To Switch or Not to Switch: Functions of Code-switching in Kazakh Direct Reported Speech

Karlygash Mukanova

School of Sciences and Humanities, Nazarbayev University

WLL 499: Languages, Linguistics, and Literature Capstone II

Course Supervisor: Pr. Eva-Marie Dubuisson

Academic Advisor: Pr. Giorgia Troiani

May 4th, 2025

Table of Contents

1. Introduction

Kazakhstan is a multi-ethnic country that promotes and uses Kazakh as a national language, Russian for interethnic, and English for international communication. This linguistic situation in Kazakhstan emerged due to the multi-ethnic composition of the state, the country's policies for preserving the rights and interests of all ethnicities living in Kazakhstan, and its history (Aliaskar, 2024; Smagulova, 2008; Syzdykbayeva, 2016).

According to Zharkynbekova & Chernyavskaya (2022), Kazakhstan has a special type of multilingualism, which is Kazakh-Russian bilingualism. Bilingualism refers to a person's ability to use their native and second languages proficiently. However, there are cases when the degree of proficiency in those languages is unbalanced, and one dominates the other (Aliaskar, 2024; Zharkynbekova & Chernyavskaya, 2022). This language dominance can be observed in Kazakhstan, where people are more proficient in Russian than in Kazakh. According to the national census collected in 2009, 94% of Kazakhstan's population reported proficiency in Russian and 74% considered themselves proficient in Kazakh (Aliaskar, 2024, p. 11). In 2021, 84% of the population reported proficiency in Russian, while 80% reported proficiency in Kazakh (Bureau of National Statistics). Although proficiency in Kazakh has increased over a decade due to the promotion and revitalization efforts of the Kazakh language and culture (Pavlenko, 2008; Smagulova, 2008; Syzdykbayeva, 2016), the use of Russian is still prominent.

The reason for such dominance of Russian on a national level, despite Kazakh being the official language of the state and the native language for many, is rooted in its history. Kazakhstan is a post-Soviet country that was affected by Soviet policies promoting Russian as a lingua franca (Aliaskar, 2024; Smagulova, 2008; Pavlenko, 2008). Russian was obligatorily taught in schools and was used in official and professional settings. The dominance of the Russian language also

increased with the settlement of Russian-speaking ethnic groups in Kazakhstan. While the government could encourage the sole use of Kazakh on a national level after gaining independence in 1991, Kazakhstan adopted a multilingual policy since it is home to various ethnic groups who speak Russian (Smagulova, 2008, Syzdykbayeva, 2016).

Code-switching is a common phenomenon in such multilingual communities. Code-switching plays a significant sociolinguistic role in forming group identity and promoting mutual understanding between speakers (Aliaskar, 2024; Auer, 1998). People often code-switch to adjust to another person's linguistic background, showing their respect and fostering better communication. The use of code-switching also depends on various factors and contexts. It can be used at schools for instructional purposes to aid an understanding of students or at home when speaking to family members. Interviews with bilingual university students revealed that they code-switch to Russian when discussing academic topics and switch to Kazakh to talk about topics related to family or culture (Smagulova, 2020, as stated in Aliaskar, 2024). This again highlights Russian language dominance in academic discourses because speakers code-switch to a more commonly used language, where it is easier to access a wide range of vocabulary.

Currently, all three languages (Kazakh, Russian, and English) are taught in schools starting primary school, so speakers develop competence in those languages from an early age. Code-switching from Kazakh to Russian or between all three languages is a common phenomenon in Kazakhstan. The speakers of Kazakh frequently mix these three languages in the conversation, especially in informal situations. Consider the following example of the conversation from the Multimedia Corpus of Kazakh Spoken Language (Filchenko et al., 2023) (see section 4):

(1) tate160522 (0:34:54)

IU	Speaker	Original	Translation
1	KAMSHAT;	*Men aitamyn-,*	I say-,
2	BIBINUR;	*Bruh.*	Bruh.
3	KAMSHAT;	*<RU> tipa:<RU>,*	like:
4		*<RU> "Nu, my ne khotim ee uvol'njat',*	"We cannot fire her,
5		*ei vsego lish' tam<RU>,*	She just has,
6		*bir apta gana qaldy."*	a week left"
7		*dep aitady da,*	she says,
8		*<EN> And I was like:*	
9		*"Fuck you." <EN>*	

In (1), the speaker tells a story about the rude receptionist and uses direct reported speech to deliver it. Direct reported speech (DRS) is the type of speech that one uses to quote another person in the conversation. It can be noticed how the speaker switches between three languages within her turn and even within a quotation. In intonation unit (IU) 1, Kamshat starts her turn with the Kazakh reporting verb *aitamyn*, which means 'I say' in English, and also adds a Russian-borrowed quotative *tipa* ('like') before the quote. Through IUs 4-5, she talks in Russian, switching to Kazakh mid-sentence (IUs 6-7). She then finishes her turn in English. This conversation serves

as an example of how speakers of Kazakh code-switch during the conversation. Given that reporting speech constructions already quote another person, code-switching seems to play an additional interactional role in the direct reported speech of multilinguals (Frick & Riionheimo, 2013). In a multi-ethnic and multilingual country like Kazakhstan, it is interesting to explore how speakers decide on which language to use when quoting another speaker, as this decision reveals information about language attitudes and linguistic identity of speakers in the community and how they interact with each other.

Thus, the current study aims to explore the role of code-switching in direct reported speech in Kazakh through conversation analysis and discourse analysis applied to the naturally occurring modern Kazakh spoken data. The study explores the following research questions:

RQ1. What triggers code-switching in Kazakh direct reported speech?

RQ2. What are the pragmatic context and communicative functions of code-switching in Kazakh DRS?

To answer the first research question, I look at both lexical and interactional means of triggering code-switching in the conversation. According to Riehl (2005), words that are shared by a linguistic community, such as discourse markers (*potom* ('then'), *eto* ('this'), etc.), can be the cause of code-switching, so the hypothesis is that the use of Russian quotatives such as *tipa* ('like') and *takoj* ('such') before the quote in Kazakh will trigger code-switching. Given the findings of Broersma et al. (2009), I also hypothesize that the use of Russian or English phrases by other speakers will not trigger code-switching of the primary speaker. Building on this analysis of trigger words and the separation of non-functional and functional code-switching, I then answer the second question, which explores code-switching functions in DRS. I anticipate that code-switching

will be mostly used to establish credibility for the provided information and ensure mutual understanding between speakers (Aliaskar, 2024).

In section 2, I review the literature on DRS and code-switching in DRS, establishing the context for further discussion of my methods (section 3), results (section 4), and discussion (section 5). The last section concludes and summarizes the main findings of the capstone. The appendix at the end includes the link to all collected data.

2. Literature Review

The first part of the literature review introduces direct reported speech (DRS), its definition, the role of DRS in conversation, and typical ways of marking DRS in multiple languages, including Kazakh. The second part defines code-switching, triggers of code-switching, and its functions in DRS.

2.1. Direct Reported Speech

2.1.1. Definition of DRS

Reported speech is used to refer to someone else's words (Voloshinov, 1973). Being one of the major types of reported speech, direct reported speech (DRS) is sometimes called direct speech, direct quotation, and direct discourse (Haberland, 1986; Yin, 2023). The canonical definition of direct reported speech is that it "reproduces the original speaker's words" and shows the situation from the original speaker's perspective by preserving all linguistic features of that speaker (Evans, 2012, p. 86).

Although it is referred to as a direct quotation, DRS is not actually direct (Evans, 2012; Mayes, 1990). It is more about "what could have been said," so talks in DRS are usually invented by speakers (Dubois, 1989; Evans, 2012, p. 73; Guerini, 2009). Yin (2023) and Evans (2012) also

state that drawing a clear distinction between DRS and indirect reported speech (IRS) is difficult because there are intermediate cases that exhibit features of both reported speech types. Despite these hardships, "canonical direct speech is widely attested," which makes it easier to exemplify DRS in many languages (Evans, 2012, p. 96).

For the purposes of this study, it is important to define two types of events in DRS. When one person quotes another person, the act of quoting and current interaction is called primary speech event (PSE) (Evan, 2012). It should be distinguished against reported speech event (RSE), which refers to the event that happened before and is currently being quoted. Respectively, I will refer to the current speaker as the primary speaker and the speaker who is being quoted as the reported speaker.

