

NEW RESEARCH ON ROMANI IN EASTERN NORTH MACEDONIA: THE MALESHEVO DIALECT

Kirill Kozhanov

Universität Potsdam
kozhanov@uni-potsdam.de

Victor A. Friedman

Chicago University
vfriedm@uchicago.edu

This article presents new data on previously undescribed Romani varieties spoken in the eastern part of the Republic of North Macedonia, specifically in the geographical region of Maleshevo. We show that this is a South Balkan dialect that differs from other South Balkan dialects spoken in North Macedonia (especially in Skopje), while sharing several features with the Romani dialects of southwestern Bulgaria. This pattern corresponds to the geographical distribution of South Slavic dialects. Following South Slavic terminology, these Romani varieties can be referred to as Maleshevo-Pirin Romani. In addition to describing typical dialectal features, this study pays special attention to borrowings from local Macedonian dialects.

Keywords: Romani, Macedonian, dialectology, language contact.

НОВО ИСТРАЖУВАЊЕ ЗА РОМСКИОТ ЈАЗИК ВО ИСТОЧНА СЕВЕРНА МАКЕДОНИЈА: МАЛЕШЕВСКИОТ ДИЈАЛЕКТ

Кирил Кожанов

Универзитет во Потсдам
kozhanov@uni-potsdam.de

Виктор А. Фридман

Универзитет во Чикаго
vfriedm@uchicago.edu

Во оваа статија се претставуваат нови податоци за претходно неописаните ромски варијанти, што се зборуваат во источниот дел на Република Северна Македонија, конкретно во географската област Малешево. Овие дијалекти, кои припаѓаат на групата јужнобалкански дијалекти на ромскиот јазик, се одликуваат со неколку карактеристики, што ги разликуваат од другите дијалекти што се зборуваат во Северна Македонија (првенствено во Скопје), но истовремено се поврзуваат и со ромските дијалекти од Југозападна Бугарија. Распределбата на овие ромски варијанти се поклопува со географската поделба на јужнословенските дијалекти и, според прифатената терминологија, можат да се наречат малешевско-пирински. Покрај опишувањето на нивните типични дијалектни карактеристики, во ова истражување посебно внимание им се посветува на заемките од локалните македонски дијалекти.

Клучни зборови: ромски, македонски, дијалектологија, јазичен контакт.

1 Introduction

Romani is an Indo-Aryan language that has been spoken in the Balkans since the Middle Ages (Matras 2002; Matras and Tenser 2020). The core of the Romani lexicon is undoubtedly of Indo-Aryan origin (see Oslon and Kožanov, in prep.), and its morphology retains several features typical of New Indo-Aryan languages (Beníšek 2020).

On their way from the Indian subcontinent to Europe, Romani speakers came into contact with Iranian and Armenian, as evidenced by loanwords (Scala 2020). However, the language was largely reshaped by contact with Medieval Greek during its development in the Byzantine Empire, where Romani speakers remained for several centuries: possibly since the 11th-12th centuries until the northern migrations, which began no later than the 15th century. During this period, Romani acquired several features typical of Balkan languages (Matras 1994; Friedman 1985, 2000), among which the use of finite subjunctive clauses instead of an infinitive serves as a characteristic example. Subsequent migrations of Roms¹ northward beyond the southern Balkans have led to the development of various dialects shaped by contact with different languages (Matras 2005).

Being spoken in southern Balkans, Romani must have been in contact with South Slavic languages prior to the 15th century, as evidenced by loanwords in those Romani dialect groups that left this area no later than the beginning of the 15th century (Boretzky 2013; Kozhanov and Oslon 2020). Historical documents mention Roms in the northern Balkans in the 14th century (Marushiakova and Popov 2001: 18–19), which is possibly the period when Roma settled in the territory of contemporary North Macedonia. However, the earliest attestations of Romani communities here are much later (for example, in Skopje, the capital of North Macedonia, in 1523; see Friedman 2017a: 30).

In North Macedonia, Romani is spoken throughout the country, with a particularly high concentration of speakers in the capital, Skopje. According to the 2021 census (https://www.stat.gov.mk/publikacii/2022/POPIS_DZS_web_MK.pdf), the Romani population of North Macedonia is approximately 49,000, or around 2.3% of the total population. Much of the Romani population resides in Skopje (approximately 20,000), particularly in the municipality of Šuto Orizari, where Romani is also recognized as an official language of administration (see Friedman [1999]).

Romani dialects spoken in North Macedonia belong to three distinct dialect groups: South Balkan, North Balkan, and South Vlax (for an overview of Romani dialectology, see Elšík and Beníšek 2020; for an overview of Romani dialects of Skopje, see Friedman 2017a). South Balkan Romani dialects (sometimes referred to as Balkan I) are primarily spoken in the southern Balkans and represent dialects that never left the Balkans. These dialects are characterized by the further contact with Greek, South Slavic, Turkish, and Albanian. There are also South Balkan dialects spoken outside the southern Balkans, such as Ursari in Romania and Crimean Romani in Ukraine and Russia. An overview of this dialect group is presented in Boretzky et al. (2008). In North Macedonia, this group is represented by Arli (the largest dialect of Skopje) and the Romani variety of Prilep (Boretzky and Cech 2016). North Balkan dialects (also referred to as Balkan II) are primarily spoken in the northern Balkans, with their center arguably in central and northern Bulgaria (for more details, see Boretzky 2000). Speakers of North Balkan dialects are usually referred to as Kovači in Skopje, where they are believed to

¹ In this article, when writing in English, we treat *Rom* as a normal English ethnonym rather than exoticizing it. Just as the English plural of *Turk* is *Turks*, not *Turkler* and that of *Magyar* is *Magyars*, not *Magyarok*, so too is the plural of *Rom* in English *Roms* (such integration into native grammar occurs in all the other languages of countries where Roms live, e.g. Macedonian *Rom-Romi*, Albanian *Rom-Romë* (or *Rrom-Rromë*), Turkish *Roman-Romanlar*). The adjective from Rom is Romani.

have migrated “from northeastern Bulgaria at some time in the distant past” (Friedman 2017a: 34). Finally, the speakers of South Vlax, referred to as Džambazi in Skopje, came to North Macedonia from Romanian-speaking territories.

