South Western and Eastern Turkic features in the Karabulak Dialect of Uzbek

Yulduzkhan Turgunova

WLL 499: Languages, Linguistics, and Literature Capstone II

Advisor: Pr. Uli Schamiloglu

May 3, 2023

Word count: 6286

1. Introduction

Some Turkic texts written between the 13th-15th centuries exhibit the linguistic characteristics of both Oghuz and Eastern Turkic (Erdal, 2015, p. 135). This phenomenon is named as *'olga- ~ bolga- dili'* (mixed language) in Turkic linguistics (Erdal, 2015, p. 135). One of the main characteristics of the *'olga- ~ bolga-'* language is the use of both South Western Turkic *ol-* and Eastern Turkic *bol-* "to become, to be " forms of the verb (Erdal, 2015, p. 135). So far, linguistic parameters of the *'olga- ~ bolga-'* language have been identified only in historical Turkic texts. However, analyzing the Karabulak Dialect of Uzbek, I have noticed that the *ol-* and *bol-* forms of the verb are used interchangeably in this dialect. This probably means that the Karabulak dialect is an *'olga- ~ bolga-'* mixed dialect. Nevertheless, other characteristics of the *'olga- ~ bolga-'* language, such as the presence of both *+GA* and *-(y)A* forms of the dative in a language, should be verified. Consequently, this paper aims to analyze Karabulak Dialect in a framework of the *'olga- ~ bolga-'* language to see if it also shows other features of mixed language listed by Erdal (2015, p. 137). Moreover, the issue of *'olga- ~ bolga-'* features in the Karabulak dialect is worth researching due to the absence of prior research on the topic and its potential to provide a new perspective to look at Turkic Dialects.

1.2. The regional information

Current research will be conducted in the Karabulak village. Karabulak is one of the largest villages in Sayram District, South Kazakhstan, with an Uzbek-speaking population (Karabulak, 2023). The village is located 45 km from the regional center of Shymkent. According to Kamiljon Abdushkuruly, Akim of Karabulak Rural District, 55 540 people live in the Karabulak village (Akhmedova, 2022). Moreover, 94% of the population who are living in Karabulak are ethnic Uzbeks. Moreover, there are fifteen Uzbek-medium schools located in the

village (Karabulak, 2023). Based on Akhmedova (2016), most people from Karabulak village are bilinguals or trilinguals who speak Uzbek, Kazakh, and Russian. In addition, it is essential to mention that the younger generation also speaks English due to the 'Trilingual Education' policy implemented in Kazakhstan (Akhmedova, 2016).

1.3. Historical background

In his article, Juliboy Eltazarov (2013, p. 281) mentioned that residents of Karabulak village are descendants of Oghuz tribes of Turks who used to speak in Oghuz language. However, due to different cultural, political, and economic factors and the dominance of Karluk (Uzbek) in the linguistic area throughout history, the Oghuz dialect was affected by the Uzbek literary language and the Karluk dialects of Uzbek. In other words, it can be said that the Karabulak dialect has occurred as the result of language integration. According to Krylova (2014), language integration is a language development process where different languages and dialects unite. This process can be voluntary (natural) or forced by powers. Furthermore, this is the main reason why residents of Karabulak village speak in the Uzbek dialect today and identify themselves as a part of the modern Uzbek community today.

1.4. Linguistic profile of Karabulak Dialect

1.4.1 Classification

The Karabulak dialect was investigated by many Russian and Uzbek Turkologists as K.Yudakhin, E.Polivanov, K.Borovkov, V. Reshetov, etc. It should be noted that Turkologists refer to this dialect as the İkan-Karabulak dialect. As this research aims to analyze only the speech of Karabulak residents, this dialect will be mentioned as the Karabulak dialect in this paper. Yudakhin divided Uzbek dialects into five main groups, and the Karabulak dialect belongs to the Northern group of Uzbek (Oʻzbek dialektologiyasi, 1978, p. 29). Like Yudakhin,

Polivanov (1933, p. 25) also mentioned that the Karabulak dialect belongs to the Northern Oghuz group of Uzbek. The speakers of this dialect also live in neighboring villages such as Ikan, Mankent, and Karamurt in South Kazakhstan, aside from Karabulak (Polivanov, 1933, p. 26 and Yildirim, 2014, p. 137). Another Turkologist, A.K. Borovkov (1953), classified Uzbek dialects differently. Firstly, he classified Uzbek dialects based on their phonetic characteristics: æ-sound group and ɒ-sound group (Oʻzbek dialektologiyasi,1978, p. 34). Based on its phonetic characteristics, Karabulak dialects belong to the æ-sound group. The Uzbek dialect which belongs to the æ-sound group has preserved synharmonism. Later, Borovkov classified Uzbek dialects based on their historical-linguistics features. In this classification, the Karabulak dialect belongs to the fourth group, which has no genetic relation with other dialects (Oʻzbek dialektologiyasi, 1978, p. 36). Moreover, the Uzbek dialects in this group have long and short vowel distinctions and genitive suffix -ing. Another Soviet Linguist, V. Reshetov, classified Uzbek dialects into three groups: Kipchak dialect, Oghuz dialect, and Karluk-Chigil-Uighur dialect (Oʻzbek dialektologiyasi, 1978, p. 37). Reshetov, opposing other linguists, thinks that the Karabulak dialect belongs to the Karluk-Chigil-Uighur group (1978, p. 41). He also outlines that this dialect has some Oghuz and Kipchak features. In general, there are different views of Turcologist about the classification of the Karabulak dialect.

