HAL
open science

Matter replication in Romani spoken in Romania

Cristian Padure, Evangelia Adamou

To cite this version:

Cristian Padure, Evangelia Adamou. Matter replication in Romani spoken in Romania. 2019. halshs-02420669

HAL Id: halshs-02420669

https://shs.hal.science/halshs-02420669v1

Preprint submitted on 20 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Matter replication in Romani spoken in Romania

Cristian Padure[1] and Evangelia Adamou[2]

[1]University of Bucharest, Romania and [2]National Centre for Scientific Research (CNRS), France

1. Introduction

The Romani language – *ʁomani t∫hib* – is spoken throughout Europe by the Roma people, variably referred to by outsiders as *Tsigan(es)*, *Gypsies*, and *Bohemians*. The lowest estimates consider that there are more than 3.5 million speakers of Romani in Europe, and more than 500,000 in the rest of the world (Matras 2005: 2). Romania, in particular, hosts the largest population of Romani speakers in Europe, with estimates from 500,000 up to 2,500,000 speakers.

Romani is an Indo-Aryan language of the Indo-European stock which is characterized by the strong influence of several of the languages with which it was or still is in contact. Borrowings were significant in the study of Romani as they allowed researches to retrace the Romani migration from the North of India to Europe (see Rüdiger 1782; Pott 1844). Roma arrived in present-day Romania from the geographical area of present-day Greece somewhere toward the end of the Middle Ages. Contact of Romani with Romanian, an Indo-European language of the Romance branch (also known as *vlah*), has been an extremely influential feature in dialectological works that led Gilliat-Smith (1915) to classify Romani varieties as 'Vlax,' exhibiting contact features with Romanian, and 'non-Vlax.' This early terminological choice is largely preserved in modern dialectological classifications of Romani.

However, the influence of language contact, in general, and Romanian, in particular, is considered by many researchers and activists as an impediment to the

standardization of Romani and its unification toward a common cultural identity (see Hancock 1988; Courthiade 1989; Sarău 2011; Djurić 2005 among others). Indeed, the creation of a unified Standard Romani was deemed necessary in the early 1970s by the members of the International Romani Union (IRU). In the 4th meeting of the IRU in 1990 in Warsaw, a standardized Romani alphabet elaborated by Marcel Courthiade was approved in order to produce educational materials, a dictionary and even an encyclopaedia. The concern for standardization needs to be viewed within the larger political project of introducing the Romani language in the education system, a project that was successfully implemented in Romania since 1998, with the use of Romani in hundreds of schools at the level of pre-school, primary and secondary education, as well as at the level of higher education at the University of Bucharest where Romani-language teachers are trained (see Sarau 2017). Indeed, the Romanian Ministry of Education has adopted the standardized alphabet, which serves for the production of the educational material at all levels.

In contrast, researchers with a background in general linguistics studied Romani within the frame of 'language ecology' according to which linguistic transfer is the outcome of individual bilingualism resulting from adaptation to the social environment and communication needs (see Boretzky 1989; Matras 1995; Bakker 1997; Adamou 2010 among others). In line with these studies, we propose an overview of 'matter replication' phenomena in Romani spoken in Romania, that is, of lexical and grammatical borrowing that involves sound-shapes of words and morphs (Matras 2009). More broadly, we follow a usage-based approach to language contact that holds that the bilingual speaker memorizes not only simple words with specific phonological attributes, but also multi-word sequences. Memorization is strengthened by higher frequency and leads to the higher likelihood for these words or constructions to be retrieved from memory and used in production as long as they are appropriate and licensed in the interactional setting (Backus 2014).

Regarding methodology, the present study is based on the analysis of a corpus of interviews with four Romani-Romanian bilinguals from Romania. The use of corpus linguistic methods for the study of contact phenomena follows on a long

tradition in variationist linguistics (see among others Poplack 2018) and, in particular, in Romani studies it follows on Adamou (2016).

In Section 2, we start by providing some background on Romani dialectology in Romania. We then turn to present the corpus in Section 3 and the results of the analysis in Section 4.