Most linguistic work on Romani in North Macedonia has focused on the Romani of Skopje. The grammars and dictionaries that have been produced locally so far (Kepeski and Jusuf 1980; Demir and Demir 2009ab; Petrovski and Veličkovski 1999; Demir, Djurić, and Demir 2010ab) are all intended mainly for pedagogical purposes. There are no linguistically oriented grammars or dictionaries. Nevertheless, fieldwork on Romani by professional linguists in North Macedonia began in the 1960s, when Austrian linguist Mozes F. Heinschink started making recordings, which are now stored in the Heinschink Sammlung at the Austrian Academy of Sciences (see Fennesz-Juhász 1996). Some of the narratives were later published in the folklore collection (Cech et al. 2009). In the 2000s, several samples were recorded using the Romani morpho-syntax questionnaire, including six recordings from Skopje and Kumanovo, which are available online (<https://romani.dch.phil-fak.uni-koeln.de/>). A glossary of Macedonian Arli, based on the Arli dialects of Skopje and Kumanovo, was also included in RomLex, a lexical database on Romani dialects (<http://romani.uni-graz.at/romlex/>). Three recordings in three dialects made by Zuzana Bodnárová in 2019 are available online as part of the VLAH (Vanishing Languages and Cultural Heritage) commission of the Austrian Academy of Sciences (<https://www.oeaw.ac.at/vlach/collections/romani/>).

Further work on the description of Romani dialects in North Macedonia is needed, especially considering the language shift observed in some communities. In this article, we present some preliminary results of fieldwork on the Romani variety spoken in the easternmost part of North Macedonia, along the border with Bulgaria (Crnik, Delčevo, Vinica). To the best of our knowledge, this variety has not been previously described, although some data have been collected from neighboring locations (for details, see Section 2).

2 Data and research questions

The data for this study were collected during several work sessions with native speakers of the dialect. Upon meeting Senada Lamovska, a native speaker of this dialect, in Skopje in March 2024, Kirill Kozhanov, one of the authors of this paper, observed that although the variety in question belonged to the South Balkan dialect group, it was distinct from the Romani dialects spoken in Skopje. With the help of Senada, who remained the main consultant on the variety, a field trip to the eastern part of North Macedonia was organized in August 2024. Recordings were made in several locations, as indicated in Fig. 1. The description was then continued in Skopje and later via Zoom.

As of now, there are five hours of free narratives in the dialect, as well as four hours of elicitation of lexical and grammatical information recorded from ten speakers. In addition to the recordings, there are fieldwork notes taken during participant observation and unrecorded conversations. These data serve as the basis for the following overview. Currently, the data are stored in the authors’ personal archive and not available online, but we are considering creating an annotated corpus of this variety in the future.

Although no fieldwork has been conducted in the named locations, Mozes Heinschink made recordings in Kočani, a nearby city. Currently, the majority of the Romani community in Kočani has shifted to Turkish. Heinschink’s data from Kočani were later used in Boretzky et al. (2008), where Kočani is included as one of the datapoints. Additionally, some recordings of Romani dialects from the other side of the border, specifically in Blagoevgrad in southwestern Bulgaria, were made by Birgit Igla (two manuscripts from 1996 and 2002 are cited by Boretzky et al. 2008: 68). In 2009, Victor Friedman recorded an interview with a

Romani speaker who was born in Štip but resided in Trabotivište near Delčevo. In the following discussion, we will use the data from Boretzky et al. 2008 for the locations in southwestern Bulgaria.

The speakers of the dialect refer to themselves as *rom* (pl. *romá*) and distinguish themselves from *gadžó* (pl. *gadžé*), the non-Romani-speaking surrounding populations. The non-Romani population can be further specified as *das* (pl. *dasá*), typically Orthodox Christian Macedonian-speaking neighbors, and *gomí* (pl. *gomjá*), usually Muslim neighbors, often with Turkish identity, regardless of home language.² The Turks are also referred to as *xor(x)áj* (pl. *xor(x)ajá*).

When referring to their dialect, the Roms call it Kovački (see an extract from an interview in the Appendix), which should not be confused with the Kovački spoken in Skopje, a North Balkan dialect, nor with the Kovači of the Ohrid region, who are Tosk Albanian speakers of Romani origin. In North Macedonia in general, most Romani speakers (90%) are Muslims.

The same Romani variety is spoken in Crnik, Delčevo, and Trabotivište. In Berovo, according to the language consultants, the Romani community predominantly speaks Macedonian. However, in 2012, during his visit to Berovo, Victor Friedman met many Romani speakers, and at least one recording was in North Balkan Romani; most Roms of Berovo also spoke Turkish. In Vinica, Roms speak both Romani and Turkish. When telling about their parents and grandparents, Roms from these locations mention Krupnik, Blagoevgrad, Simitli, and Sandanski in Bulgaria as locations where they have family. However, since the introduction of the border between North Macedonia and Bulgaria (1912–1941, 1944–present), contacts have become less frequent.



