

Variation and dialect levelling in the Romani dialect of Țândărei

DANIELE VIKTOR LEGGIO AND YARON MATRAS

The *Kangjlari* of Țândărei in southeastern Romania offer an interesting case study of the consolidation within just a few generations of a new Romani community, as a result of state-sponsored relocation and settlement between 1950 and 1980. We discuss the linguistic implications of the formation of this new community, drawing on language material from questionnaire elicitation and life history interviews among recent migrants now living in the UK, and supported by access to local ethnographic and archive material in the origin community. We show how a process of dialect levelling is underway in the Romani variety spoken by the Roma of Țândărei, which resembles cases of koineization discussed for a number of other languages in recent sociolinguistic literature. The stabilization of a particular combination of features means that the variety under discussion cannot be accommodated into current dialect classification models. This has implications for our general understanding of dialect formation in Romani. The paper also offers the very first modern, concise grammatical sketch of a Romani variety from Romania.

Keywords: Romanian Romani, Vlax Romani, koineization, dialect levelling, Țândărei

1. Introduction

Both linguistic and ethnographic approaches to Romani studies tend to postulate notions of ‘community’. In ethnography, a community can be defined variably as sharing location and socioeconomic resources (Stewart 1997), as a tight-knit group that is based around kinship (Jakoubek and Budilová 2006), or as a network of contacts based on kin, intermarriage, and ceremonial routines that can transcend location (Sutherland 1975). Linguistic analyses usually take such definitions of community for granted, adopting the label used by the population concerned, and attributing it to a ‘dialect’, which denotes a demar-

The research leading to the present publication results from MIGROM, ‘The immigration of Romanian Roma to Western Europe: Causes, effects and future engagement strategies’, a project funded by the European Union’s 7th Framework Programme under the call on ‘Dealing with diversity and cohesion: the case of the Roma in the European Union’ (GA319901).

Daniele Viktor Leggio is Lecturer in Linguistics at the School of Arts, Languages and Cultures, The University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom. E-mail: daniele.leggio@manchester.ac.uk

Yaron Matras is outgoing Editor of Romani Studies and Professor of Linguistics at the School of Arts, Languages and Cultures, The University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom. E-mail: yaron.matras@manchester.ac.uk

Romani Studies 5, Vol. 27, No. 2 (2017), 173–209

ISSN 1528–0748 (print) 1757–2274 (online)

doi: <https://doi.org/10.3828/rs.2017.10>

cated inventory of linguistic structures employed by that population. There is generally anticipation that 'dialects' are passed on from one generation to another and that variation will mirror a very gradual emergence and spread of idiolectal innovations or else influences from the dialects of neighbouring communities. Marushiakova and Popov (2004) introduce a new model of defining 'community' in which emphasis is placed on historical processes of segmentation, through which groups may drift apart through migration and the gradual loss of contact, and consolidation, whereby population groups that come into contact gradually acquire a shared sense of attachment, which becomes visible through a set of indicators pertaining to actual practices. In this paper we argue that language is one of those practices that is re-shaped through the consolidation of new communities. We offer a descriptive sketch of the Romani variety of Țândărei in southeastern Romania, placing an emphasis on variation and processes of dialect levelling and showing how the linguistic consolidation of different variants gives rise to a process of koineization.

Although the term *koiné* is well established in linguistics to refer to a common dialect, relatively few case studies exist that describe the processes of language change that lead to koineization. Siegel (1985) and Trudgill (1986) discuss koineization as a process of rapid linguistic change that is triggered when speakers of mutually intelligible varieties from different communities move together into a new location as a result of either voluntary or forced migration, and a new generation is born into that new community. Koineization is said to involve three stages (cf. also Kerswill 2002, Solheim 2009): The first is the 'Contact Phase' during which adult migrants retain their dialects. It is characterized by high inter- and some intra-individual variability. Rudimentary levelling may occur, but it tends to target structures that are less frequent. The second stage might be regarded as a 'Chaos Phase', where the first generation of speakers born into the new community lack a model for imitation in the form of a stable adult norm. This phase is characterized by considerable inter- and intra-individual variability, though extensive levelling takes place, with demographic correlates of features beginning to determine the shape of the new variety. Finally, in subsequent generations a so-called 'Focusing Phase' sets in as the new variety crystallizes and alternate realizations of structures are levelled out, with variants that are retained being re-allocated to serve distinct (socio) linguistic functions.

In his analysis of New Zealand English, Trudgill (2004) argues that the shape of new dialects can be predicted, as children born into the new community will adopt the forms that are most frequently used by adult speakers. This is based on the assumption that during colonization the speaker population was not separated by social boundaries and opportunities for social mobility were not linked to the use of structural features (cf. Kerswill 2010). Others,

however, have shown that in later periods, social factors, ideologies, and the role of standard languages shape the new varieties. Scholtmeijer (2000), for example, shows how in the Dutch polder of Haarlemmermeer and in the peat colonies in East-Drenthe and Groningen, established in the nineteenth century, koineization resulted in new dialects that resembled those spoken by the first settlers. By contrast, in the polders reclaimed during the twentieth century, Wieringermeer and Noordoostpolder, speakers converged on Standard Dutch and no new varieties emerged. Scholtmeijer argues that in the latter cases, the national language offered opportunities for social mobility and adult speakers therefore made an effort to discourage children from acquiring markedly regional features.

Similar trajectories have also been observed in England. In early cases of koineization, such as in the English Fens reclaimed during the seventeenth century (Britain 1997), the newly emerged variety became part of a dialectal continuum with its neighbouring predecessors. However, for Milton Keynes, built in the 1960s to relieve housing pressure in London, Trudgill (1986) and Kerswill and Williams (2000, 2005) show how a variety emerged that resembles Estuary English in combining features of London English with those of the standard or Received Pronunciation. Integrating a quantitative analysis of variant distribution with interviews with children and their main caregivers as well as observations on their interactions, Kerswill and Williams (2000, 2005) show how, during the early stages of community formation, when social divisions were relatively weak, adult speakers in Milton Keynes tended to converge on standard features and, since these became more frequent, children followed the model. The subsequent emergence of new social divisions led some speakers to diverge from standard forms to mark local and class-based identities. Similar processes are described by Hickey (2005) for the changes in Dublin English following the migration of rural populations to the city, and by Solheim (2009) for the speech of Høyanger, a Western Norwegian fishing village that became a town in the 1920s. Here, standardized Eastern Norwegian forms brought in by the transplanted industrial upper class were favoured in the initial stage, while hybrid forms and Western variants re-emerged in the speech of subsequent generations as markers of a local identity.

In the following we devote our attention to the case of the Romani variety spoken in the Romanian town of Țăndărei by a community that refers to itself as *Kangljari*, which emerged following the sedentarization of semi-nomadic groups of Roma between the 1950s and 1970s. In the absence of a standardized Romani language or even of any institutional use of Romani, there is no obvious target for convergence and so we witness a process of dialect levelling that is not guided by a clearly defined overt prestige norm. Moreover, since Romani speakers are bilingual and their livelihood depends largely on build-

ing relations with the majority non-Romani (in this case, ethnic Romanian) population, the choice of linguistic variants within Romani is not necessarily expected to correlate directly with economic or social mobility.

Current models of language change in Romani view dialect formation either as accompanying the break-up of links between groups, instigated through migration (the so-called dialect branching model), or as the outcome of the gradual spread of innovations from various epicentres across neighbouring communities (the so-called geographical diffusion model). Pairing linguistic with ethnographic observations, we show how the history of the community has shaped the formation of the *Kangljari* Romani dialect in a different way, bringing together different dialectal variants and triggering a process of dialect levelling.

2. Branching, diffusion and interdialectal contact in Romani

Studies in the dialectology of Romani tend to use a region-based reference grid that recognizes several dialect groups: Balkan (with two sub-divisions), Vlax (North and South), Central (North and South), Northwestern, and Northeastern, with the historical Romani dialects of the Iberian Peninsula and those of southern Italy often considered to be outliers, while the dialect of Istria is usually acknowledged as being at the crossroads (cf. Bakker and Matras 1997, Matras 2002, Boretzky and Igla 2004, Elšík and Matras 2006). Matras (2002, 2005, 2010) offers an account of the historical emergence of principal isoglosses in Romani that takes as its point of departure a reconstruction of Early Romani forms (cf. also Elšík and Matras 2006) and the postulation of subsequent innovations, emerging in a number of key epicentres and gradually diffusing across geographical space. Examples are the spread of jotation (initial segment *j-*) in northern Europe in words like the 3SG pronoun *jov* ‘he’, the generalization of inflectional forms in *-h-* (e.g. *laha* ‘with her’) in the dialects of central Europe (and their Scandinavian offspring), and the palatalization of historical *t-* in the word *cikno* ‘small’ and loss of the final segment in the abstract nominalizer *-ipen* > *-ipe*, with their epicentre in southeastern Europe. The changes are assumed to have emerged in situ, following the migration of Romani speakers from the Balkans and their settlement across central, western and northern Europe from the fifteenth century onwards. This view is supported by the existence across the Romani dialect landscape of a conservative periphery, with conservative forms such as *adava* ‘this’ and *dives* ‘day’ appearing in all geographical fringe regions of Europe, from Wales, to Spain, southern Italy, and Greece, while derived forms such as *dava* and *ada* ‘this’, or *djes*, *zis* and *di* ‘day’, are each confined to smaller regions. The geographical spread of many of these innovations is already captured in the earliest written attestations that

offer insights into dialect differences across Europe, from the late seventeenth century onwards. A further clue to the timeline is provided by the dense cluster of isoglosses referred to by Matras (2005) as the Great Divide, which follows what was the border region of the Habsburg and Ottoman Empires during the sixteenth and seventeenth centuries and is assumed to have constituted a barrier to movement and contact and thus a barrier to the spread of dialectal innovations in both directions.

For Matras, the nomenclature of dialect groups is thus merely a reference grid of convenience. Boretzky (2007), however, speculates that each of the present-day dialect groupings implied by that nomenclature emerged as an independent dialect 'branch' outside of Europe, in Anatolia, and that population groups carrying their own coherent dialects migrated and settled in Europe, resulting in the present-day distribution of dialect forms. The model seems to draw inspiration from the realization that Romani as a whole was carried into Europe by a migrant population originating ultimately in India, a theory first formulated by Johann Rüdiger in 1782 on the basis of the very first grammatical sketch of a Romani dialect (see Matras 1999). It also draws on a general Neo-Grammarians understanding of language diversification as a process of branching, and possibly also on a view of Roma as perpetual nomads. However, it seems to lack any prospect of verification, not least due to the absence of any legacy of Romani in present-day Anatolia. The nomenclature has guided Boretzky in a series of works (e.g. Boretzky 2000, 2003, and others) where the method has been to delimit the group under discussion through a pre-selection of sources, and then to take an inventory of the features found within each corpus of sources. This enumeration of features is usually carried out using maps representing the territorial spread of structures, pointedly limiting the examination of their spread to the pre-selected group. The results in fact tend to confirm a geographical spread of features, as captured in the atlas components of works such as Boretzky (2003) and Boretzky and Iglă (2004). Nonetheless, the notion of dialect 'branches' rests on the view that there is a 'prototype' of basic features or shared innovations (and retentions) that inherently define affiliation to each 'branch'. From this perspective, a dialect that fails to adhere to the prototype would defy classification.