2. Romani dialects in Romania

Matras (2013) offers a detailed update of Romani dialects in Romania based on the collection of dialectological data. The author distinguishes Romani varieties in Romania along a core North-South distribution, where the varieties of the southern area follow general dialect features of the southern Balkans and the varieties of the north-western area (around Bihor and Arad) follow dialect features of central European countries, and northern and north-central varieties demonstrating a density of specific dialectal features. This major division of Romanian Romani varieties into North and South agrees with geographical boundaries, along the Carpathian Mountains, and political boundaries that divided Romania into three provinces, with different histories: Moldavia, Transylvania, and Wallachia. For Matras, these geopolitical boundaries largely shaped social networks in the past and allowed for the diffusion of linguistic features, but are still relevant in present-day networks.

More specifically, two Romani groups occupy well-determined geographic areas: the Hungarian Roma (Romungro, Ungrika Rom or Roma Ungrika), who mainly reside in Transylvania and speak dialects of the Central branch of Romani, and the Gabor, who mainly reside in Transylvania and the Banat regions. In contrast, the Kalderaš Roma, who speak dialects of the Vlax branch of Romani, reside in many areas of Romania, possibly following migration from a common centre. At present, a significant number of speakers of the Vlax Romani dialects is also found throughout European countries and the Americas (see Adamou 2013;

Padure, de Pascale, and Adamou 2018). Kelderash is also the preferred dialect in Romani publications in Romania. The Kərəmidarja varieties are also spoken throughout the country. In the south of Romania and along the Black Sea, one finds the Ursari, who speak Romani dialects of the Balkan branch, and in the southeast, the Spoitori, who according to Matras speak Balkan Romani varieties that exhibit common linguistic features as well as linguistic features from neighbouring Romani varieties.

Official numbers of the minorities in Romania show the dispersion of Romani speaking populations in areas where they represent from roughly 1.5% of the majority population up to more than 20%, with only a minority of areas where Romani speakers represent more than 60% of the local population (see http://ispmn.gov.ro/maps/county/limba2011_roma).

3. The corpus

In this paper, we chose to analyse a corpus of life stories that has been collected and published in 2016 by the Centre of Culture and Social Research 'Romane Rodimata' (Furtună, Medeleanu, and Petrilă 2016). The corpus is based on approximately four hours of interviews with four speakers who identified their dialect as Carpathic, Kelderash, Ursari, and Spoitori thus representing a variety of Romani dialects. Table 1 presents the characteristics of the four speakers, two male and two female, all of them aged 60 and above. Though we do not have any information about the linguistic profiles of these speakers, the analysis of the interviews offers evidence that these are early bilinguals, that is, they have learnt both Romani and Romanian in childhood. These short life stories also allow us to conclude that all four speakers had an immersed experience with Romanian and Romani throughout their lives. The interviewers were Romani native speakers: Adrian-Nicolae Furtună, Daniel Samuel Petrilă, Steluța Slate, and Luiza Medeleanu.

It is well-known that the type of settings, whether formal or informal, where bilinguals interact determines the amount of contact phenomena. More specifically, the 'dominant' language is the one that fits the communication settings that belong to the public domain and power-related domains. On the other hand, the 'non-dominant' language is used in a limited number of interactions, generally by bilingual speakers, and therefore tends to be more affected by language contact. It is therefore expected that the Romani speakers who were interviewed in this project might - consciously or unconsciously - limit the use of words from the dominant language, Romanian, as the setting is more formal than everyday interactions, and the goal of the project is the documentation of Romani for younger generations and teachers. However, as Adamou (2016) has found, even in language documentation projects, speakers tend to adopt the amount of words from the current contact language that are common within their community.

The corpus amounts to 9,418 words. The edition includes the transcription in Romani, and its translation into Romanian and English, accompanied by a CD with the original video and audio recordings.

