

Romacilikanes— The Romani dialect of Parakalamos

YARON MATRAS

Descriptions of the Romani dialects of Greece have hitherto concentrated on the Vlach varieties, spoken mainly by immigrants who were expelled from Turkey in the early 1920s (cf. Iglu 1996, Messing 1987). Hardly any discussion has been devoted to the Balkan dialects spoken by Romani populations with a long history of settlement in Greece. The paper describes the dialect of the *Romacel* community of Parakalamos, in the Epirus district of Greece—a language referred to by its speakers as *Romacilikanes*. The corpus consists largely of questionnaire elicitation carried out in Epirus as part of the Romani Morphosyntax (RMS) database project. The description is also intended to serve as a programmatic contribution to the agenda of Romani linguistics: it introduces the RMS project's methodology, and demonstrates how an integrated dialectological–typological approach can provide a concise yet comprehensive outline of a Romani dialect.

Keywords: Romani language, Parakalamos, Greece, Epirus, Romacel, Romacilikanes, linguistics, dialectology

1. Introduction

Some of the earliest scholars investigating Romani had already noticed a conspicuous Greek element that is shared by all its dialects, and which includes not just lexicon, but also morpholexicon (function words) and grammatical morphemes. Miklosich (1872–1880, III–4) had consequently identified ‘Greece, or a land in which the Greek language was predominant’, as the ‘European homeland’ of the Gypsies, that is, the land in which they had lived for a prolonged period of time before dispersing into the various European regions. Despite the obvious interest in the historical Greek-speaking area (from present-day Greece to Anatolia) as the centre of diffusion of Romani populations in Europe, still little is known about the Romani dialects of either Greece itself, or Asia Minor.

After Evliya Çelebi's word list of 1668 from Thrace (see Friedman and Dankoff 1991), the earliest source on Ottoman Romani is Paspati (1870),

Yaron Matras is the Editor of *Romani Studies* and Professor at the Department of Linguistics, University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom. Email: y.matras@man.ac.uk

who describes both a Vlach and a non-Vlach variety. Subsequent work on Balkan Romani during the twentieth century has focused first on the dialects of Bulgaria, then on Vlach dialects in Albania and Bosnia, and only later on Macedonia. The first thorough description of a Romani dialect from Greece was Iglá's (1996) work on the Vlach dialect of Agia Varvara in Athens, an emigrant Romani community from Turkey (see also Messing 1987, 1991 for a dictionary and sample texts of the same variety). The opposite phenomenon, so to speak, a non-Vlach or 'Balkan' dialect of Romani originating from the vicinity of Thessaloniki, now spoken by an emigrant community in Izmir in Turkey, that of the Sepečides (Basket-Weavers), was described by Cech and Heinschink (1999). Although this dialect has been influenced mainly by Turkish, it is the only genuinely 'Greek' variety of Romani that has received elaborate attention so far.

In his pioneer monograph on the southern Balkan dialects of Romani, Boretzky (1999) draws on only two sources from Greece, namely Cech and Heinschink's description of Sepečiči, and unpublished (and apparently fragmented) material from Serres in northeastern Greece. Comparing these with Balkan dialects that are adjacent to the north—from Prilep in Macedonia, Bulgarian Erli, and the Thrace dialect described by Paspatis (1870)—Boretzky identifies several salient isoglosses that seem to separate the northern zone from the southern, or 'Greek' zone. These include: *maro* 'bread' in the north, but *mando* or *mandro* in the south; short genitive *-ko* as a variant of *-koro* in the north, but only *-koro* in the south; instrumental plural *-endzal/-endžar* in the south; preposition *andre* 'in' in the south, but reduced forms in the north; non-indicative copula *ov-* in the north, *av-* in the south; presence of *ther-* 'to have' in the south; and jotted perfective forms *kerġjum* in the north, de-jotation to *kerdom* in the south. Just how far south these features extend, however, has so far been unknown. Parakalamos Romani provides us with a test-case, and at the end of this description I shall return to its geographical and historical position and re-assess some of the salient isoglosses of the region, drawing a comparison with more recent data on the Romani dialects of Greece.

Parakalamos Romani (henceforth PR) is the language spoken by the community known as the 'Gypsy musicians' of the village of Parakalamos, near Ioannina, in the district of Epirus (Ipeiros), in northwestern Greece (see discussion of the same community in Theodosiou 2004, in this issue). The community is rather small, apparently comprising just several hundred individuals, some of whom have recently moved from Parakalamos to the district

capital, Ioannina. The families settled in the village in the 1920s, having left predominantly Muslim villages along the Albanian border, and converting to Christianity, apparently in order to escape expulsion (so-called ‘population exchange’) to Turkey. There appear to be family ties with Romani communities in Albania, which have been revived during the past decade and a half since the opening of the Greek–Albanian border. Knowledge of Albanian had generally been widespread in the region in earlier times, and like the local Greek dialects, PR also shows Albanian and Turkish influences.

In many respects, PR is a remarkably conservative variety of Romani, while on the other hand it shows some unique innovations—both suggesting a period of isolation from other Romani-speaking groups. However, it also shares some salient structural characteristics with other Romani dialects of Greece, as well as more generally with Romani dialects of the southern Balkans, confirming its affiliation with this dialect group, and historical ties between its speaker community and the larger population of settled Roma of the region. One of the outstanding markers of the group is its self-appellation, *romacel*, a label that has hitherto been known primarily from Romani groups in the western margins of Europe. When talking in Greek, the Parakalamos Romacel refer to themselves as *jifti*, setting themselves apart from the Vlax-speaking communities known in the region as *cingani*, who were expelled to Greece from Turkey during the so-called ‘population exchange’ in the early 1920s. This two-level pattern of self-ascription confirms the impression of a group that had once been part of an historical Romani population that had settled in the region (‘settled’ in the sense of occupying a position within the region’s population mosaic, notwithstanding itinerant traditions), the *jifti*, and is distinct from the more recent wave of settlers, the *cingani*; but also of a group that has developed its own, community-internal sense of identity, expressed by the exclusive adoption of the self-ascription term *romacel*.

2. The agenda of Romani linguistics

In her discussion of the role of ‘place’ in the configuration and presentation of identity among the Parakalamos Gypsies, Theodosiou (2004, in this issue) distances herself somewhat from attempts to provide ‘objective’ ethnographic descriptions, and from attempts to define straightforward demarcations among the various Gypsy groups, as well as between Gypsies and non-Gypsies. Theodosiou is not alone in attributing a contextual rather

than static character to the presentation and construction of group-identity among Gypsies. There is indeed a trend in anthropology to question and deconstruct boundaries in an almost literary interpretation of their context-bound symbolic effects on audiences. By contrast, linguistic science offers a system-oriented examination of self-contained sets of structural facts: inflectional paradigms, syntagmatic rules, phonological characteristics, lexical composition, and so on. The Parakalamos Gypsies may be both settled and nomadic, farmers and musicians, Greek and outsiders, all depending on the contrastive versus integrative effects that individuals may be able to achieve by highlighting particular angles of their no doubt multi-layered identity; but to linguists they constitute a clearly demarcated speech community. This, despite the fact that the Greek dialect of Parakalamos is known to have incorporated Romani words, and that the Romacel in turn codeswitch between Romani and Greek and would regard both as native languages. Nonetheless, Greeks, or *balame* as they are called, do not learn Romani, and it is not spoken in their families. Romani in Parakalamos is the exclusive property of the *romacel* community.