Figure 1. Map of Maleševu Romani: Locations where data were collected are marked by larger circles

² The term *gomí* in such usage is characteristic of Romani from Bosnia (see Igla 2019).

In what follows, we present several diagnostic features of these Romani varieties that help situate them within Romani dialectology. We show that this is a South Balkan dialect that shares several features with the Romani dialects of southwestern Bulgaria. Following the terminology used in South Slavic dialectology (see Friedman 1993), we refer to this variety of Romani as Maleshevo-Pirin Romani. When citing our data from eastern North Macedonia, we use the term Maleshevo Romani.

3 Preliminary results

The dialect in question undoubtedly belongs to the South Balkan dialect group. Importantly, it does not have the Romanian loanwords or sound changes typical of Vlax dialects (e.g., *čh > ś, *dl > gl, *numa(j)* ‘only’), and shows no traces of the sound change *gi-/*di- > zi-, *ki-/*ti- > ci- found in North Balkan dialect group, exemplified by Bugurdži here:

	Maleshevo	Bugurdži (Boretzky 1993)
‘song’	<i>gílí</i>	<i>zi</i>
‘day’	<i>divés</i>	<i>zis, ziés</i>
‘cheese’	<i>királ</i>	<i>cirál</i>
‘work’	<i>butí</i>	<i>bucí</i>

3.1 Phonology

The vowel system of the dialect is typical of Romani and includes the five basic phonemes: /i, e, a, o, u/. The consonant system is quite conservative, retaining the aspirated consonants /kh, ph, th, čh/ and the velars /k, g/, while introducing a few innovations, summarized below. For an overview of Romani phonology, see (Baló 2020).

3.1.1 Reflexes of Proto-Romani *ř and *nř

The reflexes of Proto-Romani *ř and *nř are always rendered as /r/ in this dialect. The development *ř > r is typical of most South Balkan dialects (Boretzky 1999: 28–29), whereas the reflex of *nř exhibits more variation in its realizations (Boretzky 1999: 28–29, map 3; Boretzky et al. 2008: 8–9, maps 3, 4). The *nř > r sound change is characteristic of Arli and Yerli dialects and contrasts with the southern part of the South Balkan dialects, exemplified in North Macedonia by the Prilep dialect, which has -nd- as a reflex of *nř (Boretzky and Cech 2016: 18).

	Maleshevo	Arli (Romlex)	Prilep (Boretzky and Cech 2016)
*ř			
‘Rom’	<i>rom</i>	<i>rom</i>	<i>rom</i>
‘stone’	<i>bar</i>	<i>bar</i>	<i>bar</i>
‘flour’	<i>varó</i>	<i>varo</i>	<i>varo</i>
*nř			
‘bread’	<i>maró</i>	<i>maro</i>	<i>mando</i>
‘egg’	<i>jaró</i>	<i>jaro</i>	<i>vando</i>

3.1.2 Retention of *s

One of the old features distinguishing Romani dialects is the alternation of *s/h* in certain morphological forms, such as the INS.SG of nominal declension, PRS.2SG/1PL in the verbal conjugation, and the initial sound in the copula forms (see Matras 1999). Maleshevo Romani consistently uses only the *-s* reflex in all these positions, which puts it in opposition to many other South Balkan or North Macedonia (Boretzky 1999: 30, map 4; Boretzky et al. 2008: 10, maps 7, 9; Friedman 2017ab). In this regard, our data align with data from Kočani and southwestern Bulgaria. Note that long present forms with *-a* are not typical for Maleshevo Romani.

		Maleshevo	Arli (Boretzky 1996: 8–9, 16–17, 25, Friedman 2017ab)
INS.SG	‘with salt’	<i>lonésa</i>	<i>-eja, -ea, -esa</i>
ACC.SG.M	‘man’	<i>manušés</i>	<i>manuše, manušes</i>
PRS.2SG	‘you do’	<i>kerés</i>	<i>kere, kereja, keres, keresa</i>
PRS.1PL	‘we do’	<i>kerás</i>	<i>kera, keraja, keras, keresa</i>
copula (PRS.3SG)		<i>si, isí</i>	<i>si, isi, hi, i</i>

3.1.3 *t' > k', *Vd' > Vg', *Cd' > Cj

The palatalized dentals *t' and *d' are typically realized as *k'* and *g'* in this dialect, a development also attested in other South Balkan dialects and typical of Macedonian, cf. for Arli (Boretzky 1996: 4) and for Prilep (Boretzky and Cech 2016: 16).

		Maleshevo	Prilep (Boretzky and Cech 2016)
*t, *t'	‘warm’	<i>tató</i>	<i>tato</i>
	‘get warm’	<i>tak'ól</i>	<i>takjol</i>
*d, *d'	‘mind’	<i>godí</i>	<i>godi</i>
	‘smart’	<i>gog'avér</i>	<i>gogja(v)er</i>

Importantly, in Maleshevo Romani, if *d' follows a consonant, it is realized as /j/. This sound change appears primarily in the past tense conjugation of the verbal paradigm with the suffix *-d-* (except for 3PL) and in the derivation of mediopassive verbs from participles, thus, appearing after sonorants. To the best of our knowledge, this change has not been previously attested in South Balkan dialects (cf. Boretzky et al. 2008: 9, map 77).

		Maleshevo	Prilep (Boretzky and Cech 2016)
‘do (PST.3SG)’		<i>kerjás</i>	<i>kerdas</i>
‘do (PST.3PL)’		<i>kerdé</i>	<i>kerde</i>
‘be done (PRS.3SG)’		<i>kerjól</i>	<i>kergjol</i>

3.1.4 *dž > dž or ž(')

The affricate *dž* can be retained but also often undergoes deaffrication and is realized as *ž* or even *ž'* such as *džal* ~ *ž(al)* ‘(s)he goes’. The existing descriptions of South Balkan dialects do not report such a change, and it could be a unique development for Romani in Eastern North Macedonia.