Although relatively little consideration has been given in research in Romani linguistics to variation within dialects, it is well established that intermarriage and other contact between groups can trigger processes of structural convergence. Matras (1994) discusses a variety of Vlax Romani which he labels, based on the alternating labels used by the speakers, as *Kelderaša/Lovara*. The speakers are members of a single family network, yet their speech patterns show inter-personal variation, partly correlating with marriage patterns as well as self-labelling preferences, but also with generation. Boretzky (1995) attributes

admixtures of forms in the Romani dialects of southeastern Europe to inter-dialectal ‘interference’. The target for convergence is invariably, according to Boretzky, the dialect of the more populous group, though all cases identified involved dialects that were essentially similar and closely related and resulted in no major structural changes. Boretzky (1995: 90) notes that there has been “less interference than might be expected [. . .] with single dialects having remained astonishingly homogeneous.” Elšík (2003) takes a broader approach in his discussion of interdialectal contact in the Czech Republic and Slovakia, noting how dialects that are said to belong to three distinct ‘branches’, Central, Vlach and Sinti (Northwestern), share a number of areal features, while more extensive dialect mixing can be observed among the more closely related Central varieties, which are spoken by the majority of the Romani population in the region. Elšík proposes that perceived boundaries stemming from different migration histories inhibit contacts across dialects ‘branches’.

Friedman (2017) also addresses perceived boundaries among pre-existing groups in his discussion of the Romani population of Skopje, Macedonia. The historically sedentary *Kovač* ‘blacksmiths’ appear at the top of a social hierarchy, followed by seven different groups of *Arli* or ‘settled’ Roma who are sub-divided based on their origin in other rural or urban communities, while the formerly peripatetic *Džambaz* ‘horse dealers’ appear at the lower end. Historically, the groups tended to reside in different parts of the town and did not intermarry, and so contacts remained limited. Inter-group contacts increased, however, as a result of urbanization after World War II, relocation following the earthquake of 1963, and the socialist state’s interventions in an effort to break down social barriers. While the dialects of the groups at the two extreme ends of the social hierarchy remained distinct, among the *Arli* groups Friedman observes instances of convergence. The dialect of the group that is most established in the town and regarded as the most prestigious shows a number of particular innovations, including replacement of the perfective marker *-g-* and *-d-*, replacement of the 1PL.PAST marker *-am* with *-em*, and loss of *-s* in word final position. The extent to which other groups adopt these features correlates with the length of their settlement in Skopje; the groups that are the most recent arrivals and who maintain an explicitly rural identity tend not to display any of the features of the dominant group.

In all these examples, preexisting group boundaries can be taken for granted at the point of inter-dialect contact and are argued to contain cross-dialect convergence. The case we present below differs in that, as far as we are able to ascertain based on life history interviews and archive research, the community formed through the coming together of individual families, subsequent intermarriage with settled and semi-nomadic Roma communities from other regions, as well as adoption of children of both Romani and non-Romani

parents from other locations. As far as we are able to tell, group boundaries and prestige relations among sub-groups have been permeable. The case of the *Kangljari* of Țăndărei thus offers an opportunity to examine a process of dialect levelling in Romani that does not seem to be constrained by rigid pre-existing group formations.

3. The Roma of Țăndărei¹

Țăndărei, a small rural town in southeastern Romania, was established in 1968 through the administrative merger of three villages – Țăndărei, Țăndărei Gară and Strachina. The town has two distinct populations of Roma. The first are descendants of *robi* ‘slaves’ who until the late nineteenth century were the property of a local landlord. They are now known by the name *țigani românizați* ‘Romanianized Gypsies’, do not speak Romani, and often marry ethnic Romanians. The larger group are Romani speakers and are known by the Romani name *Kangljari* ‘comb-makers’ or its Romanian translation *Pieptănari* and who descend from semi-nomadic populations.

The first semi-nomadic Roma settled in the Țăndărei area in the 1920s,



Figure 1: Location of Țăndărei

replacing local residents who were granted land as part of the land reforms that followed World War I, and who were therefore relocated. The Roma were recruited to provide a workforce for the then flourishing pottery industry and were settled in Strachina, close to a clay soil area. Their traditional occupation had been the production and sale of bone combs. Lists compiled by the local authorities at the onset of World War II as part of plans to deport semi-nomadic Roma to labour camps in Transnistria indicate that there were around

1. For the present section we draw on collaborative work as part of the MigRom consortium that investigated the causes of Roma migrations from Romania and their effect on the communities of origin. The research combined ethnography with archive research and statistical surveys (see MigRom 2015, Matras and Leggio 2017, Toma *et al.* 2017).

sixty Roma in Strachina at the time, and it appears that most were deported. According to one of our informants, two Roma families who had a fair complexion and were not identified by the authorities as Roma remained in the area and went into hiding. A member of one of the families was known as a *žudikator* 'arbiter', a man of authority who was called upon by Roma families to help resolve conflicts. His reputation and the fact that he survived the war undetected became a pull-factor for other Roma after the war, and Strachina became known as the *țișania* or 'Gypsy quarter' of Țândărei. While we were not able to verify this specific account, archive documents do indicate that more Roma settled in the area after the war, prompted by the restrictions on movement imposed by the communist state as well as by efforts to populate the region and develop collective agriculture and later industry. Records of the Țândărei village council indicate that by 1966, a total of 197 semi-nomadic Roma families comprising altogether 903 persons lived in Strachina; around 5 per cent worked for the state-run Agricultural Cooperative.

Țândărei was declared a town in 1968 and witnessed intensive industrialization with the establishment of a brick factory and of oil and sugar refineries in Strachina. Its population grew through the constant demand for unskilled workforce; the brick factory and the Agricultural Cooperative both provided accommodation for their employees. Testimonies indicate that many newcomers joined relatives who were already living in Strachina. Roma marriage customs usually dictate that brides relocate to join the groom's family, but it was not uncommon for grooms from other areas to relocate to Strachina as they were able to find employment there. Children were brought to Strachina as part of the *Kangljari* practice of informal adoption from both Roma and Romanian families. *Kangljari* couples who did not have male offspring often adopted boys from poorer families. Girls were also taken in, since the marriage custom of paying a bride price allowed families to benefit financially from bringing up girls. Since non-*Kangljari* girls commanded a lower bride price, adopting girls into the community expanded the opportunities for families who were less wealthy to find brides for their sons. Children who had been adopted were integrated into the community and acquired the Romani language. We met several individuals who told us that they had learned the language from their adoptive parents and siblings.

Under the communist state, these informal adoptions were referred to in local police files as cases of 'child kidnapping'. The *Kangljari* were generally represented in official documents as a savage, untamed and anti-social population and Strachina was often associated with begging and crime and was surrounded by a barrier that restricted its residents' access to the rest of the town (Toma *et al.* 2017: 82). It is probably because of this separation that ethnic Romanians were, and still are, largely unaware of the most common self-appel-

lation among the Roma of Strachina, *Kangljari*. Instead, the ethnic Romanian population refers to them as *Ursari* 'bear-tamers', a term that is widespread among other Roma communities in southern Romania. Some *Kangljari* Roma have adopted this label and use it alongside *Kangljari* to refer to their own population. The community's segregation also limited opportunities for social mobility, and economic divisions within the *Kangljari* community only emerged as a consequence of post-communist labour migrations to other countries and the investment of earnings and remittances back in Țăndărei. The most successful *Kangljari* migrants have abandoned Strachina and moved to newly built houses in the ethnically mixed areas of Țăndărei.

4. Data and method

Although Romania has the largest Romani speaking population in Europe, very few studies have been devoted to the Romani dialects of Romania. In fact, all comprehensive descriptions of Romani dialects belonging to the Vlax group, widely considered to be the dominant dialect group within Romanian Romani, are based on so-called 'diaspora' varieties that are spoken by populations that left the Romanian territories in the nineteenth century or even earlier (e.g. Gjerdman and Ljungberg 1963, Pobožniak 1964, Matras 1994, Boretzky 1994, Hancock 1995, Iglă 1996). To date, documentation of Romani from Romania is limited to collections of folklore (e.g. Constantinescu 1878, Gaster 1931, Drimba 1992) and otherwise to targeted linguistic elicitation as part of the Romani Morpho-Syntax (RMS) Database,² and there is no comprehensive description of any Romani variety from the country. Boretzky and Iglă's (2004) dialect atlas of Romani takes into consideration data from altogether five sources from Romania, all of which are collections of oral narratives. Matras (2013) discusses major isoglosses among the Romani dialects of Romania based on RMS questionnaire elicitation carried out in some forty locations. The results show a north-south divide as well as innovation zones in the trans-Carpathian areas, which are largely contained by topographical and historical political boundaries. In a more localized study with a focus on Transylvania, van den Heuvel and Urech (2014) show that the complexity of dialect variation stems in part from a series of migration waves across regions within Romania, while largely confirming Matras's (2013) observations about a split between the northern counties of Transylvania and those of the south and neighbouring Banat and an innovation zone with its centre in Mureș county.

Our study draws on elicitation using the RMS method, employed in a series of studies as the basis for descriptive sketches of Romani dialects including

2. <http://romani.humanities.manchester.ac.uk/rms/>

TABLE 1. Overview of informants

| Informant | Gender | Age at interview | Generation in Țândărei | |
|-----------|--------|------------------|------------------------|------------------|
| MG1959 | Female | 55 | 1st | Family network 1 |
| CG1962 | Male | 52 | 2nd | |
| AS1986 | Male | 27 | 3rd | |
| VT1975 | Male | 39 | 2nd | Family network 2 |
| IM1982 | Male | 32 | 2nd | |
| NT1992 | Male | 22 | 3rd | |
| MrT1995 | Male | 18 | 3rd | |
| MdT1995 | Male | 18 | 3rd | |
| SF1995 | Male | 18 | 3rd | |
| DK1980 | Female | 33 | 3rd | |
| CB1991 | Female | 22 | 3rd | |

Matras (2004) on the Romani dialect of Parakalamos in northwestern Greece, Tenser (2006) on Lithuanian Romani, and Leggio (2011) on Mitrovica Romani. The samples we use, RO-064 and RO-066, were collected in 2009 with, respectively, a male informant in his 30s and a male informant in his late teens. We also draw on open-ended interviews carried out as part of the MigRom project, which between 2013 and 2017 investigated the experiences, motivations, and ambitions of Roma who migrated from Romania to various Western European countries (see Matras and Leggio 2017). The research team in Manchester included the authors and two fieldwork assistants who were recruited among the local community of Roma migrants from Țândărei. The authors and fieldwork assistants had been collaborating since 2009 as part of a series of research and training activities in the local community of Romanian Roma migrants. The team collected life histories in Romani, with the research assistants taking the lead in conducting the interviews and the authors intervening from time to time to tease out more information from the informants. We take into consideration eleven such interviews (see Table 1).

Most of the informants belonged to two of the larger family networks residing in Manchester. Husband CG1962 and wife MG1959 and their nephew AS1986 will be referred to as Family Network 1 (FN1). VT1975, his son MdT1995, his brother IM1982 and his nephews NT1992, MrT1995 and SF1995 will be referred to as Family Network 2 (FN2). DK1980 and CB1991 were not related to these extended families or among themselves, but entertained friendly relationships with all other informants. All the informants were members of the same Pentecostal congregation established in Manchester by the oldest son of CG1962 and MG1959.

The interviews produced a total of 9.23 hours of recordings. They focused on the migratory history of the informants, their situation in Romania, the destination countries before moving to Manchester and, in migration, experiences of housing, employment, school attendance and access to services such as health care. The length of individual interviews ranges from 15 minutes to more than 3 hours. To obtain a balanced sample for the present study we therefore selected around 20 minutes from each of the longer interviews. This resulted in a corpus of 2.30 hours that yielded a total of 13,273 words, with an average of up to around 1,500 words per speaker. The sample is biased for gender and age, reflecting access and the availability of informants. Our goal is to draw on the observations to examine the extent to which structural variation along with the available historical and ethnographic information offer insights into a process of koineization that is in progress in the Romani variety spoken by the *Kangjlari* of Țăndărei.