But Romani obviously did not emerge in Parakalamos. It is a language of Indo-Aryan affiliation, and it is of no importance to this linguistic classification that many, perhaps most Romani communities have no awareness of the Indian origin of their language (nor that of their ancestors, a millennium ago), and that Indian origins may play no role whatsoever in their everyday lives or their identity narratives. The internal structure of language and its systemic coherence tell a story, albeit indirectly—and one that is not as open to context-bound interpretation as are identity narratives.

Thus while the Romacel may at times reconsider their own self-ascription, in Greek and toward the outside world, and contemplate adopting for themselves the label *cingani* (rather than *jifti*) when this is considered advantageous to their community, as Theodosiou describes, their Romani dialect remains distinct from that of the *cingani*. And so, regardless of the labels, we are able to draw a boundary between the two communities, and establish that they have two separate histories, and that they therefore constitute two separate linguistic lineages within the Romani-speaking population of the Balkans. Similarly, linguistic features shared with other communities in the region may allow us to infer that contacts existed between these speech communities, through which innovations in speech became regionally diffused. This would point to a network of social contacts among communities that are or were contiguous with a Greek-speaking population, yet also sep-

arate from this population, constituting a demarcated speech community.

In recent years, the agenda of Romani linguistics has expanded from the purely descriptive approach which had been its traditional orientation, in three novel directions. The first of those might be defined as an applied or engaged agenda, which seeks to forge partnerships with Romani-speaking communities and provide both analyses and expert advice in language-planning issues and language engineering. Linguists have been involved in activities related to language codification, standardisation, status and function elaboration, lexicographic documentation (production of dictionaries) and production of language educational material. The second is a linguistic-theoretical agenda, especially one that feeds into current discussion in contact linguistics and linguistic typology. In these areas, Romani dialects are valued as a comparative sample that can reveal natural trends in the adoption of vocabulary and grammatical structure from external sources (contact languages), and in the internal re-organisation of grammatical structure.

The third agenda point is the historical and geographical one. It is concerned with reconstructing the developments that led to the formation of present-day dialects. In the absence of historical records of earlier stages of Romani, this approach relies heavily on a contrastive analysis of dialects, and to a large extent on dialect geography, tracing the diffusion of changes through geographical space and making inferences about underlying processes of change based on the present-day geographical distribution and clustering of structural features. This latter approach is of potential interest to other disciplines as well, for the spread in geographical space of structural innovations within language presupposes contacts between members of individual speech communities. This in turn may reveal historical patterns of social contacts between Romani populations, including migrations.

3. The Romani Morphosyntax (RMS) database

In 1998, a project aimed at compiling a comparative description of Romani dialects in electronic form was launched at the University of Manchester, by the author in collaboration with Viktor Elšík.¹ The initial goal was to provide a summary of data that had been published in grammatical descrip-

1. With support from the Arts and Humanities Research Board, grants no. B/RG/AN4725/APN9447 AND B/RE/AN4725/APN11878, and later with additional support from the Open Society Institute's Roma Cultural Program.

tions and texts, analysed by grammatical categories, in a format similar to that of a descriptive grammar, but on a database, enabling the user to compare data more effectively through controlled, user-defined queries.

The initial core sample contained some forty elaborate descriptions of Romani dialects. The data were tagged in several ways: first, variants were accompanied by analytical descriptions, for instance a ‘yes’ or a ‘no’ in response to a particular question targeting a general aspect of structural variation, e.g. ‘is the definite article retained in this dialect?’ followed by fields indicating the individual forms of the definite article, representing an historical form-to-form development. Second, functions of inherited forms were encoded, allowing the user to obtain an historical form-to-function analysis, e.g. ‘which function do long forms of the present conjugation serve?’, options being ‘present-future’, ‘future’, ‘conditional’, or ‘present’. Next, function-to-form questions were included, based on state-of-the-art typological descriptions in the relevant areas; e.g. ‘how are negative indefinites expressed in the language?’ Finally, contact influences (i.e. grammatical borrowings) were tagged for source, according to the ‘depth’ of contact, representing up to three layers of historical contact languages: ‘current L2’—that spoken in the community alongside Romani; ‘recent L2’—a second language spoken only by the older generation; and ‘old L2’—a language that has had a significant impact on the dialect, but is no longer in use in the community. The database contains altogether over 5,500 fields with information on forms, or analytical questions of this kind, covering all areas of structure, with the exception of phonetics.

The project entered its second phase in 2001, when on the basis of the database structure an elaborate questionnaire—‘The Romani Dialectological Questionnaire’ (Matras et al. 2001)—with over 1,000 entries was designed, with the aim of extending the dataset to cover dialects that have not been thoroughly described so far. The questionnaire covers all areas of morpho-syntactic variation in Romani, including conjugations of all potential verb inflection classes, and a word list targeting salient variation in historical lexico-phonology. The targeted categories are usually incorporated into short sentences. The questionnaire, translated into numerous state languages, has since been in use across Europe to elicit translations from native speakers of Romani, often by Romani native speakers working as project fieldwork assistants. The responses are recorded and then transcribed onto a pre-formatted spreadsheet, where each numbered sentence is pre-tagged for the relevant grammatical-semantic categories that appear in it. This enables the

project staff to generate sub-corpora in search of particular categories or category combinations, for example, 'demonstratives' or 'relative clauses', thereby facilitating data entry into the database, and of course an overview of the structural features by category. A further advantage of a uniform questionnaire is of course the fact that numerous dialects can be compared for an identical set of sample sentences. With much of the data already transcribed but still waiting to be entered into the RMS database, in January 2004 the RMS project archive contains some 130 questionnaire-based recordings of Romani dialects from all over Europe.

The analytical working hypothesis that guides the database project is that the present-day dialects of Romani are all derived from an historical forerunner, which we call 'Early Romani' (ER), and which was spoken in the Byzantine Empire, in intense contact with Greek, in all likelihood some time between the tenth and twelfth centuries AD. Early Romani is not documented, of course, but an analytical reconstruction of much of its composition is possible by applying the standard comparative method in linguistics, and by checking the results against attested late-medieval Indo-Aryan languages.

Thus, we find in present-day Romani dialects various forms for the word 'day', including *dives*, *diveh*, *dive*, *di*, *d'ives*, *dživeh*, *džes*, *džis*, *zis*, *zes*, and more. Internal reconstruction allows us to derive all forms from an hypothesised ER **dives*, possibly with a palatalised variant **d'ives*. This is strengthened by the attestation of late Middle Indo-Aryan (MIA) *divasa*- 'day', which, given the regular sound changes that characterise Romani in comparison with MIA, would give precisely *dives*. Importantly, the methodology of ER reconstruction applied in the RMS project does not rely strictly on hypothesised derivations alone; rather, it is assumed that, by and large, ER is close enough in time and space to its present-day dialect descendants for ER forms to have been continued, sporadically and selectively, in some dialects. The challenge is therefore in the first instance to identify the most conservative form among those attested in the present-day dialects, and to ascertain whether that is the form that is likely to have given rise to the other variants. The procedure of identifying the conservative form is guided by universals and particulars of sound changes, morphological analogies, contact influences, and so on, but it is also supported by the geographical diffusion of forms. Thus, *dives* occurs in the southern Balkans, in the Baltics, in Britain, and in central Europe, while *džes* or *zis* are confined to more specific, smaller, and more coherent regions. This of course makes *dives* the more natural candidate for an ER cognate, even before sound changes, semantic shifts and the

like are applied. The reconstruction methodology is introduced and applied in more detail in Matras (2002), and in various other works in the context of the RMS project.