1.4.2 Morphological peculiarities

As this paper aims to analyze the Karabulak dialect in the framework of the 'olga- ~ bolga-' language, it is also essential to discuss the morphological features of this dialect based on existing sources. Juliboy Eltazarov, who conducted fieldwork in Karabulak village in 2013, outlined some morphological characteristics of the Karabulak dialect in his work. Based on Eltazarov (2013, p. 288), the grammatical cases in the Karabulak dialect are as follows:

Name of the case	Suffixes
Nominative	-
Genitive	*-ıŋ,-iŋ;-nı,-ni (-dı/-ti...)*
Accusative	*-ı,-i;-nı,-ni*
Dative	*-a/-ä;-gä/,-ğa; -kä,-qa,-qå*
Locative	dä /-da, -tä /-ta
Ablative	din /-tin;-dän/-dan;-tan/-tän

Table 1. Grammatical cases in Karabulak Dialect

Eltazarov (2013, p. 284) also adds that most of the Proto-Oghuz suffixes (grammatical cases) are found in the Karabulak dialect. Moreover, the Karabulak dialect has some suffixes borrowed from the Karluk language or developed as an influence of the literary Uzbek language.

It is interesting to note that the Karabulak dialect preserved its Oghuz features despite many centuries of language contact with Uzbek and different political, economic, and cultural conditions. Moreover, Eltazarov (2013, p.281) highlights that Karabulak residents are the descendants of the most ancient Oghuz tribes.

2. Literature review

2.1 Theory
Five main theories proposed by linguists explain the phenomenon of the '*olga- ~ bolga-'* language. Also, the main object of Marcel Erdal's article titled 'Explaining the olga-bolga dili' is to discuss these theories (2015, p. 135). The first theory explaining the phenomenon of the '*olga- ~ bolga-'* language was propounded by Mecdut Mansuroğlu in 1952 (Erdal, 2015, p. 139). According to him, '*olga- ~ bolga-'* texts are remnants of an early Oghuz literary language

developed under the influence of Khâqânî Turkic in the East of the Aral Sea and north of Sir Darya. Furthermore, this is the main reason for having mixed language features. This view was supported by M. Engin, Z. Korkmaz, and M. Canpolat (Tekin, 1974, p. 65). However, Şinasi Tekin (1974) argued against Mansuroğlu, saying that '*olga- ~ bolga-*' texts did not occur due to literary continuity. According to Tekin, '*olga- ~ bolga-*' texts were written by individuals who migrated to Anatolia from the East and who were not able to master the local dialect, thus wrote texts in mixed language which has both Eastern and South Western Turkish features (1974 and Erdal, 2015, p. 139). The third theory argues that '*olga- ~ bolga-*' texts were created under the influence of Khwarezm Turkic literature. Because Khwarezm Turkic literature was prevalent and appreciated in Anatolia (Erdal, 2015, p. 139). Even some Khwarezm Turkic literary works were translated into Turkish during that period (Erdal, 2015, p. 139). Gerhard Doerfer proposed the fourth explanation in 1977 (Erdal, 2015, p. 139). He argues that the '*olga- ~ bolga-*' texts were based on Khorasan Turkic dialects. It is because Khorasan Turkic people living in Eastern Iran migrated to Anatolia, and their speech before the complete acquisition of Oghuz was mixed (Erdal, 2015, p. 139). The fifth theory was proposed by Boeschoten (2004). He argues that some of the '*olga- ~ bolga-*' texts were written outside Anatolia, probably in the territory of the Golden Horde, Crimea, Caucasus, or Iran (2004, p. 222). Hence, these texts have mixed language features. Marcel Erdal (2015, p. 140) thinks that all of the explanations mentioned above are correct and reasonable. Moreover, different '*olga- ~ bolga-*' texts may have linguistic parameters of mixed language due to different factors. Marcel Erdal (2015, p. 140) also highlights that "mixed-language texts" are heterogeneous; consequently, they should be classified to determine which theory can be applied to which of '*olga- ~ bolga-*' texts.

2.2 Previous studies

Several studies discuss the issue of 'olga- ~ bolga-' mixed language by analyzing old Turkic texts. One of the first research in this field was conducted by Şinasi Tekin in 1974. In his study, Tekin (1974) analyzed old Turkic literature works such as *Feraiz Kitabi, Kissa-i Yusuf, Behçetu'l-Hada'ik*, etc. Moreover, as already mentioned above, Şinasi Tekin criticized the explanation proposed by Mansuroğlu (1952), which argues that '*olga- ~ bolga-*' texts were written in the transition period (XI and XIII century) of Old Turkish to Oghuz language and this is a reason why these texts have both South Western and Eastern Turkish features (1974, p. 67). In his article, Tekin provides some evidence to justify his critics. Firstly, Tekin outlines that *Tezkuretul - Evliyan,* which was written in the same period, has only a *bolga-* version of the verb, while based on Mansuroğlu's theory, all text written during that period should have an '*olga- ~ bolga-*' features including *Tezkuretul - Evliyan* (1974, p. 67). Additionally, Tekin mentions that texts such as *Kelile ve Dimne* and *Kisas-i Enbiya,* written in the transition period, have only Oghuz features (1974, p. 67). Consequently, Mansuroğlu's theory cannot be applied to all texts written between XI and XIII centuries. After arguing against Mansuroğlu's theory, Tekin proposed his explanation for the '*olga- ~ bolga-*' language issue. Tekin believes that '*olga- ~ bolga-*' texts were written by poets who migrated from Iran and Central Asia to Anatolia because of political pressure (1974, p. 69). Furthermore, the poets who spoke Eastern Turkish could not fully acquire the Oghuz dialect, thus consciously or unconsciously used both Eastern and South Western characteristics of Turkish in their literary works (Tekin, 1974, p. 69). For example, in the works of *Ibn Arabsah*, who has Arab origins and around 30 years lived in Turkistan and learned Chagatai, '*olga- ~ bolga-*' features can be determined (Tekin, 1974, p. 69). Later, he returned to Anatolia and started writing poetry in Oghuz Turkish, where he was not fluent. Thus,

his text has '*olga- ~ bolga-*' characteristics (Tekin, 1974, p. 69). Consequently, Tekin thinks that texts written between XI and XIII centuries have an '*olga- ~ bolga-*' feature because of the non-proficiency of some poets in Oghuz Turkish.

