered as a whole separate region of language dialect. Finally, the semiotic processes of erasure and fractal recursivity are illustrated on the basis of the language attitudes by the Kazakhs born in China towards spoken Kazakh language in China and Kazakhstan.

The community of the ethnic Kazakhs from China where I lived during fieldwork were my in-laws who were born and lived in the Tekes County of the Chinese part of the Ili region but now moved to Kazakhstan. The participants of my interview were 3 women of this community who are now living in Almaty region of Kazakhstan and belong to the Alban *uru* or ‘tribe’ of Senior Horde/Juz. All of them were born in China and have spent their majority of their lives there. Getting a basic level education in their mother tongue then written in the Arabic script

system, all of them stayed at home and dedicated themselves to taking care of children. Speaker A and C are 48- and 45-years old women with 3 and 5 children respectively. They were abducted⁵ without consensus by an acquaintance they were seeing on the streets and got married at a young age (18-20 years), while Speaker B is older (65 years old) and had an arranged marriage at the age of 15 as she lost her mother during childhood. All of them speak only in Kazakh but can understand spoken Uyghur and Mandarin a little bit.

The interview process was not set up in advance, meaning that it went naturally where I introduced the project and explained the consent form asking whether they want to take part in it. After that, the interview questions were asked by playing the audio recordings from three dialect regions of Kazakhstan, namely the Western, Southern and Northeastern. These recordings were 2 minutes long and were derived from the Multi Corpus Spoken Kazakh Language project at Nazarbayev University (Filchenko et al. 2023). This project aims to collect and record the natural spoken Kazakh language, and that is why I have chosen the real and every-day life recording from the corpus of the project that would represent each dialect region. During the interview process with my participants, I also asked additional questions for clarification trying to get more information on their language ideologies and attitudes. Following Preston's five-point method, I had prepared the map of Kazakhstan and Xinjiang region to ask the respondents to mark the borders of the dialect regions in both countries. However, an issue arose when the participants asked me ‘Where are we on the map right now? Where is Xinjiang?’. These questions signaled

⁵ Bride abduction is a cultural marriage practice where a woman is forced to marry the kidnapper due to societal force and pressure (Koishigulova et al. 2014: 4970; Werner 2009: 314). The bride theft may happen with and without the consensus of the bride, and is usually done when the bride's father refused a marriage, did not want to give his daughter to the groom's family as a bride or requested a high amount of bridewealth (Argynbaev 1978; Martin 2001; Taizhanov 1995). The women who were abducted usually do not have any choice to return back to her home and should ‘stay’ at the groom's house, or otherwise she would be shamed and blamed for bringing dishonor to her own family (Koishigulova et al. 2014: 4971).

to me that they are incapable of reading and understanding the map which presents a new problem and insight to the field of perceptual dialectology.

Investigating the perceptual dialectology and borders of regional dialects among the people who cannot recognize the map poses a new challenge. Previous works have used the map where people were asked to identify and mark the dialect borders (Alfaraz 2002; Fought 2002; Evans 2002), indicate the degree of the similarities with other regions (Sibata 1999) alongside with Preston's methodology for perceptual dialectology (1999: 350-373). Nevertheless, the study on the ethnic Kazakhs of China has revealed that these methods cannot directly be employed with people who cannot read and actually make a note on the map as the whole system and method of dialect borders in perceptual dialectology is based on pointing out the specific or general locations of the dialect borders. Hence, a new approach should be applied in this case such as examining the dialect regions orally and verbally by creating a map in the head and minds of people abstractly. For that, the questions directed on eliciting information regarding language attitudes were asked during the interviews to understand the language ideologies and dialect border perceptions of the informants. The references to specific dialect borders in their answers were described based on the names of cities, major locations or architectures, nature attractions, approximate regions, and mentions of people (i.e. relatives, famous persons) from these areas. This technique is important because it provides an alternative approach when the established methods of perceptual dialectology, which involve mapping dialect borders based on participant perceptions, are not effective.

Despite the challenges with mapping, the language attitudes displayed by the participants have shown language ideologies of purism, against Russification, and sketched some of the sociolinguistic relationship between Xinjiang and Kazakhstan. Such ideologies are created to

explain the source and meaning of the different linguistic features by justifying or rationalizing them based on the behavioral, aesthetic, and moral reasons through indexal relationships (Irvine & Gal 2000: 37). Additionally, the semiotic process of iconicity refers to the sign relationship between a language feature and linked social image (Irvine & Gal 2000: 37), which can be presented in a form of ‘pure’ Kazakh language and its speakers contrasting Kazakhs in Kazakhstan and in China. In fractal recursivity, where the same features of the group can be present in the further divided subgroup (Irvine & Gal 2000: 38; Andronis 2004: 264), the linguistic feature of the ‘pure’ Kazakh language was shown both at macro- and microlevel by projecting the prominent feature from a higher onto a lower level through a recursive framework. Additionally, Russian loanwords were indicated to be the markers of the ‘impure’ Kazakh language for the ethnic Kazakhs of China, but loanwords from Uyghur and Mandarin languages were not perceived as such, representing an erasure - the semiotic process through which a sociolinguistic phenomenon gets invisible, disregarded or ignored in one social group (Irvine & Gal 2000: 38; Andronis 2004: 264). These observations show how language ideologies influence the dialect borders and “mental maps” in the minds of the speakers further leading to a change in perception. As Woolard (2021: 2) pointed out, language ideologies are loaded with moral and political elements and represent not just how language is currently based on its features, but also how it ought to be. Similarly, my speakers described the state of the Kazakh language now used by the speakers from various regions, and also expressed their personal views regarding how ‘pure’ and ‘clean’ Kazakh should be. Based on the comments, Kazakhstan was perceived as one homogenous country with no internal linguistic variation though at microlevels differences were detected. This again illustrates how language perceptions and dialect borders were affected by

the language ideologies and were visualized in “mind map” of the participants addressing the research questions and employing the methodologies of this research work.

3.1. ‘Pure’ Kazakh language

The language ideology of ‘pure’ and ‘impure’ languages exists when there is a standardized language, and it is a mental concept and construction that occurs at the process of norm formation (Thomas 1991: 195; Brunstad 2003: 52-53). It is usually described to be against foreign elements and being closed to non-native elements including dialectics (Thomas 1991: 195; Annamalai 1979: 3-4). Purism can be identified at various linguistic levels: phonetic, syntactic, morphological, and orthographic (Brunstad 2003: 54-55). Moreover, the most common type of purism was found to be the one associated with the lexicons or the rejection of loanwords as shown by Thomas (1991: 65-66). The same ideology was detected in the language attitudes of ethnic Kazakhs born in China towards the spoken Kazakh language in Kazakhstan.

One of the frequent comments made by the speakers during interview responses was how ‘pure’ is the Kazakh language that they speak in China in comparison to how it is spoken in Kazakhstan. The terms *таза* (‘clear’) or *нақ Қазақ* (‘real Kazakh’) were used to describe the purity of the Kazakh language they speak, while to explain the spoken Kazakh in other regions besides Xinjiang the phrases such as commonly used both in C-K and K-K *шала қазақша* (‘half Kazakh’) and *басқаға ауып кеткен* (‘shifted/moved to another [language/ethnicity]’) were used.

The comment by speaker B stands out most:

[1] Біз Қытайдағылар нақ, таза қазақ екенбіз. Бұл жердегілер айтты ғой: ‘ой сендер таза қазақ екенсіңдер, [таза қазақша] сөйлейді екенсіңдер’ - деп.

Translation: We Chinese [Kazakhs] are real, clear Kazakhs. The locals told us: ‘Oh, you are clear Kazakhs, you speak clear Kazakh’.

What can be observed in [1] is the division of the language varieties based on the purity and clarity characteristics. Trying to further understand what it means to speak ‘pure’ Kazakh to the ethnic Kazakhs in China, I asked them to elaborate more on this topic. The later responses indicated that the ‘pure’ Kazakh notion means the language without the addition of Russian words or phrases. It was even commented that the words and speech of K-K sounded extremely weird to them, especially when they first came to Kazakhstan. For example, talking about the naming of the refrigerator, speaker B commented:

[2] Тоңлатқыш дейтінбіз. Сонда біз келгенде *холодильник* деген немене деп аңқаятынбыз. Немене ол қаладилілік деген десек, мынау дейді. Ой, атының жаманын-ай! Құлағымызға біртүрлі естілді.