The notion of an ER ancestor language and the ability to take into account solid assumptions about its structure provide a framework for Romani dialectology that is, like that of dialectological work on other European languages, not just descriptive (in that it describes individual dialects in isolation), but also historical and geographical, in that it relates descriptive findings to the wider context of processes of change that have occurred in Romani as a whole, and to their patterns of geographical diffusion, as documented so far. The following description of Parakalamos Romani is the first discussion to appear in print that is based exclusively on the elicitation and evaluation method of the RMS database.² It has the disadvantage of not being able to draw on a large corpus of narratives (though some were included), nor is it able to address questions relating to sociolinguistic variation within the community. It does, however, survey most relevant areas of grammatical structure, addressing specifically the outcome of general processes of divergence within Romani, and taking into account typological questions, in particular in morphosyntax.

4. Remarks on lexicon

One of the outstanding features of the PR lexicon is the autonym *romacel*, and the self-appellation of the language that is derived from it, *romacilikanes* or *romacikani čhib*. The European (Greek-derived) derivation marker in the latter, *-ikan-*, suggests that the name of the language, and possibly even the label *romacel* itself, are not pre-European, but entered the lexicon after settlement in Europe. The pre-European terms *rom* and *romni* are found in PR in the meaning ‘husband’ and ‘wife’ respectively.

Labels that are cognate with *romacel* have previously been identified only among Romani communities in western Europe. The word appears in the Basque country as *errumantxel-*, in Britain as *romanichal* (now predominantly used by or with reference to Romani clans that have emigrated to North America), in France as *romanichel*, and in Finland as *romačel*. Bakker (1999) had suggested that the label may have been an innovation of what he

2. The recordings, with altogether five speakers of different age groups, were carried out by Aspasia Theodosiou in Parakalamos and Ioaninna in December 2001 and in August–September 2002, and transcribed by the author.

calls—due to their predominance in northern Europe—the ‘Northern’ dialects of Romani. It is important to note however that in neighbouring and/or overlapping regions, the term is used alongside other self-appellations, in particular those deriving from *kale* ‘blacks’: *caló* in Spain, *kååle* in Wales, *kaale* in Finland, and formerly *kale* in Germany.

While *kale* is not attested as a self-appellation in other parts of Europe, the type *romačel* appears not only in Parakalamos, but apparently also as one of the self-appellations of the Crimean Gypsies: *urumčel* (E. Marushiakova and V. Popov, p.c., and forthc.). The hypothesis of a northern- or western-European origin of *romačel/romacel* must therefore be refuted. Significantly, the word *rom* as a designation relating to a member of the group, in some function or other, is always present in the language, irrespective of the term that is used as a default self-appellation. Among the *sinte* of Germany, for instance, *rom* and *romni* mean, as in Parakalamos, ‘husband’ and ‘wife’, and the language is called *romnes* (often alongside *sintitikes*, which, like *romacilikanes*, is a European derivation).

It is thus possible to reconstruct an original pool of self-appellations which appears to have existed among Romani-speaking populations. The oldest, base-form, is the word *rom* < *řom* < *đom*, which appears to have denoted a ‘person belonging to the group’. The ‘group’ itself, as the Indian cognate term *đom*—a cover-term for *jatis* of a particular status—suggests, had originally been a caste-like denomination sharing social status and a range of permissible occupations. From this semantic prototype derive both the kinship label ‘person belonging to the group, in matrimonial alliance’, and the group label itself, ‘person(s) belonging to the group as an identifiable collective’, from which, in turn, the name of the language, as ‘language spoken by the group’ derives.

The semantically most volatile member of this derivation chain is the label used for the collective. It is subject to replacement both by semantically meaningful, internal derivations, such as *kale* ‘Blacks’ or *manuš* ‘people’ (attested in France, and previously in Germany), or by other inherited labels or those possibly borrowed from neighbouring peripatetic groups. Examples are *gurbet* (in the Balkans; cf. the Domari-speaking *qurbāti* of northern Syria), and *sinte* (adopted by the German Gypsies around the late eighteenth or early nineteenth century, possibly as a camouflage term from neighbouring peripatetic groups; cf. Matras 1999a). In southeastern Europe especially, we find a renewal of the term for the collective favouring external labels (mainly Turkish and Romanian), which have to do with status and occupa-

tion vis-à-vis mainstream society: *yerli* ‘settled’, *ursari* ‘bear-leaders’, and so on. In some communities these new labels co-exist with the older ones. Elsewhere, they replace the older name of the collective and its members, and sometimes also the generic label for ‘person, who is also a member of the group’ (cf. in Germany *sinto* ‘a man, of the Sinti group’). They may also give rise to new labels for the language (cf. *sintitikes*, *arlikanes*, both alongside *romnes/romanes*). A replacement of the kinship term (‘member of the group, related by matrimonial alliance’) is not attested, however (cf. German Romani *sinto* ‘member of the group’ and ‘man, who is a member of the group’, but *rom* ‘husband’), making this the most stable element in the chain.

It appears therefore that *romacel* was part of a pool of labels that existed already in Early Romani, but have survived only in the margins of Europe, perhaps an attestation of relative isolation of these respective groups from other Romani populations during the past two or three centuries. A language designation based on *romacel* has so far not been attested, except for PR, and it is certainly possible, especially when one takes into consideration its Greek-derived basis *-ikan-* (available, admittedly, in Romani as a whole as an ER loan), that it emerged locally (‘locally’ meaning in this specific speech community, though the precise location cannot of course be determined, since the word is most likely to have pre-dated settlement in Parakalamos itself).

The term that is used to denote outsiders is the usual *gadžo*. Another label, *balame*, is reserved specifically for ‘Greeks’, and appears to be used predominantly not in Romani, but in a Greek-speaking context, as a counterpart to Greek *jifti* ‘Gypsy’. Albanians are called *xoraxane* ‘Muslims’, and more specifically in Greek *arvanitiki*. From the vocabulary survey of some 250 inherited words contained in the RMS questionnaire, PR emerges as rather conservative in its retention of vocabulary. Lost (or unattested) items include **čovaxani* ‘witch’, for which we find Greek *májisa*, and *rikono* ‘puppy’, for which we find *pápi*. Also missing are **manřikli* ‘cake’, for which we find *guldo* (originally ‘sweet’), **kir-* ‘to cook’, for which we find *ker-jela* ‘to make food’ (*jela* from Greek), and, interestingly, **fóros* ‘town’, an ER Greek loan which is not however found in this meaning in contemporary Greek, and for which PR has the internal creation *baro gav*, lit. ‘large village’, alongside the Greek loan *póli*.