Another study that discusses the phenomenon of the '*olga- ~ bolga-*' language was conducted by Hendrik Boeschoten in 2004. Boeschoten analyzed the *Qisas-i Rabguzi* manuscript regarding orthography, phonetics, and morphology (2014, p. 217). The author determined that the text has only an *olga-* version of the verb, characteristic of Oghuz (2014, 221). Moreover, the author identified that the pronoun *'gandü'* (self) was used in this work besides 'öz' (2014, p. 221). According to Boeschoten, even though *the Qisas-i Rabguzi* manuscript mostly has South Western Turkish features, there are some instances of mixed languages, such as the use of both forms of accusative case /yI/ and /nI/, which belongs to South Western and Eastern Turkish respectively (2014, p. 221). Thus, the Qisas-i Rabguzi manuscript can be considered an '*olga- ~ bolga-*' text.

One of the recent studies on the issue of the '*olga- ~ bolga-*' language was conducted by Fatma Koç in 2020. In her article, Koç analyzed Turkic texts written between the XIII and XV centuries to determine phonetic characteristics of the '*olga- ~ bolga-*' language and reveal numerical data. Koç analyzed the following three works in terms of phonetics: Ali's Kıssa-i Yusuf, Kitâb-ı Güzîde, and FerâǾiż Kitâbı (2020, p. 55). Firstly, the author identified all phonetic/morphological characteristics of mixed language using a screening method (2020, p. 55). After it, obtained phonetic/morphological data were classified according to their characteristics in order to reveal numerical data (2020, p. 55). The main findings of Fatma Koç's study were that both *olga-* and *bolga-* forms of verbs were used in the literature mentioned above. In Kıssa-i Yusuf, the *bolga-* form was used 27 times, while the *olga-* form was used 415

times (2020, p. 66). In Kitâb-ı Güzîde, the *bolga-* form was used one time, while the *olga-* form was used ten times (2020, p. 66). In Ferâⵁiⵁ Kitâbı, the *bolga-* form was used one time, while *olga-* form was used 41 times (2020, p. 66). The author also identified that both ablative (–dIn or –dAn) forms were used with equal weight in these texts (2020, p. 66). These results led the author to conclude that the '*olga- ~ bolga-*' texts have more Oghuz Turkic features than Eastern Turkic.

3. Hypothesis and research questions

There are several reasons why I hypothesize that Karabulak Dialect is an '*olga- ~ bolga-*' mixed language. Firstly, as mentioned in the introduction, I have noticed that people living in Karabulak village use both *olga-* and *bolga-* forms of the verb 'to be', which is one of the main characteristics of '*olga- ~ bolga-*' languages (Erdal, 2015, p. 137). Secondly, Eltazarov (2013) argues that residents of Karabulak village are descendants of the Oghuz tribes and speak in the Oghuz dialect. Moreover, he outlines that the Karabulak dialect occurred because of the integration of Oghuz and Kipchak. This means that the Karabulak dialect has both Oghuz and Karluk features or South Western and Eastern Turkic features that led to some assumptions that this dialect might be an '*olga- ~ bolga-*' mixed dialect. Considering my observations, historical development, and linguistic features of the Karabulak dialect, it is assumed that this dialect is 'olga- ~ bolga-' mixed language and may exhibit other characteristics of the mixed language listed by Erdal (2015, p. 137). Moreover, there is a high probability that the speech of people living in Iqan, Mankent, and Karamurt villages also may have linguistic parameters of the '*olga- ~ bolga-*' language. However, it is not the focus of this paper.

Research questions:

1. Is the Karabulak Dialect of Uzbek an 'olga- ~ bolga-' mixed language?

2. What are the main linguistic parameters in the Karabulak dialect which indicate that it is an 'olga- ~ bolga-' language?

3. Do Karabulak dialects have exclusively *olga-* or exclusively *bolga-* features, or both?

4. What is the frequency of use of Oghuz and Eastern Turkic features in the Karabulak dialect?

5. Are Oghuz and Eastern Turkic characteristics in free distribution?

6. Why do Karabulak dialects have mixed language features? Do the theories mentioned in Erdal's (2015) article explain this phenomenon, or should other explanations be developed?

4. Methodology

4.1. Data Collection and Sampling

This project aims to show that Karabulak dialect is the *'olga- ~ bolga-'* language. Hence, the speech of Karabulak residents was analyzed. As there is no Corpus of Uzbek Dialects spoken in Kazakhstan, natural linguistic data was collected by three different methods. Firstly, all publicly available videos on social media such as Tiktok, Youtube, and Instagram were collected using the #Karabulak or #Qorabuloq. Only the videos spoken in regional dialects and do not contain inappropriate content were selected for analysis purposes. Secondly, the snowball sampling method was used for recruiting participants who were residents of Karabulak village, native speakers of Karabulak Dialect, and above eighteen years. According to Buchstaller and Khat (2013, p. 80), snowball sampling is a method where researchers use the social network to recruit participants, and this method can be applied in locally specific research. After the recruitment, participants were asked to share their WhatsApp audios. Thirdly, two recruited participants were invited to participate in audio recording to collect natural linguistic data.