2.1.2. DRS as an interactional evidential

One can easily find DRS in narrative stories speakers introduce in everyday conversations when sharing and recalling their past experiences (Yin, 2023). Depending on the context, different linguistic devices can become stance markers, even those that are not initially considered stance markers, such as direct reported speech. Reported speech is not placed randomly and has interactional purpose and systematicity. Speakers can use DRS to indicate epistemic authority (Clift, 2006; Yin, 2023). Epistemic authority refers to "authority we ascribe to people in virtue of their favorable relation to epistemic goods" (Jäger, 2025, p. 1). In other words, it is the authority that other speakers prescribe to the narrator. By positioning reported speech in a certain order within a conversation, one can achieve this epistemic authority and deter (sometimes only temporarily) other conversation participants from offering their evaluations on the situation. Clift (2006) argues that DRS is an interactional evidential that should be separated from the "stand-alone" evidential. The stand-alone evidential is a linguistic device used to mark whether the person

heard about something from someone else or saw the situation oneself. For example, in the Kazakh language, such evidentials are grammatically marked via morphemes, as in (2).

(2) *Ol* *bar-dy* ***de-p*** ***ait-yp-ty***

3SG go-PST say-CVB2 say-EVD.PST-PST

'Apparently, he/she said that he/she went there'

The interactional evidential can be identified by considering the context, while the stand-alone evidentials are found based on form. This highlights that accounting for the context in DRS is important, especially when analyzing the interactional functions of DRS and linguistic devices within the DRS.

2.1.3. DRS marking

In this subsection, I will discuss common ways to mark DRS in a language (reporting verb, verb of accompanying action, prosody, quotatives, and pseudoquotations) and boundaries of direct reported speech. The clause before the beginning of DRS is called onset, while the end of it is referred to as offset (Bolden, 2004). Using data from the corpus of conversational Russian with 40 hours of phone conversation audio recordings and three hours of workplace conversation videos, Bolden (2004) presents different ways of beginning a quote. The first one has the SPEAKER + REPORTING VERB format, as in example (3).

(3) *Ona* *govori-l-a*

3F.SG say-PST-3F.SG

'She said'

Kazakh speakers also use this type of DRS marking. However, SPEAKER + REPORTING VERB construction in Kazakh reverses to "reporting verb + speaker" construction at the end of the quote, as illustrated by Muhamedowa (2016) in the following sentence (4):

(4) *"Qol-y-nan is kel-me-y-tin a'kim-der-den qutyl-uw kerek",* ***- de-di Elbasy.***

Hand-3SG.POSS-ABL work can-NEG-CVB1-DER.NMZ mayor-PL-ABL get.rid-INF need say-PST leader.of.the.nation.

From hand work that cannot from mayors need to get rid of said leader of the nation.

'"Get rid of mayors who cannot do their job properly," the leader of the nation said' (p. 29).

Kazakh speakers usually use verbs of saying, such as *de* ('say') in DRS. If any other verb of saying is used, it follows the conjunction *dep*. For example, one could say *dep aitty* ('said that').

The second method of reporting someone else's words is using a verb that describes an accompanying action. Bolden (2004) provides an example of such a quotation in (5).

(5) ***Astapov pozvoni-l****: "Bol'shoe tebe spasibo"*

Astapov call-PST: big.NSG you.DAT thanks

Astapov called: Big for you thanks

'Astapov called: "Thank you very much"' (p. 1083)

Here, the verb is not a verb of speech, but an action associated with the reported speech event.

Prosody can also be used to mark the onset of the quote (Bolden, 2004; Klewitz & Couper-Kuhlen, 1999). Prosody refers to acoustic features of speech, including volume, intonation, and pitch. According to Klewitz & Couper-Kuhlen (1999), prosody delimitates the boundaries of reported speech and functions similar to quotation marks in written communication. This delimitation is mostly associated with a shift in pitch range or register. For example, one person

can switch to a higher pitch when quoting the original speaker. Reported speech can also be marked by volume changes, rhythm, or all of them simultaneously.

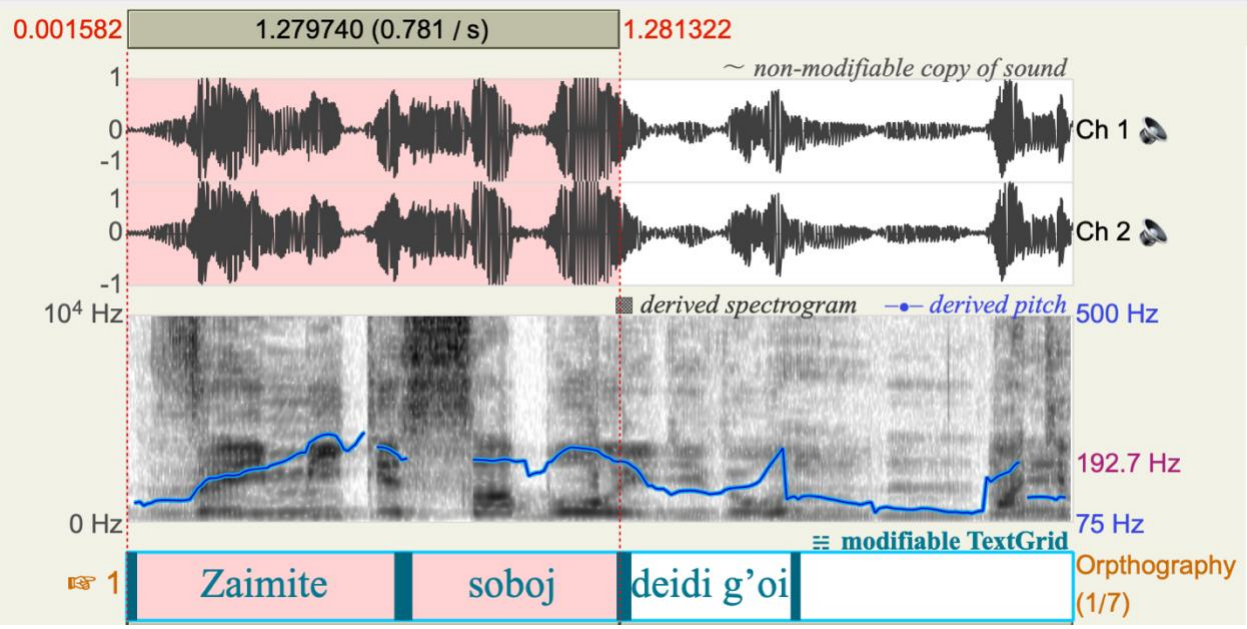

Image 1. Prosody exaggeration example in Praat

In the example above, illustrated by Image 1, the speaker uses prosody for quotation. The blue line along the spectrogram demonstrates the pitch change. The part highlighted in red is the quote, where the pitch is higher compared to the matrix clause that follows it.

Klewitz & Couper-Kuhlen (1999) also found that the choice of using prosody for marking reported speech does not depend on the presence or absence of quote-framing devices (like "she said") but rather on the stance the quoting person is taking. In most cases, prosody does not correspond to the actual boundaries of reported speech but serves as an indicator for the listener to pay attention to a following sequence of utterances. It was also found that prosody can help assign a certain set of prosodic features to a particular person in the conversation. This means that after some time, based on these features, listeners can identify who the primary speaker is referring to when quoting.

As Bolden (2004) highlighted, the onset of the quote can be marked by quotatives. Quotatives are usually grammaticalized prepositions that have acquired a quotative or a discourse function over time (Kolyaseva, 2018). A common example of such quotatives is English *like* or Russian *tipa* and *takoj* quotatives. *Tipa* emerged from the word *kind* in Russian (Lapteva, 1983, as quoted in Lazareva, 2020). Lazareva (2020) points out that both *tipa* and *takoj* are used to introduce a citation, and sometimes they can be both used in the same sentence. Such quotatives help take responsibility off the speaker because they signal that the quotation is only an approximation of what was said. In Multimedia Corpus of Spoken Kazakh Language (MultiCorSKL), I observed young women use Russian *tipa* and *takoj* when quoting others. They are not frequent in the data and usually appear with verbs of saying like *aitty* or *dedi* ('He/she said'). *Tipa* can sometimes be employed alone without these verbs before the quoted sentence, but currently it still seems to be in the process of grammaticalization in Kazakh (Mukanova & Troiani, in prep).