Conservativisms include the opposition pairs *kašt* ‘wood’ vs. *rukħ* ‘tree’, *jiv* ‘snow’ vs. *páhos* ‘ice’, and *čhon* ‘moon’ vs. *másek* ‘month’, the retention of *kaljardo* ‘soldier’, and the retention of *kam-* ‘to want’ as well as ‘to love’, contrast-

ing with the frequent extension in the southern Balkans of *mang-* ‘to beg, to demand’ to mean ‘to want’. The verb *pir-* ‘to receive, to get’ reminds us of Domari *par-* ‘to take’. Like other dialects of Greece, PR shows an innovation in the creation of a possessive verb ‘to have’, *ther-*, from ‘to hold’. Other diagnostic lexical features include the presence of *ayav-* ‘to understand’, both *thaber-* and *phaber-* ‘to burn’, *vaker-* ‘to speak’, and *phirav-* ‘to open’, the loss of **balo* ‘pig’ and use instead of the diminutive derivation *baličo* (originally ‘piglet’), and the emergence, by analogy to Greek, of *phinjov-* ‘to be called’, from *phin-* ‘to say’.

In the domain of body parts and related expressions, we find loss of the items **asva* ‘teardrop’, and **men* ‘neck’, but otherwise a rather conservative formation, by and large. Table 1 provides an overview.

In addition to Greek loans, PR shows a significant inventory of Turcisms, such as *sahati* ‘time’, *sabahi* ‘morning’, *behari* ‘spring’, *džepi* ‘pocket’, *péndžeri* ‘window’, *maxalas* ‘neighbourhood’, *fustáni* ‘dress’, *kundúra* ‘boot’, some of which will have been borrowed via Albanian, as well as Albanianisms, such as *déti* ‘sea’, *lóti* ‘teardrop’, *tózi* ‘sand, dust’. Speakers are often conscious of the latter, and describe them as *xoraxano* or *arvanitiko*. Both Turcisms and Albanianisms are quite common however in the local Greek varieties, too, and Greek may have been the source for the adoption of some of them.

TABLE 1. Body parts and related terms

eye	<i>jakh</i>	hair	<i>bal</i>
teardrop	<i>lóti, đákri</i>	bodyhair	<i>(d)zar</i>
neck	<i>kori</i>	fur	<i>morti</i>
throat	<i>kori</i>	skin	<i>cépa, morti</i>
mouth	<i>muj</i>	hide	<i>morti</i>
tongue	<i>čhib</i>	leather	<i>morti</i>
lip	<i>vošt</i>		
ear	<i>kan</i>	breast	<i>čuči</i>
arm	<i>vast</i>	leg	<i>čang</i>
hand	<i>bádza, vast</i>	foot	<i>čang</i>
elbow	<i>kóči</i>	knee	<i>kóči</i>
armpit	<i>kak</i>	hip	<i>gófos</i>
finger	<i>jísti</i>	thigh	<i>búti</i>
fingernail	<i>naj</i>		
belly	<i>por</i>	soul	<i>gi</i>
stomach	<i>stomáxi</i>	heart	<i>gi</i>

5. Phonology and historical phonology

On the whole PR shares the system that is common to Romani dialects of the Balkans, and is inherited directly from ER. The vowel system consists of five vowels, /i e a o u/, with a slight tendency toward centralisation of /o/ in some unstressed positions, and a tendency toward raising of /e/ to /i/, and to a lesser extent also of /o/ to /u/. The system of stops shows voice opposition, as well as distinctive aspiration, in the series /b d g p t k ph th kh/, with affricates /c č ħ dz dž/. Sibilants are /s z š/, with a tendency toward palatalisation of /s/, and occasional realisation of the affricate /dž/ as /ž/, as well as alternation of /dz/ and /z/: *dzet/zet* ‘oil’, *dzar/zar* ‘body hair’. Other fricatives are /f v x h/, nasals are /m n ŋ/, liquids are /l r/, and there is a palatal semi-vowel /j/. Under Greek influence, there is occasional (though rare) alternation of /b/ and /v/, as well as occasional alternation of /x/ and /h/: *kaxni, kahni* ‘hen’. The Greek fricatives /γ ç θ ð/ are usually preserved in contemporary Greek loans.

In historical perspective, then, PR preserves all of the ER phonemes, with the exception of *ř, which in some dialects of Romani is continued as a uvular or even a retroflex, but in PR merges with /r/. There appear to have been few additions to the original system, disregarding the preservation of Greek fricative phonemes in contemporary Greek loans. The stress pattern is also conservative, with stress falling on the final grammatical inflection morpheme of the word, whereby later inflectional markers, notably vocative endings, remoteness markers on the verb (-*as*), and Layer II case markers (-*ke*, -*te*, etc.) remain unstressed. As is generally the rule in Romani, nominative inflection endings of loan nouns are also unstressed: *džép-i* ‘pocket’, < Turkish *cep*, cf. *đáskalos* ‘teacher.NOM’, from Greek, but oblique *đaskalós-*. In verbs, the perfective marker -*il-* also remains unstressed, and stress falls on the syllable preceding it: *bárido* ‘he grew’ (but *barjovél* ‘grows’), *našávdiljom* ‘I was lost’ (but *našávdjováva* ‘I am being lost’), *darándiljom* ‘I feared’, etc.

Several phonological processes that came into motion already in ER, some perhaps even at an earlier, Proto-Romani, stage, have continued to shape the individual dialects in various ways (cf. Matras 2002: 64–71), and I shall now survey the behaviour of PR in respect of these developments. The prothesis of *v-* (cf. already ER *vast* ‘hand’ from **ast*), spreads in PR to *vando* ‘egg’, *varo* ‘flour’, *vasjav* ‘mill’, *vaver* ‘other’, as well as the labial *vudar* ‘door’ and *vošt* ‘lip’, but unlike other dialects it does not affect *angar* ‘coal’, *haz-* ‘to lift’, *učo* ‘high’, *jiv* ‘snow’, or the personal pronouns, *ov* ‘he’, *oj* ‘she’, *ol* ‘they’. By con-

trast, prothetic *j-* appears in the dialect only in a single word, *jiv* ‘snow’, in addition to the Pan-Romani forms (ER prothesis) *jakh* ‘eye’, *jekh* ‘one’ and *jag* ‘fire’. Prothetic *a-* is found in a limited number of words: *ares-* ‘to arrive’, *arakh-* ‘to find’, and *anav* ‘name’, but not in *lav* ‘word’, *bijav* ‘wedding’, or *šun-* ‘to hear’, thus setting PR apart from developments in the northern Balkans and in Vlax. Inherited initial *a-* is generally retained, with the exception of *mal* ‘friend’.

The ER cluster **n(d)ř* from MIA *nd* appears throughout as /nd/: *vando* ‘egg’, *kando* ‘thorn’, *pindo* ‘foot’, *mando* ‘bread’, *mindu* ‘my’ and by analogy *tindo* ‘your’ (while ER *parno* ‘white’ and *xarno* ‘short’ are continued). Isolated **ř* continues as /r/ in *varo* ‘flour’, as well as *rom* ‘husband’, *roj* ‘spoon’, and so on.

Jotation around morphological boundaries is generally preserved, but is sometimes optional: *panjesa* ‘with water’, *geljom* ‘I went’ but *dikhlom* ‘I saw’ alongside *dikhljom*, *kerdjom* ‘I did’ alongside *kerdom*. Analogous jotation is common in feminine nouns: *romacelja* ‘Romani girls’, *phenja-* ‘sister.OBL’. Occasionally, though not frequently, there is umlaut in the perfective of the 1SG: *aljom*, *alem* ‘I arrived’. In the copula, jotated segments shift from a dental to an alveo-palatal sibilant: *me išom/išjom* ‘I am’, *tu išan/išjan* ‘you are’. The trigger appears to be a similar development /si/ > /ši/ in the local Greek dialect.