Linguistic data was collected after the oral consent of participants, and ethical guidelines for linguistic research were followed. All of the recruited people participated voluntarily. A total of 57 minutes of natural linguistic data was collected, including the speech of 13 people (5 male and 8 female) on various topics and social contexts. It is also important to mention that the research participants were mainly multilingual and speak Kazakh and Russian (some English) along with Uzbek. Moreover, all participants have studied in Uzbek-medium schools where the lessons are instructed in the Standard Uzbek language, meaning that these speakers know both the Karabulak Dialect and the Standard variety of the Uzbek language.

4.2. Data Annotation

Before the annotation process, all audio and video recordings were watched and listened to, and only 38 minutes of the data was transcribed using the linguistic annotator ELAN. The language data was divided into meaningful sentences, clauses, and discourse markers before the transcription. Additionally, it was transcribed using both Cyrillic and Latin based Alphabet of Uzbek. Moreover, the selected instances of the *'olga- ~ bolga-'* phenomenon were glossed following the Leipzig Glossing rules and translated into English.
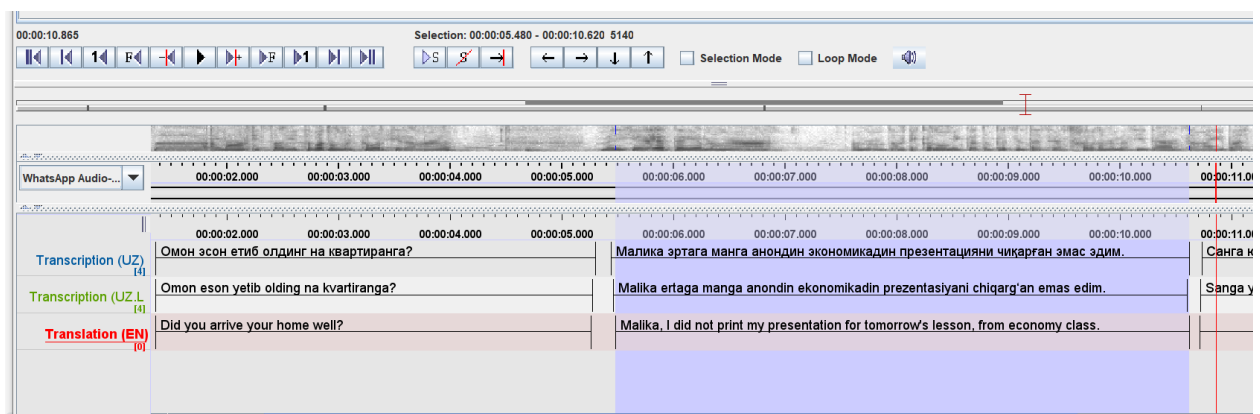
Picture 1. The transcribed material in the application ELAN

Besides, all names and personal information were changed during the transcription to protect the participants' anonymity. Also, transcription does not include non-verbal sounds such as pauses, laughter, chucklings, etc.

4.3 Coding

Approximately 400 sentences, clauses, or discourse markers were transcribed. Moreover, randomly selected 247 instances of natural data were coded to see if they showed the linguistic parameters of the *'olga- ~ bolga-'* language. The five main features of the mixed language were chosen from Erdal's article (2015, p. 137), and all selected data were checked to determine if they have these features. More specifically, selected data were checked to see if they have *olga-* or *bolga-* verb, postposition *bilan* or *ila,* accusative case, dative and ablative case. After that, sentences that do not show any characteristic of the *'olga- ~ bolga-'* language were coded as NA (not applicable). The rest of the sentences were classified based on the linguistic parameters: South Western Turkic or Eastern Turkic features. The instances of the South Western Turkic were coded as SWT, and the instances of the Eastern Turkic were coded as ET.

Data	Olga vs Bolga	Bilan vs Ila	Accusative	Dative	Ablative
Zakazat'ni shundoq bosangiz bo'di.	ET	NA	ET	NA	NA
Bir minuti ichida kaladi.	NA	NA	NA	NA	NA
To'y egalari bizardin rozi o'sin.	SWT	NA	NA	NA	ET
Alo taksi bizardin rozi bo'sin.	ET	NA	NA	NA	ET
Joyi jannatda o'sin	SWT	NA	NA	NA	NA
Xo'jayin sizga bir murojat bilan kadiydim	NA	ET	NA	ET	NA
Qizima sovchi qo'yg'an ekansan.	NA	NA	NA	SWT	NA

Table 1. The visual representation of the coded data.

After the coding process, some linguistic examples were analyzed from morphological, syntactic, phonological, and semantic perspectives to identify their pattern of use.

5. Results and Discussion

This chapter aims to discuss the main findings of the research. A total of 245 linguistic examples were analyzed, and 87 linguistic examples which do not show any features of the *'olga- ~ bolga-'* language were coded as NA. Moreover, 158 linguistic examples have at least one characteristic of the *'olga- ~ bolga-'* language mentioned by Erdal (2015, p. 137), and they are coded as SWT or ET. Thus far, analyzing the linguistic data, it can be said that Karabulak Dialect has both Oghuz and Eastern Turkic features. This means that the Karabulak Dialect of Uzbek is an *'olga- ~ bolga-'* mixed dialect.