Dubois (1989) investigated quotes with a certain structure "quote formula + hey + pseudoquotation." Pseudoquotation is also a form of direct quote, but it is constructed by the primary speaker. One can identify a pseudoquote by looking at the context. If the matrix clause (the clause surrounding the quote) includes words of uncertainty, like *seems to have*, or if a speaker uses constructions like *his eyes were saying*, then the speaker is using a pseudoquotation. Dubois (1989) provides an example of such pseudoquote as in (6):

(6) Those [undocumented workers who have applied for amnesty] are the ones we're trying to get out and saying, "Hey, it's possible for you to lift up your head and say, 'I was from there, but . . .'." (p. 349)

In the example above, the speaker says that they want workers say certain things, but these things are not actually what they have said yet. Hence, this is a pseudoquotation.

Overall, this subsection summarized different types of DRS marking, including SPEAKER + REPORTING VERB construction, the verb with an accompanying action, prosody, quotatives, and pseudoquotations. Differentiating between these types of DRS marking is important for identifying cases of direct reported speech in the data.

2.2. Code-switching

Code-switching, also called code-mixing, refers to the spontaneous use of one language inside another language (Gullberg et al., 2009; Zharkynbekova & Chernyavskaya, 2022). Contrary to previous assumptions that speakers who code-switch are incompetent language learners, it has been established that code-switching is consciously used as a communicative tool to reach an understanding between speakers and aid their interaction (Aliaskar, 2024; Auer, 1998; Zharkynbekova & Chernyavskaya, 2022). Code-switching is automatic and happens when a speaker is fluent in both languages and finds no better translation of a word in one of the languages they speak. For example, one can say "*specialisty v digital humanities*" ('specialists in digital humanities') because there is no better alternative for the English word in Russian (Zharkynbekova & Chernyavskaya, 2022, p. 430).

An important difference to account for is that of code-switching and borrowing (Zharkynbekova & Chernyavskaya, 2022). Borrowings are usually integrated into the language morphologically or phonologically, while code-switching is not. There seems to be no clear boundary between borrowings and code-switching in Kazakh compared to languages like Italian or English. This feature of Kazakh may complicate the analysis of corpus data, but it also highlights the importance of studying language phenomena from multilingual perspectives and revising bilingualism theories. Variability in use and flexibility in behavior "fundamentally govern the language systems and should not be seen as its secondary characteristics," meaning that linguistic

phenomena should be studied from a multilingual perspective (Franceschini, 1998, p. 52). The current project aims to contribute to the field by studying multilingual data and code-switching in Kazakhstani context. Another problem with studying code-switching is that it is difficult to tell which speech element belongs to what language, especially if the languages are related. Yet, this is not a problem for the multilingualism situation in Kazakhstan because the three languages are not genetically related.

2.2.1. Triggers of Code-switching

There are various approaches to studying code-switching. The primary one for this research is the sociolinguistic approach, where one focuses on the interactional aims of speakers. This is the approach the current study is taking. Another approach is the grammatical approach, which focuses on grammatical structures of code-switching possible in the data (Auer 1998). Here, the language system is of primary concern. The third approach is psycholinguistically-motivated code-switching (non-functional code-switching), which unlike the first type, does not consider code-switching as intentional. For the researchers of non-functional code-switching, code-switching just "happens" in the conversation and is triggered by trigger-words (Riehl, 2005, p. 1946). Although I am primarily interested in interactional and communicative functions of code-switching in DRS, I plan to look at triggers of code-switching to separate psychologically and sociolinguistically-motivated code-switching and understand what triggers, if anything, code-switching in Kazakh.

Among the scholars who studied triggers of code-switching, Broersma et al. (2009) study is one of the most relevant ones to our context. Drawing from corpora with naturally occurring conversations of Dutch-English and Russian-English bilinguals, Broersma et al. (2009) explore how the shift to one language influences the language choice in next utterances on the lexical level. These words are called triggers, and they are mostly nouns. The code-switching turned out to be

more likely to happen after the trigger, while clauses with trigger words were more likely to have the code-switching inside. In addition, overlap in the form and meaning of words, cognate words (words with a shared origin in both languages), and discourse connectors trigger code-switching, while, in Russian-English data, false friends (e.g. Russian *akkuratnyj* ('neat') vs. English *accurate*) are more likely to trigger it. One person's use of trigger words does not affect the code-switching of another person, meaning that the trigger word has to be produced by the speaker themselves.

If a text in a certain language is used as a collocation, which refers to a combination of two or more words that frequently appear together, then it will impact the choice of the language in code-switching (Frick & Riionheimo, 2013). For example, if a song is in a certain language, then, when quoting a phrase from that song, one may be tempted to keep it in the original form. Guerini (2009) also points out that speakers code-switch when they want to be as close as possible to the original speaker's utterance. This function will be called "preserving the original quote" function for the purposes of this study. In section 4, I will elaborate on this and other functions mentioned in this section and provide examples from the corpus.

2.2.2. Functions of Code-switching in DRS

In addition to "preserving the original quote", code-switching can also create "interactional meaning," establishing the context for engaging in various activities, such as assessing the voice persona or setting a context for the speech event (Frick & Riionheimo, 2013, p. 566; Riehl, 2005; Zharkynbekova & Chernyavskaya, 2022). Such interactional functions include "invoking a voice persona" and stance-taking.

One interactional function of code-switching is "invoking voice persona" (Frick & Riionheimo, 2013, p. 578). This means that speakers using code-switching are distancing themselves from what the original speaker said. DRS is often used in third-party complaints in

order to “save their face” and avoid responsibility for the reliability of their quote (Troiani, 2023). Thus, this interactional function of code-switching in DRS identified by Frick & Riionheimo (2013) is further strengthened by the fact that DRS is encountered in a context in which speakers are trying to save face.

In addition, code-switching in DRS is used for stance-taking. Stance-taking refers to “ways that people position themselves in conversation” (Kiesling, 2022, p. 410). Speakers can use various constructions and linguistic strategies, including code-switching, to indicate their emotional position or position in terms of politeness and certainty. Kiesling (2022) notes two types of stance: epistemic and affective (interactional). The epistemic stance indicates the state of one’s knowledge about the situation. A primary example of this is the use of evidentiality. The affective stance is about feelings and emotions about something or someone. Similarly, Guerini (2009) points out that language choice reveals attitudes toward language variety. When code-switching, speakers show their affiliation or disaffiliation to the quoted group and other emotions pertinent to the object of the conversation.

Code-switching also helps relieve the cognitive load on a speaker and aids the flow of the conversation. Evidence from writing in a foreign language suggests that, when writing in their L2, students’ cognitive abilities are overloaded, as they are translating their thoughts from their native language to another language, which slows down communication processes (Nawal, 2018). Writing, unlike speaking, allows for more time to process the information and revise the written material. Hence, it provides less burden on the writer compared to speaking, where a person has to retrieve the word from memory, process it, and articulate it fairly fast (Liu, 2020). In these cases, such cognitive load can prompt code-switching (Liu et al., 2023) in order to lessen the burden by

retrieving words that are much easier to access. I am also considering such function of code-switching, which I called “ease of access” in my data.

As was established, code-switching plays an additional interactional role in DRS, which is important to consider in the Kazakh language. Given the multilingual context Kazakh speakers live in, it is important to see what role code-switching has in Kazakh direct reported speech. Muhamedowa (2016), who wrote an extensive work on Kazakh grammar, devotes only a small section about Kazakh DRS, even though it is a frequent phenomenon in naturally occurring conversations. By studying the data from the spoken corpus of MultiCorSKL (Filchenko et al., 2023), I am planning to expand the existing literature on Kazakh direct reported speech and code-switching. I base my categorization of instances in the corpus on the literature above. I explained each function used in this research in detail in the methods section.

