

Code-switching among Muslim Roms in Bulgaria*

HRISTO KYUCHUKOV

Abstract

This article presents the language of Muslim Roms, who form a trilingual community in the northeastern part of Bulgaria, discussing language mixing between Romani, Turkish, and Bulgarian, applying the model of Myers-Scotton (1993). Two different Roma linguistic groups are observed: speakers of Vlax (Zagondzi) and non-Vlax (Xoraxani). The non-Vlax language, which is full of Turkish loanwords, is more prestigious than the Vlax language, which is more conservative, but both show almost the same patterns of language mixing. The study argues for the connection between language mixing and language prestige and looks at the question: Why do Roma speakers use elements from other languages when speaking Romani?

1. Introduction

According to the last census in Bulgaria (1992), the total percentage of Roms is 3.4% (approximately 300,000) of a population of 8.5 million in Bulgaria. However, neither the Bulgarians nor the Roms believe this number; everyone seems to think that there are many more Roms, but they do not know exactly how many. Other sources (for example, Liegeois 1994; Marushiakova and Popov 1993) estimate the total is about 800,000, which seems closer to the truth. Some Roms are Muslim and the others are Orthodox Christians, but the proportion is unknown. During the census, many Muslim Roms identified as Turks and many Christian Roms identified as Bulgarians, this is one reason why the official number of Roms is so low. It is supposed that the total number of the Muslim Roms is approximately 350,000 to 400,000 (Kyuchukov 1997). Until now, there has been no comprehensive study on the culture of Muslim Roms in Bulgaria; there are some observations regarding their language

(Kyuchukov 1995), but, generally, their language and culture have not been described.

The Muslim Roms in Bulgaria are settled in areas with a large Turkish-speaking population; most concentrated in northeastern and southern Bulgaria. According to Gilliat-Smith (1913–1914), the Gypsies in Sofia followed the Muslim religion at the beginning of the century, but during the last seventy to eighty years, apparently many have converted to Christianity. The following groups of Muslim Roms may be distinguished according to their identity: (a) Roms who normally speak only Romani (although they know Turkish and Bulgarian) and who identify as Roms. They call themselves *Xoraxane Roma*, (Turkish/Muslim Roma) and speak Vlach or non-Vlach dialects; (b) Roms who intensively mix Romani with Turkish and Bulgarian in everyday communication, and who identify as Roms (they call themselves *millet* or *čingene* “Gypsy”); (c) Roms who use only Turkish in everyday life (and rarely Bulgarian and Romani, although they may speak some Romani dialect) and who identify themselves as Roms (also called *millet* or *čingene*); (d) Roms who speak only Turkish and cannot speak Romani. They identify as Roms (they call themselves *č-arale* or *usta milletli*, (which is a polite term for “Gypsy”); (e) Roms speaking only Turkish who identify as Turks. All live in separate areas — the so-called “Gypsy quarters” — and all have their own customs and traditions.

These Roms are Muslims, but they are not very strict; they celebrate only the important Muslim holidays, *Ramadan* and *Bayram*; the boys are circumcised; when someone dies, the Muslim mullah is called. However, they drink alcohol and eat pork, although these are forbidden. They celebrate some Christian holidays such as Christmas and Easter.

Today all Vlach and non-Vlach Muslim Roms speak Turkish, which has higher prestige than Bulgarian or Romani. Their Turkish differs from the Modern Standard version, because of surrounding language influence. Speakers of Turkish, whose first language is a Vlach dialect, speak Turkish with a Romani accent; the Romani influence easily recognized; speakers of Turkish whose first language is a non-Vlach dialect speak it more comfortably and without much influence from Romani; however, any native non-Roma can immediately recognize that a speaker is not a native Turk, since a Romani accent is always present, as well as loanwords from Bulgarian.

The Romani dialects used by Muslim Roms are divided into Vlach and non-Vlach. Vlach in Bulgaria are *Kalderaš* (copper maker), *Laxo*, *Kalaydži*, *Zagondži*, *Kalburdži* (‘sieve maker’), and contain more loanwords from Rumanian and fewer from Bulgarian or Turkish. The non-Vlach are

Drandar, Erlij (settled), Xoraxani (Muslim), and most have many Turkish loanwords. Sometimes two or three dialects or subdialects are spoken in one location. Nevertheless, there are no communication problems, Vlax speakers usually know non-Vlax as well and can easily switch. If a non-Vlax person comes to a Vlax family, the family speaks the dialect of the guest out of politeness; however, it is not the other way around, because the Vlax dialect has a lower prestige than the non-Vlax.