3.1.5 *št > šč

Another unique change in this Romani variety, not reported for other South Balkan dialects, is the optional realization of the cluster *št as /šč/. This development probably replicates the variation in the pronunciation of št ~ šč in the local Maleshevo-Pirin Macedonian dialects (cf. Friedman 1993: 302):

‘four’	<i>ščar</i>	<i>štar</i>
‘wood’	<i>kaščá</i>	<i>kaštá</i>
‘can’	<i>aščí</i>	<i>aští</i>

3.2 Morphology

The morphology of Maleshevo Romani is typical of South Balkan dialects (for an overview, see Boretzky 1999: 37–126; Boretzky et al. 2008: 11–34). In the following discussion, we present several features that form isoglosses within South Balkan dialects and are thus important for determining the position of Maleshevo Romani within this group.

3.2.1 Personal pronouns

Maleshevo Romani has the following system of personal pronouns:

	SG	PL
1	<i>me</i> (obl. <i>man</i>)	<i>amén, amé</i> (obl. <i>amén</i>)
2	<i>tu</i> (obl. <i>tut</i>)	<i>tumén</i> (obl. <i>tumén</i>)
3M	<i>ov</i> (obl. <i>les</i>)	<i>olá</i> (obl. <i>len</i>)
3F	<i>oj</i> (obl. <i>la</i>)	

One of the features that distinguishes the South Balkan dialects is the form of third-person pronouns. Maleshevo Romani represents a unique combination for the Romani dialects of North Macedonia (Boretzky 1999: 56–60, maps 22, 24; Boretzky et al. 2008: 16, maps 32–33).

3.2.2 Borrowed noun accommodation

South Balkan Romani dialects exhibit various suffixes for the accommodation of loan words, a phenomenon that shows considerable diversity across other Romani dialects as well (Elšík 2020: 168–170). In Maleshevo Romani, the following endings are primarily used for the singular and plural accommodation of borrowed nouns:

		SG	PL	
masculine nouns	‘city’	<i>-os</i>	<i>-ja</i>	↔ Mac. <i>град</i>
	‘celebration’	<i>grádos</i>	<i>grádja</i>	↔ Mac. <i>празник</i>
masculine nouns	‘phone’	<i>práznikos</i>	<i>práznikja</i>	↔ Mac. <i>телефон</i>
	‘blacksmith’	<i>-i</i>	<i>-(i)ja</i>	↔ Mac. <i>ковач</i>
	‘student’	<i>telefóni</i>	<i>telefónja</i>	↔ Mac. <i>студент</i>
feminine nouns	‘student’	<i>kováči</i>	<i>kováčja</i>	↔ Mac. <i>балтија</i> (↔ Turk. <i>balta</i>)
	‘axe’	<i>studénti</i>	<i>studéntija</i>	↔ Mac. <i>фамилија</i>
feminine nouns	‘family’	<i>-a</i>	<i>-es</i>	
	‘axe’	<i>baltíja</i>	<i>baltíjes</i>	
	‘family’	<i>famílíja</i>	<i>famílíjes</i>	

Borrowed masculine nouns receive the suffixes *-os* (pl. *-ja*) or *-i* (pl. *-(i)ja*). The distribution is phonetic: stems ending in *-v-*, *-f-*, *-r-*, *-l-*, *-m-*, *-n-*, *-t-*, *-č-* typically take the ending *-i*, while the rest employ the suffix *-os*.

When comparing the distribution of accommodation suffixes with other South Balkan dialects (cf. Boretzky 1999: 52–54, maps 19, 20; Boretzky et al. 2008: 12–13, maps 15–17), Maleshevo Romani presents a contradictory picture, highlighting its distinct position among the Romani dialects of North Macedonia. On the one hand, it differs from other Romani dialects of North Macedonia and instead aligns with the dialects of southwestern Bulgaria. This is evident in the singular suffix *-os* of masculine nouns (Boretzky et al. 2008: 12–13, map 6) and the plural suffix *-es* of feminine nouns (Boretzky et al. 2008: 12–13, map 17). This is consistent with other languages in the region, where the dialects of eastern North Macedonia align with those of Pirin Macedonia (southwestern Bulgaria). On the other hand, the accommodation suffixes *-i* (pl. *-(i)ja*), found in this dialect, is typical of other Romani dialects in North Macedonia, but is also present in the bordering Romani dialects of western Bulgaria (see Boretzky et al. 2008: 13, map 16).

3.2.3 PRS.3SG suffix **-el > -ol*

Maleshevo Romani has typically the PRS.3SG ending *-ol* (vs. Common Romani *-el*), e.g., *dikhól* ‘see’, *khelól* ‘dance’, *sikavól* ‘teach’, *phúdol* ‘blow’ etc. This variant *-ol* is present alongside *-el* in Arli (Boretzky 1996: 18), and in Prilep, *-ol* seems to be a dominant variant (Boretzky 1999: 81, map 42; Boretzky and Cech 2016: 45). According to Boretzky (1999: map 42), the *-ol* change in PRS.3SG is attested only in the South Balkan dialects of North Macedonia and not in those of Bulgaria or Greece.

Since the change **-el > -ol* is not observed in *khél* ‘dance (IMP.2SG)’ (another diagnostic word, **šél* ‘100’, is obsolete in Maleshevo Romani), this innovation must be morphological—the suffix *-ol* spreads from the paradigm of intransitive or passive verbs, as in *susl’ól* ‘get wet’, *kerjól* ‘get done’ etc., to other verbs. On the other hand, a similar change is attested in the noun *dovól* ‘God’ (vs. Common Romani *devél*), but it is probably an unrelated phonetic innovation triggered by the unique combination **-eve-*.