Features	South Western Turkic	Eastern Turkic	Total
Olga- vs bolga-	68	20	88
Postposition	0	9	9
Accusative case	2	38	40
Dative case	7	41	48
Ablative case	0	35	35

Table 2. Statistical data.

5.1. The verb *olga-* and *bolga-*

A total number of 88 instances of the *olga-* or *bolga-* form were identified from the analyzed linguistic data. Moreover, 68 linguistic examples have an *olga-* form of the verb, a South Western Turkic feature. In comparison, 20 of them have a *bolga-* form, which is an Eastern Turkic feature. This means that the Oghuz form of the verb 'to be' is more frequently used by speakers of the Karabulak Dialect.

In the following sections, structure and function of the *olga-* or *bolga-* forms will be analyzed.

5.1.1 Main verb of the sentence

It was found that both *olga-* or *bolga-* forms can be the main verb of the sentence. For instance, in the following sentences, *olga-* form is the main verb of the sentence.

(1) *Zavod-ni   eng   yaxshi   ishchi-lar-i-din   bir-i   oʻl-di-m.*
Factory-GEN   most   good   worker-PL-SG-ABL   one-?   be-PST-1SG
(I was one of the best employees of the factory.)

(2) *Bugun   jonazas-i   oʻ-di.*
Today   funeral-3SG   be-PST
(Today was his funeral)

(3) *Qaysi   dars   oʻl-di ?*
Which   lesson   be-PST ?
(What lesson was it?)

In addition, the main function of the *olga-* the verb in given sentences is to inform about the state/condition of the subject. Similarly, the *bolga-* form can also be used for the same purpose and be the sentence's main verb, as illustrated in example (4).

(4) *Oʻziz-da   ne   boʻ-yatti ?*
Self-LOC   what   be-IMPF
(What about you?)

Hence, it can be claimed that *olga-* or *bolga-* forms appear in free distribution when they are the sentence's main verb. However, the *bolga-* form of the verb is less frequent than the *olga-* form in this context.

5.1.2 Converb clause

Moreover, both *olga-* and *bolga-* forms can be inflected by the suffix *-ip* and function as a converb. For instance:

(5) *Mashu   zavod-qa   oddiy bir   ishchi   oʻ-b   kir-gan-man.*
Worker   factory-DAT   simple one   worker   be-CONV   enter-PST-1SG
(Two years ago I came came to this factory as an ordinary worker)

14

(6) *Kelin* **bo'-b** *tush-gan-din keyin oydaq ovqat ye-ma-di-g'u*
Bride be-CONV fall-PST-ABL after good food eat-NEG-PST-?
(After you came as a daughter-in-law, we did not eat any good food)

Considering that in the provided examples, both *olga-* and *bolga-* have the same grammatical structure and function, it can be said that they occur in a free variation. Nevertheless, it should be noted that the *olga-* form of the converb is more common in the annotated data.

5.1.3 Imperative sentences

Additionally, it is identified that both *olga-* and *bolga-* forms of the verb can be used in imperative sentences. Examples (7) and (8) illustrate this phenomenon.

(7) *To'y-lar muborak* **o-'sun**.
Wedding-PL blessed be-2SG.IMP
(May your marriage be happy)

(8) *Bizar-din rozi* **bo'-sin**.
3PL-ABL glad be-IMP
(May he will be satisfied by us)

As shown above, both verb forms are inflected by the imperative suffix -s(I)n and perform the same function, which means that they appear in free variation in this environment. However, the *olga-* form has a higher frequency of use in imperative sentences.

In paragraphs 5.1.1, 5.1.2, and 5.1.3, it was discussed that *olga-* and *bolga-* forms are in free distribution when they are the main verb of the sentence, converb, or have an imperative function. This probably means that the only difference between them is their frequency of use. Maybe there is some phonetic explanation for this phenomenon? However, based on the dataset, it cannot be verified; more data is needed to prove this assumption.

5.1.3 Semantic change

In order to find the reason why Karabulak residents use the *bolga-* form, each sentence with it was analyzed from a syntactic, morphological, semantic, and functional perspective. It was determined that the *bolga-* a form of the verb experienced a semantic shift. Hence it has various functions. Moreover, it was found that in a particular discourse, only the *bolga-* form is used by Karabulak residents. This means that *bolga-* forms have complementary distribution in particular cases. In the examples provided below, the *bolga-* form appears in a set environment.

(9)   *Ha, **bo'di!** Tepaga chiqabenglar!*
        (Yes, OK! Come upstairs!)

(10)   *Xo'p, **bo'ldi**, xo'p!*
        (OK, all right, OK!)

(11)   ***Bo'ldi, bo'ldi!** Erta-ga-chay bu ishlarni qib qo'yama.*
        (OK! OK! I will do this works by tomorrow)

(12)   **Bo'ldi**, go'rkov bilan gaplashib ko'rayik!
        OK! Let's talk with gravedigger!

Firstly, in the given examples, the *bolga-* form performs as a discourse marker, not as a verb. Secondly, the primary function of the *bolga-* form in the examples is to mark agreement. In other words, *the bolga-* form functions as the English words 'Ok,' 'All right,' or the Turkish word 'Tamam,' 'Peki.' This shows that *bolga-* forms have various semantic meanings in the Karabulak Dialect of Uzbek and can function as a discourse marker. However, no instance of *olga-* form functioning as a discourse marker was identified from the linguistic dataset. Consequently, it can be stated that the *bolga-* form is in complementary distribution when it is a discourse marker.

In addition, the *bolga-* form can be used as a discourse marker to show disagreement, as

in the following examples:

(13)   ***Bo'ldi!*** *Yig'lavermachi!*
        (That's enough. Stop crying!)