3. Methods

My study features qualitative research that combines conversation analysis and discourse analysis to get a better understanding of the context of conversations in the corpus data and interactions within them. I will use a conversation analysis (CA) method because I use naturally occurring data and focus on participant interactions and how code-switching helps them achieve their interactional goals. One of the advantages of CA is that it divides conversation into orderly sequences of talk (Baxter, 2018). This way, it will not be just a text but an ordered sequence of utterances that differentiates one’s words from those of others. Given that I am dealing with Kazakh-Russian-English code-switching, this orderliness is particularly useful when reporting the results. Since the use of several languages may create chaos during transcription, making it harder to differentiate the boundaries of speech, dividing a conversation into different turns will allow me to solve this problem. Most importantly, CA allows us to account for interactional functions of

grammar, which is the primary purpose of my study. In addition to CA, I utilize discourse analysis (DA). DA combines both micro- and macroanalytic approaches to language. This means I can look at particular constructions in the talk and their use and functions in a given context and discourse (Baxter, 2018). This is useful for answering my research questions that look at both triggers of code-switching (microanalytic approach) and the function of code-switching in the interaction (macroanalytic approach). Although there is a trend for choosing experimental methods to study code-switching, they cannot entirely substitute natural language data from a corpus (Gullberg et al., 2009), so that is why I will be working with the corpus.

Data Collection

The data comes from the MultiCorSKL (Filchenko et al., 2023), a corpus compiled at Nazarbayev University that consists of around 218 hours of naturally occurring conversations in modern Kazakh in various settings and with people from different demographic groups. I selected a total of three hours (180 minutes) of conversation and noted 47 instances of code-switching in DRS. In total, I selected five different recordings for 20 minutes to diversify the data, except for one recording, which I analyzed entirely (80 minutes). The reason for that was that the recording contained a lot of code-switching and target trigger words, such as *tipa* and *takoj*.

The metadata for each recording can be found in Table 1.

Recording	**Length**	**Speakers**	**Degree of intimacy between speakers**
tate160522	80 minutes	3 women	Friends
akak230921	20 minutes	3 women and 1 man	Family (primarily: mother and daughter)

zhtu190522	20 minutes	2 women and 1 man	Family
dakh070722	20 minutes	2 women	Mother and daughter
bise310722	20 minutes	3 men and 1 women	Family
tonu190622	20 minutes	Mostly women	Relatives (a big gathering)

Table 1. Metadata about recordings

It can be seen that all conversations are either between friends or family members. Regarding tonu190622 and dakh070722 recordings, the full metadata was not available, as participants could refuse to fill out parts or the entirety of the form about themselves. Therefore, some information about these two recordings was deduced from the conversations and notes from the recorders.

Data Coding

I used the Excel sheet (see Appendix) for storing and coding the data. Each column contained the following information:

		Sentence		Language				Trigger		
Name of the file	Timestamp	Quote	Matrix clause	Quote	Matrix clause	DRS type	Function of CS	Yes/No	Initiator	Notes
tate160522	00:15:59.00	можно просто екі төрт жүз курс алуга и все.	они сказали	Russian/Kazakh	Russian	Speaker+reportin	Accessibility (lang	Yes	Primary s…	
tate160522	0:23:08	семьям өзі сондай высокий,	типа,	Kazakh/Russian	Russian	tipa	Accessibility (lang	Yes	Primary s…	trigger of one code-switching - adjective
tate160522	0:35:53	ол, ол типа, старый адамдардың ауыруы	деп,	Russian	Kazakh	Reporting verb de	Accessibility (lang	No		
tate160522	00:40:01.00	(TSK) менің жігітім бар, ол маған предложение жасайды, мартта	А, ну, мен, сыныптастарым, он бірінші класста жүргенде айтатын ... олар шықпай қалды в итоге.	Kazakh	Kazakh	Reporting verb	Accessibility (lang	No		

Image 2. Data coding conventions

As indicated in Image 2, after I selected the data with code-switching, I put the timestamp of the quote in the table. I divided the quote into two parts: the original quote and the quote-framing device, which is indicated as a matrix clause. I also indicated the language or languages used inside

the quote and in its matrix clause. I have done this to be able to identify when and how code-switching may be triggered in direct reported speech. Whether the trigger had happened and who triggered it was also marked in separate columns. I assigned types of direct reported speech and the function of code-switching in that dialogue. Regarding types of direct reported speech, some possible variations I identified in the literature were the SPEAKER + REPORTING VERB format, quoting using a verb of accompanying an action, prosody, pseudoquotations, and quotatives. The last column is dedicated to notes that I added while analyzing and collecting the data.

Data Transcription

While I look at already completed transcriptions by corpus creators, for a fuller understanding of the context, I listened to the original recordings and transcribed necessary parts of those conversations using Discourse-Functional Transcription (DFT) conventions. I also segmented the data into intonation units, each of which was placed on a separate line. The symbols for various features of language and conventions used to differentiate between languages are shown in Table 2.

Symbol	**Meaning**
Tate160522 (0:23:08)	Filename (timestamp)
@@@	Laughter
[ ]	Simultaneous speech of two speakers
<RU>...<RU>, <EN>...<EN>	Indication of code-switching to Russian or English. Only phrases in Russian or English will be marked with <RU> or <EN>,

	respectively. Everything outside these signs is in Kazakh.
( )	For describing the behaviour/emotions

Table 2. Transcription conventions based on DFT used in the current study

Data Analysis

For data analysis, I looked at the code-switching inside the quote or before the quote in transcriptions and recordings to understand the context of the conversation and assign a preliminary function to a conversation. I created a list of functions based on the literature and my analysis of conversations, as shown in Table 3. In the next section, I will discuss and provide more examples for each function.

Function	Definition	Example
Stance-taking	Refers to any instance where code-switching conveys an emotional message and/or stance.	I **aitady vot tipa** "Senderdin' gruppalaryn' <RU> **takoj, ubityj**"<RU> dep aitady da. And says like "your group is so dead," she says.
Preserving the original quote (Guerini, 2009; Zharkynbekova & Chernyavskaya, 2022)	Refers to situations where the primary speaker switches to another language because they want to preserve what the reported speaker said in the original form, establishing credibility.	<RU> Potom <RU> aitady <RU> **"No my ej nichego ne mozhem delat', potomu cho ona beremennaja**"<RU> deidi de. Then she says "We cannot do anything to her because she is pregnant."
Invoking voice persona (Frick & Riionheimo, 2013; Troiani, 2023)	Refers to instances where the speaker does not want to take accountability for their words, indicating that the	Sony <RU>vot vykladyvaet.<RU> **Birden'e** zhazady sosyn <RU>vot **"S novym godom"**<RU> dep**.**

	information is only the approximation of what was said.	She uploads that. Then writes **something** saying **"Happy New Year"**
Ease of access (Liu et al., 2023; Nawal, 2017; Riehl 2005)	Refers to instances when a person switches to another language to ease the mental load and choose words that are easily accessible.	Men aitamyn: "Men eshkimdi **netpeimin-** <RU> **ne zastavljaju nikogo**<RU>.

Table 3. Functions of CS in DRS

4. Results

This section discusses triggers of code-switching and functions of code-switching and answers my initial two research questions.

4.1. Triggers of Code-Switching

The first research question was "What triggers code-switching in Kazakh direct reported speech?". My hypothesis regarding triggers of code-switching was that the use of Russian-borrowed quotatives, *tipa* and *takoj*, would trigger the code-switching. Out of 14 cases of use of some forms of *tipa* and *takoj* in DRS, only 4 instances triggered code-switching to Russian. All trigger words were uttered by the primary speaker. Consider the example (1) mentioned in the introduction:

(1) tate160522 (0:34:54)

1	KAMSHAT;	*[Men aitamyn],*	I say,
2	BIBINUR;	*["Bruh."]*	"Bruh."
3	KAMSHAT;	*<RU>* ***tipa****<RU>,*	like,

4	*<RU>"Nu, my ne khotim ee uvol'njat',*	"We cannot fire her,
5	*ei vsego lish' tam<RU>,*	She just has,
6	*bir apta gana qaldy."*	a week left"
7	*dep aitady da,*	she says,
8	*<EN>* ***And I was like,*** *<EN>*	
9	*<EN>"Fuck you."<EN>*	

In (1), Kamshat quotes a worker at the reception of a clinic she interacted with. After the quotative *tipa*, she switches to Russian but finishes the quote in Kazakh, which seems to influence the ending of the quote *dep aitady da*. Then she expresses her emotions in English in IU 9 while using the English quotative *like* to introduce this quote. With this, she prepares her listeners that the next thing she will say is in English. Thus, in this case, not only Russian-borrowed quotatives but also quotatives in any language might have influenced the choice of the language for her next utterance. One may assume that such use of a quotative in a particular language prompts the speaker to continue talking in that language. However, apart from this example and three similar instances, *tipa* and *takoj* quotatives did not seem to trigger the use of Russian inside the quote, while English quotatives were no longer used in the conversation to test the hypothesis further. This aligns with the literature that the choice to code-switch is not easily influenced by trigger words and mostly depends on the speaker's conscious choices (Riehl, 2005). While there were instances where the speaker used Russian adjectives, such as *vysokij* ('tall') inside the quote following *tipa*, these instances were probably caused by the fact that the person forgot an equivalent word in Kazakh.