Many Muslim Roms converted to Christianity over the last twenty to thirty years. The best examples are Roms in Sofia who do not celebrate *Ramadan* and *Bayram*, although when someone dies, they still invite a mullah. In childhood, they observed Muslim traditions more strictly, but today have ceased to do so and participate in Christian celebrations. The younger generation do not know Muslim traditions, because their elders do not practice them. In areas with a large Turkish population (e.g., in the northeast), Muslim traditions are still alive, but Christianity has a strong influence, so Roms there celebrate Christmas and other Christian holidays as well.

A new phenomenon among Muslim Roms is Pentecostal Christian religion, which has, in a way, a positive influence since those who became followers are strict; the men stop drinking alcohol, stop fighting, and stop stealing. Those who belong to Pentecostal churches have the freedom to use any language they wish to practice their religion. It is normal to use Turkish, Romani or Bulgarian during meetings. In contrast, traditional Muslim customs are only in Arabic or Turkish; it is impossible to use Bulgarian or Romani during meetings.

The reason for the shifts to Christian and Pentecostal religion among Muslim Roms is the negative attitude of the majority towards Muslims and their religion. During the Communist regime, they had no opportunity to practice their religion, and after 1989, Christian and Pentecostal Christian churches became more active.

2. Theoretical background

Borrowing is frequent, and borrowed words from language A are phonologically and morphologically integrated into language B. For instance, among Romani speakers, many words were borrowed from the language of the majority (Bulgarian, Swedish, Russian) where they lived, these words formally adapted and integrated into Romani. There is code-switching if speakers know two languages (A and B) using words or phrases from language A speaking language B, when these words are not adapted phonologically and morphologically. Reasons for code-switching

are diverse: discourse strategy, identification with group, lack of knowledge of the word in the first language, etc.¹

Myers-Scotton (1993) developed a Matrix Language Frame (MLF) model of code-switching, where the speaker who code-switches is dominant in one language, called the Matrix Language (ML), while the other, is termed the Embedded Language (EL). Between ML and EL, a fundamental distinction must be made. Three types of constituents can occur in code-switched sentences: (a) ML + EL constituents; (b) ML islands, which draw grammatically and lexically on the ML only; and (c) EL islands, which draw grammatically and lexically on the EL only. Speakers are more proficient in the ML, making ML morphemes are more frequent than EL in ML + EL constituents, and ML islands are more common than EL islands. Two fundamental principles govern the structure of ML + EL constituents: (a) the morpheme-order principle: the ML determines the morpheme order in ML + EL constituents; and (b) the system-morpheme principle: the ML supplies all productive system morphemes, that is, productive inflections and function words, in ML + EL constituents.

Everyday communication of Muslim Roms presents many patterns of the MLF model, and very few researchers explain why. Borrowing from Turkish by Muslim Roms in Balkan countries (especially Turkey) is described by Heinschink (1989); Friedman (1989) investigated Turkish loanwords in Balkan Romani dialects, showing many Turkish words entering Romani.

3. Empirical study

Data was collected in Northeast Bulgaria in a village, Gradinarovo and a town, Provadia, informants from Gradinarovo speaking *Xoraxani* (non-Vlax), informants from Provadia speaking *Zagondži* (Vlax). Eleven informants were interviewed, resulting in 11 hours of recordings, there were three males and eight females; the speakers of each dialect are presented in Table 1.

The interviews were in the informants' homes in informal situations. They were asked to tell autobiographical stories in Romani, not Turkish,

Table 1. *Informants in the study*

	Informants	<i>Xoraxani</i>	<i>Zagondži</i>
Male	3	2	1
Female	8	5	3

which they speak as well. Only one refused to speak in Romani, and the conversation with her was in Turkish. Informants were between thirty and sixty years old, almost all having a basic level of education (1st to 8th grade in Bulgarian schools).