Translation: We call it a refrigerator. So, when we came [to Kazakhstan], we were bewildered, questioning and trying to understand what is *холодильник*. What is *холодильник*? Then [they said] it is this one [Speaker B pointed to the refrigerator in the kitchen]. Oh, what a bad name! It sounded really weird to our ears.

What Speaker B was talking about is how the refrigerator is named in Kazakhstan and China. In Kazakhstan, the Russian loanword of *холодильник* is used, while among ethnic Kazakhs in China, it is either a loanword from Mandarin, [bɪŋfæŋ] (冰箱), or *тоңлатқыш* (*freeze-GIsh*). That is why Speaker B was struggling at the moment when she came to Kazakhstan to understand what is meant by *холодильник* as it is in Russian, and she does not understand this language. Thus, she admits that it sounded extremely weird and bad to her. This description already suggests that mixing Russian loanwords in the speech in Kazakh can be an example of iconization, the misrecognition of mixing, which indexically points to the situation of language policy and contact in the former Soviet republic, as an ‘essential’ – and unwanted – characteristic of the language or speakers. In other words, the code-switching and use of loanwords from Russian are perceived as ‘not clear’ and ‘not real’ Kazakh. In addition, it applies not only to Kazakh but also to other ethnicities who use Russian. For example, when trying to

describe how the Kazakhstan people speak the ‘corrupted’ Kazakh language (*шала қазақша сөйлеу*), speaker A highlights that it is not just about the Kazakhs in Kazakhstan but all the nationalities and ethnicities living in Kazakhstan:

[3] Қазақстанда негізі бәрі орысқа ауып кеткен ғой. Өзбектер де орысқа ауып кеткен. Қырғыз тілі де орысқа ауып кеткен. Ал бізде қырғыз қырғызша, ұйғыр ұйғырша сөйлейді, қазақ - қазақ, қытай - қытай, моңғол - моңғол.

In Kazakhstan everything was shifted to the Russian language. Uzbeks have shifted to Russian. Kyrgyzs have shifted to Russian, too. While in our [region of China] Kyrgyz speak Kyrgyz, Uyghurs speak Uyghur, Kazakhs [speak] Kazakh, Chinese [speak] Chinese, Mongolians [speak] Mongolian.

What we can see based on [3] is the elements of the Russification of Kazakh speech due to the socialist ideology of nationalities where each ethnicity is equated to a national language. To be specific, speaker A believes that each ethnicity is described and associated with only one language. This could be found in the policies and reforms implemented during the Soviet Union when nationalities were encouraged to establish their own independent and autonomous cultures. The main explanation for that was seeing the language as a tool for societal organization, a means of fostering unity among nations, and seen by political leaders as an essential prerequisite for achieving economic and cultural progress (Slezkine 1994: 429-431). The same policy is present in the People's Republic of China which considers all nationalities or *minzu* equal in terms of status and educational and cultural needs (Harrel 1993: 97-100). In both cases, the important aspect is that the nationalities should possess their own language under the socialist regime. Thus, the comments by Speaker A coming from China, a socialist country, can be the traces of these projects where they believe that each nationality is described based on the language too.

Historically, the socialist state wants to declare one language as a lingua franca or a means of communication between ethnicities and nationalities. For the USSR it was Russian,

while for the PRC it is Mandarin Chinese or Putonghua – a standardized version of the Beijing dialect. The minority languages and ethnicities were forced to learn and assimilate to the dominant language by making them bi- or even multilingual in some cases (Dwyer 2005: 34-35). The consequences of it can still be demonstrated based on the perceptions of the speakers of the same language from a different region. To be detailed, comment [3] proves that to the person who does not speak Russian, the spoken Kazakh language in Kazakhstan feels like shifted to the Russian language with more code-switching and phrases. That illustrates the russification of the Kazakhs in Kazakhstan as a consequence of the USSR regime where efforts were made to involve the Russian language as much as possible. This resulted in a massive influx of Russian terms not just in everyday life but also in education, science, and technology (Pavlenko 2008: 281). It even led to the creation of dictionaries and social media pages where newly coined terms and Kazakh slang using both Kazakh and Russian words were listed reflecting the bilingual reality of most of the country's citizens (Eisenberg 2021: 147).

However, for external observers such as ethnic Kazakhs of China, the ubiquity of Russian appears 'impure', reflecting an underlying ideology of language being equal to ethnicity. Being the policy implemented by the socialist regime, where each ethnicity should be associated with one language, the speakers were explicitly commenting on this feature not being present in Kazakhstan and all the ethnicities and their languages becoming similar or shifted to the Russian language and culture.

3.2. The erasure of regional differences

The interview part of my fieldwork included the recordings that represent three dialect regions of Kazakhstan and were played to the participants with the aim of getting language

attitudes and perceptions by the ethnic Kazakhs from China. Responses to the interview questions with more details can be seen in Table 3.1.

	West	South	Northeast
Speaker A	West (the words seem to be from Western region);	South (sounds are similar to South region)	Almaty
Speaker B	Shifted to Russian or somewhere else; But there is no explanation as to how it was specifically shifted (e.g. no definite characteristics);	South (words are similar to South region)	Almaty
Speaker C	Almaty; No difference from our speech. More Kazakh words are used.	Almaty; Speaks quickly	Almaty; Words are mixed (Kazakh-Russian) - shala-Kazakh

Table 3.1 The comments by the ethnic Kazakhs born in China towards the speech from three regions of Kazakhstan

The speech from the Western region was immediately noticed by Speaker A, and I believe it is because I was in contact with her longer in comparison to Speakers B and C (I am from the West side of Kazakhstan and speak this region's dialect). However, speaker B was not sure about the region and described it in general as being shifted to the Russian language. When I clarified what she meant by that, she responded by saying that it is somehow shifted and unsure if it is towards Russian or other languages, but she knows that it is changed even though she cannot specifically point out the characteristics of the differences and changes. On the other hand, as can be seen from Table 3.1, speaker C described all the speeches she heard from the recordings as being from Almaty. She also pointed out that there is no difference in how the speaker in the recording was speaking and how she [Speaker C] would speak every day. Later, while we were talking about the speech from the Western region, it was commented that the people from this side use more Kazakh words meaning that Russian words are less used during the speech. Thus, the speech from the Western region was not clearly associated with certain

ideologies, and each speaker had different perceptions towards it, and sometimes contradicting (Speaker B vs. Speaker C about the use of more Russian vs. Kazakh words). What should be highlighted is that all the participants admitted that they started to differentiate the speech from the Western region only after talking with me and taking me as a *келін* or ‘daughter-in-law’ to their tribe. But before that, they wouldn’t be able to guess specifically if the person is from the Western region as they had no idea or clue what the main features of the speech from this region are.

The only region that my participants were to indicate correctly was the speech from the South region given that the speaker in the recording was from Shymkent city on the South side of Kazakhstan. In other words, they were able to detect the speech of their own region even though the speech came from a non-Almaty person. Based on Table 3.1, the speakers mentioned that the words and phrases in the recording were similar to their way of speaking. Speaker C was sure that the speaker was from Almaty. Moreover, later she commented that specifically people from Shymkent speak extremely quickly and do not pronounce the ending of the words: *сөздің артын жұтып қояды* (lit: swallow the endings of the words). Earlier it was shown that based on the calculation of the speaking rates based on syllable per second from the dialect regions of Kazakhstan that the Southerners speak faster than Western and Northeastern speakers (Bizhanova 2021). Speaker C's comment corroborates this trend, as she observed a similar phenomenon. Hence, the ethnic Kazakhs from China were able to detect the speech from the South region - the place where they are living now.

On the other hand, the speech from the speaker of the Northeastern region was defined as the speech from someone who lives in Almaty. It was commented by all speakers that local people of the Almaty region speak this way, and they see no difference. Speaker C even added

that the speaker in the recording is *Shala-Kazakh* – a term used to describe people who code-switch to the Russian language a lot. The phrase *Shala-Kazakh* literally means ‘half-Kazakhs’, and historically meant the people who come from interethnic marriages of Kazakhs and Tatar, Uyghur, and others (Fedorov 1903: 280). However, today this meaning has been altered and typically refers to young, urban, Russified Kazakhs who are unable to speak Kazakh or are semi-speakers of Kazakh (Foster 2017: 5-6). Moreover, this term is usually employed to show the “impurity” of the speech and indicate that the way of speaking is heavily influenced by the Russian language even in the speech of Kazakhs from Kazakhstan (Foster 2017: 5-27). Thus, *Shala-Kazakh* can be found in Almaty and this is how their way of speaking is described.