The ER phonemes /s/ and /h/ are thought to have been interchangeable in intervocalic position in grammatical endings—the conjugation endings of the 2SG and 1PL, *-esa* and *-asa* respectively, the remoteness tense marker **-asi*, and the instrumental singular case endings, *-esa* and *-asa*—as well as in parallel sets of the copula. Dialects that have generalised forms in *-h-* (often shifting to *-j-*, or being deleted altogether) are found to the northwest of Parakalamos, among both Arli (Balkan) and southern Vlax varieties in northern Macedonia, Kosovo and Serbia, but are also reported from Serres, in northeastern Greece (Sechidou 2002). Overlapping with the *h*-zone to the northwest is also a region in which word-final /s/ is aspirated (*dives* ‘day’ > *diveh*), and a more contained zone in which we find aspiration of /s/ in pre-consonantal position (*leske* ‘for him’ > *lehke*).

In PR we find variation, which may well represent the original ER state of affairs. Spirantisation is optional in grammatical endings in intervocalic position, as well as in the remoteness marker *-as* (inherited from a marker **-asi* in which the sibilant was in intervocalic position). With some speakers we find only marginally forms like *džakerahas* ‘we were waiting’, *ama tu t'avehas idžara* ‘if you had come yesterday’. Alternation appears therefore

to be individual, and irregular rather than context-dependent, and in the comparison between speakers' responses to the questionnaire elicitation we often find variants such as *tu na kamesa te džas ti poli* 'you didn't want to go to town' alongside *tu na kameha te džas ti poli*; *theresasi dikhlo* 'you would have seen her' alongside *therehasi dikhlo*; *njek čikesa* 'with a hammer' alongside *njek čikeha*; *isi očhavdo o čhavo kuvertasa?* 'is the child covered with a blanket?' alongside *očhávđilo o čhavo e kuvertaha?*

Alternation is also found in word boundaries, in inflectional endings such as the past tense of the 3SG *mukhljas* 'he left', but occasionally *mukhlah ov* 'he[is the one who] left', the short present of the 1PL *prepi te džas othe* 'we must go there' alongside *prepi te džah othe*, or the third-person copula, *ova si mo kher* 'this is my house' alongside *ova hi mo kher*. There are also occasional instances of alternation in preconsonantal position: *soske* 'why' beside *sohke*; *e ruveskoro dand* 'the wolf's tooth' beside *e ruvehkoro dand*.

Through Greek influence, there is a rather young tendency toward palatalisation and word-specific fricatisation of /g/ in positions preceding /i/, e.g. *g'isti* alongside *jisti* 'finger', *g'ilavava* alongside *jilavava* 'I sing', and a strong palatalisation of /ki/ to /k'i/ or even /çi, či/ in *k'il* alongside *čil* 'butter'. On the whole, however, stops are maintained in the relevant positions, as in *kiral* 'cheese', *kin-* 'to buy', *g'ili* 'song', *g'iv* 'wheat', *gi* 'soul'. With dentals in comparable positions there is no palatalisation: *dives* 'day', *godi* 'mind', *buti* 'work'. The one exception in *čiknoro* 'small, little' (< **tiknoro*), possibly a contamination with *čika* 'a little, few'.

Marginally, there is a tendency toward reduction of /v/ in intervocalic position: *lav* 'word', PL. *laimata* < **lav-imata*, *džuvel* 'woman', OBL. *džula*. This will have led to the word-specific contraction in *del* 'God' (< **devel*). Clusters are preserved in *grast* 'horse', *kaxni* 'chicken', *čergeni* 'star', *angrosti* 'ring'. The raising of vowels alluded to above has led to an actual phonemic shift only in *phin-* 'to say' (< **phen-*).

6. Nominal forms

6.1. Noun derivation and inflection

The diminutive marker *-or-* is not found in a large number of words. In fact, its most frequent attestation is in *čikoro* 'small', an adjective. The rather unique marker *-in-/ilin* deriving names of fruit trees is attested in *ambrolin* 'pear tree' and *phabilin* 'apple tree'. There is no attested genitive derivation of either ordinary or abstract nouns. The nominalisation marker is *-ibe*, which

derives nouns both from verbs (*khelibe* ‘dance’, *bašalibe* ‘music’) and from adjectives (*džungalibe* ‘ugliness’, *guldibe* ‘sweetness’). Greek-derived *-imos* is not attested.

Layer II case endings, which are suffixed to nominal and pronominal oblique forms (Layer I endings), are the usual dative *-ke/-ge*, ablative *-tar/-dar*, locative *-te/-de*, and instrumental *-sa*, the latter showing regular voice assimilation in the plural, and elsewhere following /n/, to *-za* (*olenza* ‘with them’, *manza* ‘with me’). The genitive has ‘short’ forms in *-k-* on possessive pronouns—*olesko* ‘his’, *olako* ‘her’ *olengo* ‘their’—and ‘long’ forms in *-kVr-* in nouns, usually, though not always, showing introflexion (i.e. the adjectival inflectional ending of the genitive is copied into the vocalic segment that is internal to the genitive morpheme *-kVr-*): *e ruveskoro dand* ‘the wolf’s tooth’, *o danimata e ruveskere* ‘the wolf’s teeth’.

The vocative forms are M. *-éa* for vocalic stems and *-a* for consonantal stems (*phuréa!* ‘old man!’ to *phuro*, *phrála!* ‘brother!’ to *phral*), F. *-e* (*phurije* ‘old lady!’, *phéne!* ‘sister!’), and PL. *-ale(n)* (*čhajalen!* ‘girls!’, *phenjale!* ‘sisters!’, *čhavale!* ‘boys!’). Masculine nouns that end in a consonant take the vocative ending in the singular only, and their regular (Greek-derived) plural ending in the vocative plural: *phrála!* ‘brother!’, but *phralimata!* ‘brothers!’. The words for ‘mother’ and ‘father’ both show the same vocative ending, *-e*, indicating an extension of the feminine form to the masculine in this case: *dáde!* ‘father!’, *dáje!* ‘mother!’

The outstanding feature of PR nominal declension is the tendency to adopt Greek-derived (so-called ‘athematic’ or ‘xenoclitic’) nominative inflection markers into the class of consonantal masculine nouns. There are two expressions of this tendency. The first is in the regular adoption of *-imata* as the plural of consonantal masculine nouns: *vast* ‘hand’, PL. *vastimata*, *dives* ‘day’ PL. *divesimata*, *berš* ‘year’, PL. *beršimata*. The second is in the shift in class affiliation of some consonantal masculine nouns of pre-European origin to the class of European nouns in *-i*, expressed by the addition of an unstressed *-i* class inflection marker: *kóči* ‘knee’ (< **khoč*), *kóri* ‘neck’ (< **koř*). The vocalic class of pre-European masculine nouns continues the ER inflection pattern (see Table 2). The small class of masculines in *-oj* is a mixed class.