Turkish on various linguistic levels influences both the Vlax and non-Vlax dialects: (a) On the phonological level, in Romani, there are Turkish high front vowels *o* and *u* in both Vlax and non-Vlax; consonants with aspiration are from Turkish *köşe* or *ker* instead of *kher* (*house*). Most Turkish loanwords are phonologically and morphologically adapted to Romani; often some phonemes or morphemes are lacking from the Turkish word or are metathesized (e.g., *ilaşlardI* from standard Turkish *ilaçlıyordu* ‘treat’ and *ilaçlırdi* in local Turkish. (b) Lexically, there are borrowings from Turkish (for example *keçi-s* ‘goat’ from Turkish *keçi* ‘goat’. (c) Syntactically, word order could be SOV as in Turkish or SVO as in Bulgarian.

(a) (*Zagondži* dialect: Vlax)

Mir dej'da ando gav pirel-dI giva kidel-dI
 my mother and in a village walk 3sg. PT corn gather-3sg. PT
 ‘And my mother was walking in the village and gathered corn.’

In example (a), the *-dI* suffix in *pirel-dI* ‘walked’ and *kidel-dI* ‘gathered’ is Turkish past tense (PT); Turkish word order is SOV. In example (b), note that the word order is SVO, that of Bulgarian.

(b) (*Xoraxani* dialect: non-Vlax)

Me mang-av-as ek čaes
 I want-1sg. PT one/a boy
 ‘I was in love with (a) boy.’

Vlax is more conservative and has fewer borrowings from Turkish compared to non-Vlax. Typical non-Vlax dialects to have a high number of Turkisms, and for Romani speakers who do not speak Turkish, it can be difficult to communicate with speakers of non-Vlax. However, even without borrowings, there are some differences.

The sentence: ‘Where are you going?’ in Vlax (*Zagondži*) will sound like: *Kaj džas?* while in non-Vlax (*Xoraxani*) will sound like: *Kariga žasa?* In this particular Vlax (*Zagondži*) dialect, every single *ž* is changed to *dž*, except for names of people. The non-Vlax (*Xoraxani*) contains only the consonant *ž*. Although *ž* is typical for Vlax and *dž* for non-Vlax, there is overgeneralization, typical for languages in contact. For both, the influence and changes come from Turkish.

In the data, there is no single case of intersentential code-switching (CS); all the code-switching of the informants’ is intrasentential, interest-

ing because speakers never “break” sentence borders when speaking Romani.

The designated words in the following examples are non-Romani, illustrating code-switching, borrowing and triggering, with Romani, Turkish and Bulgarian:

A. Lexical code-switching from Bulgarian into Romani

- (1) (*Xoraxani*)
Znači mangela te čija
 means-it want-3sg PRES f. your daughter
 ‘It means that she wants your daughter.’
- (2) (*Zagondži*)
Nezavisimo save gaveste arakadjol
 does no matter which village-in is-3sgPRES
 ‘It does not matter in which village he is.’
- (3) (*Xoraxani*)
Ajšake može te dav la
 Ayshe-to (I) can to give her
 ‘I can give her to Aysha.’
- (4) (*Zagondži*)
Mir day, sašto amare romnja ...
 my mother also our women
 ‘My mother and also our women ...’
- (5) (*Xoraxani*)
Ame kidija sijamas amaro semejstvo
 We like that were our family
 ‘We, our family, were like that.’
- (6) (*Zagondži*)
Me kerav i večerja
 I do-1sg PRES the dinner
 ‘I prepare the dinner.’

The words *znači*, *nezavisimo*, *može*, and *sašto* are from Bulgarian without changes. The examples show CS can occur at the beginning or in the middle of sentences. The main reason for CS here is lexical gaps, the speakers switching to Bulgarian to fill them. Also in *Xoraxani*, there are no words for ‘it means’ and ‘I can’. The known words for ‘it means’ and ‘I can’, from other dialects such as *dikel pe*; *phenel pe* and *me asthi*, are typical for *Vlax* and are unknown to speakers of non-*Vlax*. The words *semejstvo* and *vecerja* are Bulgarian nouns used in *Vlax* and non-*Vlax*,

speakers not knowing the Romani word for 'family' but knowing the Bulgarian word. Why they prefer the Bulgarian word is unclear. The speaker who used the Bulgarian word for dinner knows both Romani and Turkish, but the reason to use Bulgarian is to make fun of what she is doing (observed from the context of the interview).

The speakers of Vlax also code-switch, and places of CS are almost the same, the explanation for CS by the Vlax is also the lack of knowledge of Romani words. Both groups (Vlax and non-Vlax) do not use Turkish words for CS because these words are too abstract or unknown.