The obvious tendency that was detected as a result of the interview is that all the speakers used Almaty as a reference point in their descriptions. Not only was the recording of northeasterners misrecognized by my participants as someone who is from Almaty, but to the speaker C speech from all three regions were from the same city or region where she lives. As she said: “People speaking in such a way could be found in Almaty too”. Here, we can observe how Almaty is viewed as a metropolitan place where every kind of speech can be detected. It might be true given that Almaty is the ex-capital and the largest city of Kazakhstan. Every year students and migrants from various regions come to this city making its population diverse. That’s why one can find different ways of speaking in Almaty as it encompasses people from every side of Kazakhstan. In other words, it reveals a significant trend where speakers associate diverse speech patterns with Almaty, perceiving it as a metropolitan center where various linguistic variations merge. This further leads to the broader attitude or ideology of prioritizing

urban centers like Almaty as standard bearers of linguistic norms, potentially leading to the erasure of internal linguistic diversity within Kazakhstan.

One of the questions that I asked my participants was: ‘Did you know any of the dialect regions in Kazakhstan before moving here?’ All of the speakers answered immediately saying ‘No’, and that they did not have any idea about internal linguistic differences. Also, during the interview about the southern regions, speaker C asked me which cities belong to the southern and northeastern regions, indicating that she didn't even know major landmarks. The fact that participants were not aware of the existence of dialect regions and even regional distinctions within Kazakhstan before moving here highlights a lack of recognition and understanding of internal linguistic differences. This lack of awareness emphasizes the notion of a dialect-less Kazakh language, where internal linguistic variations are not acknowledged or noticed. This ignorance of regional dialects and linguistic diversity suggests an ideology of homogeneity, where there's a tendency to view Kazakhstan as linguistically uniform and without internal variation.

During the interviews, the constant comparison of my participants was between ‘we’ and locals/Kazakhstan, where the former is the ethnic Kazakhs in China. Participants constructed a primary difference between large regions of Kazakhstan and Xinjiang as totally separate and unified even though both are huge with many internal variations. For example, as a daughter-in-law I was called *местный келін* which includes the Russian *местный* (‘local’) and Kazakh *келін* (‘daughter-in-law’) words. The Russian term is applied and restricted to an area or neighborhood where we live, and in this case Almaty. In the beginning, I wondered why they called me that if I was from the Western region of Kazakhstan and not from Almaty. An even more interesting question was why they called me ‘local’ using Russian but not Kazakh,

especially if they do not understand the Russian language? It turns out that the concept of locality itself, as my participants explained to me, was related to the whole country of Kazakhstan, no matter from which region you originally came. In other words, being born in Kazakhstan automatically makes a person - *местный* ('local'). Hence, the whole country of Kazakhstan was viewed as one unified region with a homogenous structure. This could also be supported by the responses given to me by my interlocutors when they informed me that before coming to Kazakhstan, they did not know anything about how the Kazakh language is spoken here not mentioning its variations, and only heard that it became too Russianized. So, while in Xinjiang participants recognize differences and features of the Kazakh dialects, in Kazakhstan they perceive the varieties of the Kazakh language as all the same.

This is an example of erasure, specifically the disregard of internal differences in language ideology, where diverse linguistic expressions within a country are overshadowed by a perception of linguistic uniformity of homogeneity (Irvine & Gal 2000: 38). The distinction in perception between language differences in Xinjiang and Kazakhstan reveals that at a micro level, participants are able to recognize and perceive differences in language, as evidenced by their awareness of dialects and variations in Xinjiang. However, at a macro level, these differences appear to be erased or minimized, particularly with limited exposure to linguistic diversity like in the case of my speakers. Participants' "mental maps" of Kazakhstan reflect a homogenous view, where the entire country is perceived as one unified region with a more or less similar linguistic structure. This can be contrasted with their recognition of diverse dialects and variations in Xinjiang suggesting an ideology that ignores internal linguistic differences within Kazakhstan. In perceptual dialectology, the works done on English, Japanese, and Turkish have shown a difference in the perceptions of the dialects and dialect borders (Lance 1999; Inoue

1999a; Demirci & Kleiner 1999), while only the young generation viewed Madrid as a homogenous area (Fernández & Fernández 2002: 317). This indicates the presence of the language ideologies in the mental maps of my speakers - ethnic Kazahs of China - while recognizing the dialect borders and seeing Kazakhstan as a whole homogenous area.

In summary, these findings illustrate the potential erasure of internal linguistic variation within Kazakhstan, driven by an ideology that prioritizes urban centers like Almaty as linguistic standards and neglects the rich diversity of regional dialects. This creates a perception of a homogenized, dialect-less Kazakh language, overlooking the linguistic complexity present within the country. While participants demonstrate an ability to perceive linguistic differences at a micro level, their perceptions at a macro level illustrate the homogenous view of language within Kazakhstan. This reflects broader attitudes and ideologies that prioritize linguistic uniformity while erasing internal differences, particularly in contexts of limited exposure to linguistic diversity.

3.3. Fractal recursivity

Delving more into the topic of ‘pure’, ‘clean’ and ‘real’ Kazakh language, the Xinjiang Kazakhs are divided into regional dialects where each dialect is perceived differently based on the levels of purity. In other words, the areas of Xinjiang Autonomous Region are further considered to see which community participants speak ‘cleaner’ Kazakh. That is an example of the mechanism of fractal recursivity in language ideologies, where the difference perceived at a higher level is repeated again at lower levels. In this situation, first, as it was shown and discussed before based on the interview, the notion of ‘clean’, ‘pure’ and ‘real’ Kazakh was

applied to Kazakhstan and Xinjiang at a macro-level, and then this same feature and framework will be used to distinguish better spoken Kazakh in Xinjiang region among Altai, Ili and Tarbagatay areas.

All participants during my interview mentioned that Altai people speak more ‘real’ Kazakh than they themselves [speakers A, B, C] do, while people from the Ili region use lots of Uyghur words for objects and person names. For example, speaker A commented on what Altai people have told her about how her community from the Ili region speak:

[4] Speaker A: Біз жақтағы қазақшаға тағы ұқсамай кетеді. алтай нақ қазақша сөйлейді. Шәуешек Алтай Үстірт уық район. Біздікі тағы сәл ұйғырға ауып кеткен. Алтайлар солай айтады. Алтайдықы таза қазақ.

Sometimes our [Kazakh] is not similar to Kazakh. Altai [people] speak real Kazakh. Shaueshek, Altai, Ustyurt are cold regions. Ours are shifted to Uyghur [language]. Altai people told us so. Altai’s one is the clear Kazakh [language].

Additionally, Speaker B even gave an example of how they use Uyghur words instead of Kazakh ones:

[5] Speaker B: Біздікі тағы сәл ұйғырға ауып кеткен. Леген ұйғырша, жалпақ қазақша жұқа табақ.

Ours is shifted a little bit to Uyghur [language]. *Леген* (Uyghur ‘big dish’) is in Uyghur, *жұқа табақ* (Kazakh ‘thin dish’) is in Kazakh.

These language attitudes tell us not just about how the ethnic Kazakhs from the Ili region perceive their own way of speaking, but also how others such as people in Altai perceive their [ethnic Kazakhs from the Ili region] way of speaking. As excerpts [4] and [5] have shown Altai people tend to speak ‘real’ Kazakh because they do not incorporate any Uyghur words which is again a language ideology where purity is eliminating non-native elements. Since Altai people are not perceived to have such ‘outside’ features, their Kazakh is described as ‘purer’ than the one in the Ili region.

During my fieldwork among the ethnic Kazakhs born in Tekes county of the Ili region, I noticed a list of ‘unfamiliar’ words and phrases to me which were later found to be loanwords from the Uyghur language. Table 3.2 presents this preliminary list of these words which I have collected.