Table 1. Descriptive representation of the Romani-Romanian bilinguals with respect to age, gender, dialect and group self-identification, place of residence, and profession

Name	Viorica Marin[1]	Valerică Stănescu	Mihai Puncă	Viorica Mimiş
Age at the moment of recording	64 years old	75 years old	83 years old	60 years old
Gender	Female	Male	Male	Female
Dialect and group self-designation	Richinara (Ursari)	Kelderash	Carpathian	Spoitori (Tinsmith)
Place of residence	Buzău, Buzău County	Piteşti, Argeş County	Glodeni, Mureş County	Giurgiu, Giurgiu County
Profession	Tailor and merchant	Writer	Basket weaver	Tinsmith and merchant
Total amount of words	2820	2586	2120	1892

4. Results

The analysis of the corpus shows, as predicted, that the great majority of the conversation is conducted in Romani, with a majority of Romani verbs and occasional use of Romanian verbs. Romanian verbs are integrated into Romani morphology and they typically require the use of so-called 'loan verb adaptation markers.' Consistent with observations in the distribution of the forms of loan verb markers in Matras (2013), we note the form *-isar-*, which is usually associated with

[1] We do not anonymize the interviewees following the original publication.

the Vlax dialects, for the speaker of Kelderash, as in *skriisardem* 'I wrote' from Romanian *a scrie* 'to write,' and *träisaraw* 'I live,' from Romanian *a trăi* 'to live,' and for the speaker of Ursari, as in *ïncepisalăm* 'we started' from Romanian *a începe* 'to start.' We find the South Balkan *-zis-* loan verb adaptation marker for the speaker of Spoitori, *ïncelezisajlem* 'I understood,' from Romanian *a înțelege* 'to understand.'

We also note the occasional use of Romanian lexico-grammatical constructions, which, as noted by usage-based approaches, seem to be memorized by the speakers as such and are most likely considered to be more expressive or context-appropriate (e.g., *Când într-o parte, când în alta* 'I just bounce around between them,' *Era un om de regretat* 'He was worth mourning,' *parcă nu mă trăgea inima* 'it was like my heart told me not to…').

As noted in other language contact settings, age is consistently provided in the social-dominant language, in this case Romanian (e.g., *şapteşcinci de ani* 'seventy five years old,' *şapteşpe ani* 'seventeen years old,' *paişpe ani* 'fourteen years old'). Romanian is also used for dates (e.g., *Cam până şaizeci şi şapte* 'Around (19)67,' *în şaptezeci* 'in (19)70,' *în treizeci şi doi* 'in (19)32'), and temporal descriptions (e.g., *10 ani fără 8 luni* '10 years, but without 8 months'). The connection of birth dates and age to the socially dominant language is reported more generally in the literature of language contact as it is related to the fact that giving one's date of birth generally takes place in an institutional context and is therefore associated to the language of the institutions (Matras 2009).

In keeping with this observation, institutional realities, referential meanings, and Romanian-related activities also frequently draw from Romanian: for example, Romanian is used in school-related domains, as in (*a doua) şcoală* 'the secondary school,' *uniformă dă elev,* 'school uniform,' *cu emblema la mână* with an emblem on her arm'; for medical realities, as in *cam cu dureri de cap* 'with a sort of a headache' *să răspundă la tratament* 'if he responded to the treatment'; to refer to shops and describe housing, as in *La magazia dă tăbăcărie care* 'At the tannery warehouse', *cu patru camere, două băi* 'with four rooms and two bathrooms'; and for technical terms such as *cu vopsea gonflabilă – cauciucată* 'with rubber paint.'

A closer look at the corpus also allows for the identification of a variety of Romanian 'utterance modifiers' that are introduced into the Romani speech of the interviewees. Utterance modifiers refers to a broad category including conjunctions, tags, fillers, interjections, and focus particles that according to Matras (1998, 2009) occur in intense communicative negotiation and are therefore more vulnerable to the cognitive pressure that bilinguals face in speech production. For instance, the Romanian connective particle *atunci* 'then' is used by a single speaker on seven occasions (with no Romani equivalents); see examples from (1) to (7). It is found mainly clause initially, on one occasion following a marker of responsive progression such as 'well,' and in a single use, illustrated in (5), following the temporal clause. The Romanian marker of temporal succession, *atunci* 'then,' seems to be an old borrowing that is also found among Lovari speakers who no longer speaker Romanian (Matras 1998).