5. Phonology

The dialect shows the five basic vowels that are common in Romani dialects, /i e a o u/. The vowels /ə/ and /i/ are found in Romanian loanwords such as *kərbun* ‘coal’ Rom. *cărbune*, *məritisal-* ‘to marry’ Rom. *a se mărita*, the indicative complementizer *kə* Rom. *că* and *kurind* ‘soon’ Rom. *curând*. Among all informants /ə/ is found within the inherited lexicon in word-final, unstressed position (*andə* ‘in-DET.OBL’, *khatə* ‘here’), often alternating with /e/ (*berš ~ bərš* ‘year’, *ək ~ ək* ‘one’, subjunctive complementizer *te ~ tə*) and occasionally with /i/ (*inkl ~ ənkl* ‘go out’, *beršind ~ bəršind ~ bəršənd* ‘rain’).

The inventory of consonants similarly mirrors that common in Romani, with stops /b d g p t k ph th kh/, affricates /c č ċ dž/, sibilants /s z š/, fricatives /f v x/, nasals /m n/ and liquids /l r/. Both semivowels /j w/ are also present, with /w/ occurring as an alternative realisation of /v/ in intervocalic and word-final position (see below).

Vowel assimilation is systematic among all our informants in *šoro* < **šero* ‘head’ and *soro* < **savoro* ‘all’. Fronting of the vowel in the diphthong *-aj* > *-ej* (in *dej* ‘mother’, *čej* ‘daughter’) is also ubiquitous. Vowel fronting also takes place in past tense person concord markers as a result of historical jotation at morpheme boundaries. However, while all informants show fronting in the 1SG.PAST marker (*-*jom* > *-em*) we find variation in past tense verbs and copula for the 2SG (*-jan* ~ *-ean*) and 1PL (*-jam* ~ *-eam*):

- (1) a. CG1962 *pučhel tut* “so relidža *šjan*?” “pentekostalo”, “pe tu *sean*?”
He asks you “what faith are you?” “Pentecostal”, “and you are?”

- b. IM1982 *bešljam, deam te thoas ko semafori*
We stayed, we started to clean [windshields] at traffic lights.

Alveolar plosives in selected lexemes show light palatalization across all speakers (*pakja* < **patja* ‘believe’; *bukja* < **butja* ‘work.PL’; *gjes* < **dives* ‘day’; *kiro* ~ *kjo* < **tiro* ~ **to* ‘yours’). All speakers show affrication to *cikno* < **tikno* ‘small’:

- (2) a. MG1959 *o Jesus te binikuvintil kjo dad, kje dadoroske phraloren*
May Jesus bless your father, your father’s brothers.
b. IM1982 *neas amen bukja, amare dada munčinas kəl manuš gjesenca*
We didn’t have jobs, our fathers worked daily for that man.
c. MRT1995 *pakjau kə me sem singuro rom kaj si but prjeteni ingleši*
I believe I’m the only Rom who has many English friends.
d. mdT1995 *ni kerdem ni ekh munka pentru kə semas ciknoro*
I didn’t do any work because I was young

Our informants are also consistent in the retention of /s/ in pre-consonantal position (e.g. *leske* 3SG.DAT, *ivendesko* ‘winter.GEN.M’). In word-final position, /s/ is retained in word stems (*gjes* ‘day’; *pes* REFL) but shows variation in inflectional items: Interview data show that, for the remoteness marker all speakers strongly favour *-as*, while they favour *-a* for the 3SG.PAST marker (Table 2).

TABLE 2. Final /s/ in inflectional items, variation in interviews

| | | % | N |
|-------------------|------------|-------|-----|
| Remoteness marker | <i>-as</i> | 91.96 | 698 |
| | <i>-a</i> | 8.04 | 61 |
| 3SG.PAST | <i>-as</i> | 26.02 | 32 |
| | <i>-a</i> | 73.98 | 91 |

- (3) a. VT1975 *i me džawas, familja, cinoro sima me*
I too went, [with the] family, I was young.
b. CG1962 *ni phendja ni kə si rom, phejo!*
He never said that he was Rom, sister!
c. CG1962 *ou bistardjas ekh autorizacija*
He forgot a permit

Truncation of /a-/ in words such as *akana* ‘now’, *av-* ‘to come’ and *andre* ‘in’ is unattested in the Țândărei dialect. At the same time, /a-/ prothesis is systematic across all speakers (*anav* < **nav* ‘name’, *abjav* < **bjav* ‘wedding’, *ašun-* < **šun-* ‘to hear’). Prothetic /v-/ is rarely found, but where it occurs it alternates with /u-/ (‘hand’ *vast* in RO-064, *uast* in RO-066):

- (4) a. RO-066 *uasde e čhaves*
Lift the child!
- b. RO-066 *muro dad vasdja le čhaes ande urdon*
My father lifted the child on the wagon

Among all speakers, /v/ also shows variation with /u/ in word-final position: *gav* ~ *gau* ‘village’, *anav* ~ *anau* ‘name’ (see also (2c) for 1SG.PRES -*av* ~ -*au* and (3c) for 3SG.M *ov* ~ *ou*). When followed by voiced stops /k t/, word-final /v/ is often assimilated to /p/: *phenap tuke* ‘I will tell you’, *kadap kher* ‘this house’. Similar variation is also found in intervocalic position, as shown in (3)a and 4, with /v/ ~ /w/ ~ /Ø/.

Consonantal de-affrication in words such as *džukel* > *žukel* ‘dog’ and *čhavo* > *šavo* ‘son’ is rarely attested in our sample, although most speakers showed some degree of variation (Table 3). Two speakers in particular showed a markedly distinct behaviour. CB1991 used de-affricated forms exclusively. During the interview she explained that she was born in Buzău, about 100 kilometres to the northwest of Țăndărei, and defined herself as *Ardžintari*. Asked about her dialect, she explained that her relatives do not speak like the Roma from Țăndărei and that she had learnt the speech of Țăndărei after marrying. Similarly SF1995, although born in Țăndărei and defining himself as *Kangljari*, reported that his mother is a “true Roma woman” from Țăndărei, but that his father is “Roma but also Russian” (suggesting an origin in the Russian-speaking regions on

TABLE 3. Consonantal de-affrication, variation in interviews

| | | Family Network 1 | | | | | | | | | | | |
|----------------------|-----|------------------|-----|--------|----|--------|----|--------|---|--------|---|--------|----|
| | | Sample total | | MG1959 | | CG1962 | | AS1986 | | DK1980 | | CB1991 | |
| | | % | N | % | N | % | N | % | N | % | N | % | N |
| <i>čh</i> > <i>š</i> | yes | 16.25 | 26 | 5.88 | 1 | 23.53 | 4 | 0.0 | 0 | 0.00 | 0 | 100 | 11 |
| | no | 83.75 | 134 | 94.12 | 16 | 76.47 | 13 | 100 | 9 | 100 | 6 | 0.0 | 0 |
| <i>dž</i> > <i>ž</i> | yes | 13.75 | 44 | 9.09 | 5 | 2.86 | 1 | 0.0 | 0 | 0.00 | 0 | 100 | 9 |
| | no | 86.25 | 276 | 90.91 | 50 | 97.14 | 34 | 100 | 8 | 100 | 7 | 0.0 | 0 |

| | | Family Network 2 | | | | | | | | | | | |
|----------------------|-----|------------------|----|--------|----|--------|----|---------|----|---------|----|--------|----|
| | | VT1975 | | IM1982 | | NT1992 | | MrT1995 | | MdT1995 | | SF1995 | |
| | | % | N | % | N | % | N | % | N | % | N | % | N |
| <i>čh</i> > <i>š</i> | yes | 16.67 | 1 | 6.67 | 1 | 0.00 | 0 | 0.00 | 0 | 10.53 | 2 | 31.58 | 6 |
| | no | 83.33 | 5 | 93.33 | 14 | 100 | 22 | 100 | 19 | 89.47 | 17 | 68.42 | 13 |
| <i>dž</i> > <i>ž</i> | yes | 3.33 | 1 | 2.50 | 1 | 22.22 | 8 | 0.00 | 0 | 0.00 | 0 | 51.35 | 19 |
| | no | 96.67 | 29 | 97.50 | 39 | 77.78 | 28 | 100 | 47 | 100 | 16 | 48.65 | 18 |

the Romanian–Moldovan border). He usually showed *čh* but used *š* more frequently than the other speakers. His deviation from the general pattern was even more pronounced for *dž* > *ž* as he used the two variants almost at the same rate.

Across all speakers, the Early Romani cluster **nř* is systematically retained as /rn/ in *arno* ‘egg’, *pirno* ‘foot’, and *marno* ‘bread’.

6. Nominal forms

6.1. Abstract nominalizer

The Greek-derived abstract nominalizer *-imos* was used systematically with all roots, except by MG1959, who in a single case employed pre-European *-ipe*. VT1975 and IM1982 employed pre-European *-ibe* with *xa-* ‘to eat’. Greek-derived *-imata* (PL) and *-imasa* (OBL) were employed with all roots, including *xa-*.

- (5) a. MG1959 *phenau o čáčimos, mangawas mangipe*
I’m telling the truth, I used to ask for handouts.
- b. IM1982 *pokinasas amari kirja, xabe*
We used to pay for our rent, food.
- c. RO-064 *ašundem le nevimata kathar le čhave*
I heard the news from the boys.
- d. MD1995 *te ažutil le familjan, te inkjarel amen xamasa*
To help the families, to bring us food.

6.2. Noun inflection

Layer I markers for pre-European nouns match the inflection classes reconstructed by Elšik (2000a) for Early Romani (Table 4). With European nouns, *u*-masculine nouns have been assimilated into the *o*-masculine class. The Romanian-derived NOM.M.PL marker *-ur* ~ *-uri* ~ *-urja* (this variation is common across all speakers) has replaced *e* both in the European *o*- and *i*-classes. With some borrowed nouns such as *prjeteno* ‘friend.M’ and *mašina* ‘car.F’, the Romanian markers *-i* (M.PL) and *-e* (F.PL) are also used.

Layer II case markers are, as usual, affixed to the oblique (Table 5). We found no instances of /s/ > /h/ in intervocalic (instrumental *-sa*) or pre-consonantal (dative *-es-ke*) positions.

6.3. Demonstratives, deictics and definite articles

The system of demonstratives shows a preference for the stem in *k_d-*, with reduced forms in *ko-/ka-* occurring marginally (Table 6). We find the conservative inflection M. *-va*, F. *-ja*, though reduced inflection forms are also common, thus M. *kadava* alongside *kada*, F. *kadaja* alongside *kadja/kaja*. In

TABLE 4. Layer I inflection markers

| Pre-European | Example | Nominative | | Oblique | |
|--------------|-----------------------|------------|--------------------|---------|-------|
| | | Sg. | Pl. | Sg. | Pl. |
| Ø-masculine | <i>kher</i> 'house' | Ø | -a | -es- | -en- |
| o-masculine | <i>čhavo</i> 'boy' | -o | -e | -es- | -en- |
| i-masculine | <i>pani</i> 'water' | paj | N/A | paes | N/A |
| Vj-masculine | <i>rašaj</i> 'priest' | -Vj | -a | -es- | -en- |
| Ø-feminine | <i>džuv</i> 'louse' | Ø | -(j)a | -ja- | -jan- |
| i-feminine | <i>piri</i> 'pot' | -i | -ja | -ja- | -jan- |
| Vj-feminine | <i>dej</i> 'mother' | -Vj | -ja | -ja- | -jan- |
| European | Example | Sg. | Pl. | Sg. | Pl. |
| o-masculine | <i>foro</i> 'town' | -o | -ur/-uri/-urja, -i | -os- | -en- |
| i-masculine | <i>autobuzi</i> 'bus' | -i | ur/uri/urja | N/A | N/A |
| a-feminine | <i>mašina</i> 'car' | -a | -(j)a, -e | -a- | N/A |

positions adjoining voiceless stops, the inflectional ending *-v(a)* may assimilate to *-p* (*kadap kher* 'this house'). The usual vowel distinction appears in the demonstrative stems, separating reference to visible or perceivable entities (*-a-*) from that to verbally expressed, discourse-based entities (*-o-*). There are a few instances of forms denoting specificity, which draw on an initial vowel: *akadja trebja* 'this [specific] work'. In this way, we have some evidence for the overall retention of a four-term opposition system (cf. Matras 2002: 103–12).