	Phrase	IPA	Meaning
1	беш	bɛʃ	five
2	тошу	tɔʃɯw	to finish
3	тоштыру	tɔʃtɯɾɯw	to fill up
4	борлау	borlaw	to fatten
5	сақаю	saqaɯw	get better
6	қаттау	qattaw	to fold (clothes)
7	тақсы	taqsɯ	plate
8	леген	legen	big dish plate
9	дәзімал	dæzɪmal	iron
10	саңзы	saŋzɯ	fried dough
11	ләңмән	læŋmæn	lagman
12	дөт	døt	stupid
13	тетік	tetik	agile, tireless
14	қадыржам	qadɯɾʒam	careless
15	сәт	sæt	ugly
16	қыжалат	qɯʒalat	shy, embarrassed
17	әмәләттә	æmælættæ	actually
18	удыл	uwdɯl	straight
19	тәмсел	tæmsel	similarity, comparison
20	қалай мақан	qalaj maqan	mess, disorder
21	бағақ	baʒaq	celebration invitation
22	бопты	boptɯ	ok, let it be, fine
23	аям күні	aɯam kɯni	one day before Eid

Table 3.2. The loanwords from the Uyghur language used in the speech of ethnic Kazakhs born in China (compiled by author 2023)

From Table 3.2, we can see that the Uyghur loanwords are not only naming of the objects, foods and household furniture, but also the verbs which were changed phonologically. Even though some words have Kazakh translations, the ethnic Kazakhs in the Tekes region preferred using the Uyghur alternative. Speaker C shared her story about the word *удыл* which means ‘straight’ in the Uyghur language:

[6] Бір таксиге отырып, ‘больницаның жанынан удыл жүре береңіз’-депін. Удыл деген сол ұйғыршадан ауып кеткен шығар бізге. Такси айдайтын үлкен кісі ашуланып: ‘Удыл деген қандай сөз? Удыл деп айтушы ма еді? Қазақсың ғой сен, тура жүр деп айтпайсың ба?’-деді. Содан қазір тура жүр деп үйреніп алдым. Бұл жерде удыл деп айтпайды екен.

I got into a taxi and said: ‘Drive *удыл* (Uyghur ‘straight’) next to the hospital.’ Probably *удыл* (Uyghur ‘straight’) was shifted to us from Uyghur. The old taxi driver got angry: ‘What kind of word is *удыл* (Uyghur ‘straight’)? Does anyone say *удыл* (Uyghur ‘straight’)? You are Kazakh, shouldn’t you say go *тура* (Kazakh ‘straight’)?’ So now I learned how to say *тура* (Kazakh ‘straight’). Apparently, here they [local/Kazakhstani people] don’t say *удыл* (Uyghur ‘straight’).

This illustrates again how there is a border between the Kazakhstan and Xinjiang region for the ethnic Kazakhs in China as the Kazakhstani people are referred to as ‘they’ or ‘locals’. Also, it should be noted that speaker C like Speaker A and B is admitting that some words in her speech are from the Uyghur language. But in this case too, the existence of loanwords – despite their naturalness given cultural similarity and geographic proximity – is understood as ‘polluting’ the Kazakh language. Speaker C specifically mentions learning the Kazakh version instead of Uyghur so that locals will understand her, and so that she will have ‘clean’ Kazakh.

This example of fractal recursivity when the same linguistic or social phenomenon is repeated at the lower level, in this case based on perceptions of purity, shows how language ideologies function and impact individuals' mental maps. As Irvine and Gal (2000: 38) show, fractal recursivity allows us to find the meanings based on the discursive and cultural resources that are used to claim certain “communities”, identities, roles, and levels of contrast with a

cultural field. Similarly, it is important to understand the language attitudes and ideologies behind the perceptions towards the dialects or varieties of the language in order to construct the indexical associations so that the specific semiotic processes happening in the society can be understood. Based on the notion of purity where non-native elements were explained to be “polluting” the Kazakh language, the dynamic interplay could be observed between overarching language ideologies and individual linguistic practices . Thus, it demonstrates how ideologies operate at both macro and micro levels to shape perceptions and behaviors related to language and dialects.

3.4. Erasure related to the ‘purity’ of the Kazakh language

Unexpectedly however, given the speakers’ emphasis on purity and the dislike of Russian and Uyghur loans which I often observed during my fieldwork, one important fact that my participants born in China did not mention was that their speech contained not only Uyghur but also Mandarin Chinese loanwords. I was questioning if they would notice and mention that during our interview, and even tried to comment on that saying in [2] that they call the refrigerator [biŋfæŋ] (冰箱). Moreover, during my fieldwork, I have noted down examples of Mandarin Chinese words that are used in everyday life conversation which can be seen in Table 3.3 below.

	Phrase	IPA	Meaning	In Chinese orthography
1	жозы	dʒozɯ	table	桌子
2	вошы	wofɯ	bedroom	卧室
3	цїжуан	tsiɽeuwan	tile	瓷砖
4	биңшәң	biŋfæŋ	refrigerator	冰箱
5	биңгуи	biŋgui	freezer	冰柜
6	кәушан	kæuʃan	oven	烤箱

7	зілэшүй	zilæfyj	water that comes by itself	自来水
8	шәйдау	ʃæjdæw	a special type of big knife	菜刀
9	сәй	sæj	vegetable	菜
10	пінтяузі	pintjauzɿ	chinese vermicelli	粉条子
11	ганпан	ganpan	meal with steamed rice	干饭
12	бәй сәй	bæj sæj	chinese cabbage	大白菜
13	чін сәй	teɪŋ sæj	celery	芹菜
14	жу сәй	ʒuw sæj	leek	韭菜
15	жизі	ʒyjzy	tangerine	橘子
16	биян	bijan	rhinitis	鼻炎
17	диэнштай	diɪnʃtɑj	telechannel	电视台
18	мота	mota	motorbike	摩托
19	джоңго	dʒoŋgo	China	中国
20	ханзу	hanzu	Han Chinese	汉族
21	жеже	ʒeʒe	sister	姐姐

Table 3.3. The loanwords from the Mandarin language used in the speech of ethnic Kazakhs born in China

The loanwords from Mandarin contain loanwords similar to Uyghur such as the food, objects, furniture, and household wares. Even though their speech contained loanwords from Mandarin as well, they were referring to themselves as people who speak ‘purer’ Kazakh because the one in Kazakhstan has incorporated lots of Russian words. Nevertheless, there was erasure of the presence of Mandarin loanwords which, based on their language ideology, should make their Kazakh ‘less clear and pure’ as it contains non-native elements. However, this did not happen, and it is a semiotic process of erasure when the linguistic element or feature is eliminated. It does not matter if it is happening intentionally or unintentionally, and targeting non-native elements as threats to the language may happen selectively (Woolard & Schieffelin 1994: 64). It might be that ethnic Kazakhs from China have already incorporated Mandarin or

Uyghur loanwords into their speech due to their exposure and experience in China, these loanwords may not stand out to them as much. However, their lack of familiarity with the Russian language makes Russian loanwords more noticeable and potentially perceived as markers of linguistic impurity. Because Russian is non-native to them or differs significantly from their everyday speech, the presence of Russian loanwords may catch their attention more readily and index a deviation from what they consider to be "pure" Kazakh language. This phenomenon reflects the role of linguistic exposure and familiarity in shaping perceptions of language purity and impurity. It suggests that individuals are more likely to notice linguistic deviations that are foreign or unfamiliar to them while being less sensitive to deviations that they have already assimilated into their linguistic repertoire. Furthermore, it may even lead to the erasure of the linguistic characteristics, namely the presence of Uyghur and Mandarin loanwords, while the use of Russian words and phrases indicates the perceived impurity of the language.

Due to the space and word limit, the investigation of the loanwords from Uyghur and Mandarin Chinese languages was not included in this work. However, further study on this topic by looking at the phonological changes would benefit in the future as it might provide a new insight and perspective on the spoken Kazakh language by Kazakhs born in China. Additionally, conducting the interview to more participants from various regions of Xinjiang would provide more information on the language attitudes and ideologies of the Kazakhs in China.