Examples (7) and (8) demonstrate the same phenomenon, but the code-switching is from Turkish:

(7) (*Xoraxani*)

Onbes-tane bajce sine amen
fifteen CLASS pigs are to-us
'We have fifteen pigs.'

(8) (*Xoraxani*)

Mo dad evlendi, baška romna lijas
my father married-3sg PT another woman get-3sgPT
'My father got married to another woman.'

Here, speakers prefer to use Turkish words because they sound more prestigious; the speakers code-switch to Turkish, although they know the Romani words. See examples (9) and (10):

(9) (*Zagondži*)

A onar in-keren
and they not-do-3 pl. PRES kindness
'And they are not kind.'

(10) (*Zagondzi*)

Kaj ingjareh man çocum
where bring-2sg. PRES me son-my
'Where do you bring me my son.'

In the data, there are few examples of Bulgarian-Turkish-Romani code-switching; they are from the non-Vlax (*Xoraxani*), which seems more "open" to CS, as there are no examples in the data of simlilara CS in Vlax:

(11) (*Xoraxani*)

Samo tIkIysIn to muj
only close-2sg PRES your mouth
'You only close your mouth.'

In example (11), *samo* is Bulgarian and *tIkIysIn* Turkish; the rest is Romani, a kind of CS very typical for *Xoraxani*.

B. Patterns of borrowings from Bulgarian

(12) (*Xoraxani*)

Ekta gerda-ja komplek-ci lijas
 and necklace-pl. set-sg buy-3sgPT
 'And she bought a set of necklaces.'

(13) (*Zagondzi*)

Savoren e prazni-ča kerah
 all the holiday-pl do-1pl. PRES
 'We celebrate all holidays.'

Here Bulgarian words are adopted to Romani: *erdaya komplekci* in example (12) from the Bulgarian *gerdan* 'necklace' and *komplekt* 'set'; however, Romani endings for plural in this dialect *-ja* and for singular *-ci* are added. In example (13), the Bulgarian word is *praznik* 'holiday'; this dialect's ending for plural form is *-ča*, from the Bulgarian, *-ci* for plural.

C. Patterns of borrowings from Turkish

(14) (*Xoraxani*)

Alti-jeski koro lijas
 gold-from bracelet buy-3sg. PT
 'A bracelet of gold was bought.'

(15) (*Zagondzi*)

Mor dad džambaz-llk-i kerel-di.
 my father horse-selling make-3sg. PT
 'My father was a horse-dealer'.

The borrowings from Turkish are much more frequent in *Xoraxani* than *Zagondži*, and in the data, it was difficult to find Turkish borrowings in *Zagondži* speech. In example (14), the word borrowed from Turkish is *altın* 'gold', but there are changes, mostly in endings (*-jeski*). In example (15), borrowed from Turkish is *cambazlık* 'horse selling'; however, the word is adapted to Romani with the ending *-i*.

Triggering is a process where a border-word stimulates code switching, and because of the same pronunciation in both languages, it is very easy to switch from one language to another.

D. Triggering from Romani to Turkish

(16) (*Xoraxani*)

Mo rom Varna'da iş-ley
 my husband in Varna work-3sg. PRES
 'My husband works in Varna'.

Here, the border word is *Varna'da*. In the *Xoraxani* dialect, instead of the local Turkish word *işley*, the word *kerla xIzmeci* could be used, meaning 'works', but probably has the same pronunciation of the word in *Varna* (*Varna'da* [locative]), the reason for using the Turkish word *işley*.

4. Findings

Analyzing the data, I was trying to find factors that influence code-mixing. First, there are *social factors* such as prestige of the used language and the participants. Some speakers changed their normal speech because I was a participant and knew they were being recorded. Other informants tried to show better knowledge of Turkish, using more Turkish borrowings. *Xoraxani* dialect speakers have negative attitudes to Romani spoken by *Zagondžiji*, also negative attitudes to Bulgarian. In the hierarchy of languages (according to their system of values), Turkish ranks highest, one reason that parents teach their children Turkish from a very early age, and not as much Romani or Bulgarian.