TABLE 5. Layer II markers

| | Singular | | Plural | |
|--------------|-------------|---|-------------|--|
| Dative | <i>-ke</i> | M: <i>čhaveske</i> F: <i>čhejaske</i> | <i>-ge</i> | M: <i>čhavenge</i> F: <i>čhajange</i> |
| Genitive | <i>-k</i> | M: <i>čhavesk</i> F: <i>čhejak</i> | <i>-g</i> | M: <i>čhaveng</i> F: <i>čhejang</i> |
| Ablative | <i>-tar</i> | M: <i>čhavestar</i> F: <i>čhejatar</i> | <i>-dar</i> | M: <i>čhavendar</i> F: <i>čhejandar</i> |
| Instrumental | <i>-sa</i> | M: <i>čhavesa</i> F: <i>čhejasa</i> | <i>-ca</i> | M: <i>čhavenca</i> F: <i>čhejanca</i> |
| Locative | <i>-te</i> | M: <i>čhaveste</i> F: <i>čhejate</i> | <i>-de</i> | M: <i>čhavende</i> F: <i>čhejande</i> |

- (6) a. mDT1995 *kadap kher nea amaro, kadava si gadžako kaj bešas pe rente*
 This [visible] house isn't ours, it (this [visible]) belongs to
 the woman and we are renting it.

- b. RO-066 *primisardem kala lulugja katar mure phej*
I got these [visible] flowers from my sister.
- c. RO-064 *tehara va terminiu akadja trjeba*
Tomorrow I will finish this [perceivable, specific] work .
- d. RO-064 *prindžanav le lenge deja akadale džuvlenge so le štarengje*
I know the mother of these [visible, specific] four women.
- e. MG1959 *sas po timpo kodova, phejo, le Ceauceskosko timpo*
It was during that [aforementioned] time, sister,
Ceausescu's time.
- f. CG1962 *murro dad sas nakhado kal gadžo. Haj kodolaski ni mudarde murre dades*
My dad moved to the Gajo. And for that [aforementioned] reason they didn't kill my dad.
- g. RO-064 *haj kodo manuš džanelas sar te lačharel lə instrumente*
And that [aforementioned] man knew how to repair instruments.

TABLE 6. Demonstrative stems,
variation in interviews

| | % | N |
|------------------|-------|-----|
| <i>kad-/kod-</i> | 71.76 | 155 |
| <i>ka-/ko-</i> | 28.24 | 61 |

A curious feature is the occasional replication of the discourse-based or anaphoric demonstrative stem vowel in *-o-* in the inflectional ending of the masculine demonstrative form *kod-o* alongside *kodo-va/ko-va*. As we have not come across a counterpart form in the set denoting visible/perceivable entities (thus no **kad-o*) nor a counterpart feminine form (thus no **kod-i* or **kad-i*) we attribute the form *kodo* to a progressive vowel assimilation, possibly motivated by the regressive assimilation that we find in *šoro* < **šero* 'head' and incipiently also in the 1SG.POSS *moro* < *muro*.

Alongside the demonstratives we find the deictic expressions *kathe* 'here' and *kothe* 'there' (alongside single tokens of *orde* 'there') and the comparison deixis *adiki* and *kadiki* 'such, in this way'.

Examples (4b) and (5d) show how, like most dialects of the northern parts of Romania, Țăndărei Romani shows retention of *l-* in the definite article, deriving from the article's origin in the Early Romani demonstrative **ola* (cf. Matras 2002: 96ff.) (see Table 7). In the oblique, *le* is preferred by all our informants (89.45%, 178 tokens in interviews), although it occasionally alternates with *e* (10.55%, 21 tokens, see example 4a). Forms with *-l-* were also encountered

TABLE 7. Definite article inflection

| | NOM | OBL |
|------|--------------|--------------|
| M.SG | <i>o, əl</i> | <i>e, le</i> |
| F.SG | <i>i, e</i> | <i>e, le</i> |
| PL | <i>e, əl</i> | <i>e, le</i> |

in the NOM. M.SG and NOM.PL, with *əl* alternating with *o* and *e* respectively. Occasionally, *le* was employed with NOM.PL (see 5c):

- (7) a. MG1959 *anas ma o timpo te maj. . . Phendem o căcimos!*
There came the time for me to. . . I said the truth.
- b. CB1991 *haj ənke əl šao din prima kasaturje murre phralesko, dešupanž*
And with the son from my brother's first marriage, [it is]
fifteen.
- c. NT1992 *haj te aven sar si sa e lumja*
And that they might become like the other people.
- d. AS1986 *akana maj skimbojsajle əl vremuri, maj skimbojsajle əl timpuri*
Now the times have changed a lot, the times have changed a lot.

6.4. Personal, reflexive and possessive pronouns

While in both RMS samples the conservative 3rd person nominatives (*ov/oj/on*) were used exclusively, we identified instances of *v*-prothesis in the interview data (Table 8). Conservative forms were strongly favoured (86.49%, 64 tokens), but three speakers (CG1962, AS1986 and SF1995) also employed transitional forms with *w*- (9.46%, 7 tokens) and forms in *v* (4.05%, 3 tokens, all produced by SF1995).

TABLE 8. Personal and reflexive pronouns

| | NOM | OBL | Short OBL | DAT | INST | POSS |
|------|-------------|---------------|-------------|----------------|----------------|-------------------|
| 1SG | <i>me</i> | <i>man-</i> | <i>ma</i> | <i>mange</i> | <i>manca</i> | <i>mVr(r)-/m</i> |
| 2SG | <i>tu</i> | <i>tut-</i> | <i>tu</i> | <i>tuke</i> | <i>tusa</i> | <i>kir-/k(j)-</i> |
| | M: (w)ov | M: les- | | M: leske | M: lesa | M: lesk- |
| 3SG | F: (w)oj | F: la- | M: le | F: lake | F: lasa | F: lak- |
| 1PL | <i>ame</i> | <i>amen-</i> | <i>ame</i> | <i>amenge</i> | <i>amenca</i> | <i>amar-</i> |
| 2PL | <i>tume</i> | <i>tumen-</i> | <i>tume</i> | <i>tumenge</i> | <i>tumenca</i> | <i>tumar-</i> |
| 3PL | (w)on | <i>len-</i> | | <i>lenge</i> | <i>lenca</i> | <i>leng-</i> |
| REFL | | <i>pes-</i> | <i>pe</i> | <i>peske</i> | | <i>pesk-/p-</i> |

TABLE 9. 1SG.POSS, variation in interviews

| | M (NOM) | | F (NOM) | | OBL | |
|---------------|---------|----|---------|----|--------|----|
| | % | N | % | N | % | N |
| <i>mur(r)</i> | 92.78 | 90 | 64.62 | 42 | 53.26 | 49 |
| <i>mər(r)</i> | 0.00 | 0 | 30.77 | 20 | 46.74 | 43 |
| <i>mir(r)</i> | 0.00 | 0 | 4.62 | 3 | 0.00 | 0 |
| <i>mor(r)</i> | 7.22 | 7 | 0.00 | 0 | 0.00 | 0 |
| Total | 100.00 | 97 | 100.00 | 65 | 100.00 | 92 |

First- and second-person possessives are formed by attaching a possessive suffix to the pronoun base form (Table 8 and 9). The Early Romani distinction between the 1SG marker *-*nr̥-* and the 2SG *-*r-* is lost in Țândărei Romani, with 1SG.POSS also taking *-r-* and showing variable lengthening of the trill. The vowel in 1SG.POSS is /u/ and in interview data we have noticed instances of vowel assimilation to the adjectival inflection, or introflection: M *mur(r)-o* > *mor(r)-o*, F: *mur(r)-i* > *mər(r)-i* ~ *mir(r)-i*, OBL *mur(r)-ə* > *mər(r)-ə*. Introflection appears to follow the hierarchy: OBL > F(NOM) > M(NOM).

For 1SG, 2SG and reflexive possessives we also find reduced forms *m-*, *k(j)-* and *p-* respectively:

- (8) a. RO-066 *kadaja si muro kher*
This is my house.
- b. CG1962 *morro dad sas ekh rom parno, blondo*
My father was a white Rom, blonde.
- c. Mdt1995 *bešav məre familijasa, məre dadesa, məre dasa*
I live with my family, with my dad, with my mom.
- d. SF1995 *me kamap te avel murri firma, mirri kompanja*
I'd like to have my business, my company.
- e. MG1959 *džanel mo Del te xoxavaw*
My God knows if I lie.
- f. RO-066 *trebu te akcjonis pala e parerea kiri*
You must act according to your own mind.
- g. MG1959 *"thaj kaj si ki dej haj kjo dad?"*
"And where are your mother and father?"
- h. Mdt1995 *kamawas anda lende kə aštil keren peski buki maj laši*
I wish for them that they can improve their own job.
- i. RO-064 *əl čhaore kadala haj čheorori kadaja barile kə pi baba*
This small boy and this small girl grew up with their aunt.

TABLE 10. Indefinite pronouns. Manner indefinites unattested in our data

| | Specific | Negative | Free choice | Universal |
|------------|---|------------------------|-----------------|---|
| Determiner | <i>niște</i> | <i>ni(j)ek(h)</i> | N/A | <i>soro, fjesa, fjesaro</i> |
| Person | <i>(da)khoniva</i> (NOM), <i>(da)khanikas</i> (OBL) | | | <i>sea lumja</i> <i>urkaste/orkastar</i> |
| Thing | <i>dajči(k)</i> | <i>khanč(i)</i> | <i>dajči</i> | N/A |
| Location | <i>khatinende</i> | <i>ni(j)ek(h) than</i> | N/A | N/A |
| Time | <i>orkana</i> | <i>ni(j)ek(h) data</i> | <i>vreodata</i> | <i>sea timpo</i> <i>fjesare data</i> |

6.5. Interrogatives

Interrogatives, also functioning as relative pronouns, are largely based on the historical sets in *k*- and *s*- (see Matras 2002: 112). The *k*- interrogatives are *kon* ‘who’, *kaj* ‘where’ and occasionally ‘why’, *kana* ‘when’ and *kozom* alongside *kabor* ‘how much/many’. The *s*- interrogatives are *so* ‘what’ and *sar* ‘how’. Țăndărei Romani shows a distinction between the interrogatives for goal *anda* (*so*)*ste* ‘to what end?’, combining the inherited locative preposition *anda* and the locative of *so*, and for reason *kə* and *pentru kə* ‘for what reason?’, both borrowed from Romanian.

6.6. Indefinites

The negative determiner, negative thing indefinite and most person and location indefinites in Țăndărei Romani are conservative forms. They are formed by combining the Early Romani indefinites **kha-/khaj-* and **či*, the indefinite particle **ni* and the numeral (*j*)*ek(h)* ‘one’ (cf. Elšik 2001, Matras 2002: 115). The universal determiner *soro* < **savoro* ‘all’ is also inherited. The remaining indefinites are either borrowed or combine borrowed and inherited materials, as common across Romani dialects (cf. Elšik 2001) (Table 10). Borrowed indefinites include *niște* < Rom. *niște* ‘some’, *fjesa* < Rom. *fiecare* ‘every’ (occasionally combined with inherited *soro* forming *fjesaro*), and *vreodata* ‘any time’. Inherited *sea* ‘all, entire’ is combined with Romanian *lumja* ‘people’ and *timpo* ‘time’ to form the corresponding universal forms. Romanian *data*, also ‘time’, combines with *ni(j)ek(h)* and *fjesare* to form, respectively, the negative and universal time indefinites. The Romanian marker *or-* is combined with interrogative *kana* ‘when’ to form the specific time indefinite, and with the locative (*kaste*) or instrumental (*kastar*) person interrogative to form the person universal indefinite. A further marker, *da-*, was found occasionally with the Person indefinite *dakhoniva* ‘somebody, nobody, anybody’ and systematically in the Thing indefinite *dajči* ‘something, anything’.