3.5. Conclusion

This chapter delves into the findings concerning the language attitudes and perceptions of Chinese-born ethnic Kazakhs towards various Kazakh language variations based on my

fieldwork experience. These insights were acquired through interviews conducted with participants of three women from the Ili region of Xinjiang, where they expressed their opinions on different dialects spoken in both Kazakhstan and China. The process involved asking questions about the language beliefs, pinpointing dialect regions on a map, and eliciting feedback on recordings featuring speech in various Kazakh dialects. This methodology was followed based on Preston's five-point methods, but due to the inability of my participants to recognize the map, the mapping technique failed to be accomplished. Nevertheless, the language attitudes and ideologies were detected regarding the purity of the Kazakh language, where it is spoken and the presence of the semiotic processes in the language attitudes.

The 'purity' of the language, according to the language ideology system, is when the language avoids non-native elements as much as possible when the standardized version of the language exists. This attitude was also present among the ethnic Kazakhs of China who believed that they speak *таза* ('clear') or *нақ Қазақ* ('real Kazakh') in comparison to the Kazakhs in Kazakhstan. Upon the interview, it was apparent that by these terms it meant that the Kazakh language does not include Russian words or phrases. So, for them, the Kazakh language that code-switches a lot to the Russian language was a representation of 'impure' Kazakh which is an example of a semiotic process of iconization. Moreover, the participants pointed out that it was not just the Kazakh language in Kazakhstan that was Russified, but also the languages of other ethnicities (e.g. Uyghur, Kyrgyz, Uzbek) underwent the same process and were shifted to the Russian language and culture by becoming 'not clean' or 'not real'. Speakers' comments, originating from socialist China, reflect these ideologies, where ethnicity is linked to language.

During the interview phase, recordings representing three distinct dialect regions of Kazakhstan were played to participants. The goal was to elicit language attitudes and perceptions

from ethnic Kazakhs from China about the regional differences in spoken Kazakh language. As a result, the Western region was recognized by only one speaker, and in general all speakers admitted that they were not aware of the specific linguistic differences of this area before coming into contact with me, the Western region speaker. The speech from the South region was guessed correctly and emphasized that the speech was similar to how participants and locals would speak in everyday life (participants are living in Almaty region now which belongs to the south region of Kazakhstan). However, the speech of the speaker from the Northeastern region was characterized as resembling that of someone residing in Almaty. One of the participants mentioned that it is a speech of *шала қазақ* ('half-Kazakh'), the term which signifies the person who is heavily influenced by the Russian language and thus has "impure" Kazakh speech with Russian words and phrases.

The main trend observed from the responses was the fact that in half cases my participants associated the speech of the people in the recordings from other regions with the city of Almaty. They were specifying that 'a person with such speech could be found in Almaty' meaning that this city is a metropolitan that gathers various kinds of people with different styles and ways of speaking the Kazakh language. Additionally, the whole country of Kazakhstan and Almaty were together perceived as a whole homogenous region before ethnic Kazakhs from China came into contact with Kazakhstan's people suggesting the erasure of internal linguistic diversity within Kazakhstan. This lack of recognition displays the concept of a Kazakh language lacking distinct dialects with internal linguistic differences. Although participants display an awareness of linguistic nuances on a smaller scale, when it comes to a broader level, these distinctions have faded away, especially having been exposed to limited linguistic diversity as

seen with my participants. Consequently, participants' perceptions of Kazakhstan form homogeneous country as a singular region with a uniform linguistic structure.

Further investigation even showed that inside Xinjiang there are more dialect groups and regions which were assessed again based on the 'purity' of the language. My participants, who originally were born and raised in Tekes county of the Ili region identified their own speech as 'less pure, clean and real' in contrast to the spoken Kazakh language in the Altai area. The reason for that was the Kazakh language that my participants use, as they believe, was shifted to the Uyghur language. The assessment of the language based on the 'purity' quality at such a micro-level inside Xinjiang entails that it is a semiotic process of fractal recursivity where patterns of language use and beliefs about language are repeated on different (in this case, lower) scales or levels within a society or culture (Altay, Ili and Tarbagatay areas). Hence, it reveals that the language ideologies operate at both macro and micro levels influencing the perceptions and language attitudes which further impacts the mental maps of the dialect regions and borders in the minds of the folks.

Nonetheless, there was one feature of the spoken Kazakh language of ethnic Kazakhs born in China that was not mentioned by my participants during the interview – the presence of Mandarin loanwords. Despite my attempt to point out this phenomenon to them, none of my three participants mentioned that similar to Uyghur words, they have incorporated the words and phrases from the Mandarin Chinese language into their speech through phonological changes. During my stay with the ethnic Kazakhs from China, I noted down the list of loanwords from Uyghur and Mandarin languages, and I have presented it as a table in the paper and as an appendix at the end of the work. Loanwords from both languages were used to name the objects, house furniture and wares, and foods of each culture, but still only the Uyghur language was

mentioned by my participants to be the indicator of ‘pureness’ among the dialect regions of Xinjiang. This refers to the ideological mechanism of erasure when certain linguistic elements or practices are eliminated or surpassed. In this case, it was the presence of loanwords from Mandarin Chinese that were not noted in the process of assessing the ‘purity’ of the spoken Kazakh language in China and Kazakhstan when the Kazakhstani version was downgraded due to the use of Russian words, but Chinese stayed as ‘purer’ indicating how this semiotic process shaped the participants’ perception and ‘pureness’ of the Kazakh language varieties.

Because of constraints on space and word count, this study did not explore in depth the origins and adaptations of loanwords from Uyghur and Mandarin Chinese. Yet, future research focusing on phonological shifts could offer fresh perspectives on spoken Kazakh among Kazakhs born in China. Moreover, interviewing additional participants from diverse regions in Xinjiang could yield further insights into language attitudes and ideologies among Kazakhs in China by contributing to the studies of perceptual dialectology.

Chapter 4. Conclusion

This study aims to identify linguistic differences in the spoken Kazakh in China and explore the language attitudes and perceptions of ethnic Kazakhs from China towards Kazakh language variations. The paper begins with a literature review, examining Kazakh language varieties in Kazakhstan and relevant studies on today's spoken Kazakh. It then delves into the historical, political, and sociolinguistic context of ethnic Kazakhs in China to provide background information. Additionally, existing research on spoken Kazakh in China is reviewed to identify gaps, and how my study tries to contribute to the studies in this field. Then the research questions and methodologies are outlined, explaining how linguistic and perceptual dialectology methods were adapted to my studies. Namely, first, the 'dialect' and 'language' terms were clarified, and the semiotic processes in language ideology were shown indicating the types and cases. Then it explains it aims to identify and compare the linguistic differences on the phonetic, lexical, and meaning levels by contrasting the spoken Kazakh language varieties of Kazakhstan and China. The second goal of the research - discussing the language ideologies and attitudes of the ethnic Kazakhs to the spoken Kazakh varieties includes analyzing comments based on the recordings and statements given by the speakers during the fieldwork.

The second chapter highlights the differences between spoken Kazakh in China and Kazakhstan, focusing on phonetic variations, new lexemes, phrases that have different meanings in comparison to spoken Kazakh in Kazakhstan, and specific dialect words. Data was collected from fieldwork in the Almaty region with ethnic Kazakhs from China (Ili region), supplemented by a list of commonly used words from Altai and Tarbagatay regions by Professor Aynur Abish. Phonetic differences in Kazakh spoken in China resemble those of the Southern regions in Kazakhstan, including shifts from [s] to [ʃ], the use of the affricates instead of the postalveolar

fricatives, and the variations (addition or deletion) in sounds. These observations align with previous research findings and include words unique to Kazakhs in China, dialectal lexemes, and words with different meanings compared to Kazakhstan. Notable findings include additional meanings for certain words, varied usage of auxiliary verbs, and the presence of taboo words not found in Kazakhstan. Even though similarities exist between spoken Kazakh in China and regions of Kazakhstan, distinct lexical and dialectal differences are evident based on the results of my work. It's important to note that Kazakh spoken in China also exhibits regional dialects with linguistic variations that were not considered in this study.