The code-switching was sometimes triggered because of the use of common words in both languages, such as *waltz* (Riehl, 2005; Frick & Riionheimo, 2013). In (7), Bek did not use Russian in their turn, yet Aisha started her turn with the Russian word *vypusknoj*. One possibility is that the word *val's*, which is the same in both languages, was interpreted as the Russian one by the primary speaker. That is why they switched to Russian inside the quote.

(7) zhtu19052 (0:05:39)

1	BEK;	*Ne istep otyr olar?*	What are they doing?
2		***Val's*** *pa?*	Waltz?
3	AISHA;	*Vypusknoj.*	Prom.
4		*A vypusknoj emes,*	Oh, not prom,
5		*"Poslednij zvonok,*	"The last bell,
6		*Podgotovka k linejke,*	the preparation for the school event,
7		*Tam uzhe shkol'nyj val's."*	and there is already school waltz",
8		*Birden'e dedi.*	he said something.

Apart from this case illustrated by (7), where another speaker triggered the code-switching of the primary speaker, mostly the primary speaker's use of trigger words initiated the code-switching in my data. This confirms Broersma et al.'s (2009) statement; the fact that some other speakers used the trigger word does not influence the primary speaker's code-switching.

In summary, it was found that *tipa* and t*akoj* rarely trigger code-switching from Kazakh to Russian. However, sometimes using frequent Russian words or words used in both languages can influence the switch to Russian. This highlights the importance of considering the interactional purposes of code-switching to another language because even if trigger words can help switch to another language, code-switching still primarily depends on the choice of the speaker.

4.2. Functions of Code-Switching in Kazakh DRS

Regarding the functions of code-switching in DRS in Kazakh, I found four different functions of code-switching in DRS in the corpus data. Figure 1 below summarizes the distribution of functions of code-switching in the recordings. I identified functions like invoking voice persona (14), accessibility (10), stance-taking (11), and preserving the original quote (12). I will discuss each case in detail in the following subsections.

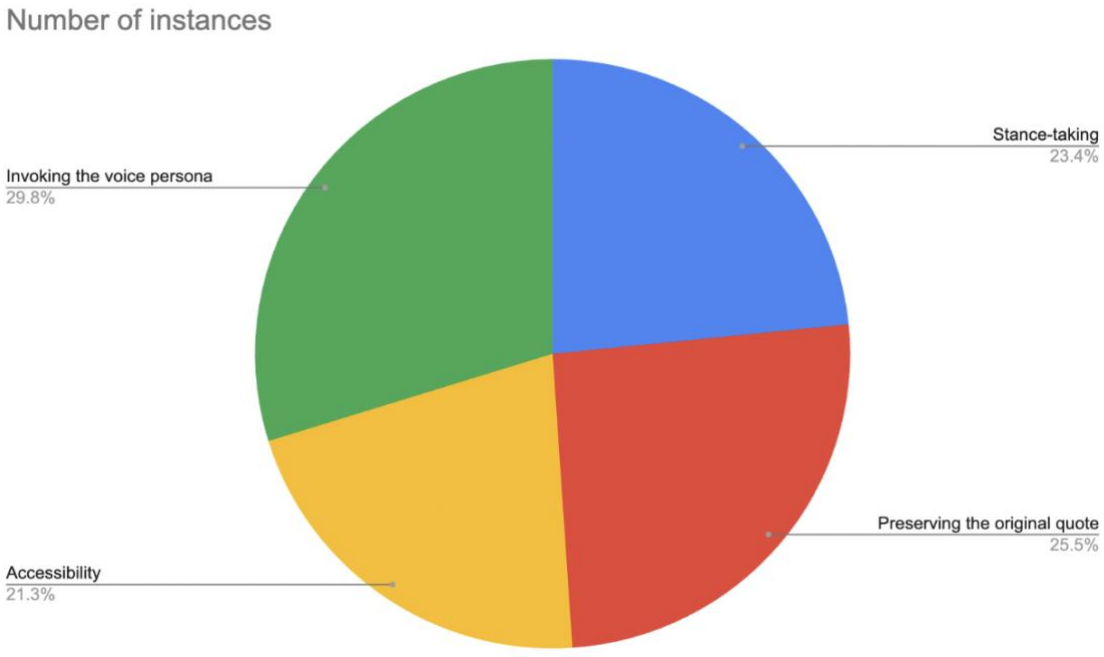

Figure 1. Distribution of code-switching functions

All of these functions were distributed among the instances quite fairly, with the "invoking the voice persona" function being the most common among the four. This function refers to situations where the speaker does not want to take full responsibility for their words, conveying the message for the listener that this information is only an approximation of the reported speech. This way, a speaker acknowledges that they do not recall everything from the reported speech event, which increases their credibility. Hence, it aligns with my initial hypothesis that code-switching will be mostly used to establish credibility for the quoted information and foster mutual understanding and trust between speakers. "Preserving the original quote" was the second common function of code-switching, accounting for one-fourth of all instances.

Invoking voice persona

The "invoking voice persona" function of code-switching refers to various situations, where a person is distancing oneself from what's being said in the quote. They use it to take responsibility away from themselves, as they cannot fully recall the conversation, so they are letting another person know about this through the use of code-switching. It can also refer to situations where the quoted dialogue did not actually happen, so they are approximating one's feelings or thoughts.

In this part of the conversation between three friends, students are discussing a Chinese class they took from a professor and their opinions about it. Then, they talk about how the professor invited her students to a party to celebrate Chinese New Year. After this, Bibinur starts talking about her Chinese professor's Instagram account and how she posts stories there:

(8) tate160522 (0:22:20)

1	BIBINUR;	*O'zinin' <RU> kvartirasynda,*	In her apartment,
2		*cho-to vyrashivaet,*	grows something,
3		*kazhetsya.<RU>*	it seems.
4		*Sony <RU> vot vykladyvaet, <RU>*	So she posts it,
5		***Birden'e** zhazady sosyn,*	And writes something,
6		*<RU> vot "S novym godom"<RU> dep,*	Saying "Happy New Year,"
7	KAMSHAT;	*He*	(Sounds of admiration)
8	AIGUL;	*Hmhm*	(Sounds of admiration)
9	BIBINUR;	***Birden'e-birden'e.***	something like this.

Bibinur says that her professor grows some plants and posts them on her Instagram, writing a caption "Happy New Year," which other speakers find cute. The quote here is introduced using a word that describes a behavior *zhazady* ('She writes') and ends with the verb *de* ('say'). Although Bibinur used Kazakh to introduce the quote, the quote itself is in Russian. Given the context, the original sentence was presumably in Chinese, but, knowing that her listeners may not know it, Bibinur switched to Russian to indicate that these words are only an approximation and/or a translation of the original. She also lets the listeners know that this quote is only an approximation

by using words of uncertainty before and after the quote, such as *birden'e-birden'e* ('something-something'). Thus, code-switching to Russian might convey clues to listeners that the quoted words are only the translation of the original and indicate the speaker's consideration of her listeners' language skills. This way, Bibinur avoids additional questions about the translation of the phrase, which allows her to continue her talk.