7. Verbs

7.1. *Verb derivation and valency alternation*

Țândărei Romani shows the usual Romani patterns of synthetic morphology to derive verbs and to alter valency (cf. Matras 2002: 119ff.). The infix *-av-* is used to derive transitives from a number of obsolete roots (*gar-av-* ‘to hide something’, *xox-av-* ‘to lie’, *sikh-av-* ‘to teach’), from nouns like *gili* ‘song’ (*gil-av-* > *gil-ab-* ‘to sing’) from intransitive verbs like *dara-* ‘to fear’ (*dar-av-* ‘to frighten’) and from transitive roots like *ker-* ‘to do’ (*kir-av-* ‘to cook/boil’). In the latter case, the *-i/-e-* variation in the root indicates the presence of a lexicalized form, though *kir-av-* is rarely used and (*pre*)*get-isar-* < Rom. *a* (*pre*)*găti* ‘to cook (to prepare)’ is favoured instead.

The infix *-ar-* also derives transitives from obsolete roots: *bist-ar-* ‘to forget’, *phab-ar-* ‘to burn’, *mund-ar-* ‘to kill’, *put-ar-* ‘to open’. While in most dialects *-ar-* is regularly employed as a de-adjectival and de-nominal marker, we have only found two instances of this use: *dil(j)-ar-* ‘to drive someone crazy’ < *dil-* ‘crazy’ and *xolj-ar-* ‘make someone angry’ < *xoli* ‘anger’. Most verbs that show *-ar-* in other dialects are replaced in Țândărei Romani by Romanian borrowings (e.g. *murdar-i-* ‘to pollute’ < Rom. *a murdări* instead of **mel-jar-* < *mel-* ‘dirty’) or by inherited verbs showing semantic expansion (*xa-* ‘to eat’ instead of *dand-ar-* < *dand* ‘tooth’ for ‘to bite’, *gil-ab-* ‘to sing’ instead of **baš-al-* for ‘to play an instrument’).

Intransitives and mediopassives are derived using *-(j)o(v)-*, from the existential auxiliary *ov-* ‘to become’: *bar-(j)o(v)-* ‘to grow’ < *bar-* ‘big’, *phur-(j)o(v)-* ‘to get old’ < *phur-* ‘old’, *mat-(j)o(v)-* ‘to get drunk’ < *mat-* ‘drunk’, *king-(j)o(v)-* ‘to get wet’ < *king-* ‘wet’, *lol-(j)o(v)-* ‘to become red’ < *lol-* ‘red’. The infix *-(j)o(v)-* is also used to derive intransitives from obsolete roots and from the past tense stem of transitives: *ker-d-(j)o(v)-* ‘to become’ < *ker-* ‘to do’, *gara-d-(j)o(v)-* ‘to hide oneself’ < **gara-*, *bučh-(j)o(v)-* ‘be called’ < **bučh-*, *sikh-(j)o(v)-* ‘to study/learn’ < **sikh-*, *phab-(j)o(v)-* ‘to burn’ < **phab-*, *phuk-(j)o(v)-* ‘to swell’ < **phuk-*. A further intransitive marker *-áv-* derives from another verb of motion and state turned auxiliary, *av-* ‘to come’, but was only encountered with *xoli* ‘anger’ > *xolj-áv-* ‘get angry’.

7.2. *Loan-verb adaptation*

Transitive loan verbs are integrated into Țândărei Romani using *-isar-*, combining Greek-derived *-is-* and the inherited valency marker *-ar-*. We found this marker only with past tense verbs, while present tense transitives only displayed the reduced marker *-i-*. Intransitive loan verbs take *-i-* in the present

tense but display *-isajl-* (Greek *-is-* combined with the past tense of the intransitive marker *-áv-*). Romanian reflexive and impersonal verbs are adapted using *-(j)o(v)-*:

- (9) a. CB1991 *anda mande te skriu romanes nea phares*
For me writing Romani is not difficult.
- b. IM1982 *gjes koleste, gjes koleste, gjes. . . te daštín te supravecin*
A day at this [field], a day at that, a day. . . so they can survive.
- c. IM1982 *kə sikhiljam te traisailjam maj mišto*
Because we had learnt that we could live better.
- d. MG1959 *o primari le gavesko ažutisardja lake da, lake dades*
The mayor of the village helped her mother, her father.
- e. AS1986 *skimbon sa pakjaw*
They are all changing, I believe.
- f. NT1992 *Man plačol ma kathe andi kadaja cara*
I like it here, in this country.

7.3. Person-inflection markers

In the present tense, Țăndărei Romani shows the typical consonantal and vocalic inflectional classes and the set of mediopassive, contracted forms that are common in Romani dialects (cf. Matras 2002: 136ff.). In the consonantal class the vocalic components of the concord markers (*-a-* for 1st persons, *-e-* for all others) remain unaltered. In the vocalic class, the vocalic component of the concord markers is assimilated to the root vowel: *-a-* in inherited verbs and *-i-* in loan verbs. With 1SG loan verbs *-v* is occasionally dropped (see Example 11a). Concord marker vowels are assimilated to contracted derivations in *-(j)o(v)-* following the person hierarchy identified by Matras (2002: 137) 3SG/PL > 2PL > 2SG > 1SG/PL; thus only the 1st persons retains the distinction between valency and concord markers (Table 11).

TABLE 11. Person concord markers

| | Present | | | Perfective |
|-----|--------------|-----------|----------------|-----------------|
| | In consonant | In vowel | Mediopassive | |
| 1SG | <i>-av</i> | <i>-v</i> | <i>-juv-av</i> | <i>-em</i> |
| 2SG | <i>-es</i> | <i>-s</i> | <i>-jos</i> | <i>-an</i> |
| 3SG | <i>-el</i> | <i>-l</i> | <i>-jol</i> | <i>-as ~ -a</i> |
| 1PL | <i>-as</i> | <i>-s</i> | <i>-juv-as</i> | <i>-am</i> |
| 2PL | <i>-en</i> | <i>-n</i> | <i>-jon</i> | <i>-en</i> |
| 3PL | <i>-en</i> | <i>-n</i> | <i>-jon</i> | <i>-e</i> |

The perfective concord markers match those encountered in the Vlax dialects of the southern Balkans (see below), with umlaut in the 1SG *-jom* > *-em*, the selection of 2SG *-an* rather than *-al* and substitution of 2PL *-an* for *-en* by analogy to the 2PL.PRES (Matras 2002: 144ff.). As discussed above, the 3SG marker for intransitive verbs shows variation *-as* ~ *-a*, with *-a* being the more frequent (see Table 2, Examples 3b, 3c).

3SG intransitives retain the conservative, adjectival inflection (M *-o/F -i*) (Table 12) although in interview data *-ja(s)* was also employed occasionally:

- (10) a. IM1982 *a la urma geljasas ko semafori, ko thojmos*
later [she] went to the traffic light, to clean [windshields].
b. IM1982 *gelo murro dad pi džermanja*
My father went to Germany.
c. CG1962 *avili te mangeltar love*
she had arrived to ask for money.
d. SF1995 *avilja andəl fetešti*
[my father] arrived in Fetești.

TABLE 12. Intransitive 3SG.PAST, variation in interviews

| | | % | N |
|----------------|---------------|-------|----|
| 3SG.PAST | <i>-o/-i</i> | 92.65 | 63 |
| concord marker | <i>-ja(s)</i> | 7.35 | 5 |

7.4. Past-tense inflection classes

The perfective classes found in Țândărei Romani match the class re-assignment hierarchy postulated by Matras (2002: 139) to account for the transition from the Proto-Romani split between the *-t-* and *-l-* classes to the situation in the current dialects. A reflection of the historical marker *-t-* is retained only with stems ending in /r l n v/ and shows voice assimilation, resulting in *-d-*. The stem *džan-* ‘to know’ shows partial re-assignment to the *-l-* class (*džan-gl-* ‘know.PERF’) across all our informants except RO-066, where it shows full re-assignment (*džan-l-*). All other verbs have been re-assigned to the *-l-* class, although *kam-* ‘to want’ shows variation between *-l-* and *-n-* in RO-064.

Intransitive derivations in *-(j)o(v)-* including loan verbs, verbs of motion and change of state (*av-* ‘to come’, *ušt-* ‘to stand up’, *dža-* ‘to go’), psych verbs (*asa-* ‘to laugh’) and mono-consonantal *l-* ‘to take’ take the adjectival ending *-il-* as perfective marker. Person concord markers attach directly or through a glide insertion to the mono-consonantal stem *d-* ‘to give’ (*d-em* ~ *d-j-em* ‘I gave’). The lexicalized alternations found across other Romani dialects (Matras 2002:

143) are also attested in our data: *peĭ-* to *per-* ‘to fall’, *mul-* to *mer-* ‘to die’, *sut-* to *sov-* ‘to sleep’ and *ruj-* to *rov-* ‘to cry’.

7.5. Copula

Among all our informants the Early Romani option selection between *s-* and *h-* as the copula root (cf. Matras 2002: 229ff.) is systematically resolved in favour of *s-*. Historically a perfective stem, it combines with the perfective person concord markers with the exception of 3rd persons, which continue forms in *si* (cf. Matras 2002: 145) (Table 13). The past tense copula is realized by attaching the remoteness marker *-a(s)* (see Table 2, Example 3a for details on the realization of /s/) to the present tense copula. The past-tense form *sas* ‘be.PAST.3SG/PL’ shows vowel assimilation and systematic retention of /s/. The root *av-* ‘to come’ acts as a suppletive for the realization of the future tense copula.

TABLE 13. Copula inflection

| | |
|-----|-----------------------------|
| 1SG | <i>s-em</i> ~ <i>s-im</i> |
| 2SG | <i>s-jan</i> ~ <i>s-ean</i> |
| 3SG | <i>si</i> |
| 1PL | <i>s-jam</i> ~ <i>s-eam</i> |
| 2PL | <i>s-en</i> |
| 3PL | <i>si</i> |

We mentioned above how historical jotation in morpheme boundaries is responsible for the variation in the realization of 2SG and 1PL markers (see example 1) and for vowel raising in 1SG *-om* > *-em*. In interview data we have noticed that in the copula this morpheme also underwent vowel raising to *-im* among some speakers. Overall, *sem* was slightly preferred (Table 14). However,

TABLE 14. Vowel raising in 1SG copula, variation in interviews

| | Sample total | | Family Network 1 | | | | | | | | | |
|------------|--------------|-----|------------------|----|--------|----|--------|----|--------|---|--------|---|
| | | | MG1959 | | CG1962 | | AS1986 | | DK1980 | | CB1991 | |
| | % | N | % | N | % | N | % | N | % | N | % | N |
| <i>sem</i> | 59.43 | 104 | 33.3 | 11 | 30.00 | 6 | 7.69 | 1 | 100 | 6 | 54.55 | 6 |
| <i>sim</i> | 40.57 | 71 | 66.7 | 22 | 70.00 | 14 | 92.31 | 12 | 0.00 | 0 | 45.45 | 5 |

| | Family Network 2 | | | | | | | | | | | |
|------------|------------------|---|--------|---|--------|----|---------|----|---------|---|--------|---|
| | VT1975 | | IM1982 | | NT1992 | | MrT1995 | | Mdt1995 | | SF1995 | |
| | % | N | % | N | % | N | % | N | % | N | % | N |
| <i>sem</i> | 57.14 | 4 | 60.00 | 9 | 82.76 | 24 | 100 | 25 | 80.00 | 8 | 66.67 | 4 |
| <i>sim</i> | 42.86 | 3 | 40.00 | 6 | 17.24 | 5 | 0.00 | 0 | 20.00 | 2 | 33.33 | 2 |

members of Family Network 1 show a preference for *sim*, particularly the younger member, AS1986. It is important to note that MG1959 reported being born in Moldova and gave *Lejaša* as her group affiliation. Among the members of Family Network 2, all born in Țândărei, *sem* is preferred, and this preference increases among the younger speakers. SF1995, whose father is originally from the border region with Moldova (see above), is an exception and his preference for *sem* is relatively weak.