In an earlier description of South Balkan dialects, Boretzky (1999: 26) analyzes, most likely incorrectly, cases such as *sovol* ‘sleep’, *dol* ‘God’ in Gnjilane Arli as a result of Vlax influence.

3.2.4 Verb conjugation in the past tense

Maleshevo Romani has the following past tense endings:

	SG	PL
1	<i>-om</i>	<i>-am</i>
2	<i>-an</i>	<i>-en</i>
3	<i>-as</i>	<i>-e</i>

Two past tense suffixes are relevant to the variation within South Balkan dialects: the 1SG and 3SG suffixes. Maleshevo Romani PST.1SG suffix *-óm*, as in *dikh’óm* ‘I saw’, *kerjóm* ‘I did’ etc., is also attested in the Romani dialects of southwestern Bulgaria. In contrast, other South Balkan dialects of North Macedonia commonly have *-um* (Boretzky 1999: map 43; Boretzky et al. 2008: 27, map 79).

The PST.3SG suffix forms an East-West divide within South Balkan dialects. In the South Balkan dialects of North Macedonia, the typical suffix is *-a*, whereas *-ás* is present in the

Romani dialects of Bulgaria, including its southwestern part (Boretzky et al. 2008: 28, map 8, 80).

In Romani, past tense formation involves not only personal endings but also different past stem suffixes, which are distributed across verbal paradigms (for an overview of Romani verbal inflection, see Elšík 2020: 159–163). South Balkan dialects exhibit considerable variation in past tense stems and suffixes used by particular verbs (Boretzky et al. 2008: 63, maps 63–76). Some relevant verbs are exemplified in Maleshevo Romani in contrast to the Prilep dialect:

	Maleshevo	Prilep (Boretzky and Cech 2016)
‘give’	<i>dinj-</i> (3PL <i>dindé</i>)	<i>den-</i> , <i>dend-</i>
‘rise’	<i>uštinj-</i> (3PL <i>uštindé</i>) <i>uštindil’-</i> (3PL <i>uštindile</i>)	<i>uštind-</i>
‘go out’	<i>iklístil’</i> (3PL <i>uklístile</i>)	<i>iklist-</i>
‘laugh’	<i>asanj-</i> (3PL <i>asandé</i>) <i>asándil’-</i> (3PL <i>asándile</i>)	<i>asand-</i> , <i>asandil-</i>

Although there are similarities in the past stem suffixes found in the South Balkan dialects of North Macedonia, Maleshevo Romani occupies again a unique position. It differs not only from other Romani dialects of North Macedonia but also from the available data on Romani dialects in southwestern Bulgaria (see maps 63, 69, 70, 74 in Boretzky et al. 2008 for the verbs ‘give’, ‘rise’, ‘go out’, and ‘laugh’ respectively).

3.3 Lexis

Several lexical isoglosses (including the choice of lexemes but also phonetic changes typical of specific lexemes) create a West-East divide among South Balkan dialects.

	Maleshevo	Arli (Romlex)	Prilep (Boretzky, Cech 2016)
‘can’	<i>ašti</i> , <i>aščí</i>	<i>šaj</i>	<i>možin-</i>
‘shoe’	<i>tiráx</i>	<i>kundra</i>	<i>kundura</i>
‘iron’	<i>sas</i> (obl. <i>sases-</i>)	<i>sastrn</i>	<i>sastrîn</i> , <i>sastêrn</i>
‘name’	<i>aláv</i>	<i>anav</i>	<i>anav</i>
‘chicken’	<i>khajní</i>	<i>khani</i>	<i>khajni</i>
‘small’	<i>tiknó</i>	<i>tikno</i>	<i>tikno</i>

As shown, the lexemes ‘can’, ‘shoe’, ‘iron’, and ‘name’ differ from those found in the South Balkan dialects of North Macedonia. However, according to Boretzky et al. (2008: maps 92, 120, 118, 140), the same variants are attested in the Romani dialects of southwestern Bulgaria.

Conversely, when comparing maps 122 and 153 in Boretzky et al. (2008), the Eastern Macedonian data align with other Romani dialects of North Macedonia and differ from those on the Bulgarian side, which have *kaxni* and *cikno*, respectively.

3.4 Borrowing

In addition to the lexicon shared by most varieties of Romani, this dialect features several possibly late Greek loanwords typical of South Balkan dialects but absent in other dialect groups (cf. *prandél* ‘marry’, *naborénol* ‘become sick’). It also includes more recent borrowings from Turkish and dialectal Macedonian (specifically the Maleshevo-Pirin dialect).

The influence of local Macedonian is evident in several loanwords. For instance, *bóbos* ‘beans’ is likely borrowed from the local Macedonian dialect, which has *боб*, in contrast to Standard Macedonian *зрав* (MDA 2024: map 46). Similarly, *dúma* ‘word’ corresponds to Eastern Macedonian *дума*, whereas Standard Macedonian uses *збор*.

In the following discussion, we present several examples of further borrowings from Macedonian.

3.4.1 Borrowing of pronouns

The inherited Romani interrogative pronoun *kon* ‘who’ is replaced in the direct form by the Macedonian *koj*, while the Romani declension is retained, as seen in *kas* ‘whom (ACC)’, *káske* ‘(DAT)’ etc. This replacement is also attested in the Arli and Prilep dialects of Romani (Boretzky 1999: 67).