In this part of the conversation, Baqyt talks about her arguments with Raqima and Maqsat about which of the two cities, Oskemen and Astana, is better:

(9) tonu190622 (0:09:53)

1	BAQYT;	*<RU> Ja govorju <RU> "Oskemennen zhaqsy zhoq.*	I say, "There is nothing better than Oskemen."
2		*Ketemin Oskemenge."*	"I'll go to Oskemen."
3		*Ui boldy,*	There was a house,
4		*Ony osylai satylyp ketip keldim osy zhaqqa.*	we sold it and came here.
5		*"Zhoq, Maqsat,*	"No, Maqsat,
6		*Oskemen zhaqsy, myna Astanany qoi.*	Oskemen is good, but forget about this Astana.
7		*Myna appaq shan'y,*	This white-white dust,
8		*Myna nemene mine, osy, mine, Astana, Oskemende bolsa,*	All of this, whatever this is, here in Astana, but in Oskemen,

9	*Oz elim, oz zherim.*	My people, my land.
10	*Ba'ri meni tanyp."*	Everyone knows me."
11	*Raqima <RU>* ***chto-to*** *<RU>: "Nemene sen Oskemen?*	Rakhima said **something:** "Why are you always about Oskemen?
12	*<RU>zarabotal<RU> Oskemende <RU>svoe legkije<RU>" dep shyg'a keldi.*	you earned your lungs (cancer) in Oskemen", she said.

Baqyt switched from Kazakh to Russian before quoting Raqima to indicate that the following information will only be an approximation of her actual words. The use of words of uncertainty, such as *chto-chto* ('something') in IU 11, enhances this message to her listeners (Frick & Riionheimo, 2013).

In (10), Kamshat talks about an employee who caused stress for her and how then a psychologist came to talk with her after that.

(10) tate160522 (0:34:42)

1	KAMSHAT;	*<RU> Potom,<RU>*	Then,
2		*Magan-,*	To me-,
3		*menimen,*	With me,
4		*Psiholog soilesedi de,*	A psychologist talks,
5		***Aitady tipa,***	and says like,

6		*<RU> "Vot, tak ne poidet." <RU>*	"No, it will not work."
7		*Dep* ***"Andai-myndai",***	Says "something-something",
8		*Dep aitady da.*	She says.

Here, IU 6 refers to the original speech of the reported speaker, given the dynamics at our university, where the administration usually talks in Russian. Kamshat utilizes this to establish the credibility of her words. Then she adds *andai-myndai* ('something-something') in Kazakh, which seems to exhibit an approximative function, indicating that the primary speaker is not willing to provide further details on this, as it is not important. Other essential information that follows after this appears to be mostly in Russian, while Kazakh seems to indicate the voice of the primary speaker.

Preserving the original quote

This function refers to situations where the primary speaker switches to another language because they want to preserve the authenticity of the quote of the reported speaker and increase the credibility of what they are saying. It also helps distinguish more clearly between the primary speech event and the reported speech event.

The example of this function can be seen in (11). This conversation happens while parents are looking through a website to shop online to buy some clothes.

(11) zhtu190522 (0:12:12)

1	BEK;	*Myna,*	That,
2		*Menin' keshe,*	Yesterday I was,

3		*Rieltordyn' nesin qarap otyrmyn,*	watching that things of that estate agents,
4		*Storisterin,*	stories,
5		*Kilen' zhan'a,*	All that,
6		*Ne aldynda…*	in front of that…
7	AISHA;	*Ka'dimgi kvartira ma?*	The real apartments?
8		*Zhoq, zher ui ma?*	Or house?
9	BEK;	*Kvartira.*	An apartment.
10		*Kilen' zhan'agy:*	All that:
11		*"On million,"*	"ten million",
12		*<RU>"**Tsenasy snizhena,**"<RU>*	"lower price,"
13		*"Tog'yz zharym" dep shyg'ady sondai.*	"Nine and half, pops up like this."
14	AISHA;	*Iii… qursyn.*	Iii… whatever.

Here, Bek talks about the declining price for housing and quotes a real estate agent's stories about the lower price for an apartment. These words, like in IU 12, are also the ones one can frequently encounter in the real estate market. So, Bek code-switches to Russian to preserve the original wording of the agent's stories to establish the credibility of their source and indicate that they, indeed, read this.

The same function is also present in the following recording (12). This conversation happens in the setting of a family gathering, where one woman pours qymys for everybody and tells various stories related to this traditional drink. She actively engages in a conversation with other speakers who come to her for a drink.

(12) tonu190622 (0:01:34)

1	AZHAR;	*"Biz sapyru kerek" deidi.*	"We have to sweep," they say.
2		*Sebebi bunda <RU> kislota <RU> bar goi.*	Because there is acid inside.
3		*<RU> Kislota. <RU>*	Acid.
4		***Orystar aitady g'oi,*** *<RU>* ***"spirt est'"*** *<RU> dep.*	As the Russians say, "There is alcohol."
5		*Sonda sapyrg'anda mynau <RU> kislotasy <RU> ketedi.*	So when you sweep, that acid goes away.
6	BAQYT;	*<RU> Slyshish', Tansholpan? <RU>*	Do you hear me, Tansholpan?
7	AZHAR;	*Zhumsarady.*	It becomes soft.

In this excerpt, Azhar talks to a younger relative and another older woman, explaining why they sweep qymys. In IU 4, she quotes a saying of Russians about alcoholic beverages. By switching to Russian, Azhar is preserving the original quote to establish credibility of her knowledge and the originality of the quote.

Stance-taking

This function of code-switching is used to refer to any instance where code-switching is trying to deliver an emotional message and/or stance. In the following recording (13), three university students talk about their classes at the university, frequently code-switching between Kazakh, Russian, and English.

(13) tate160522 (0:11:06)

1	KAMSHAT;	*<RU> I<RU> aitady <RU>* ***vot*** ***tipa*** *<RU>,*	And then [the professor] says,
2		*"Senderdin' gruppalaryn' <RU>* ***takoj,***	"Your group is so,
3		***ubityj****"<RU> [dep aitady] da.*	dead."
4	BIBINUR;	*[@@@@]*	@@@

In (13), the switch to Russian inside and before the quote carries an emotional meaning to make other speakers laugh and create a certain image about the professor. It is highly unlikely that the person talking in the original quote, who is also probably a foreign professor, will use such expressions similar to *your group is so dead*. The code-switching to Russian in this case seems to communicate to the listener that the quote is not an accurate representation of what was said but just an approximation. The quotative *tipa* in IU 1 before the sentence also seems to prepare the listener for this and expresses a certain stance towards the professor that is amplified and supported by the use of *takoj ubityj* ('so dead') expression. The speaker tries to make other speakers accept a certain attitude towards the professor.

The conversation (14) is similar to the example above, as they both evoke and convey a positive stance towards the situation, which may have been misinterpreted without the switch to another language.

(14) tonu190622 (0:01:20)

1	AZHAR;	*Sayagul keshe berip zhiberipti alty litr.*	Apparently, Sayagul yesterday send six litres.
2		*Sosyn to'rt litrin alyp keldi.*	And from that she brought four litres.
3		***"Ba'rin <RU> zabiraite" <RU>,*** *<RU>ja govorju.<RU>*	"Take everything", I say.
4	BAQYT;	@	@

In IU 3, Azhar code-switches to Russian mid-sentence when she quotes herself. She might have chosen Russian to indicate that this is a joke, and she was actually pleased with the gift from Sayagul. This also means she did not mean to say *take everything away*. That is why Baqyt laughed at it, as she understood it as an exaggeration, which was conveyed through the use of code-switching and prosody.

Code-switching can also be used to convey negative emotions. In this part of the dialogue in between a mother and her daughter, they are discussing the gathering of money for a classmate's funeral.

(15) dakh070722 (0:06:36)

1	ZINAIDA;	*Boka da salmaji qojiupty g'oi,*	Boka didn't put the money too,
2		*"Ketsinshi a'ri" dep.*	saying, "screw them all".
3		*"Aqsham zhoq" dep aitypty.*	Apparently, he said "I do not have money."
4	ZHANAR;	*Sony aitam.*	That's what I am saying.
5	ZINAIDA;	*Myna uly ija,*	This son, yes,
6		*<RU>* ***"prjam sprashivaet"*** *<RU>,*	he asks,
7		*<RU>* ***"Zastavljaet"*** *<RU> deidi.*	Obliges, he says.
8	ZHANAR;	*Ija.*	Yes.