7.6. Verb negators

Various verb negators are attested in our data. The inherited negator *na* is used by all speakers for non-indicative negation (see Example 11d) while only a single speaker in the sample, SF1995, used it also in indicative function (Table 15). For indicative negation we otherwise find the forms *ni* and *či*, with a general preference for *ni*. Among speakers born before the mid-1980s, only CG1962 produced a single token of *či*. By contrast, speakers born in the 1990s show a higher usage of *či*.

- (11) a. MRT1995 *kə te č̣i inteledži ma maj mišto*
because I don't understand [English] very well.
b. MRT1995 *akana ni kerav nijek munka*
Right now I don't have any job.
c. SF1995 *na žanawas ingleso, ni žanawas te keraw*
I didn't know English, I didn't know what to do.
d. SF1995 *kamawas [. . .] te na čoren*
I wish [. . .] they wouldn't steal.

TABLE 15. Indicative lexical verb negation, variation in interviews

| | Family Network 1 | | | | | | | | | | | |
|-----------|------------------|-----|--------|----|--------|----|---------|----|---------|----|--------|----|
| | Sample total | | MG1959 | | CG1962 | | AS1986 | | DK1980 | | CB1991 | |
| | % | N | % | N | % | N | % | N | % | N | % | N |
| <i>na</i> | 2.23 | 5 | 0.00 | 0 | 0.00 | 0 | 0.00 | 0 | 0.00 | 0 | 0.00 | 0 |
| <i>ni</i> | 85.27 | 191 | 100 | 46 | 92.86 | 13 | 100 | 8 | 100 | 10 | 5.26 | 1 |
| <i>či</i> | 12.50 | 28 | 0.00 | 0 | 7.14 | 1 | 0.00 | 0 | 0.00 | 0 | 94.74 | 18 |
| | Family Network 2 | | | | | | | | | | | |
| | VT1975 | | IM1982 | | NT1992 | | MRT1995 | | MDT1995 | | SF1995 | |
| | % | N | % | N | % | N | % | N | % | N | % | N |
| <i>na</i> | 0.00 | 0 | 0.00 | 0 | 0.00 | 0 | 0.00 | 0 | 0.00 | 0 | 21.74 | 5 |
| <i>ni</i> | 100 | 17 | 100 | 23 | 92.31 | 24 | 87.50 | 21 | 100 | 14 | 60.87 | 14 |
| <i>či</i> | 0.00 | 0 | 0.00 | 0 | 7.69 | 2 | 12.50 | 3 | 0.00 | 0 | 17.39 | 4 |

TABLE 16. 3SG copula negators

| | | % | N |
|------|----------------|-------|----|
| PRES | <i>naj</i> | 60.98 | 50 |
| | <i>nea</i> | 23.17 | 19 |
| | <i>nja</i> | 15.85 | 13 |
| PAST | <i>nas</i> | 67.69 | 44 |
| | <i>ni sas</i> | 24.62 | 16 |
| | <i>naj sas</i> | 3.08 | 2 |
| | <i>nea sas</i> | 3.08 | 2 |
| | <i>neas</i> | 1.54 | 1 |

Unique 3SG.PRES copula negators are also attested in our data (Table 16). All our speakers employ the unique present negator *naj* and strongly favour the unique past negator *nas* over the non-unique *ni sas*. In a handful of cases they employ compromise forms in which the present negator is followed by the inflected form of the copula (*naj sas*, *nea sas*) and we even find a single token of an innovative unique past negator *neas*.

The realization of the unique present negator and of these compromise forms shows variation in the treatment of the palatal glide. It would seem that, following metathesis *naj* > *nja*, the negator is affected by the same process observed in the 2SG/1PL.PAST markers (see above), leading to the emergence of *nea*.

7.7. Modals

As in other Balkan languages, modal verbs are followed by a complementizer and an inflected verb. Most modals in Țăndărei Romani are borrowed from Romanian, with *trebu(l)/trobul* < *trebui* ‘must, need, should’ remaining uninflected as in the donor language (Table 17).

TABLE 17. Modal verbs

| | Form | Inflected | Origin |
|--------------------|------------------------|-----------|-----------|
| must, need, should | <i>trebu(l)/trobul</i> | no | Romanian |
| can | <i>(da)ști-</i> | yes | Inherited |
| cannot | <i>naști-</i> | yes | Inherited |
| want | <i>kam-</i> | yes | Inherited |
| begin | <i>inčeposar-</i> | yes | Romanian |
| stop | <i>termini-</i> | yes | Romanian |
| | <i>opri-</i> | yes | Romanian |

- (12) a. CB1991 *dar trobul te sikhjov i sar la candarei kə murro manuš si candareano*
But I had to learn [Romani] like [it's spoken] in Țândărei
because my husband is from Țândărei.
- b. RO-066 *daštis te džas manca?*
Can you go with me?
- c. RO-066 *naštıl te dikhel ma*
She can't see me.
- d. RO-064 *kana deskjardja o udar ənčepisalo te del o brašınd*
When he opened the door, it started to rain.

While tense formation follows the conservative, inherited pattern, a Romanian borrowing infiltrates the transitional domain of tense and modality, with the particle *va* (derived from the Romanian inflected future auxiliary in the 3SG) often employed for the affirmational future (*tehara va terminiu akadja trjeba* 'tomorrow I shall [definitely] finish this work').

8. Aspects of syntax

Relative clauses are introduced by *kaj*, from the interrogative *kaj* 'where', both following animates and inanimates and with generic expressions. Occasionally, *sao* < **savo* 'which' is used with inanimates. Resumptive pronouns are used when the head noun takes a role other than subject in the relative clause:

- (13) a. RO-066 *djan ande le čhaves kaj rovel othe?*
Did you hit the boy who is crying there?
- b. RO-064 *ni ačhilo ma khanč kaj štı tə dav tut*
I have nothing left to give to you.
- c. RO-066 *arakhljan i džanta kaj rodesas lan?*
Did you find the bag you were looking for?
- d. RO-064 *čordem o sastre sao sas kolektime*
I stole the metal that had been collected.

Similarly, embeddings are introduced by the relevant interrogative. In the case of the conditional 'whether', the subjunctive complementizer *te* is used:

- (14) a. RO-066 *ni džanau kana te džap khere*
I don't know when to go home.
- b. RO-064 *phende mange so kerdjan agjes*
Tell me what you have done today!
- c. RO-066 *kamauas te phučav les anda soste kerdja kadaja*
I want to ask him why he did this.

- d. RO-064 *phendem le domoske te kamela te ael ki abjau*
I asked the teacher whether he wants to come to the wedding.

The particles *kaj* and *te* also act as complementizers: *kaj*, in alternation with the Romanian equivalent *că* > *kə*, introduces factual statements while *te* introduces non-factual ones, such as those expressed by the complements of modal verbs (see also Example 12):

- (15) a. RO-066 *ašundem kə aver rom bešel kathe*
I've heard that another man lives here.
b. RO-066 *sintusailo kaj si dajči lasa greuo*
He could feel that something was wrong with her.
c. RO-066 *te arakhlo o kher nas kate*
If he had found the house, he wouldn't be here.
d. MR1995 *dakə si man interneto dav duma lenca prin feisbuk*
If I have Internet [connection] I speak with them on facebook.

The conjunction *te* also introduces adverbial conditional clauses, but in this function it competes with *dakə* < Rom. *dacă* 'if'. Inherited adverbs are in competition with borrowed ones and with analytical expressions in other adverbial clauses: *kana* 'when' ~ *ando timpo kaj* 'in the time that'; *pala* ~ *dupə če* < Rom. *după ce* 'after'; *kaj* ~ *pentru kə* < Rom. *pentru că* 'why/because'. Other clauses are introduced exclusively by inherited or borrowed adverbs: *dži kaj* 'until'; *anglal* 'before', *sar* 'how', *fera te na* < Rom. *fără* 'without'; *afera* < Rom. *a fără de* 'except':

- (16) a. RO-064 *ame gilabastas gilja ando timpo kaj bešasas ko kampo*
We sang songs while we were staying in the field.
b. RO-066 *ək phuri sas la griža mandar kana simas ciknoro*
An old woman looked after me when I was little.
c. RO-066 *pala kadaja e škuala inčeposardem te munčiv and ek fabrika*
After I left school, I started working at a factory.
d. RO-064 *dupə či ənkləm i škoala inčepisajlem te munčiu andə ekh fabrika*
After I left school, I started working at a factory.
d. RO-064 *naštiv te sikhjuvau pentru kə ažutiu murrə dan*
I cannot study, because I help my mother.
e. RO-066 *dem les lake ke sas len nevoja lendar*
I gave it to her because she needed it.

TABLE 18. Adverbs

| Adverb | Origin | Meaning |
|-------------------------------|-----------------------------|-----------------|
| Local | | |
| <i>ando</i> | Inherited | inside |
| <i>avri</i> | Inherited | outside |
| <i>opre</i> | Inherited | up, above |
| <i>tele</i> | Inherited | down, below |
| <i>andi faca, angla</i> | Inherited + Rom., Inherited | in front |
| <i>pala(l), parpale</i> | Inherited | behind |
| <i>paša, apropime</i> | Inherited, Romanian | besides, nearby |
| <i>maškar</i> | Inherited | in between |
| <i>vizavi tar</i> | Rom. + Inherited | opposite |
| <i>maškar</i> | Inherited | through |
| <i>anda u žuro, pa o žuro</i> | Inherited + Rom. | around |
| <i>katar</i> | Inherited | away, from |
| <i>dural</i> | Inherited | from far |
| Temporal | | |
| <i>agjes</i> | Inherited | today |
| <i>tehara</i> | Inherited | tomorrow |
| <i>araki</i> | Inherited | yesterday |

TABLE 19: Borrowed prepositions

| Preposition | Meaning |
|----------------------|----------------------------------|
| <i>de</i> | from, out of, in (also temporal) |
| <i>la</i> | to/at |
| <i>andi direkcja</i> | towards (spatial) |
| <i>ando žuro</i> | past |
| <i>de anda</i> | since |

Borrowing of adverbs and the combination of inherited and borrowing material is not limited to the items listed above, as shown in Table 18. All the local adverbs in the table also function as prepositions alongside inherited prepositions *ke/te/pe* ‘at/to’, *anda* ‘from/through’, *karing* ‘towards’, and *dži* ‘until’. The inventory of prepositions is further enriched by more Romanian borrowings, listed in Table 19.

The dialect shows conservative word order rules. In elicitation we found no evidence of final positioning of the finite verb in declarative clauses, while verb-subject inversion is found in presentative constructions, in narrative openings, possessive constructions and foregrounded propositions expressing unexpected events (see 17).

- (17) a. RO-064 *avri si jekh čavoro*
There is a small boy outside.
- b. MG1959 *barjardja ame murri dej sar de o del*
My mother raised us, as God allowed.
- c. DK1980 *sas ma mərre phral khate*
I had my brothers here.
- d. CB19991 *sas murro dad nasfalo, kerdjas ekh infarto*
My father was ill, he had a heart attack.
- e. CG1962 *mule mange duj phra*
Two of my brothers died.

Perhaps the most distinctive feature pertaining to word order is the frequent postpositioning of attributes (both nominal and adjectival), a product of contact with Romanian:

- (18) a. RO-064 *i džuvli kaja cikni si la trin šave*
This little woman has three children.
- b. RO-064 *ek romni phuri sas la griž amuri kana simas cəno*
An old woman looked after me when I was little.
- c. RO-064 *o rom kaj avilo ko avjau si les ek mašina nevi*
The man who came to the wedding has a new car.
- d. RO-064 *o dad le šavesko sa es ək barba farte bari*
The boy's father had such a big beard.