Speaker: Stănescu

1. **Atunci** äk ćhej phendă manqä kadja… 'Then a girl told me…'[2]
2. **Atunci** me träbulas te źaw kaj śkoàla. 'Then I had to go to school.'
3. Apoi **atunci** phendem laqä… 'Well, then I told her…'
4. **Atùnći**, kana śundem gälem k-o diriźìnto 'haj phendem lesqä… 'Then, when I heard I went to the Principal and I told him…'
5. Kana sa' amenqä bok **atuncea** khälas amenqä. 'When we had hungry then we dance…'
6. **Atuncea,** *chiar* kadja kärdăs. 'Then, he did exactly like that…'
7. **Atùnći** oj so phendă manqä. 'Then, she told me that…'

[2] All the examples in the paper are from (Furtună, Medeleanu, and Petrilă 2016). The words under discussion are in bold; words from Romanian that are not the focus of discussion are in italics; words from Romani are in plain font.

The Romanian *după aia*, *dupa ce*, and *pe urmă* 'after that' is used among all four speakers to structure the discourse, whether with a Romani or a Romanian verb. *După* is used as a temporal adverb in a single case; see example in (18). The Romani equivalent in these cases is used by one speaker; it is the locative-temporal *palal* (see (27) and (28)).

Speaker: Marin

8. **După aia,** kerdŏm *paraśuta.* 'After that, I did the parachute…'
9. **După aia,** uśtinŏm. 'After that, I woke up…'
10. Haj **după aia**, sas amen *śablònă.* 'And after that, we had the model.'
11. **După aia**, so maj kerdăm... 'After that, what we did…'
12. **După aia,** so *màrfes* maj kerdăm… 'After that, what articles we did…'
13. **După aia** dinăs amen jekh *tristețea*… 'After that, they gave us a sadness…'
14. **După ce** *pornisalăm* ta gelăm k-o xurdorre… 'After we start and went to the little children…'
15. **După ce** prandesalŏm... 'After I got married…'
16. **După ce** *începisalăm* te bitinas maj but *màrfes*… 'After we start to sell more merchandise…'
17. **După ce** gelo, lăs man jekh rojpe. 'After he went, I start to cry…'
18. Ke sas **după** Ceauşescu... 'That was after Ceausescu…'
19. **După aia**, uśtinŏm. 'After that, I woke up…'
20. **După asta**, kerdăm *blùzes*. 'After this, we did blouses…'
21. **Pe urmă** pale gelŏm. 'After that, I went again.'

Speaker: Stănescu

22. **După ce** kărdăm e *baràka*... 'After we did the booth…'
23. **După aceea** telărdăm oćal, and-e aver fòro. 'After that, we left, in another city.'
24. Haj **pă urmă** voj delas dùma le *kolezenca.* 'And after that, she was speaking with the colleagues.'

Speaker: Mimiş

25. **După ce** inklŏm… 'After I left…'
26. **Pe urmă** *aflisardăs* o părinte... 'After that, the father found out…'
27. **Palal** so ćinas o kuzum thos o rat… 'After cutting the lamb we put the blood...'
28. Aj **palal** trin zis… 'And after three…'

The Romanian utterance modifier *pentru că* 'because' is used by one speaker; see (29) and (30). *Pentru* is also used with the meaning 'for,' as in (31), (32), and (33), in variation with Romani *anda*, in (34).

Speaker: Marin

29. **Pentru că**, *ani de muncă*, na-s les… 'Because, working years, they have not…'
30. **Pentru că** geldŏm lan *prea* mïndro leça... 'Because I was very good with him…'
31. **Pentru** el… 'For him…'
32. Haj *regretiskerdăm* **pentru** el... 'And I was sorry for him…'
33. Sar kaj phenesas ke othe si **pentru** *ultima dată*... 'Like he was there for the first time…'

Speaker: Stănescu

34. Pa' so träbul te keras **anda** rrom… 'About what we have to do for the Roma people…'

The Romanian discourse particle *parcă* 'like,' is relatively frequent (see (35) and examples from (39) to (45)). The Romani *sar* 'like' can be used in some contexts as in (36), (37), and (38).