Other factors influencing language mixing are *linguistic*, Romani, Bulgarian and Turkish have been in contact for centuries and have influenced each other. In all three languages, one can find different stages of words from each language. Romani has many borrowings, Bulgarian and Turkish, but especially those two dialects discussed here borrowed more from Turkish. They may be organized into the following: *phonological* *ö* and *ü*, sounds from Turkish; *morphological*, in both dialects, there are borrowed morphemes from Turkish, such as: the possessive *-m* (*Devla-m* 'my God'); past tense-*dI* (*pirel-dI* 'walked'); locative-*da* (*Varna'da* 'in Varna'); and *lexical* in both dialects, words from Turkish are adapted to Romani.

5. Conclusion

The two dialects discussed in this article are non-Vlax (*Xoraxani*), which shows language mixing and a higher prestige, and Vlax (*Zagondžji*) with less mixing and lower prestige. Actually, the language and culture of the Muslim Roms in Bulgaria are not very different from those Muslim Roms in the other Balkan countries (Romania, Albania, Macedonia, Greece), clear from publications available, although their number is very limited (Heinschink 1989; Friedman 1989). The following conclusions could be made:

- (a) The Muslim Rom speakers switch to Turkish and Bulgarian while speaking Romani because those three languages are connected. Switches can be at the beginning, in the middle and at the end of sentences, such switches typical for *Xoraxani* and *Zagondzi* dialects. Usually speakers switch to Bulgarian because they do not know the Romani or Turkish substitutions, and switch to Turkish for more prestige. For the *Xoraxani*, it is typical Bulgarian and Turkish code-switching, but it is not typical for the *Zagondži*, although there are these kind of patterns as well.
- (b) In both dialects, there are borrowings from Bulgarian and Turkish, but there are no sentences with mixed Bulgarian and Turkish borrowings.

It seems that the main reason for CS among Muslim Roms is language prestige. Although speakers know the Romani words, they prefer to use Turkish, giving them higher prestige. The feeling that one may not be accepted stimulates speakers for more code-switching and borrowings from Turkish; moreover, this becomes a reason for language shift from Romani to Turkish for many. From this perspective, Romani dialects spoken by Muslim Roms could be determined as an interlanguage to target the Turkish language.

University of Veliko Turnovo St. St. Cyril and Methodius

Notes

- * I would like to thank Peter Bakker from the Linguistics Department of the Aarhus University for his ideas and help, and Yaron Matras from the Linguistics Department of Manchester University for his comments and suggestions on previous versions of the paper.
1. Linguists usually differentiate between “code-switching *within* and code-switching *between* sentences. Intrasentential code-switching is studied mainly by theoretical linguists and psycholinguists. Sociolinguists place less emphasis on the distinction between inter- and intrasentential code-switching. They are in general more concerned with the reasons why bilinguals code-switch at all” (Backus 1992: 1).

References

- Backus, Ad (1992). *Patterns of Language Mixing. A study of Turkish-Dutch Bilingualism*. Wiesbaden: Otto Harrassowitz.
- Friedman, Victor (1989). Toward defining the position of Turkisms in Romani. In *Jazik i Kultura Roma* [Language and culture of the Roms], Said Balic (ed.), 251–267. Sarajevo: Nasa kniga.

- Gilliat-Smith, Bernard (1913–1914). A ninth Bulgarian Gypsy folktale. *Journal of the Gypsy Lore Society, New Series* 7(1), 37–41.
- Heinschink, Mozes (1989). Language and culture of the Izmir basket-weavers. In *Jazik i Kultura Roma* [Language and culture of the Roms], Said Balic (ed.), 103–111. Sarajevo: Nasa knjiga.
- Kyuchukov, Hristo (1995). The Turkish dialects of Muslim Roms (Gypsies) in Bulgaria. *Journal of Turkology* 2(2), 305–307.
- (1997). *Psiholingvisticni aspekti na rannija bilingvizam* [Psycholinguistic aspects of early bilingualism]. Sofia: Balkan Foundation for Cross-Cultural Education and Understanding.
- Liegeois, Jeane-Pierre (1994). *Roma, Gypsies, and Travelers*. Strasbourg: Council of Europe.
- Marushiakova, Elena; and Popov, Veselin (1993). *Ciganite v Balgarija* [The gypsies of Bulgaria]. Sofia: Club 90.
- Myers-Scotton, Carol (1993). *Social Motivations for Codeswitching*. Oxford: Oxford University Press.