In Maleshevo Romani, as elsewhere, the system of indefinite and negative pronouns is fully borrowed from Macedonian, cf. *něšto* ‘something’, *níšto* ‘nothing’, *nikade* ‘nowhere’ etc.

3.4.2 Borrowing of prepositions

Maleshevo Romani employs several prepositions, the most frequent being *ki* (*ko* before masculine nouns) ‘at, in, on’ and *tari* (*taro* before masculine nouns) ‘from’. These prepositions can also be used to mark dative/locative and ablative with nouns, which is a general tendency in the Romani of North Macedonia.

Additionally, this dialect has borrowed several Macedonian prepositions. The prepositions *u* ‘in’ (again, from the local Macedonian dialect) and *od* ‘from’ appear only with place names, as in (1).

- (1) *u Skópje da phir-j-án pal ko rom-á?*
 in Skopje also walk-PST-2SG across DAT Rom-PL
 ‘did you also go to Roma in Skopje?’

Other Macedonian prepositions commonly used in this dialect include *za* ‘for’, *prez* ‘during’, *pred* ‘before’, *protiv* ‘against’, cf. (2–3), where the Macedonian preposition *za* governs the dative form of the Romani personal pronoun and a singular noun, respectively.

- (2) *nanáj bút vréme ili za túke*
 NEG.be.PRS.3SG much time and for 2SG.DAT
 ‘there is not much time [left] for you too’
- (3) *prandesál-i za mi papús-ke*
 get_married.PST-3SG.F for my grandfather-DAT.SG
 ‘she got married to my grandfather’

3.4.3 Borrowing of Macedonian preverbs

Several Macedonian preverbs, such as *iz-*, *po-*, *pre-*, *za-*, are borrowed and used with inherited Romani verbs, cf. (4), where the Macedonian preverb *iz-* is used with the Romani verb *naš-* ‘run’ (cf. the Macedonian verb *узбегза* as the source of this derivation).

- (4) *lésker-i čháj iz-naš-l-ás u avstrálija*
 his-SG.F daughter.NOM.SG PVB-run-PST-3SG in Australia
 ‘his daughter ran away to Australia’

3.4.4 Borrowing of the Macedonian comparative prefix *po*

A typical Balkan Slavic influence, found in many South Balkan dialects of Romani but absent in other Romani dialects outside the Balkans, is the use of the prefix *po-* to derive comparative adjectives and adverbs (Boretzky 1999: 55), as illustrated in (5).

- (5) *garáv-tut po-telé*
 hide.IMP.SG-2SG.RFL COMP-low
 ‘hide lower’

3.4.5 Borrowing of object indexes

An interesting feature of this dialect, which, to the best of our knowledge, has not been previously reported, is the sporadic use of Macedonian object indexes with Romani verbs, as in (6), where *-go* is an object index on the verb (traditionally referred to as an object clitic).

- (6) *dikhl’óm-go*
 see-PST.1SG-OBJ.3SG.M
 ‘I saw him.’

According to our data, the borrowed object indexes *-go* and *-gi* are commonly used in alternation with the Romani personal pronoun forms *les* and *len*, respectively, whereas the feminine singular pronominal object is represented only by the Romani pronoun *la*. As far as our data can tell, the borrowed index is not used co-referentially with another noun in the same noun phrase (i.e. as “clitic doubling”).

4 Discussion

The Romani dialect of eastern North Macedonia (the Maleshevo region) belongs to the South Balkan dialect group and exhibits several distinguishing features that set it apart from other South Balkan Romani dialects spoken in North Macedonia—the Arli dialects of Skopje and Kumanovo, as well as the Prilep variety. As shown in this study, several features of this dialect are shared with the Romani dialects spoken in the neighboring region of southwestern Bulgaria.

The common development of Romani dialects in eastern North Macedonia and southwestern Bulgaria is further supported by close family ties between Romani communities in this area. The special position of these Romani dialects corresponds to the dialect division of South Slavic (the Maleshevo-Pirin dialects). At the same time, other features of this Eastern Macedonian Romani are either unique or shared with the South Balkan dialects of North Macedonia. This places it in a special position, situated on the border of the East-West divide within South Balkan Romani dialects (Boretzky et al. 2008: 47–48; cf. also Sechidou 2011: 89–95).

The dialect is in close contact with the local Turkish and Macedonian dialects. Further documentation is needed, as language shift to Turkish or Macedonian is currently observed in Romani communities of this region.

Text in Maleshevo Romani

Recorded by Kirill Kozhanov in Crnik, North Macedonia, from a male Romani speaker, born in Crnik in 1957.

amé sijám akaná od crnik, ama mir-ó páp-us,
 1PL.NOM be.PRS.1PL now from Crnik but my-NOM.SG.M grandfather-NOM.SG
amé, amar-ó potékl-os sí od délčevo
 1PL.NOM our-NOM.SG.M origin-NOM.SG be.PRS.3 from Delčevo
 ‘we are now from Crnik, but my grandfather, us, our origin is from Delčevo’

[KK: *soske ale athe?* ‘why did they come here?’]

athé al-ó za me nen-áke.
 here come.PST-3SG.M for my grandmother-DAT.SG
 ‘he came here because of my grandmother’
a inače aménde, amar-ó potékl-os sí taro délčevo,
 but otherwise 1PL.LOC our-NOM.SG.M origin-NOM.SG be.PRS.3 from Delčevo
cárevo sélo vaker-ól-as-pes pó-rano, xoraj-á vaker-én-as les, e.
 Carevo Selo call-3SG-IMP-F RFL COMP-early Turk-NOM.PL call-3PL-IMPF 3SG.ACC.M, yeah
 ‘But otherwise, our origin is from Delčevo. It used to be called Carevo Selo—that’s what the Turks called it, yeah’