(14)   *E,* ***bo'di!*** *Yetadi!*
        (That's enough. Stop!)

Using the *bolga-* form inflected with the past tense suffix *-di*, the speaker wants to highlight his/her disagreement and stop the conversation. Interestingly, *olga-* forms never occur in the same environment. This means that *bolga-* forms have complementary distribution in this discourse. Moreover, it was identified that the *bolga-* form also has a restrictive meaning. And it can function as a discourse marker, as illustrated below:

(15)   *Shu birinchi balkon-g'a-chay oparib ko'ysang* ***bo'ldi.***
        (You have to **just** accompany them to the balcony.)

(16)   *Zakazat'ni shundoq bosangiz* ***bo'di.***
        (You have to just press the (deliver) button.

(17)   *Shuni belgilaysiz dag'in* ***bo'di.***
        (You have to **just** click that.)

In these examples, the *bolga-* form is used to restrict an action. In other words, the *bolga-* form behaves as English 'just' or as Turkish 'Yeterli.' Moreover, the conditional suffix *-sa* is attached to the main verb of the sentences (15) and (16). In this environment, *bolga-* form is in complementary distribution. Because no evidence of *olga-* forms were found from the dataset in the same condition.

In addition, based on the transcribed data of Eltazarov (2013, p. 299), Karabulak residents also use the *volga-* form of the verb 'to be.' Nevertheless, no instances of the *volga-* form was identified in my dataset. It should be noted that most of the participants of Eltazarov's

research are older people who are 70 and above. In comparison, the speech of the younger group was analyzed in this research. It is likely that the *volga-* form is used only by the older generation, and it is rare among young people.

5.2. Postpositions *bilan* and *ila*

According to Erdal (2015, p. 136), in the South Western Turkic languages, the sound /b/ is dropped in the *bolga-* verbs. Consequently, the postposition *bilan*, which has a similar phonetic form as *bolga-* verb, is also expected to have initial consonant deletion and be used as ila/ilen says Erdal (2015, p. 136). Knowing that the Karabulak dialect has both *olga-* and *bolga-* forms, it is significant to check which form of the postposition is used by Karabulak residents. From the analyzed data, only 9 sentences out of 245 with the postposition *bilan* were identified. Moreover, there is no sentence with the *ila/ilen* form of the postposition. This indicates that the postposition *bilan* did not have an initial consonant deletion despite the sound /b/ being dropped in the *bolga-* form. Some sentences with postposition *bilan* are provided below:

(18)   *Inson-lar*   *bilan*   *yaxshi*   *muloqat-ta*   *o'l-a-di.*
        Person-PL   with   good   relationship-LOC   be-?-PST
        (They are in a good relationship with the people)

(19)   *Siz-ga*   *bir*   *murojat*   *bilan*   *ka-diy-dim*
        You-DAT   one   request   with   come-PST-1SG
        (I came to you with one request)
    So why is the sound /b/dropped in *bolga-* word but remains in the postposition *bilan*? It might be an interesting question to further research. Eltazarov (2013, p. 289) also mentioned that there is also a *vilän* form of the postposition used by Karabulak residents. However, this form is not found in the linguistic dataset. It is expected that, as a *volga-* verb, the postposition *vilän* is also used only by older generations.

5.3. Accusative Case

Erdal (2015, p. 137) mentioned that it is also significant to check if the language uses the +(y)I form of the suffix as in Turkish or the +(n)I form as in Azerbaijani and Turkmen. Moreover, only the suffix +(n)I is used in written Uzbek. From the linguistic data, it was identified that 38 words were infected by the suffix +(n)I, which is a feature of Eastern Turkic. Additionally, only 2 instances of the accusative suffix +(y)I determined, which is characteristic of the South Western Turkic dialects. For example, in the following sentences accusative suffix +(n)I attached to the object of the sentences :

(20)   *U*   *dars-im-**ni***   *o'-ti-m.*
       3SG   lesson-1SG-ACC   instruct-PST-1SG
       (I instructed that my lesson)

(21)   *Man*   *slar-ga*   *bu*   *video-**ni***
       1SG   2PL-DAT   this   video-ACC
       (I am posting this video)

However, in examples (22) and (23), the South Western Turkic accusative suffix +(y)I is used by Karabulak residents.

(22)   *Ayb-im-**i***   *bekit-adigan*   *bo'l-di.*
       Guilt-1SG-ACC   hide-PTCP.NPST   be-PST
       (It will hide my guilt)

(23)   *Shu*   *akkaunt-im-**i***   *TikTok*   *blok-qa*   *tiq-ib*   *tashlap-di.*
       That   account-1SG-ACC   TikTok   block-DAT   put-CNV   through-PST
       (TikTok blocked my that account)

From the examples given above, it can be seen that the suffix +(y)I always comes after the 1st person possessive suffix. Nevertheless, there is no specific condition for the suffix +(n)I. This means suffix +(y)I have complementary distribution and suffix +(n)I in free variation. At this stage, no explanation can be provided for this phenomenon. More data is needed for more

accurate analysis, or there is no reasonable explanation. The main difference between the South Western and Eastern form of accusative suffix is their frequency of occurrence. The Eastern Turkic accusative +(n)I is more frequent in the Karabulak dialect.