The third chapter explores the language attitudes and perceptions of ethnic Kazakhs born in China towards various Kazakh language variations based on fieldwork results. Insights were gleaned through interviews with three women from the Ili region of Xinjiang, where they shared opinions on different dialects spoken in both Kazakhstan and China after listening to the recording from three regional dialects of Kazakhstan. The methodology involved questioning participants about their language beliefs, attempting to pinpoint dialect regions on the map, and eliciting feedback on recordings featuring speech in various Kazakh dialects. While the mapping technique failed due to participants' difficulty in recognizing the map, the interviews revealed attitudes and ideologies regarding the purity of the Kazakh language and its spoken forms in different regions. According to language ideology, the "purity" of a language is when the speakers try to avoid non-native elements, especially when a standardized version of the language exists. In my case, ethnic Kazakhs from China expressed a preference for *маза* 'clear' or *нақ Қазақ* 'real Kazakh' language that does not have any influences from Russian language and culture. This was contrasted with what they perceived as "impure" Kazakh spoken in Kazakhstan due to frequent code-switching with Russian. Moreover, participants identified

specific dialect regions in Kazakhstan based on recordings, associating most speech patterns with Almaty city as metropolitan. However, their perceptions of Kazakhstan as a homogeneous region show limited awareness of linguistic diversity within the country, and it suggests a homogenization of Kazakhstan before the ethnic Kazakhs from China encountered its people, highlighting an erasure of internal linguistic diversity. This lack of recognition reflects the concept of a Kazakh language lacking distinct dialects despite participants' awareness of linguistic nuances on a smaller scale. Further investigation revealed similar assessments of dialect purity within Xinjiang, with those participants from Tekes county (Ili region) perceiving their own speech as “less pure” compared to Altai. This assessment at a micro-level reflects a semiotic process of fractal recursivity, wherein language ideologies influence perceptions and attitudes, impacting mental maps of dialect regions. Moreover, as I have shown in my work, while participants acknowledged Uyghur loanwords, they did not mention Mandarin loanwords “polluting” the Kazakh language, indicating an ideological erasure of certain linguistic elements. This suggests how semiotic processes shape perceptions of linguistic purity within Kazakh language varieties.

Due to the time and space limit, it was not possible to focus on loanwords and calques from Uyghur, Mandarin, and Russian languages, but future research can analyze these phonological shifts and interview participants from diverse regions of both Kazakhstan and China to further explore language attitudes among Kazakhs in Kazakhstan and China, and how it might affect the language perceptions.

5. References

- Abish, Aynur. 2016. *Modality in Kazakh as Spoken in China*. Wiesbaden: Harrassowitz. <https://doi.org/10.2307/j.ctvc770j9>.
- Abish, Aynur & Éva Csató. 2011. Recent developments in Kazak as spoken in the People's Republic of China. *Turkic Languages* 15(2): 275-290.
- Alfaraz, Gabriela. 2002. Miami Cuban Perceptions of Varieties of Spanish. In: Long & Preston (eds.), *Handbook of perceptual dialectology*, Volume 2. 1-13. Amsterdam: John Benjamins.
- Andronis, Mary. 2004. Iconization, Fractal Recursivity, and Erasure: Linguistic Ideologies and Standardization in Quichua Speaking Ecuador. *Texas Linguistic Forum* 47: 263-269.
- Annamalai, E. 1979. *Language movements in India*. CIIL Conferences and seminars series, Volume 5. Mysore: Central Institute of Indian Languages.
- Argynbaev, K. 1978. Marriage and marriage rites among the Kazakhs in the nineteenth and early twentieth centuries. In: Weissleder (Ed.) *The nomadic alternative: modes and models of interaction in the African-Asian deserts and steppes*. 331-41. The Hague: Mouton.
- Awde, Nicholas, William Dirks & Umida Hikmatullaeva. 2002. *Uzbek-English, English-Uzbek Dictionary and Phrasebook: Romanized*. New York: Hippocrene Books.
- Benson, Linda & Ingvar Svanberg. 1998. *China's Last Nomads: History and Culture of China's Kazaks*. New York: An East Gate Book. <https://doi.org/10.4324/9781315285214>.
- Benson, Erica. 2003. Folk Linguistic Perceptions and the Mapping of Dialect Boundaries. *American Speech* 78(3): 307-330.
- Bizhanova, Moldir. 2021. "Can you speak slowly, please?" *The speaking rate of dialects and perceptions towards it in Kazakhstan*. Nur-Sultan: Nazarbayev University. BA Capstone.
- Blommaert, Jan. 2008. Artefactual Ideologies and the Textual Production of African Languages. *Language Communication* 28(4): 291-307. doi:10.1016/J.LANGCOM.2008.02.003.
- Brunstad, Endre. 2003. Standard language and linguistic purism. *Sociolinguistica - International Yearbook of European Sociolinguistics* 17(1): 52-70. doi:10.1515/9783110245226.52
- Buchholtz, Mary & Nancy Bermudez, Victor Fung, Lisa Edwards and Rosalva Vargas. 2007. Hella Nor Cal or Totally So Cal?: The Perceptual Dialectology of California. *Journal of English Linguistics* 35 (4): 325-352. <https://escholarship.org/content/qt6492j904/qt6492j904.pdf>

- Chambers, Jack & Peter Trudgill. 1998. *Dialectology*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511805103>.
- Demirci, Mahide & Brian Kleiner. 1999. The Perception of Turkish Dialects. In: Dennis R. Preston (ed.), *Handbook of perceptual dialectology*, Volume 1. 263-284. Amsterdam: Benjamins.
- Dotton, Zura & John Doyle Wagner. 2008. *A Grammar of Kazakh*. Almaty.
- Dwyer, Arienne M. 2005. *The Xinjiang Conflict: Uyghur Identity, Language policy, and Political Discourse*. Washington: East-West Center Washington.
- Eisenberg, Leora. 2021. Janasozdik: codifying an inclusive Kazakh slang. *Bulletin of L.N. Gumilyov Eurasian National University*, 136(3): 147-159. <https://doi.org/10.32523/2616-7255-2021-136-3-147-159>
- Evans, Betsy. 2002. Attitudes of Montreal Students Towards Varieties of French. In: Long & Preston (eds.), *Handbook of perceptual dialectology*, Volume 2. 71-95. Amsterdam: John Benjamins.
- Fernández, Juliana Moreno & Francisco Moreno Fernández. 2002. Madrid Perceptions of Regional Varieties in Spain. In: Dennis R. Preston (ed.), *Handbook of perceptual dialectology*, Volume 2. 295-320. Amsterdam: John Benjamins.
- Filchenko, Andrey, Giorgia Troiani, John W. Du Bois, Gulnar Sarseke, Akyl Akanov, Moldir Bizhanova, Nikolay Mikhailov, Tansulu Temirbekova, Bibarys Seitak, and Zhansaya Turaliyeva. 2023. *Multimedia Corpus of Spoken Kazakh Language*. mcskl.nu.edu.kz
- Foster, Hannah Lofton. 2017. *Imagining the Shala-Kazakh: codeswitching and satire in Post-Soviet Kazakhstan*. The University of Texas at Austin. MA Thesis.
- Fought, Carmen. California Students' Perceptions of, You Know, Regions and Dialects? In: Long & Preston (eds.), *Handbook of perceptual dialectology*, Volume 2. 113-135. Amsterdam: John Benjamins.
- Harrell, Stevan. 1993. Linguistic Hegemony in China. *International Journal of the Sociology of Language* 103: 97-114. University of Washington.
- Inoue, Fumio. 1999a. Classification of dialects by image: English and Japanese. In: Dennis R. Preston (ed.), *Handbook of perceptual dialectology*. 147-159. Amsterdam: John Benjamins.
- Inoue, Fumio. 1999b. Subjective dialect division in Great Britain. In: Dennis R. Preston (ed.), *Handbook of perceptual dialectology*. 161–176. Amsterdam: John Benjamins.
- Irvine, Judith & Susan Gal. 2000. Language ideology and linguistic differentiation. In: P. V. Kroskrity (ed.), *Regimes of language: Ideologies, politics, and identities*. 35-84. Santa Fe: School of American Research Press.