Speaker: Marin

35. **Parcă** nu mă trăgea inima… 'Was like I had no desire to…'
36. **Sar** kana phenes ke *cineva* delas man *cuțite*... 'Like one says that someone gave me knifes…'
37. **Sar** kana phenes ajakha ke sas man *presimtiries*… 'Like one says that I had feelings…'
38. **Sar** kana *däspärcïs* tu' te dadestar 'haj te dejatar… 'Like when you separate of your father and mother…'

Speaker: Stănescu

39. **Parcă** dikhaw les, **parcă** sï angla mande. 'Like I (can) see him, is like he is in front of me.'
40. **Parcă** nakhälas andar murro ilo korènto… 'Like electricity passed by my heart…'
41. **Parcă** sïm maśkar lende. 'Like I am among them.'
42. **Parcă** sïm and-e lende. 'Like I am among them.'
43. Haj **parcă** mande sas o *ćokàno* and-o vast. 'And like I had the hammer in my hand.'
44. **Parcă** sas kaj e śkoàla *dă când îi lumea.* 'Like he was in school for ever.'
45. **Parcă** *traisarawas* me kodă *viàca* lenqi. 'Like I was living their life...'

Other examples of Romanian sentence particles are *deci* 'so' and *asa ca* 'so,' and we note some uses of the adversative marker *dar* 'but.'

The interviewees also used the Romanian adverb *foarte* 'very' with adjectives, adverbs, and verbs (examples (46), (48), (49), (50), (51), (53), (54), (57) and (58)), in variation with Romani *but* (the latter being more frequent with nouns; see examples from (59) to (64)). Sometimes we find a combination as in *foarte but* 'very much' in (47), (52), (55), and (56), replicating the Romanian *foarte mult* 'very much.'

Speaker: Marin

46. Sas **foarte** laćho 'It was really nice...'
47. Haj *suferiskerdŏm* **foarte** but 'And I suffered very much…'

48. Li woj sas **foarte** laćhi ‘She, also, was very nice…’
49. **Foarte** cumsecade ‘Very nice…’
50. *ïmpaciskerdŏm* man laça **foarte** śukar ‘We understood each other very well…’
51. **Foarte** phares ‘Very difficult…’

Speaker: Stănescu

52. Haj i wo paśolas **foarte** but paśä mande… ‘And he, also, closed very much to me…’
53. Fală man **foarte** nasul… ‘I was very sorry…’
54. Sas **foarte** *apropime* mandar… ‘He was very close to me…’
55. *Plaćas* man **foarte** but… ‘I liked very much…’
56. Haj *plaćala*’ man **foarte** but… ‘And I was enjoying very much…’

Speaker: Mimiş

57. Ama me *ïncelezisajlem* **foarte** bine lença ‘But I was very well with them…’
58. **Foarte** śukar semas la început ‘At the beginning I was very good…’

Speaker: Puncă

59. Haj **but** phrala hamas… ‘And we were lots of brothers…’
60. **But** zene hamas ‘We were lots of people…’
61. Hin **but** nùma’ ùngri hin ‘We were a lot, we were only Hungarians…’
62. Thaj k-ä **but** thana ‘And in a lot of places…’
63. Has dùme **but**… rromane ‘There were lot of words…Romani words…’
64. Apo’ has **but** źuvlă, save śukar źilabanas… ‘Well, there was lots of women that were singing nice…’

5. Conclusion

To conclude, the analysis of the corpus of life stories reveals the use of Romanian elements as typically encountered among bilinguals, that is, in a way where the components of the socially dominant language are integrated into the speech of the minority language. This is the case for specific realities around which the speakers interact in the socially dominant language, in this case Romanian, as used to be the case at school but also in domains such as health and other professional activities. Uses of Romanian material in Romani discourse are further noted for the so-called 'utterance modifiers,' that is, in the case where speakers may experience some cognitive pressure during the on-line organization of their discourse. Many uses at the level of utterance modifiers can be 'conventionalized,' that is, shared by the various members of the bilingual speech community, and as such constitute frequent elements that are activated in the bilingual's mind. In sum, the use of Romanian in the interviews of these Romani-Romanian bilinguals does not reflect any deficit in their capacity to fully express themselves in Romani. Rather, the replication of material from the socially dominant language, Romanian, reflects the experience of the Roma speakers in a society where they have been using Romanian in their everyday interactions with outsiders, throughout their lives, and where the other members of their community have been doing the same for several generations.

References

Adamou, Evangelia. 2010. Bilingual speech and language ecology in Greek Thrace: Romani and Pomak in contact with Turkish. *Language in Society* 39(2): 147–171.