5.4. Dative Case

Erdal (2015, p. 137) outlines another main feature of the *'olga- ~ bolga-'* language is using both +GA and +(y)A forms of the dative case. A total number of 41 +GA forms of the dative suffix was identified from the linguistic data. Despite the fact that the Eastern form of the dative is more common, there are also 7 instances of the South Western Turkic +(y)A . The +GA form of the dative attached to the word in the following examples:

(24) *Samalyot-i-**g'a***   *katta   sumka-siz   o-g'an-idi-m.*
      Airplane-3SG-DAT   big   bag-NEG   take-PST-PST-1SG
      (I bought the plane ticket without big luggage)

(25) *Man   slar-**ga**   bu   video-ni   chiq-ar-ib   tur-ip-man.*
      1SG   2PL-DAT   this   video-ACC   post-AOR-CNV   stand-CONV-1SG
      (I am posted this video for you)

In addition, based on linguistic examples, it was identified that the +(y)A form of the dative only attaches to nouns in the first person singular. For instance, in the following examples preceding suffix of dative +(y)A is first singular +(I)m :

(26) *Keyin   as-im-**a***   *ka-b   qol-di.*
      Then   mind-1SG-DAT   come-CNV   stay-PST
      (Then it came to my mind)

(27) *Qiz-im-**a***   *sovchi*
      Daughter-1SG-DAT   sovchi
       (Sovchi for my daughter)

However, there are no conditions for the Eastern dative +GA. This means that the Eastern form of dative is in free distribution, and the South Western form is in complementary

distribution. Besides, examples (28), (29), and (30) also can be an instance for the South Eastern form of the dative case. However, it should be noted that some scholars may gloss it differently. For instance, the word *maŋa*, in example (28), was divided into two significant parts: maŋ-a, where the maŋ is stem (1SG) and the *-a* is a dative suffix.

(28)   *Odam-lar*   *maŋ-a*   *podpischik*   *o'l-yati.*
       Person-PL   1SG-DAT   follower   be-PRT
       (People are becoming my followers)

(29)   *Saŋ-**a***   *yub-or-ib*   *ber-sa-m*
       2SG-DAT   send-AOR-CNV   give-COND-1SG
       (If I will send it to you)

(30)   *Omon-eson*   *yet-ib*   *o-di-ng*   *na*   *kvartira-ŋ-**a**?*
       Good   arrive-CNV   take-PST-2SG   INT   apartment-3SG-DAT
       (Did you get to your apartment well?)

5.5. Ablative Case

The next feature of the *'olga- ~ bolga-'* language mentioned in Erdal's (2015,137) article is the use of both South Western and Eastern Turkic forms of the ablative. The 35 instances of the ablative case were found in the dataset. Moreover, all ablatives have +dIn, a linguistic feature of Eastern Turkic. Some examples of it illustrated below:

(31)   *Bir*   *odam*   *birov-**din***   *qarz*   *o'-b*   *qo-sa*
       One   person   someone-ABL   debt   be-CNV   stay-COND
       (If some person loans money from someone)

(32)   *To'y*   *ega-lar-i*   *bizar-**din***   *rozi*   *o'-sin*
       Wedding   owner-PL-3SG   1PL-ABL   glad   be-3.IMP
       (May the weddingers will be pleased by us)

Hence, this means that the Karabulak dialect does not have the South Western form of the ablative based on analyzed data.

6. Further Discussion and Interpretation

It has been proved that the Karabulak dialect of Uzbek is an '*olga- ~ bolga-'* mixed dialect. As mentioned in the introduction, the characteristics of the '*olga- ~ bolga-'* language was only identified in ancient Turkic writings. Considering that the Karabulak dialect of Uzbek is a modern Turkic dialect with mixed language features, it can be claimed that '*olga- ~ bolga-'* features not only emerge in historical Turkic texts but also it can be found in Modern Turkic dialects. Moreover, it is unknown whether Turkic spoken languages in the 13th-14th centuries also had '*olga- ~ bolga-'* features, and it cannot be identified because of the unavailability of recordings. However, this research was based on spoken language. So, there is a probability that not only written language in the 13th-14th century, but also spoken language had '*olga- ~ bolga-'* features. If this assumption is correct, this means that the theory proposed by Şinasi Tekin in 1974 is not accurate. Tekin (1974) argued that old Turkic texts have an '*olga- ~ bolga-'* feature because poets did not fully learn the Oghuz language and thus used some Eastern Turkic linguistic features in their works. To clarify, according to Tekin (1974), '*olga- ~ bolga-'* texts were produced due to an individual language characteristic of a writer, and it is not collective. Consequently, Tekin's explanation might be inconsistent. Moreover, Tekin's theory cannot explain why the Karabulak dialect of Uzbek has '*olga- ~ bolga-'* features.

So, why do Karabulak dialects have mixed language features? There are two possible explanations for this phenomenon. Firstly, if Karabulak residents are descendants of Oghuz tribes who used to speak in the Oghuz dialect and today's Karabulak dialect occurred as an integration of Oghuz and Karluk, as mentioned by Eltazarov's article (2013), it can be argued that the Karabulak dialect has mixed language features as a result of language contact. In this case, the *olga-* verb, the Oghuz feature, is a more ancient form. And, Karabulak residents borrowed the

bolga- form from the Karluk language. There is also another possible explanation. If the *bolga-* the verb is a more ancient form of the verb used by Karabulak residents, it can be assumed that the dialect has mixed language features as a consequence of phonological change where the initial consonant /b/ was dropped over time. And, Karabulak residents use both the *olga-* and *bolga-* forms as they are in the transition period where the /b/ sound is dropped; however, they could not fully adopt a new version. These two explanations are at the hypothesis level. In order to prove these hypotheses, the most archaic version of the verb 'to be' should be identified.