Feminine nouns in *-i*, such as *bibi* ‘aunt’, *gili* ‘song’, *luludi* ‘flower’, and so on retain jotation in the oblique and plural. Jotation also appears for the noun *suv* ‘needle’, PL. *suvja*, a development that Elšík (2000a) reconstructs already for ER. Analogous jotation continues in PR to include animate feminine nouns that end in consonants, such as *phen* ‘sister’ PL. *phenja*, *džuvel*

TABLE 2. Declension classes of pre-European nouns

Gender	Example	NOM.SG	OBL.SG	NOM.PL	OBL.PL
m.	<i>čhavo</i> ‘child’	-o	-es-	-e	-en-
M.	<i>murš</i> ‘man’	–	-es-	-imata-	-en-
M.	<i>šošoj</i> ‘rabbit’, <i>roj</i> ‘spoon’	-oj	-(oj)es-	-a	-(oj)en-
F.	<i>gili</i> ‘song’	-i	-ja-	-ja	-jen-
F.	<i>phen</i> ‘sister’, <i>suv</i> ‘needle’	–	-ja-	-ja	-jen-
F.	<i>jakh</i> ‘eye’, <i>džuv</i> ‘flea’	–	-a-	-a	-en-

‘woman’ PL. *džuvlja*, but does not appear to include inanimates, nor animals or bodyparts, cf. *džuv* ‘flea’ PL. *džuva*, *jakh* ‘eye’ PL. *jakha*.

The attested athematic or European (‘xenoclitic’) classes are masculines in *-os* (*radjos* ‘radio’, *kafenijos* ‘café’), by far the most common class of loans, masculines in *-i* (*dukjáni* ‘shop’, *péndžeri* ‘window’), which tends to encompass primarily Turcisms and Albanianisms, feminines in *-a* (*míza* ‘ant’, *kárta* ‘letter’), and a rare class of masculines in *-as* (*maxálas* ‘neighbourhood’). They form their oblique by stressing the final inflectional segment that allows stress in the pre-European component: *pučhljom e daskalós* ‘I asked the teacher’, *aftokindóske* ‘for the car’, *kapelósa* ‘with a hat’. Plural formation is usually with *-a* for the class of masculines in *-os* and *-i* (*aftokindos* ‘car’ PL. *aftokinda*, *klíði* ‘key’ PL. *klíðja*), and in *-es* for the class of feminines in *-a* (*míza* ‘ant’ PL. *mízes*), though Greek plurals are commonly retained (cf. *práymata* ‘things’). In the plural oblique, loan nouns are integrated into the inherited inflection pattern. Greek nouns may also appear as insertional switches, with no adaptation at all.

6.2. Adjective derivation and inflection

The adjectival-like prefix *bi-* is productive, as in *bitukoro* ‘without you’ (note the secondary genitive formation, modelled on the third person *-koro*). Attested adjectival derivation markers are *-al-* as in *džungalo* ‘ugly’, *-an-* as in *ladžani* ‘shy’, *bengani* ‘upset’, *darano* ‘frightening’, *-un-* as in *kaštuni* ‘made of wood’, and with European loans *-ikan-* as in *polikani* ‘golden’. Ethnicity is generally expressed through *-ikan-* (*romacilikano* ‘Gypsy’) or *-itik-* (*arvanitiko* ‘Albanian’).

Adjectives take the nominative inflectional endings M. *-o*, F. *-i*, PL. *-e*, and full nominal case inflection in the other cases. The latter is a unique feature

of PR, paralleled only in the Dolenjski dialect of Romani in southern Slovenia (Cech and Heinschink 2001: 159), in North Russian Romani (Wentzel 1980), and in Lithuanian Romani (Tenser 2003): *kamama e parnes e grastes* 'I want the white.OBL horse.OBL', *dživdilom jek bareste khereste* 'I lived in a big.LOC house.LOC', *tho te vast tatesa panjesa* 'wash your hands with hot.INSTR water.INSTR'. Predicative adjectives agree in nominative gender: *o kher e du-jengoro e phralengoro isi but čikoro* 'the house of the two brothers is very small'.

Borrowed adjectives may show an insertion *-n-* or Greek-derived derivation in *-ikan-*, and inflect like inherited adjectives: *i polikani angrusti* 'the golden ring'. Quite commonly, however, Greek-derived adjectives retain their Greek inflection: *dikhlem je periargo kher to gav* 'I saw a strange house in a village', *ov isines but plusjus* 'he was very rich', *o čhave mi bibjakere isi plusi* 'my aunt's children are rich', *ov si etimos kana šan tuja etimi* 'he is ready when you are ready'.

The comparative is formed analytically with the preposed particle *mo*, an Albanian loan: *olesko kher si mo baro mi kherestar* 'his house is bigger than my house'. The superlative is the determined form of the comparative: *o mo čikoro čhavo* 'the youngest son', *o mo lačo xabe* 'the best food'.

The numerals 1–10 and 100 are those common in Romani, with no changes (*jekh, duj, trin, štar, pandž, šov, epta, oxta, enja, deš, šel*). The word for 'half' is *opaš*. Between 11–19, the conjunction *-u-* is used: *deš-u-jekh* 'eleven'. In combinations above 20, no conjunction is employed (*trijanda-duj* 'thirty two'). Tens, beginning in 20 (i.e. 20, 30, 40 etc.) are all borrowed from Greek, as is 1000. Like other adjectives, numerals too take full case inflection: *šundom oleske mi dujendar mi phralendar* 'I heard about him from my two sisters', *kamamas te džavas ti poli olenza i trinenza i muršenza* 'I wanted to go to town with those three men', *džanav e daja olengere štarengere čhajengere* 'I know the mother of those four girls'.

There are however a number of exceptions. In the accusative, the nominative form of the numeral is used: *therav duj phenja* 'I have two sisters', *i dživel therel trin čhavan* 'the woman has three sons'. In temporal expressions, numerals are also exempted from case agreement: *dživdilom to gav pandž beršenge* 'I live in the village for five years'. The numeral '1' finally, which may also serve as an indefinite article, may either show full case agreement, as in *jekhesa kaljardesa* 'with one/a soldier', or just the generic adjectival oblique endings, M. *-e*, F. *-i (-e)*, PL. *-e*, as in *oj phirela palal jeke muršestar* 'she is walking behind a man'.

6.3. Personal, reflexive and enclitic pronouns

Deictic personal pronouns are *me* 'I' (oblique *man-*) and *tu* 'you' (oblique *tu-*, accusative generally also *tu*, seldom *tut*), with emphatic extensions *mejla tuja*, and in the plural *amen* 'we' (oblique *amen-*), and *tumen* 'you.PL' (oblique *tumen-*). The emphatic forms are used for contrast: *ov si etimos kana šan tuja etimi* 'he is ready when you are ready'. The first person singular has an object clitic variant *ma*: *tu dikhljan ma* 'you saw me'. Occasionally, shortened enclitic forms for the plural deictic pronouns are also encountered: *ov dikhljah ame* 'he saw us', *ov dikhljas tume* 'he saw you.PL'.

The possessive deictic pronouns are *mindo* 'my', *tindo* 'your', and their short variants *mo* and *to* respectively, *amaro* 'our' and *tumaro* 'your.PL'. Possessives take adjectival inflection: *mo čhavo* 'my son', *mi phen* 'my sister', *me phralimata* 'my brothers'. However, only long forms of the possessives appear to take full case inflection: *therav mindes dženes* 'I have a relative [lit. my person]', *i istoria isi njekeske mindeske dženeske* 'the story is about a relative of mine'. Short possessives take the generic adjectival oblique endings, M. -e, F. -i (-e), PL. -e: *tu aljan me khereste* 'you came to my house', *geljom te dikhav mi dajá* 'I went to see my mother', *čumiz ti phenja!* 'kiss your sister!', *tho te vast!* 'wash your hands!'. Doubling of possessive pronouns in combination with other adjectives is common: *me duj me phenja* 'my two sisters', *mo baro mo phral* 'my big brother'. The combination of long and short possessive is used to emphasise exclusive (reflexive) ownership: *sigo ka me therav but love ka te kinav mindo mo aftokindos* 'soon I will have enough money in order to buy my own car'.