- Kagan, Arik. 1999. *Shamanism, Culture and the Xinjiang Kazaks: A Native Narrative of Identity*. University of Washington. ProQuest Dissertations and Theses Global. Doctoral dissertation.
- Kara, David Somfai. 2002. *Kazak*. Munich: Lincom Europa.
- Koishigulova, D.M. & A. Ospanova, A. Syrgakbaeva, O. Abdimanuly. 2014. Expression of an Agreement vs. Disagreement in the Course of Kidnapping of the Bride among Kazakhs. *Procedia - Social and Behavioral Sciences* 116: 4970-4976. <https://doi.org/10.1016/j.sbspro.2014.01.1057>
- Kim, U Gene. 2018. *Ethical Management of Speech among Kazak Nomads in the Chinese Altai*. Michigan: University of Michigan. Doctoral Dissertation.
- Kim, Ujin [=U Gene]. 2020. The Quest for the ‘Best’ Language for Modern Xinjiang: Language Ideologies of Practicality and Aesthetics. *Journal of Multilingual and Multicultural Development* 43 (8): 787-795. <https://doi.org/10.1080/01434632.2020.1778708>.
- Kirchner, Mark. 1998a. Kazakh and Karakalpak. In: Johanson & Csató (eds.), *The Turkic Languages*. 318–332. London: Routledge.
- Kirchner, Mark. 1998b. The Language of the Kazaks from Xinjiang: A Text Sample. In: Benson & Svanberg (eds.), *The Kazaks of China: Essays on an Ethnic Minority*. 189-199. Stockholm: Almqvist & Wiksell International.
- Krippes, Karl. 1996. *Kazakh Grammar with Affix List*. The University of Michigan: Dunwoody Press.
- Kurath, Hans. 1939. *Handbook of the Linguistic Geography of New England*. Providence, Rhode Island: Brown University Press.
- Labov, William. 1972. The social stratification of (r) English in New York City Department Stores. *Sociolinguistic Patterns*:. 43-69. Philadelphia PA: University of Pennsylvania Press.
- Lance, Donald. 1999. Regional Variation in Subjective Dialect Divisions in the United States. In: Dennis R. Preston (ed.), *Handbook of perceptual dialectology*. 283-314. Amsterdam: John Benjamins.
- Mariou, Eleni. 2020. Narratives about ‘Homeland’, Heritage, Languages and Belonging: A Case of ‘Return’ Migration. *Linguistics and Education* 56 (2020): 1-9. <https://doi.org/10.1016/j.linged.2019.100793>.
- Martin, V. 2001. *Law and custom in the Steppe: the Kazakhs of the Middle Horde and Russian colonialism in the nineteenth century*. London: Curzon.
- McCollum, Adam, & Si Chen. 2021. Kazakh. *Journal of the International Phonetic Association* 51 (2): 276-298. <https://doi.org/10.1017/S0025100319000185>

- Muhamedowa, Raihan. 2016. *Kazakh: A Comprehensive Grammar*. *Kazakh: A Comprehensive Grammar*. Routledge: <https://doi.org/10.4324/9781315738239>.
- Myers-Scotton, Carol. 2006. *Multiple Voices: An Introduction to Bilingualism*: 16-34. Oxford: Blackwell Publishing.
- National Bureau of Statistics of China. 2022. China Statistical Yearbook. <https://www.stats.gov.cn/sj/nds/2022/indexeh.htm>
- Noda, Jin. 2019. Kazakh Migrants and Soviet-Chinese Relations during the 1940s: A Background of Xinjiang Refugees to the Middle East. In: Jin Noda & Ryosuke Ono (eds.), *Emigrants/Muhacir from Central Asia to the Middle East: The Case of Xinjiang (1940s–1950s)*. 25–42. Tokyo: Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Pavlenko, Aneta. 2008. Multilingualism in Post-Soviet Countries: Language Revival, Language Removal, and Sociolinguistic Theory. *International Journal of Bilingual Education and Bilingualism*, 11 (3-4): 275-314, DOI: [10.1080/13670050802271517](https://doi.org/10.1080/13670050802271517)
- Payne, Thomas. 1997. *Describing Morphosyntax: A Guide for Field Linguists*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511805066>.
- Preston, Dennis. 1999. A language attitude analysis of regional US speech: is northern US English not friendly enough? *Cuadernos de Filología Inglesa* (8): 129-146.
- Preston, Dennis (ed.). 1999. *Handbook of perceptual dialectology*, Volume I. Amsterdam: Benjamins.
- Salimjan, Guldana. 2017. Debating Gender and Kazakhness: Memory and Voice in Poetic Duel Aytis between China and Kazakhstan. *Central Asian Survey* 36 (2): 263-280. <https://doi.org/10.1080/02634937.2017.1281219>.
- Salimjan, Guldana. 2021. “Mapping Loss, Remembering Ancestors: Genealogical Narratives of Kazakhs in China.” *Asian Ethnicity* 22 (1): 105-120. <https://doi.org/10.1080/14631369.2020.1819772>.
- Schwartz, Henry G. 1992. *An Uyghur-English dictionary*. Center for East Asian Studies. Bellingham: Western Washington University.
- Silverstein, Michael. 1979. Language structure and linguistic ideology. In: Clyne, Hanks, & Hofbauer (eds.), *The Elements: A Parasession on Linguistic Units and Levels*. 193-247. Chicago: Chicago Linguistic Society.
- Slezkine, Y. 1994. The USSR as a Communal Apartment, or How a Socialist State Promoted Ethnic Particularism. *Slavic Review* 53(2): 414–452. <https://doi.org/10.2307/2501300>

- Spencer, Chen. 2018. Visualizing language ideologies and verbalizing perceived linguistic boundaries: The case of Mandarin Chinese in contemporary Taiwan. *Global Chinese* 4(1): 1-21.
- Sussman, Nan M. 2010. *Return Migration and Identity: A Global Phenomenon, a Hong Kong Case*. Hong Kong: Hong Kong University Press. <http://www.jstor.org/stable/j.ctt1xwfd>.
- Thomas, George. 1991. Linguistic purism. London & New York: Longman. Sibata, Takesi. 1999. Consciousness of Dialect Boundaries. In: Preston (Ed.), *Handbook of perceptual dialectology*, Volume 1. 39-63. Amsterdam/Philadelphia: John Benjamins B.V.
- Vajda, Edward. 1994. Kazakh Phonology. In: Edward Kaplan & Donald Whisenhunt (eds.), *Opuscula Altaica: Essays Presented in Honor of Henry Schwarz*. Bellingham, WA: Western Washington University.
- Werner, Cynthia, Celia Emmelhainz & Holly Barcus. 2017. Privileged Exclusion in Post-Soviet Kazakhstan: Ethnic Return Migration, Citizenship, and the Politics of (Not) Belonging. *Europe - Asia Studies* 69 (10): 1557-1583. <https://doi.org/10.1080/09668136.2017.1401042>.
- Werner, Cynthia. 2009. Bride Abduction in Post-Soviet Central Asia: Marking a Shift Towards Patriarchy through Local Discourses of Shame and Tradition. *The Journal of the Royal Anthropological Institute* 15(2): 314–331. <http://www.jstor.org/stable/20527710>
- Wikle, Thomas & Guy Bailey. 2019. Mapping Language Variation and Change in the USA and Canada. In: Brunn & Kehrein (eds.), *Handbook of the Changing World Language Map*. 1-11. https://doi.org/10.1007/978-3-319-73400-2_107-1
- Woolard, Kathryn & Bambi Schieffelin. 1994. Language Ideology. *Annual Reviews* 23: 55 – 82.
- Woolard, Kathryn. 2021. Language Ideology. In: Stanlaw (Ed.), *The International Encyclopedia of Linguistic Anthropology*: 1-21. Hoboken: Wiley & Sons.
- Zhang, Qi, & Ting Yang. 2021. Reflections on the Medium of Instruction for Ethnic Minorities in Xinjiang: The Case of Bilingual Schools in Urumqi. *International Journal of Bilingual Education and Bilingualism* 24 (1): 50–71. <https://doi.org/10.1080/13670050.2018.1442409>.

Kazakh sources:

- Äşimbaeva N., S. Aqaev, Q. Rısbegenova, J. Mankeeva, B. Äbilqasımov. 2007. *Qazaq ädebı tiliniñ sözdigi* [Dictionary of the Kazakh literary language]. (Vol. 4). Almatı: Arıs baspası.
- Cüyerqulova B. S. Japaqov, T. Januzaqov, O. Jumabaeva, M. Jolşaeva, B. Qapalbekov, Q. Ğabıtıxan, Q. Esenova, N. Äşimbaeva. 2006. *Qazaq ädebı tiliniñ sözdigi* [Dictionary of the Kazakh literary language] (Vol. 3). Almatı: Arıs baspası.