By switching to Russian in IUs 6-7, Zinaida tries to convey a negative emotion and attitude towards the situation that the reported person felt. The reported speaker was not pleased with that classmate's son's behavior. Hence, instead of retelling the whole situation, Zinaida chose to use Russian words that would describe this in a more concise and informative manner to the interlocutor.

Ease of access

The following excerpt (16) contains an example of the ease of access function, which refers to instances where a person forgets a word in one language, so they switch to another. This seems to happen due to the dominance of Russian in Kazakhstan, which makes some Russian words more easily available in the mental lexicon compared to Kazakh words. This relieves the

mental load on a speaker and aids the flow of the conversation. So, instead of looking for the right word in a current language and slowing down the response time, one could say it in the language that is more convenient for them (Liu et al., 2023). As they also know that other speakers will understand them, given the shared knowledge and belonging to the same community, multilingual speakers are prone to this type of code-switching even outside of DRS.

In (16), Bibinur continues talking about her Chinese professor. She describes an instance when she went to her office to do a make-up quiz. This is what she tells other speakers about her interaction with the Chinese professor:

(16) tate160522 (0:22:59)

1	BIBINUR;	*<RU>I potom<RU> magan qarap,*	And then she looks at me,
2		*<RU> "Mozhno ya sproshu vopros"<RU> deidi.*	"Can I ask a question?" she says.
3	KAMSHAT;	*Mhm.*	Mhm
4	BIBINUR;	*<RU> Takaja "Ok."<RU>*	I was like "Ok".
5		*"Sen tipa,"*	"You like,"
6		*(CLEARS THROAT)*	
7		*<RU>"Pochemu ty takaja vysokaja?"<RU> deidi*	"Why are you so tall?" she says.
8	AIGUL;	*[@@@@]*	

9	KAMSHAT;	*[@@@@]*	
10	BIBINUR;	*[@@]*	
11		*<RU>Ja takaja<RU>*	I was like
12		*"Men bilmeimin men"*	"I do not know I-"
13		*"<RU>**Sem'jam**<RU> ozi sondaj <RU> **vysokij**<RU>."*	"My family members are tall."

In IU 13, the speaker uses the words *sem'jam* ('family') and *vysokij* ('tall') in Russian, despite mostly talking in Kazakh inside the quote. I presume these were the words that she used more frequently and, thus, it was more accessible to use. In addition, Bibinur probably knew that other speakers would understand these words without translation.

In (17), we can notice the same ease of access function. Mother (Zeinep) asks about her daughter's (Aruzhan) boyfriend while they are working on completing reports for Zeinep's work.

(17) akak230921 (0:08:36)

1	ZEINEP;	*Taksovat' ete me?*	Does he work as a cab driver?
2	ARUZHAN;	*Zhoq.*	No.
3	ZEINEP;	*Endi?*	If no, then what?
4	ARUZHAN;	*Keshke dejiin sabaq onda.*	He has classes all day.
5	ZEINEP;	*A?*	What?
6	ARUZHAN;	*Sabaq deimin.*	I say classes.

7	ZEINEP;	*Aaa.*	
8	ARUZHAN;	*“Zhiyrmasy” deidi.*	It says, “The twentieth”.
9	ZEINEP;	*Qiyn dei me?*	Does he say it is difficult?
10		*A?*	Hey?
11	ARUZHAN;	*Olarda, oibai,*	Oh my, also,
12		*Olardyn’ bir semestrlaryn,*	Their one semester,
13		*Ekige bolip tastag’an g’oi <RU>eshe<RU>.*	is divided into two.
14		*Olarda uzhe,*	They already are,
15		*Rubezhka bolyp zhatyr tipa bizde,*	having their midterm week like the one we have,
16		*“Zheti aptadan kejiin,*	“After the seventh week of study,
17		*<RU>**reitingovaja nedelja**<RU> sijaqty birna’rse bolady g’oi,*	like that rating week thing that we have.
18		*Sondai bir na’rse” deidi-deidi-deidi.*	Something like this”, he says-says-says.

Since it is hard for Aruzhan to look for alternatives for *reitingovaja nedelja* (‘the rating week’) in Kazakh, she code-switches to Russian inside the quote. She uses the word she frequently encounters at the university and is something her mother has heard before and will understand. I

considered IUs 16-18 as part of one quote, as this is also probably how her boyfriend explained their system to Aruzhan as well, so these words were part of the reported speech event.

An example of this code-switching function in DRS can also be observed in the following conversation in (18) where they were discussing their relationship with their daughter-in-law. Zinaida talks about how she stopped expecting help from her daughter-in-law, and when she asked her why she leaves quietly in the morning, she stated that she is ok with not bothering or forcing anyone to help her.

(18) dakh070722 (0:14:09)

1	ZINAIDA;	*Sony ana zholy mag'an aitady:*	She tells me that last time:
2		*"Tan' aitpai erte ketushi edin'iz undemei" dep.*	"You used to leave silently early in the morning", she says.
3		*Men aitamyn:*	I say,
4		*"Men eshkimdi* ***netpeimin-***	"I don't do-
5		*<RU>* ***ne zastavljaju nikogo****<RU>.*	I don't force anyone.
6		*Ozimnin' qolym-ajag'ym bar.*	I have my own arms and legs.
7		*<RU> Mogu spokojino<RU>*	I can freely,
8		*Avtobusta bolsa,*	If on a bus,
9		*Avtobusqa otyramyn."*	I will take a bus."

Although Zinaida starts quoting herself in Kazakh in IU 4, she quickly realizes that she does not know the right word to convey her message in Kazakh, so she finishes her turn in Russian, as illustrated in IU 5.

5. Discussion

Bringing my results together, the first part was about the code-switching before the quote. I established that *tipa* and *takoj* quotatives rarely trigger code-switching. The reason is that these quotatives are in the process of grammaticalization in Kazakh and are currently perceived as Kazakh quotatives (Mukanova & Troiani, in prep). That is why they and other trigger words did not trigger CS. This finding also aligns with the literature that the choice to code-switch is not easily influenced by trigger words and depends on the speaker's conscious choices (Riehl, 2005). This highlights the interactional role of code-switching in the conversation.

As for the code-switching inside the quote and its interactional functions in DRS, all four functions from the literature were identified in Kazakh DRS. Out of four functions, "ease of access" is a cognitive function, while others are interactional. These interactional functions accomplish two main tasks: expressing stance and reducing responsibility. "Preserving the original quote" and "invoking the voice persona" functions both accomplish the same task of reducing responsibility, but they do that in a nuanced way. When using the "preserving the original quote" function, a speaker admits that these words are not theirs but authentic and close to the original. Meanwhile "invoking the voice persona" function highlights that these words are not exactly accurate, so the speaker is distancing themselves from what is being quoted.

6. Conclusion

In conclusion, this capstone project explored how code-switching is triggered and its function in direct reported speech in Kazakh spoken conversation. The findings confirm that code-switching is triggered by the primary speaker. The hypothesis that Russian quotatives trigger code-switching turned out to be not entirely true for my dataset, as they rarely triggered code-switching from Kazakh to Russian. This proves that code-switching is intentional and depends on the speaker's choice to code-switch.

The data from the MultiCorSKL corpus (Filchenko et al., 2023) did not include many instances of code-switching in DRS. In cases where code-switching was used, it was mostly used to reduce the responsibility for the quoted material, establishing credibility and mutual trust between speakers. Overall, I identified four functions of code-switching, which are invoking the voice persona, preserving the original quote, conveying the emotional message (stance-taking) and "ease of access" function.

As for limitations, firstly, I could not analyze more data and full recordings of conversations due to time constraints. The analysis of more corpus data might have resulted in discovering more functions of code-switching in DRS or a slightly different distribution of the current functions. It could also have made my findings more generalizable. Future research could look at more data from the MultiCorSKL corpus (Filchenko et al., 2023) and other sources. Secondly, the social context of the recordings did not vary significantly. Even though the data came from speakers from various age groups and parts of Kazakhstan, it mostly included conversations between family members and only one of them was between friends. Future research could potentially look at DRS in various social contexts, including conversations between acquaintances, friends, classmates, and colleagues, to name a few.