The third person (anaphoric) pronouns are *ov* 'he', *oj* 'she', and the archaic *ol* 'they'. This matches the series of assumed Proto-Romani remote demonstratives **ova*, **oja*, **ola* which had been shortened to serve as anaphoric pronouns (see discussion in Matras 2002: 106–12). In ER, the 3PL appears to have had a variant **on*, which is the form that is continued in most Romani dialects outside the southern Balkans, while in the region itself we find both *on* and *ol* (cf. Boretzky 1999: 230). The conservatism of the PR set coincides with the retention in PR of the old demonstrative set *ova*, *oja*, *ola* (see 6.4), alongside the renewed forms *okova*, *okoja*, *okola*. The oblique third-person pronoun set is also conservative, showing retention of the initial vowel: M. *oles-*, F. *ola-*, PL. *olen-*. Possessive anaphoric pronouns are based on the short genitive formation in *-k-*, and are often, under Greek influence, preceded by a definite article: *o olesko kher isi but baro* 'his house is very big'.

The reflexive third-person pronoun is *pes-*: *džanen pes o njek vaveresa* ‘they know one another’, *theren olenge práymata pal peste* ‘they have their goods with them’. The reflexive has the possessive form *po* that is common in the southeastern European dialects of Romani: *na boruse te arakhelas po kher* ‘he couldn’t find his (own) house’.

Subject clitic pronouns *-lo*, *-li*, *-le* are retained, as in most other Romani dialects of southeastern Europe, only with copula predications, where they accompany free-standing pronouns: *akate si-li oj!* ‘here she is!’.

The extraordinary feature of the PR pronominal system is the presence of unstressed third-person enclitic object pronouns *-os*, *-i*, *-ele* (Table 3). The system is fully productive, and although free-standing object pronouns/demonstratives can be used as deictic forms (*dikhava oles* ‘I see that one’), cliticisation (*dikhavos* ‘I see him’) is the preferred option for anaphoric reference to direct objects, including reference switches: *olako dad na mukhelos te phandreveli* ‘her father won’t let him marry her’.

The cliticisation of object pronouns in Romani has so far been known only from Abruzzian Romani, where it is thought to have been influenced by the local Italian dialects. At first glance, it appears attractive to view the object clitic pronouns of PR as a local innovation, which might have derived through attachment to the verb of the following full object pronoun M. **oles* > *-os*, and PL. **ole* > *-ele*, with the feminine form changing from original **ola* > **-a* to *-i* by analogy to feminine nominal endings (in the noun and adjective inflection as well as the definite article). This scenario remains a likely possibility, although unlike the Abruzzian case there is no obvious trigger in a contact language which might have set the development in motion.

Another, clearly more remote and less attractive scenario, is that we are dealing with extremely archaic, Proto-Romani pronominal clitics, of the type attested in Domari (cf. Matras 1999b): *karda* ‘he did’, *kard-os-is* ‘he did it’ (cf. *kard-om-is* ‘I did it’). Here, the clitic *-os* stands for the nominative form

TABLE 3. Third-person enclitic object pronouns

Meaning	...him	...her	...them
‘I see ...?’	<i>dikhav-os</i>	<i>dikhav-i</i>	<i>dikhav-ele</i>
‘I saw ...?’	<i>dikhljom-os</i>	<i>dikhljom-i</i>	<i>dikhljom-ele</i>
‘s/he sees ...?’	<i>dikhel-os</i>	<i>dikhel-i</i>	<i>dikhel-le</i>
‘s/he saw ...?’	<i>dikhljias-os</i>	<i>dikhljias-i</i>	<i>dikhljias-ele</i>

that represents the third-person agent (cf. *-om* for the first-person agent), while *-is* stands for the oblique form representing the third-person object. Such patterns are common among the Indo-Iranian frontier languages, and the cliticisation of subject pronouns to the old particples to form the perfective conjugation in Romani (*kerd-j-om* 'I did', *kerd-j-as* 's/he did') attests to the fact that Proto-Romani had undergone a similar development (cf. Matras 2002: 145–51). The Domari third-person subject and object clitic *-os* resp. *-is* derives, like the Romani 3SG subject ending *-as*, from the MIA enclitic object pronoun *-se*. The same form resembles the PR 3SG.M. object clitic *-os*, though if this etymology were to be adopted, the emergence of F. *-i* and PL. *-ele* would still have to be explained as analogies: to the nominal-adjectival inflection, in the case of the feminine singular form *-i*, and, more of a challenge, to the oblique plural demonstrative *olen* with subsequent dropping of the final consonant and progressive vowel assimilation, in the case of the plural form *-ele*.

The first scenario has two obvious advantages: It is more economical, in that it derives two of the forms directly from attested, postposed object pronouns, with relatively little speculation in respect of the underlying erosion or sound simplifications. And, as an innovation rather than an archaism, it is easier to defend in the broader geographical context of Romani dialects, where PR stands almost alone in showing this type of development, and alone, so far, in showing this specific pattern of forms.

6.4. *Demonstratives*

The emergence of demonstratives in Proto-Romani and ER saw a renewal of the old set, M. **ava/ova*, F. **aja/oja*, PL. **ala/ola*, through prefixing of the local deixis *adaj*, *akaj* 'here' and *odoj*, *okoj* 'there', to a four-term system: *adava*, *akava*, *odova*, *okova* etc. (see Matras 2002: 103–12). The latter are the forms that were inherited into the dialects from ER. They are often retained in the southern Balkans, as well as in the extreme periphery (British, Iberian, and southern Italian Romani). PR maintains the four-term system that is by and large typical of Romani, but shows a rather unique combination of archaic forms in *ava/ova* etc. for the 'default' set ('this', 'that'), and forms in *akava/okova* etc. for the 'specific' set ('this one here', 'that one there') (Table 4). The ER set in **ada-* has apparently disappeared, but a trace of it is left in the expression *pe ada* 'therefore'.