9. Conclusion: Variation, classification, and koineization

The *Kangjlari* dialect of Țăndărei shows a checkered pattern of features, which we will address first in respect to their distribution among speakers in the sample, then in respect to their geographical and group-specific attestation among the Romani dialects of Romania, and finally in relation to existing dialect classification models.

All speakers in our sample show historical umlaut in *čhej* 'girl' and *dej* 'mother' and regressive assimilation of the vowel to the historical jotated segment in 1SG.PAST *-em/-im* < **-jom*, as well as prothetic *a-* (*abjau* > **bijav* 'wedding'), vowel assimilation across word stem and inflection in *šoro* < **šero* 'head' and accompanying consonant reduction in *soro* < **sa(v)oro* 'all', continuation of historical **-nř-* as *-rn-* in *marno* 'bread' etc., affrication in *cikno* < **tikno* 'small' and palatalization in *kjiro* < **tiro* 'your', syllable reduction in *gjes/djes* < **dives* 'day', absence of *s* > *h* in grammatical paradigms (including case inflections and copula stem) and retention of *-s-* in inflectional positions between vowel and consonants. There is also consistent use of the 2SG.PAST *-an* and 2PL.PAST *-en* and of the set of loan verb adaptation markers in *-i-* (present),

-isar- (past transitive) and *-isajl-* (past intransitive), demonstrative stems in *k_d-* and a conservative inflection pattern (*kadava*, *kadaja* etc.), deictics *kathe/kothe* ‘here/there’ and (*k*)*adiki* ‘such’, negative indefinite *khanč(i)*, loan plural markers in *-urj(a)*, and a general preference for the abstract nominalizer *-imos* and its oblique form *-mas-*. Overall, then, Țăndărei Romani is not a random contemporary mixture of forms but a consolidated variety.

Variation is found in the de-affrication in *džukel* ‘dog’ > *žukel* and *čhej* > *šej* ‘girl’, in the secondary raising of the vowel in *sim* ‘I am’, in the reflection of jotation in the 2SG and 1PL copula forms (*sj-* and *se-*), incipient prothesis in *ov* > *wov* ‘he’ etc. and forms like *uast* alongside *vast* ‘hand’, the retention of *-l-* in the definite article, the reduction of *-as* to *-a* in the remoteness marker and 3SG.PAST, use of the indicative negators *či* and *ni* and the past-tense copula negator *naj sas*, *ni sas*, *neas*, *nas*, and *nea sas*, use of *kozom* ‘how much’ alongside *kabor*, use of the reduced possessive pronouns 1SG *mo/mi*, 2SG *kjo/ki* and REFL *po/pi* and introflexion in the 1SG.POSS *murre* > *mərrə*, *murro* > *morro*, *murri* > *mirri*, and retention of adjectival inflection in 3SG past-tense intransitive verbs (e.g. *ail-i* ‘she arrived’). For some of these variants we found either an overwhelming or a clearly noticeable preference across the corpus. These include the retention of *-s* in the remoteness marker, 1SG.POSS *murro* with masculine nominative nouns, and a preference for adjectival inflection in 3SG past-tense intransitive verbs, all of which are present in around 90% of relevant tokens. Other preferred variants include the indicative verb negation in *ni* (ca. 85% of tokens), retention of affricates (and absence of de-affrication) in *džukel* ‘dog’ and *čhej* ‘girl’ (over 80%), and loss of *-s* in 3SG.PAST (ca. 70%). The forms *naj* and *nas* are used for copula negation present and past respectively in more than 60% of tokens in the corpus, while variation between *sem* and *sim* ‘I am’ tends to be balanced. Only for some variants is there clear evidence of individual preferences among speakers; thus CB1991 is fairly consistent in using the indicative verb negation in *či* while a number of other speakers show no tokens at all of this form (Table 15), and several speakers show no tokens of de-affrication (Table 3). There is thus, overall, evidence of a general tendency toward consistency and levelling of variants, suggesting that the dialect is in an advanced stage of transition between the ‘Chaos Phase’ and a ‘Focusing Phase’.

Many of the core features (those with high consistency and relatively low variation) align themselves with forms that are attested in the Romani dialects of the Banat and Transylvania regions in Romania, especially among population groups known as *Kurturari*, *Kelderaš*, and *Gabor* (cf. Matras 2013, van den Heuvel and Urech 2014). This is consistent with the strong North Vlax profile that the dialect seems to have, especially in regard to conservative traits but also in regard to some innovations (see below). Geographically, a number

of features of the *Kangljari* dialect represent developments that are diffused across Romania. These include the presence of prothetic *a-* (*abjau* > **bijav* ‘wedding’), which is ubiquitous except for conservative retention pockets of mainly *Romungre* speakers close to the Hungarian border in Transylvania (cf. Matras 2013: 218), the 2SG.PAST *-an* with a similar distribution (contrasting with pockets of *-al* in *Romungre* in Transylvania), and the affrication in *cikno* < **tikno* ‘small’ (widespread in Romania except in some *Ursari* varieties in the extreme south). The demonstrative stems in *kad-* are found across Romania with the exception once again of *Romungre* in the Northwest and *Ursari* in the south. Most Romani dialects of Romania also favour *-s-* in inflectional paradigms and in pre-consonantal position, though a shift to *-h-* is attested in a cluster of Romani dialects in the county of Mureș in central Romania as well as in Timiș in the west, where it is part of a continuum of dialects showing the same development in neighbouring Serbia (cf. Matras 2013: 221–2). The copula stem generally has *s-* apart from some dialects of Transylvania (cf. Matras 2013: 224). The de-affrication in *džukel* ‘dog’ > *žukel* is attested across the country except for conservative retention pockets (cf. Matras 2013: 219).

Țăndărei Romani also shows a number of key features that are shared more specifically with developments that are typically attested in Banat and Transylvania (cf. Matras 2013: 216–23), far from its current location. These include the umlaut in *čej* ‘girl’ and *dej* ‘mother’ incipient de-affrication in *čej* > *šej* ‘girl’, the reduction of *-as* to *-a* in 3SG.PAST (except retention pockets), while on the other hand the tendency toward retention of *-as* in the remoteness marker also agrees with the Romani dialects of Transylvania, except for a cluster of dialects of the *Gabor* group centred in the county of Mureș. It should be noted however that variation within a single dialect is not uncommon in the *Kelderari*-type varieties of North Vlax (cf. Gjerdman and Ljungberg 1963, Boretzky 1994). Further typical of Banat and Transylvanian Romani are variation in forms like *uast* alongside *vast* ‘hand’, palatalization in *kjiro* < **tiro* ‘your’, the use of reduced 2SG possessive forms *kjo* alongside long forms *kjiro*, syllable reduction in *gjes/djes* < **dives* ‘day’, and the retention of *-l-* in the definite article. The use of loan verb adaptation markers in *-i-* (present), *-isar-* (past transitive) and *-isajl-* (past intransitive) is a feature of the Vlax-type varieties of Transylvania, while the *Romungre* varieties show *-in-*, and retention of *-isar-* in the present tense is more common in the south of the country. The regressive assimilation of the vowel to the historical jotted segment in 1SG.PAST *-em/-im* < **-jom* appears in various pockets in Romania in Banat and neighbouring Hunedoara and Arad counties, with only recent diffusion into Transylvania (van den Heuvel & Urech 2014: 60), but it also correlates with the group affiliation *Kelderari* irrespective of location (Matras 2013: 214), suggesting an origin in Banat and retention in mobile or formerly nomadic groups. All this seems

to point to a founder population, or parts thereof, originating in a community from the broader Banat-Transylvania area.

At the same time, other core features of Țândărei Romani align themselves with those that are more widely attested in the south of the country (cf. Matras 2013: 214ff.): These include the continuation of historical **-nř-* as *-rn-* in *marno* ‘bread’ (contrasting with *-nř-* *-nr-* or *-ndr-* in the north), the 2SG.PAST in *-en* (contrasting with *-an* or *-al* in the north), the indicative negator *ni* (contrasting with *na* or *či* in the north) the interrogatives *kozom* and *kabor* (contrasting with *kici* and *sode* in the north), the retention of conservative inflection on demonstratives (contrasting with the adoption of adjectival inflection in many of the dialects of Transylvania), the absence of a clear consonantal prothetic segment in *ov* ‘he’ etc., and the use of the reduced 1SG possessive pronoun *mo* alongside the long form *murro*. Many of these agree with the patterns attested for the southern Romanian Romani dialects of the *Ursari* and *Spoitorja*.

We now come to the issue of classification of the Țândărei Romani dialect. The question is usually phrased in terms of Neo-Grammian diversification models, which in light of models of geographical spread of innovations, contact and convergence are often considered out of date (see section 2). We address the question not least in order to demonstrate just how problematic linear models of dialect sub-branching can be. Our principal point of reference here is Boretzky’s (2003) mapping and classification model of the Vlax dialects of Romani. To begin with, we note that many of the features described above for Țândărei Romani are consistent with (the broader pool of) descriptions of Vlax Romani, and that several features match what Boretzky (2003: 87ff.) considers to be the main diagnostic features of Vlax, notably historical umlaut in **daj > dej* ‘mother’, short genitive forms in *-ko-* etc., plural formation in *-uri(a)*, comparatives in *maj*, indefinites in *khon-/khan-*, negation particles in *ni* and/or *či*, copula negation in *naj*, lexical verb extension in *-tar*, 1SG past-tense and copula inflection in *-em*, and loan verb adaptation in *-isar-*. Contrasting with those is the dominance in Țândărei Romani of *(w)ov* ‘he’ etc., with just incipient prosthesis, whereas Boretzky considers full prosthesis to *vov* etc. to be constitutive of Vlax Romani.

Boretzky (2003: 93ff.) further lists prototypical differences between North and South Vlax (NV and SV, henceforth), which we summarize in Table 20 and compare with the forms attested for Țândărei Romani. It is clear that Țândărei Romani fails to adhere consistently to the prototype of either NV or SV. In some cases, it patterns with NV; in others, it patterns with SV; and in others still, it shows a mixture of both, with variants corresponding to each of the prototypes. The picture can be differentiated to some degree by considering more specifically the distribution of innovations as opposed to conservative retentions. NV innovations shared by Țândărei include the raising of the vowel

TABLE 20. North and South Vlax features following Boretzky (2003), and their presence in Țăndărei

| | North Vlax | South Vlax | Țăndărei |
|---------------------------------|-----------------------|--------------------------|-------------------------------|
| Vowel in 'head' | <i>šero, šəro</i> | <i>šoro</i> | <i>šoro</i> |
| Abstract nominalizer | <i>-imo(s)/ -ipen</i> | <i>-ipen</i> | <i>-imos, (-ipen)</i> |
| Preposition 'from' | <i>anda</i> | <i>andar</i> | <i>anda</i> |
| -l- in definite article | <i>-l-</i> | no <i>-l-</i> | <i>-l-, no -l-</i> |
| Enclitic pronouns <i>-lo</i> | <i>-lo</i> | – | – |
| Reflexive pronouns | <i>pesko</i> | <i>po, piro</i> | <i>pesko, po</i> |
| Copula 1SG | <i>șim/sim</i> | <i>sem</i> | <i>șim/sim, sem</i> |
| Copula 2SG | <i>san</i> | <i>sen</i> | <i>șjan, sean</i> |
| Negated copula 3SG | <i>nas</i> | <i>naj-sa</i> | <i>nas, ni-sas, (naj-sas)</i> |
| Negated lexical verb | <i>či</i> | <i>ni, in</i> | <i>ni, či</i> |
| 3SG intransitive past | <i>-a(s) + -o/-i</i> | <i>-o/-i</i> | <i>-o/-i, -a(s)</i> |
| 2PL past tense | <i>-an</i> | <i>-en</i> | <i>-en</i> |
| 1SG mediopassive | <i>-uvav</i> | <i>-jav</i> | <i>-uvav</i> |
| Future tense marker | <i>-a</i> | <i>ka</i> | <i>va</i> |
| Causative/transitive | <i>-av-</i> | <i>-al-</i> | <i>-av-</i> |
| Reduplicated deixis 'here' etc. | <i>katka</i> | – | – |
| 'Day' | <i>djes, gjes</i> | <i>dive, give, džive</i> | <i>djes, gjes</i> |
| 1SG possessive | <i>muro</i> | <i>munro/ mərno</i> | <i>muro</i> |

in the 1SG copula *sim* 'I am', the introduction of the lexical verb negator *či*, loss of medial *-n-* in **munro* > *muro* 'my', and syllable reduction in **dives* > *djes, gjes*. SV innovations shared by Țăndărei Romani are the labialisation of the vowel in *šero* > *šoro* 'head', the reduction of the reflexive *pesko* > *po*, the (partial) raising of vowel through jotation in *sean* 'you are', and the raising of the vowel through analogy in the 2PL.PAST *-en*, as well as the loss of a number of conservative features still found in NV, such as the enclitic pronoun *-lo*.