[KK: *a ko delčevo but roma živinenas?* ‘Did many Roma live in Delčevo?’]

pa i akaté sí, enjavardeš kherá isí, romá.
 so and here be.PRS.3 ninety house-NOM.PL be.PRS.3 Rom-NOM.PL
sámo o terné gelé po... germánija ker-én but-í
 only ART young-NOM.PL go.PST-3PL to Germany do-PRS.3PL work-ACC.SG
othé za-astar-d-é, isí bút mir-í familij-a,
 there PVB-catch-PST-3PL be.PRS.3 much my-NOM.SG.F family-NOM.SG
amén sijám lésa škól’sk-a drugár-ja
 1PL.NOM be.PRS.1PL 3SG.INS.M school-NOM.PL friend-NOM.PL
othé amar-ó dialékt-i povéče čalav-ól ko dasikan-ó,
 there our-NOM.SG.M dialect-NOM.SG more hit-PRS.3SG to Macedonian-NOM.SG.M
ko gadž..., gadžikan-ó dialékt-i.
 to non-Romani-NOM.SG.M dialect-NOM.SG
 ‘Well, there are 90 Roma houses here. Only the young ones went to Germany to work and settled there. I have a lot of family there. He and I are school friends. Our dialect is closer to Macedonian, to the non-Roma, non-Romani dialect’

[KK: *athe ili othe?* ‘here or there?’]

akaté.
 ‘Here’.

[KK: *a ko delčevo sar?* ‘and how is it in Delčevo?’]

ísto, amar-ó dialékt-i. Mátku sí, mátku isí rázlik-a
 same our-NOM.SG.M dialect-NOM.SG little be.PRS.3 little be.PRS.3 difference-NOM.SG
 ‘It’s the same, it’s our dialect. There is just a small difference’

[KK: *naprimer so?* ‘what, for example?’]

pa naští mé te vaker-áv tú-ke só... avká si xári neprimétlivo.
 so cannot.PRS 1SG.NOM SBJ tell-1SG 2SG-DAT what this be.PRS.3 little inconspicuous
athé, bérovo sámó isí, avér čhib mátku ili othé, xári,
 here Berovo only be.PRS.3 other language.NOM.SG little or there little
ama isto amar-ó kováčk-o dialékt-i si,
 but same our-NOM.SG.M Kovačko-NOM.SG dialect-NOM.SG be.PRS.3
čalav-ól ko gadžikan-ó dóka.
 hit-PRS.3SG to non-Romani-SG.M that
amar-é phur-é ačká sikav-d-é amén, i ačká al-ó.
 our-NOM.PL old-NOM.PL this.way teach-PST-3PL 1PL.ACC and this_way come.PST-3SG.M
 ‘I can’t tell you what... It’s somewhat inconspicuous. Here, only in Berovo the language is a
 bit different, but it’s the same our Kovačko dialect, it goes to the non-Romani. Our old ones
 taught us so, and this is how it came’

[KK: *a soske vakeren kovačko dialekti?* ‘and why is it called Kovačko dialect?’]

sóske? naprimer, athé sá siné kováč-ja, e rom-á.
 why for_example here all be.PST3 blacksmith-NOM.PL ART Rom-NOM.PL
léskor-o dád, mir-ó pápus,
 his-NOM.SG.M father.NOM.SG my-NOM.SG.M grandfather.NOM.SG
amén sijám da bašaln-é, bašal-ás-as,
 1PL.NOM be.PRS.1PL also musician-NOM.PL play.music-1PL-IMPF
čalgidžija, m-o dád siné čalgidžij-es, mi familij-a
 musician my-SG.M father.NOM.SG be.PST.3 musician-NOM.PL my-SG.F family-NOM.SG
cél-o.
 whole-NOM.SG
 ‘Why? For example, all the Roma here were blacksmiths. His father, my grandfather. And we
 are also musicians. We played music, my father was a musician, and my whole family.’

Bibliography

- Baló, M. A. (2020). Romani Phonology. In Y. Matras and A. Tenser (eds.). *The Palgrave handbook of Romani language and linguistics*, 119–153. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-28105-2_5
- Beníšek, M. (2020). The historical origins of Romani. In Y. Matras and A. Tenser (eds.). *The Palgrave handbook of Romani language and linguistics*, 13–47. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-28105-2_2
- Boretzky, N. (1993). *Bugurdži: Deskriptiver und historischer Abriss eines Romani-Dialekts*. Berlin: Harrassowitz.
- Boretzky, N. (1996). Arli: Materialien zu einem südbalkanischen Romani-Dialekt. *Grazer Linguistische Studien*, 46: 1–31.
- Boretzky, N. (1999). *Die Verwandtschaftsbeziehungen zwischen den südbalkanischen Romani-Dialekten: Mit einem Kartenanhang*. Frankfurt am Main: Lang.
- Boretzky, N. (2000). South Balkan II as a Romani dialect branch: Bugurdži, Drindari, and Kalajdži. *Romani Studies* 10(2): 105–183. <https://doi.org/10.3828/rs.2000.5>
- Boretzky, N. (2013). Lexikalische Slavismen im Romani. *Zeitschrift für Balkanologie*, 49: 10–46.
- Boretzky, N., Cech, P., and Iгла, B. (2008). *Die südbalkanischen Dialekte (SB I) und ihre innere Gliederung: Analyse und Karten*. Graz: Institut für Sprachwissenschaft, Universität Graz.