- Janpeyisov, Erbol & Köbey Xusayın, Älimxan Jünisbek, Nurjamal Oralbaeva, Seyilbek İsaev, Aytbay Ayğabilulı, Mirzatay Serğaliev, Raqış Ämir, and Nurgeldi Wäliev. 2002. *Qazaq Grammatikası. Fonetika, Sözjasam, Morfologiya, Sintaksis*. [Kazakh Grammar. Phonetics, Word Formation, Morphology, Syntax.]. Astana: Qazaqstan Respwblıkası bilim jäne ğılım mınıstrlıgı, A. Baytursınulı atındağı til bilimi İnstıtwtı.
- Januzaqov T., S. Omarbekov, Ä. Jünisbek, B. Mominova, Ğ. Änes, B. Äbilqasımov, J. Mankeeva, J. Qoñıratbaeva, Ğ. Qalıev. 2006. *Qazaq ädebı tiliniñ sözdigi* [Dictionary of the Kazakh literary language] (Vol. 1). Almatı: Arıs baspası.
- Jañabekov, A., & T. Januzaq, K. Küderinova, R. Şoybekov, A. Fazıljanova, B. Äbilqasımov, A. Üderbaev, A. İmanğazına, S. İmanberdieva, Q. Rısbergenova. 2013. *Qazaq ädebı tiliniñ sözdigi* [Dictionary of the Kazakh literary language] (Vol. 15). Almatı: Arıs baspası.
- Mamyrbekova, Gulfar. 2017. *Qazaq tilindegi arab, parsı sözderiniñ tüsindirme sözdigi* [An explanatory dictionary of Arabic and Persian words in the Kazakh language]. Almatı: Memlekettik tildi damıtıw İnstıtwtı.
- Mäsımxanulı Düken. 2004. Qıtaydağı qazaqtar jäne olardıñ tilindegi keybir erekşelikler [Kazakhs in China and some of their language uniqueness] *Terminologiyalıq xabarşı* (2): 23-27. Astana: Qazaqstan Respwblıkası Mädeniet, aqparat jäne sport mınıstrlıgınıñ Til komıtetı.
- Mustafaulı, Serik (2003). *Turaqtı tirkesterdegi jergilikti erekşelikter* (Qıtaydağı qazaqtar tili materialı boyınşa) [Local peculiarities in fixed phrases: (according to the material of the Kazakh language in China)]. Almatı : Qazaq wñiversiteti.
- Naqısbekov, O. 1982. *Qazaq tiliniñ oñtüstik govorlar tobı* [Southern dialect group of the Kazakh language]. Almatı: Qazaq SSRiniñ “ĞILIM” baspası.
- Qalıev, Bayınqol. (2014). *Qazaq tiliniñ tüsindirme sözdigi* [Explanatory dictionary of the Kazakh language]. Almatı: Memlekettik Tildi Damıtıw İnstıtwtı.
- Qalıev Ğ., S. Bızaqov, O. Naqısbekov, S. Mırzabekov, A. Jañabekova, A. Fazıljanova, Q. Küderinova, B. Süyerqulova. 2006. *Qazaq ädebı tiliniñ sözdigi* [Dictionary of the Kazakh literary language] (Vol. 2). Almatı: Arıs baspası.
- Qazaqstan Respwblıkası Strategiyalıq josparlaw jäne reformalar agenttigi, and Ulttıq statistika byurosı. 2022. *2021 Jılğı Qazaqstan Respwblıkası Xalqınıñ Ulttıq Sanağınıñ Qoritındıları* [Results of the 2021 National Census of the Republic of Kazakhstan]. Nur-Sultan.
- Sawranbaev, Niğmet & Jamalbek Bolatov. (1957, March 14). Qazaq tiliniñ jergilikti erekşelikterin zertteu [Study of local features of the Kazakh language]. *Socialistik Qazaqstan*. Orınbor.
- Talaspäeva, J. & N. Tursınğalı. 2021. *Dästür men jañaşıldıq toğısındağı til jäne ädebiet mäseleleri: xalıqaralıq ğılımı konferenciya materialdarı* [Problems of language and

literature at the intersection of tradition and innovation: materials of the international scientific conference]. 248-252. Nur-Sultan: L.N.Gwmliev atındağı Ewraziya ultıq wniwersiteti.

Russian sources:

Abdrahmanov, A., B. Abilhasimova, K. Ajdarbekova, F. Aliev, Z. Ahmetzhanova, S. Bizakov, N. Demesinova, A. Imanbaev, Zh. Mankeeva, O. Nakisbekov, K. Netalieva, G. Nurgazieva, S. Omarbekov, N. Ongarbaeva, K. Rysbergenova, R. Urekenova, Zh. Utesbaeva, Sh. Sarybaev, B. Sulejmenova, B. Hasanov, K. Husain. 2002. *Kazahsko-russkij slovar': okolo 50000 slov* [Kazakh-Russian dictionary: around 50000 words]. R. G. Syzdykova, K. Sh. Husain (Eds.). Almaty: Dajk-press.

Amanzholov, Sarsen. 1997. *Voprosy Dialektologii i Istorii Kazahskogo Jazyka* [Questions of Dialectology and the History of Kazakh Language]. Almaty: Sanat.

Davydov, I. 1929. O probleme iazykov v prosvetitel'noi rabote sredi natsionalnostei [On the problem of languages in educational work among nationalities]. *Prosveshchenie natsional'nostei*: 18-29.

Fedorov, D. 1903. Opyt" voenno-statisticheskogo opisaniya Iijskogo krajaxъ [Experienc of a military-statistical description of the Ili region.]. Volume 1. Izdanie Shtaba Turkestanskogo voennago okruga: Tashkent".

Taizhanov, G.E. 1995. *Kazahi: Istoriko-jetnograficheskoe issledovanie*. [Kazakhs: Historico-ethnographic research]. Almaty: Kazakhstan Publishers.

Kambarov, N.K. (Kamberi). 2005. *Ujgurskie bljuda* (Ujgur taamliri) [Uyghur foods]. Bishkek.

Kibirova, Sh. & Cunvazo Ju. 1961. *Ujgursko-russkij slovar'* [Uyghur-Russian dictionary]. Alma-ata: Izdatel'stvo Akademii Nauk Kazahskoj SSR. <https://kazneb.kz/kk/bookView/view?brId=1573108&simple=true#>

Levitskaja, L.S. & G.F. Blagova, A.V. Dybo, D.M. Nasilov, E. A. Poceluevsckaja. 2003. *Jetimologicheskij slovar' tjurkskih jazykov*. Obshhetjurkskie i mezhtjurkskie osnovy na bukvy «L», «M», «N», «P», «S» [Etymological dictionary of Turkic languages. Common Turkic and inter-Turkic bases on the letters “L”, “M”, “N”, “P”, “S”] (Volume VII). A.V. Dybo (Ed.). Moskva: Vostochnaja literatura.

Radlov, V. V. 1893a. *Opyt slovarja tjurkskih narechij : slovar' (Tom 1: Glassnye, Knuzs 1)* [Experience in the dictionary fo Turkic dialects (Vol. 1: Vowels, Book 2)]. Sankt-Peterburg : Tipografija Imperatorskoj Akademii Nauk.

Radlov, V. V. 1893b. *Opyt slovarja tjurkskih narechij : slovar' (Tom 1: Glassnye, Knuzs 2)* [Experience in the dictionary fo Turkic dialects (Vol. 1: Vowels, Book 2)]. Sankt-Peterburg : Tipografija Imperatorskoj Akademii Nauk.

Radlov, V., V. 1911. *Opyt slovarja tjurkskih narechij : slovar'* [Experience in the dictionary fo Turkic dialects] (Vol. 4). Sankt-Peterburg: Tipografija Imperatorskoj Akademii Nauk.Sankt-Peterburg.

Sevortjan, V. 1978. *Jetimologičeskij slovar' tjurkskih jazykov. Obshhetjurkskie i mezhtjurkskie osnovy na bukvu «B»*. [Etymological Dictionary of Turkic Languages. Common Turkic and Inter-Turkic stems beginning with letter «B»] (Vol. 2). Moscow: Nauka.

Other sources:

Doerfer, Gerhard & Semih Tezcan. 1980. *Wörterbuch des Chaladsch (Dialekt von Charrab)* [Khalaj dictionary (Charrab dialect)]. Budapest: Akadémiai Kiadó.