Adamou, Evangelia. 2013. Replicating Spanish *estar* in Mexican Romani. *Linguistics* 51(6): 1075–1105.

Adamou, Evangelia. 2016. *A corpus-driven approach to language contact: Endangered languages in a comparative perspective*. Boston & Berlin: Mouton de Gruyter.

Backus, Ad. 2014. Towards a usage-based account of language change: implications of contact linguistic for linguistic theory. In Robert Nicolaï (ed.), *Questioning language contact: limits of contact, contact at its limits*, 91–118. Leiden: Brill.

Bakker, Peter. 1997. Athematic morphology in Romani: the borrowing of a borrowing pattern. In Yaron Matras, Peter Bakker, and Hristo Kyuchukov (eds.), *The Typology and Dialectology of Romani*, 1–26. Amsterdam: John Benjamins.

Boretzky, Norbert. 1989. Zum Interferenzverhalten des Romani. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikation Forschung* 42: 357–74.

Courthiade, Marcel. 1989. La langue romani (tsigane) : Évolution, standardisation, unification, réforme. In Istvân Fodor and Claude Hagège (eds.), *Hamburg Language Reform: History and Future*, vol. IV, 87–110. Hamburg: Helmut Buske.

Djurić, Rajko. 2005. La langue rromani standard. *Etudes Tsiganes* 22: 66–89.

Furtună, Adrian-Nicolae, Maria Luiza Medeleanu, and Daniel Samuel Petrilă (eds.). 2016. *Ŵivipnasqe rromane paramìćă; Poveşti în graiurile romilor; Life stories in different Romani dialects*. Popeşti Leordeni: Dykhta! Publishing House.

Gilliat-Smith, B. J. 1915. Report on the Gypsy Tribes of North East Bulgaria. *Journal of the Gypsy Lore Society* 2(9): 1–55, 65–109.

Hancock, Ian. 1988. The development of Romani linguistics. In Mohammed Ali Jazayery and Werner Winter (eds.), *Languages and Cultures*, 183–224. Berlin and New York: De Gruyter Mouton.

Matras, Yaron (ed.). 1995. *Romani in Contact: The History, Structure and Sociology of a Language*. Amsterdam: John Benjamins.

Matras, Yaron. 1998. Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36(2): 281–331.

Matras, Yaron. 2005. The Classification of Romani Dialects: A Geographic-Historical Perspective. In Barbara Schrammel and Dieter W. Halwachs (eds.), *General and Applied Romani linguistics*, 7–26. Munich: Lincom.

Matras, Yaron. 2009. *Language Contact*. Cambridge: Cambridge University Press.

Matras, Yaron. 2013. Mapping the Romani Dialects of Romania. *Romani Studies* 23 (2): 199–243.

Padure, Cristian, Stefano De Pascale, and Evangelia Adamou. 2018. Variation between the copula si "to be" and the l-clitics in Romani spoken in Mexico. *Romani Studies* 28: 263–292.

Poplack, Shana. 2018. *Borrowing: Loanwords in the Speech Community and in the Grammar*. Oxford: Oxford University Press.

Pott, August Friedrich. 1844. *Die Zigeuner In Europa Und Asien: Ethnographisch-linguistische Untersuchung, Vornehmlich Ihrer Herkunft Und Sprache*. Halle: Heynemann.

Rüdiger, Johann C. C. 1782 [1990]. Von der Sprache und Herkunft der Zigeuner aus Indien. In *Neuester Zuwachs der teutschen, fremden und allgemeinen Sprach-kunde in eigenen Aufsätzen, 1. Stück. Leipzig*, 37–84. Hamburg: Buske.

Sarău, Gheorghe. 2017. *Istoricul studierii limbii rromani în învățământul universitar din România (1992-2017)*. Bucureşti: Universitatea Bucureşti.

Sarău, Gheorghe. 2011. Fonetikane akcidèntură (rimosarimata, nakhavimata, paruvimata). In Veljko Kajtazi (eds.), *Zbornik Svjetski Dan Romskog Jezika - Proceedings of World Day of the Romani Language*, 45–48. Zagreb: Kali Sara.