Despite such limitations, the study contributes to the exploration of modern Kazakh language in the multilingual context and how code-switching is normally used and is triggered in such settings. It proves that code-switching is intentional and is not an indication of incompetency in a language. The interactional functions of code-switching signify that code-switching plays a significant sociolinguistic role in promoting mutual understanding between speakers, helping them to establish trust between each other and fostering better communication within the community with a shared linguistic background. It also sheds light on the linguistic situation in the country, such as Russian language dominance in Kazakhstan based on the "ease of access" function of code-switching.

References

Aliaskar, A. S. (2024). Challenges of bilingualism: Code-switching and language dominance in the context of Kazakhstan. *BULLETIN of the Eurasian Humanities Institute, 2,* 7–15. https://doi.org/10.55808/1999-4214.2024-2.01

Auer, P. (Ed.). (1998). *Code-switching in conversation: Language, interaction and identity*. Routledge.

Baxter, J. (2018). Discourse analytic approaches to text and talk. In L. Litosseliti (Ed.) *Research Methods in Linguistics* (2nd ed., pp. 227–257). Bloomsbury Academic.

Bolden, G. (2004). The quote and beyond: Defining boundaries of reported speech in conversational Russian. *Journal of Pragmatics*, *36*(6), 1071–1118. https://doi.org/10.1016/j.pragma.2003.10.015

Broersma, M., Isurin, L., Bultena, S., & Bot, K. D. (2009). Triggered code-switching. In L. Isurin, D. Winford, & K. D. Bot (Eds.), *Multidisciplinary approaches to code switching* (pp. 103-128). John Benjamins Publishing Company.

Buchstaller, I. (2001). He goes and I'm like: The new Quotatives re-visited. https://www.researchgate.net/publication/237237465_He_goes_and_I'm_like_The_new_Quotatives_re-visited

Bureau of National Statistics. (2021). National composition, religion and language proficiency in the Republic of Kazakhstan. https://stat.gov.kz/en/national/2021/

Clift, R. (2006). Indexing stance: Reported speech as an interactional evidential. *Journal of Sociolinguistics, 10*(5), 569–595. https://doi.org/10.1111/j.1467-9841.2006.00296.x

Dubois, B. L. (1989). Pseudoquotation in current English communication: “Hey, she didn’t really say it.” *Language in Society, 18*(3), 343–359. https://doi.org/10.1017/s0047404500013646

Evans, N. (2012). Some problems in the typology of quotation: A canonical approach. *Canonical Morphology and Syntax*, 66–98. https://doi.org/10.1093/acprof:oso/9780199604326.003.0004

Filchenko, A., Troiani, G., Du Bois, J. W., Sarseke, G., Akanov, A., Bizhanova, M., Mikhailov, N., Temirbekova, T., Seitak, B., & Turaliyeva, Z. (2023). “Multimedia Corpus of Spoken Kazakh Language.”

Franceschini, R. (1998). Code-switching and the notion of code in linguistics. In P. Auer (Ed.), *Code-switching in conversation: Language, interaction and identity* (pp. 51-72). Taylor & Francis Group.

Frick, M., & Riionheimo, H. (2013). Bilingual voicing: A study of code-switching in the reported speech of Finnish immigrants in Estonia. *Multilingua, 32*(5), 565–599. https://doi.org/10.1515/multi-2013-0028

Guerini, F. (2009). Polyphony in multilingual interactions: Code-switching in reported speech among Ghanaian immigrants in Bergamo. *Linguistica e Filologia.* http://hdl.handle.net/10446/384

Gullberg, M., Indefrey, P., & Muysken, P. (2009). Research techniques for the study of code-switching. In B. E. Bullock & A. J. Toribio (Eds.), *The Cambridge Handbook of Linguistic Code-switching* (pp. 21–39). Cambridge University Press.

Haberland, H. (1986). Reported speech in Danish. In F. Coulmas (Ed.). *Direct Speech and Indirect Speech* (pp. 219-253). Mouton de Gruyter.

Jäger, C. (2025). Epistemic Authority. In J. Lackey & A. McGlynn (Eds.), *Oxford Handbook of Social Epistemology.* Oxford University Press.

Kiesling, S. F. (2022). Stance and stancetaking. *Annual Review of Linguistics, 8*, 409–426. https://doi.org/10.1146/annurev-linguistics-031120-121256

Klewitz, G., & Couper-Kuhlen, E. (1999). Quote-unquote? The role of prosody in the contextualization of reported speech sequences. *Pragmatics, 9*(4), 459–485. https://doi.org/10.1075/prag.9.4.03kle

Kolyaseva, A. F. (2018). The 'new' Russian quotative tipa: Pragmatic scope and functions. *Journal of Pragmatics*, 128, 82–97. https://doi.org/10.1016/j.pragma.2018.03.004

Lazareva, V. (2020). Типа и Такой в разговорной речи: Функциональные параллели / Tipa and Takoj in colloquial language: functional parallels. *Europa Orientalis.*

Liu, H. (2020). A socio-cognitive approach to code-switching: From the perspective of a dynamic usage-based account of language. *International Journal of Bilingual Education and Bilingualism, 23*(10), 1270–1299. https://doi.org/10.1080/13670050.2018.1441260

Liu, H., Liu, Z., Yuan, M., & Chen, T. (2023). The effect of cognitive load on code-switching. *International Journal of Bilingualism, 28*(3), 513-530. https://doi.org/10.1177/13670069231170142

Mayes, P. (1990). Quotation in spoken English. *Studies in Language, 14*(2), 325-363. https://doi.org/10.1075/sl.14.2.04may

Muhamedowa, R. (2016). *Kazakh: A comprehensive grammar.* Routledge.

Mukanova, K., & Troiani, G. (2024). *Code-switching and Functions of Emerging Quotatives in Kazakh: The Case of "Tipa" and "Takoj"* [Manuscript in preparation]. Department of Language, Linguistics, and Literature, Nazarbayev University.

Nawal, A. F. (2018). Cognitive load theory in the context of second language academic writing. *Higher Education Pedagogies, 3*(1), 385–402. https://doi.org/10.1080/23752696.2018.1513812

Pavlenko, A. (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. https://doi.org/10.1080/13670050802271517

Riehl, C. M. (2005). Code-switching in bilinguals: Impacts of mental processes and language awareness. In J. Cohen (Ed.). *Proceedings of the 4th International Symposium on Bilingualism* (pp. 1945-1959). Cascadilla Press.

Smagulova, J. (2008). Language policies of kazakhization and their influence on language attitudes and use. *International Journal of Bilingual Education and Bilingualism, 11* (3&4), 440–475. http://dx.doi.org/10.1080/13670050802148798

Syzdykbayeva, R. (2016). The role of languages in developing plurilingual identities in Kazakhstan. *NUGSE Research in Education, 1*(1), 15-19. Retrieved from nugserie.nu.edu.kz.

Troiani, G. (2023). 'I played it like this': Voice quality and complaint management in Italian conversation. *Studi Italiani di Linguistica Teorica e Applicata,* 56-72.

Voloshinov, V. N. (1973). Exposition of the problem of reported speech. In L. Matejka & I. R. Titunik (Eds.). *Marxism and the Philosophy of Language* (pp. 115-124). Seminar Press.

Yin, M. (2023). A review of research on direct reported speech. *Journal of Linguistics and Communication Studies, 2*(2), 143–150. https://doi.org/10.56397/JLCS.2023.06.19

Zharkynbekova, S. K., & Chernyavskaya, V. E. (2022). Kazakh-Russian bilingual practice: Code-mixing as a resource in communicative interaction. *RUDN Journal of Language*

Studies, Semiotics and Semantics, 13(2), 468–482. https://doi.org/10.22363/2313-2299-2022-13-2-468-482

Appendix

Collected Data

Dataset: https://docs.google.com/spreadsheets/d/1bkrNS6ksWUvDFbbtyo1YxyYQiyETlMLG-5Od1k15Xbs/edit?usp=sharing