- Boretzky, N. and Cech, P. (2016). *Der Romani-Dialekt von Prilep/Makedonien*. Graz: Karl-Franzens-Universität Graz.
- Cech, P., Heinschink, M. F., and Halwachs, D. W. (eds.). (2009). *Kerzen und Limonen. Momelja hem limonja. Märchen der Arlije*. Klagenfurt/Celovec: Drava Verlag.
- Demir, L. M. and Demir, F. (2009a). *Gramatika e romane čhibakiri*. Skopje: Romano kulturakoro thaj edukativakoro centro “Darhia”.
- Demir, L. M. and Demir, F. (2009b). *Gramatika na romskiot jazik*. Skopje: Romski kulturen i edukativen centar “Darhia”.
- Demir, L. M., Djurić, R. and Demir, F. (2010a). *Lavustik romani-makedonikani*. Skopje: Romano kulturakoro thaj edukativakoro centro “Darhia”.
- Demir, L. M., Djurić, R. and Demir, F. (2010b). *Rečnik makedonsko-Romani*. Skopje: Romski kulturen i edukativen centar “Darhia”.
- Elšík, V. (2020). Romani morphology. In Y. Matras and A. Tenser (eds.). *The Palgrave handbook of Romani language and linguistics*, 155–186. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-28105-2_6
- Elšík, V. and M. Beníšek. (2020). Romani dialectology. In Y. Matras and A. Tenser (eds.). *The Palgrave handbook of Romani language and linguistics*, 389–427. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-28105-2_13
- Fennesz-Juhász, C. (1996). Tondokumente europäischer Roma. Die „Sammlung Heinschink“ im Phonogrammarchiv der Österreichischen Akademie der Wissenschaften. In J. S. Hohmann (ed.), *Handbuch zur Tsiganologie*, 311–320. Frankfurt: Lang.
- Friedman, V. A. (1985). Balkan Romani modality and other Balkan languages. *Folia Slavica*, 7: 381–389.
- Friedman, V. A. (1993). Macedonian. In B. Comrie and G. G. Corbett (eds.). *The Slavonic languages*, 249–305. London/New York: Routledge.
- Friedman, V. A. (1999). The Romani language in the Republic of Macedonia: Status, usage, and sociolinguistic perspectives. *Acta Linguistica Hungarica*, 46(3–4): 317–339.
- Friedman, V. A. (2000). Romani in the Balkan Linguistic League—Synchrony and diachrony. In Ch. Tzitzilis and Ch. Symeonides (eds.). *Valkanikē Glossologia: Synkhronia kai diakhronia. Balkanlinguistik: Synchronie und Diachronie*, 95–105. Thessaloniki: University of Thessaloniki.
- Friedman, V. A. (2017a). Seven varieties of Arli: Skopje as a center of convergence and divergence of Romani dialects. *Romani Studies*, 5(27): 29–45. <https://doi.org/10.3828/rs.20172>
- Friedman, V. A. (2017b). The Arli of Skopje: A perceptual dialectological approach. In K. Kozhanov, M. Oslon and D. W. Halwachs (eds.). *Das amen godi pala Lev Čerenkov. Romani historija, čhib taj kultura*, 210–220. Graz: treffpunkt sprachen – Forschungsbereich Plurilingualismus.
- Igla, B. (2019). Balkanski paraleli pri zaemane ot romski: *zǎdže* < *gadžo* i *γκόμεινα* < *gomí*. *Sāpostavitelno Ezikoznanie*, 44(1): 37–44.
- Kepeski, K. and Jusuf, Š. (1980). *Romani gramatika – Romska gramatika*. Skopje: Naša knjiga.
- Kozhanov, K. and Oslon, M. (2020). Romani and Slavic contact. In M. L. Greenberg (ed.). *Encyclopedia of Slavic languages and linguistics online*. Brill. https://doi.org/10.1163/2589-6229_ESLO_COM_032594
- Marushiakova, E. and Popov, V. (2001). *Gypsies in the Ottoman Empire: A contribution to the history of the Balkans*. Hatfield: University of Hertfordshire Press.
- Matras, Y. (1994). Structural Balkanisms in Romani. In N. Reiter (ed.). *Sprachlicher Standard und Substandard in Südosteuropa und Osteuropa*, 195–210. Wiesbaden: Harrassowitz.
- Matras, Y. (1999). *s/h* alternation in Romani: An historical and functional interpretation. *Grazer Linguistische Studien*, 51: 99–129.
- Matras, Y. (2002). *Romani: A linguistic introduction*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511486791>
- Matras, Y. (2005). The classification of Romani dialects: A geographic-historical perspective. In D. W. Halwachs, B. Schrammel and G. Ambrosch (eds.), *General and applied Romani linguistics*, 7–26. Munich: Lincom Europa.
- Matras, Y., and A. Tenser (eds.). (2020). *The Palgrave handbook of Romani language and linguistics*. Cham: Palgrave Macmillan.

- Oslon, M. and Kožanov. K. (in prep.). *Ėtimologičeskij slovar' cyganskogo jazyka (obščecyganskaja leksika)*. <https://rromanes.org/pub/ЭЦЦЯ.pdf> (last access on 25 March 2025).
- Petrovski, T. and Veličkovski, B. (1999). *Makedonsko-romski i romsko-makedonski rečnik*. Skopje: Vorldbuk.
- Scala, A. (2020). Romani Lexicon. In Y. Matras and A. Tenser (eds.). *The Palgrave Handbook of Romani Language and Linguistics*. Palgrave Macmillan: Cham. https://doi.org/10.1007/978-3-030-28105-2_4
- Sehidou, I. (2011). *Balkan Romani: The dialect of Ajos Athanasios/Greece*. Munich: Lincom Europa.