Many of the forms listed in Table 20 that show agreement with NV are conservatismisms, including the retention of the *-l-* definite article, of the nominalizer *-imos*, or the 1SG mediopassive *-uvav* and transitive *-av-*. This prompts the question whether Țăndărei Romani is historically a NV dialect that has acquired SV innovations, along with non-Vlax features from other southern Romanian varieties (such as interrogative *kozom/kabor* and pronominal forms in *ov*). Boretzky (2003: 99ff.) lists a number of cases in which superficial mixtures are attested, often pertaining to the presence of word-forms, which he attributes to inter-dialect interference. Boretzky does not offer any conclusive statements about the time and place of the formation of the Vlax dialects other than to suggest that they may well have been formed before contact with Romanian, nor about the time and place of the split between NV

and SV, other than to suggest that this will have been in Romanian territory, long before the mid-nineteenth century. An intriguing question is whether Țândărei Romani is a missing link between the Northern and Southern sub-branches of Vlax.

Boretzky (2003, *ibid.*) also points to the possibility that during the period of Romani enslavement in the Romanian principalities, individual Romani clans may have re-located and thus come under the influence of dialects belonging to the other sub-branch, giving rise to mixtures. A further hypothesis might therefore be that Țândărei Romani represents a case of inter-dialect interference, where a NV dialect has come into contact with a neighbouring SV variety. The dialect's present-day geographical position affords this hypothesis some plausibility, as some of the SV features are in fact widespread in southern Romania irrespective of dialect 'branch', i.e. they also occur in non-Vlax varieties such as those of the *Spoitorja* and *Ursarja*. This might lend itself to a further explanation, namely that Țândărei Romani is a displaced NV variety that settled in the south and absorbed features of neighbouring dialects, not necessarily limited to SV, resulting in the appearance of a NV–SV hybrid. Our principal argument against such a theory is the limited time frame for contact with neighbouring varieties, given that some 'southern' features, such as negation in *ni* and verb inflection 2SG.PAST in *-en*, are 'core' features that are used consistently by most speakers.

While we are, at this stage of our research, still unable to trace the precise origin of the majority of *Kangjlari* families or to try to reconstruct the dialectal forms that they brought with them when they re-located to Țândărei, we believe that the most plausible explanation for the emergence of a variety that combines this particular inventory of features, given their attested geographical and group-specific distribution outside of Țândărei, is as the outcome of a process of dialect levelling. The structural profile, coupled with attested contemporary inter- and intra-person variation, the timeline of the emergence of this community and individuals' history of re-location into it, as well as our knowledge about this process both from archives and from personal testimonies (some of which are alluded to briefly above), lead us to conclude that the *Kangljari* dialect is a product of recent koineization. This has both theoretical and methodological implications for work in Romani dialectology. In theoretical perspective, it offers an alternative both to the 'genetic' view of historical sub-branching of dialects through a process of perpetual diversification, propelled by the breaking away of sub-groups through migrations, and to the well-evidenced model of the diffusion of innovations across geographical space through chains of neighbouring communities. In methodological perspective, our findings underline the importance of combining linguistic sampling in a speech community with historiographical and ethnographic research.

References

- Bakker, Peter, and Matras, Yaron. 1997. Introduction. In: Matras, Yaron, Bakker, Peter, and Kyuchukov, Hristo, eds. *The typology and dialectology of Romani*. Amsterdam: John Benjamins. vii–xxx.
- Boretzky, Norbert. 1994. *Romani. Grammatik des Kalderaš-Dialekts mit Texten und Glossar*. Wiesbaden: Harrassowitz.
- 1995. Interdialectal interference in Romani. In: Matras, Yaron, ed. *Romani in contact. The history, structure and sociology of a language*. Amsterdam: Benjamins. 69–94.
- 2000. South Balkan II as a Romani dialect branch: Bugurdži, Drindari, and Kalajdži. *Romani Studies*, 5(10): 105–83.
- 2003. *Die Vlach-Dialekte des Romani. Strukturen – Sprachgeschichte – Verwandtschaftsverhältnisse – Dialektkarten*. Wiesbaden: Harrassowitz Verlag.
- 2007. The differentiation of the Romani dialects. *Language Typology and Universals* 60: 314–35.
- and Iglă, B. 2004. *Kommentierter Dialektatlas des Romani. Teil 1: ergleich der Dialekte; Teil 2: Dialektkarten mit einer CD-ROM*. Wiesbaden: Harrassowitz.
- Britain, David. 1997. Dialect contact and phonological reallocation: “Canadian Raising” in the English Fens. *Language in Society* 26: 15–46.
- Constantinescu, Barbu. 1878. *Probe de limba și literatura Țiganilor din România*. București: Tipografia Societății Academice Române.
- Drimba, Vladimir. 1992. *Imprumuturi românești din limba țiganească*. *Studii și Cercetări Lingvistice* 43: 173–90, 253–78.
- Elšik, Viktor. 2000a. Romani nominal paradigms: Their structure, diversity and development. In: Matras, Yaron and Elšik, Viktor, eds. *Grammatical relations in Romani*. Amsterdam/Philadelphia: John Benjamins Publishing Company. 9–30.
- 2001. Word-form borrowing in indefinites: Romani evidence. *Sprachtypologie und Universalienforschung* 54 (2): 126–47.
- 2003. Interdialect contact of Czech (and Slovak) Romani varieties. *International Journal of the Sociology of Language* 162: 41–62.
- Gaster, Moses. 1931. Rumanian Gypsy Folk-tales. *Journal of the Gypsy Lore Society Third Series*, 10(4): 153–71; 12(4): 166–89; 15(1): 10–20; 15(4): 160–81; 17(3): 58–66.
- Gjerdman, Olof, and Ljungberg, Erik. 1963. *The language of the Swedish Coppersmith Gipsy Johan Dimitri Taikon: Grammar, texts, vocabulary and English word-index*. Uppsala: Lundequist.
- Friedman, Victor A. 2017. Seven varieties of Arli: Skopje as a center of convergence and divergence of Romani dialects. *Romani Studies*, fifth series, 27: 29–45.
- Hancock, Ian F. 1995. *A handbook of Vlax Romani*. Columbus, OH: Slavica.
- Hickey, Raymond. 2005 *Dublin English. Evolution and change*. Amsterdam: John Benjamins.
- Iglă, Birgit. 1996. *Das Romani von Ajia Varvara. Deskriptive und historisch-vergleichende Darstellung eines Zigeunerndialekts*. Wiesbaden: Harrassowitz.
- Jakoubek Marek and Budilová, Lenka. 2006. Kinship, social organisation and genealogical manipulations in Gypsy osadas in eastern Slovakia. *Romani Studies*, fifth series, 16: 63–82.

- Leggio, Daniele V. 2011. The dialect of the Mitrovica Roma. *Romani Studies*, fifth series, 21: 57–113.
- Kerswill, Paul. 2002. Koineization and accommodation. In: Chambers, Jack K., Trudgill, Peter and Schilling-Estes, Natalie, eds. *The handbook of language variation and change*. Oxford: Blackwell. 669–702.
- 2010. Contact and new varieties. In: Hickey, Raymond, ed. *Handbook of language contact*. Oxford: Blackwell. 230–51.
- , and Williams, Ann. 2000. Creating a new town koine: Children and language change in Milton Keynes. *Language in society* 29: 65–115.
- — 2005. New towns and koineisation: Linguistic and social correlates. *Linguistics* 43(5): 1023–48.
- Marushiuakova, Elena and Popov, Veselin. 2004. Segmentation vs. consolidation: The example of four Gypsy groups in CIS. *Romani Studies*, fifth series, 14: 145–91.
- Matras, Yaron. 1994. *Untersuchungen zu Grammatik und Diskurs des Romanes. Dialekt der Kelderaša/Lovara*. Wiesbaden: Harrassowitz.
- 1999. Johann Rüdiger and the study of Romani in eighteenth century Germany. *Journal of the Gypsy Lore Society*, fifth series, 5: 89–116;
- 2002. *Romani: A linguistic introduction*. Cambridge: Cambridge University Press.
- 2004. Romacilikanes: The Romani dialect of Parakalamos. *Romani Studies*, fifth series, 14: 59–109.
- 2005. The classification of Romani dialects: A geographic-historical perspective. In: Schrammel, Barbara, Halwachs, Dieter. W. and Ambrosch, Gerd, eds. *General and Applied Romani Linguistics*. Munich: Lincom Europa. 726.
- 2010. *Romani in Britain: The afterlife of a language*. Edinburgh: Edinburgh University Press.
- 2013. Mapping the Romani dialects of Romania. *Romani Studies*, fifth series, 23: 199–243.
- , and Leggio, Daniele. V., eds. 2017. *Open borders, unlocked cultures. Romanian Roma migrants in Western Europe*. London: Routledge.
- MigRom (2015) Report on the Extended Survey, Romanian Institute for Research on National Minorities June 2015. http://migrom.humanities.manchester.ac.uk/wp-content/uploads/2015/08/Yr2report_Cluj.pdf. Accessed 7 Aug. 2017.
- Poboźniak, Tadeusz. 1964. *Grammar of the Lovari dialect*. Kraków: Państwowe wydawnictwo naukowe.
- Scholtmeijer, Harm. 2000. Language in the Dutch Polders: Why dialects did not mix. In: Mattheier, Klaus, ed. *Dialect and migration in a changing Europe*. Frankfurt am Main: Peter Lang. 123–44.
- Siegel, Jeff. 1985. Koines and koineization. *Language in Society* 14: 357–78.
- Solheim, Randi. 2009. Dialect development in a melting pot: The formation of a new culture and a new dialect in the industrial town of Høyanger. *Nordic Journal of Linguistics* 32(2): 191–206.
- Stewart, Michael. 1997. *The time of the Gypsies*. Boulder, CO: Westview.
- Sutherland, Ann. 1975. *Gypsies: The hidden Americans*. Prospect Heights, IL: Waveland.
- Tenser, Anton. 2006. *Lithuanian Romani*. München: Lincom Europe.

- Toma, Stefania, Tesăr, Catalina and Fosztó, Lazsló. 2017. Romanian Roma at home: Mobility patterns, migration experiences, networks, and remittances. In: Matras, Yaron and Leggio, Daniele V., eds. *Open borders, unlocked cultures. Romanian Roma migrants in Western Europe*. London: Routledge. 57–82.
- Trudgill, Peter. 1986. *Dialect in contact*. Oxford: Blackwell.
- 2004. *New-dialect formation: The inevitability of colonial Englishes*. Edinburgh: Edinburgh University Press.
- van den Heuvel, Wilco and Urech, Evelyne. 2014. Romani dialect variation in Transylvania: Migration and diffusion. *Romani Studies, fifth series*, 24: 43–70.