As discussed in section 1.4, there are different opinions about the classification of the Karabulak Dialect of Uzbek. One scholar says that it belongs to the Northern Oghuz group of Uzbek, and another one claims that the Karabulak dialect has no genetic relation with Uzbek. Eltazatov argues that it is the Oghuz language. Maybe the Karabulak dialect of Uzbek should be classified as an '*olga- ~ bolga-*' mixed dialect.

7. Conclusion

The main object of the study was to determine whether the Karabulak Dialect of Uzbek is the '*olga- ~ bolga-*' language or not. After the analyses of natural linguistic data, it was found that the Karabulak Dialect of Uzbek is a mixed language with both South Western and Eastern Turkic characteristics. The main outcome of the research was that Eastern Turkic features are more common in the Karabulak Dialect. Moreover, it was found that *bolga-* forms mainly occur in a set environment. Moreover, the *bolga-* forms function as discourse markers to show agreement, disagreement, and restriction. Another finding is that the *olga-* form always performs as the main verb or converb of the sentence. There are some cases when the *bolga-* form also

acts as the main verb or converb. However, no explanation for the case can be provided at this stage of research. It was also revealed that the South Western form of the accusative and dative appear in complementary distribution. Furthermore, no instance of South Western forms of ablative and postposition were determined.

Even though this study reached significant results, the research has some limitations. One of the study's main limitations is restricting the sample size to young people between the ages of 18-45 and not considering the more elderly population. Based on Eltazarov's research, the elderly population uses the three different versions of the 'to be' verb, including *the volga-* form, not determined in the present study. Moreover, there might be some mistakes during the transcription process, as transcription is a subjective process. Another limitation of the study is considered only five features of the '*olga- ~ bolga-*' language. For more accurate analysis and interpretation, it is advisable to consider all characteristics of the mixed language provided in Erdal's article (2015). Additionally, it is recommended to analyze highly interactional conversations in further research. It is because in monologues, *bolga-* form functions only as a verb or converse, while in dialogues, it may have more functions and semantic meanings. It is also suggested to analyze the Karabulak dialect of Uzbek from the lexical perspective. Because during the analysis process, some words which are not used in Standard Uzbek were found. They might be Oghuz equivalents of Karluk words.

There are some unanswered questions that can be the focus of future research. It would be interesting to examine old Turkic texts and check if *bolga-* forms are also used as discourse markers. Additionally, this research only analyzed the speech of Karabulak residents. In further studies, the Uzbek dialect spoken in other villages as Mankent, Ikan, and Karamurt, can be included.

7. Reference

Akhmedova, K. (2022). Қазақстанды етпен қамтыған Қарабұлақ.

https://adilet-gazeti.kz/қазақстанды-етпен-қамтыған-қарабұла/

Akhmedova, Z. (2016). Кўҳна ва навқирон Қорабулоқ.

Boeschoten, H. (2014). "Some Notes on 'Mixed' Written Western Oghuz Turkic." In E. Csató, B. Utas, and C. Jahani (eds.), Linguistic Convergence and Areal Diffusion. https://ebookcentral-proquest-com.ezproxy.nu.edu.kz/lib/astana/reader.action?docID=19 9 400#

Buchstaller, I. & Khattab, G. (2013). Population Samples. In Podesva & Sharma

(Eds.). Research Methods in Linguistics. Cambridge University Press.

https://sharifling.files.wordpress.com/2014/10/5_population-samples.pdf

Erdal, M. (2015). Explaining the olga-bolga dili. Turkish Language, Literature, And History Travelers' Tales, Sultans, And Scholars Since The Eighth Century, Ed. Bill Hickman and Gary Leiser, Routledge, 135-146. https://www.taylorfrancis.com/books/edit/10.4324/9781315750705/turkish-language-liter ature-history-bill-hickman-gary-leiser

Eltazarov, J. (2013). GÜNEY KAZAKİSTAN'DA AZ KONUŞURLU BİR OĞUZCA DEĞİŞKE ÜZERİNE. Tehlikedeki Diller Dergisi, 3(3), 275-307. https://dergipark.org.tr/en/download/article-file/305242

Karabulak (3 February 2023). Retrieved from Wikipedia. https://en.wikipedia.org/wiki/Karabulak_(South_Kazakhstan

Koç, F. (2020). KARIŞIK DİLLİ ESERLERİN FONETİK ÖZELLİKLERİ ÜZERİNE BİR DEĞERLENDİRME. Milli Kültür Araştırmaları Dergisi, 4(1), 55-68. https://dergipark.org.tr/tr/download/article-file/1135367

Krylova, M. (2014). Интеграция и дифференциация как процессы развития языка: история, современное состояние, перспективы. Филология и литературоведение, (2), 2-2.https://philology.snauka.ru/2014/02/665

Polivanov, E. (1933). Узбекская диалектология и узбекский литературный язык. Ташкент: Узгосиздат.

Reshetov V. & Shoabdurahmonov Sh. (1978). Oʻzbek dialektologiyasi. Tashkent.

Tekin, Ş. (1974). 1343 Tarihli Bir Eski Anadolu Türkçesi Metni ve Türk Dili Tarihinde 'Olǧa-bolǧa' Sorunu. Türk Dili Araştırmaları Yıllığı - Belleten , 21-22 (1973-1974), 59-157. Retrieved from https://dergipark.org.tr/tr/pub/belleten/issue/38478/446183

Yildirim, H. (2017). ÖZBEK TÜRKÇESİ AĞIZLARI ÜZERİNE BİR DEĞERLENDİRME. *Dil Araştırmaları*, *11*(20), 121-145.978. https://dergipark.org.tr/tr/pub/dilarastirmalari/issue/60166/872300