Under Greek influence, PR demonstratives are usually accompanied by a definite article in the position immediately preceding the determined noun:

TABLE 4. PR demonstratives

		M.		F.		PL.	
		NOM.	OBL.	NOM.	OBL.	NOM.	OBL.
Regular	Proximate	<i>ava</i>	<i>ales-</i>	<i>aja</i>	<i>ala-</i>	<i>ala</i>	<i>alen-</i>
	Remote	<i>ova</i>	<i>oles-</i>	<i>oja</i>	<i>ola-</i>	<i>ola</i>	<i>olen-</i>
Specific	Remote	<i>akava</i>	<i>akales-</i>	<i>akaja</i>	<i>akala-</i>	<i>akala</i>	<i>akalen-</i>
	Proximate	<i>okova</i>	<i>okoles-</i>	<i>okoja</i>	<i>okola-</i>	<i>okola</i>	<i>okolen-</i>

ova o čhavo našadah but love ‘that boy lost a lot of money’, *kamesa te khes olenza i trinenza i romaceljenza?* ‘do you want to play with those three Gypsy girls?’, though this is subject to variation, cf. *ova kher isi mo paše* ‘this house is closer’, *ola čikore čhave ale but dural* ‘those small boys cam from far away’. Like PR adjectives, demonstratives take full case inflection: *dos nje kotor guldibe oles e čhaves* ‘give a piece of cake to this boy’, *o kondos oleskero čhaveskero si pharaldo* ‘this boy’s shirt is torn’. Note that demonstratives that are inflected for the genitive case, are distinguished from anaphoric possessive pronouns through the use of long forms of the genitive in the first, as in *oleskero* ‘that one’s’, and short forms in the second, as in *olesko* ‘his’. Demonstratives may take full case inflection even if case is expressed on the noun by a preposition, rather than a Layer II case ending: *kapjos dživel aleste to kher* ‘somebody is living in this.LOC house’.

6.5. Interrogatives

Here we find the usual conservative forms *kon* ‘who’, *so* ‘what’, *sar* ‘how’, *kaj* ‘where’ and *kana* ‘when’, as well as *soske* ‘why, for what reason’. The quantity interrogative is *abor* ‘how much’.

6.6. Indefinites

Indefinite expressions are characterised on the one hand by the conservative retention of the set of forms in *-n(j)i-* and *-muni* (cf. Elšík 2000b) and the ontological markers *k-* and *č-* in the set of negative indefinites: *konjek*, oblique *kanjikas-* ‘nobody’, *čumuni* ‘nothing’, *katemuni* ‘nowhere’. The specific determiners for ‘some’ are *čika* and *xandi*, both originally meaning ‘a little’. The specific and universal indefinites are otherwise largely borrowed from Greek: *káti* ‘something’, *kápo* ‘somewhere’, *kápote* ‘sometime’, *kápos* ‘somehow’, *káthe* ‘every’, *káthe forá* ‘every time’, *pánda* ‘always’, *poté* ‘never’. For ‘somebody’ we find Greek *kápjos*, but in the oblique the inertied *njeke-* (< **ni-jekhe-*

‘any-one’), and for ‘everybody’ the combination *káthe džene* ‘every person’. Free-choice indefinites include *saro* ‘anybody’ (from ‘all’), and for ‘anything’ inherited *čumuni* alongside Greek *káti*.

The set of specific indefinites is employed to express propositional entities that are specific known (*arakhljom káti* ‘I found something’), specific unknown (*kápjos dživel aleste to kher alá na džanav kon isi* ‘somebody lives in this house, but I don’t know who it is’), conditional (*ama te dikhes káti, phin mange* ‘if you see somebody, tell me’) and irrealis (*ov kamelas te arakhe-las njekesa akate* ‘he wanted to meet with somebody here’). The negative series is used in questions (*avela čumuni?* ‘is something happening?’), and in direct and indirect negation (*čumuni na alo* ‘nothing happened’, *na džanav kanjikas akate* ‘I don’t know anybody here’).

6.7. Definite and indefinite articles

Alongside *jek* < ‘one’, the more widely used indefinite article is the renewed *njek*, apparently from *n(j)i-jekh* ‘any-one’. The appearance of the indefinite article is variable, and it is often redundant when introducing a non-topical entity: *oj bičhavdas mange lil* ‘she sent me a letter’. As in attributes, there is a tendency toward full case inflection of the indefinite article *njek*: *i istoria isi njekeske mindeske dženeske* ‘the story is about a relative of mine’. However, full (Layer II) case inflection remains optional, and alternates with generic oblique adjectival inflection: *o čikoro čhavo garádilo pala njek rukhestar* ‘the little boy hid behind a tree’, *ov mardas i kahnja njek čhurjasa* ‘he killed the chicken with a knife’. With the indefinite *jek*, no Layer II markers are found, and even the generic adjectival oblique is optional: *oj phirela palal jek muršestar* ‘she is walking behind a man’, but *dikhlem jek grastes to veš* ‘I saw a horse in the woods’, *kana šomas terni dživdilom jek bareste khereste ti poli* ‘when I was young I lived in a large house in the city’.

The definite article shows three distinct forms, with identical formation in the M.SG and PL. on the one hand, and in the F. nominative and oblique, on the other (Table 5).

Table 5. Definite articles

	NOM.	OBL.
M.SG	o	e
F.SG	í	í
PL.	o	e

Conservative consonantal oblique forms in *-l-* are not attested; however, the corpus has one recorded token of a conservative nominative form in *ov del* ‘God’, perhaps lexically conditioned. Greek articles may accompany Greek lexical insertions: *rotindom to daskalo* ‘I asked the teacher’, *oj našavdja but džene oleste ton polimos* ‘she lost many relatives in that war’.

7. Verb inflection

7.1. Valency marking

The inherited valency-increasing morphemes *-av-* and *-ar-/-er-* are lexicalised, and no longer productive in PR. The older of the two is more widespread and appears, as in other dialects, in verbs like *sikhav-* ‘to show’, *phirav-* ‘to open’, as well as in analysable but nevertheless lexicalised verbs such as *čarav-* ‘to feed’ (from *čar* ‘grass’), *darav-* ‘to frighten’ (cf. *dara-* ‘to fear’), *nakhav-* ‘to cross’ (*nakh-* ‘to pass’), *gilav-* ‘to sing’ (*gili* ‘song’). The younger form is attested more rarely, e.g. in *amen phaberdam e jela* ‘we burned the food’, *dinjaresa ma* ‘you drive me crazy’. The only derivations with *-ker-* are *vaker-* ‘to speak’ and *džaker-* ‘to wait’, both MIA formations rather than derivations with the late ER **-aker-*. The productive causative is expressed analytically: *oj kerdasos te našel* ‘she made him run away’.

The mediopassive, on the other hand, remains, as in the contact language Greek, extremely productive. The pattern is, as in ER, to add *-jov-* to the adjectival, nominal or verbal stem: *kana o dand e čhaveskere barjovena o vošt šuvljovena* ‘when the child’s teeth grow the lips swell’, *tavjovela o pani?* ‘is the water boiling?’. In this fashion, mediopassives can also be derived from transitive derivations, by attaching *-jov-* to the perfective transitive stem: *našadjovava* ‘I am getting lost’, past *našávdilom*, *garavdjovava* ‘I am hiding (itr.)’, past *garávdiljom*.

7.2. Loan verb adaptation

ER is assumed to have incorporated Greek tense-aspect markers along with borrowed Greek stems, resulting in a partial adaptation of Greek inflection patterns. These remained productive even after the breakaway from Greek-speaking communities/ territories, and served to incorporate verbal loans from other contact languages. This is known in some studies as ‘athematic verb morphology’ (Bakker 1997, Hancock 1995). However, the original inventory was severely simplified and reduced, with dialects choosing among the

various markers, some opting selectively for Greek-derived markers, some for pre-European valency markers, some for combinations. The overall distribution pattern of loan verb adaptation markers in the present-day dialects is geographical (see Matras 2002: 128–34), with dialects in the geographical margins showing greater diversity of forms, and so greater conservatism.

In PR, the most widespread loan verb adaptation marker is *-Vz-*: *yrazava* ‘I write’, *vojθizava* ‘I help’, *parakalizava* ‘I thank’, *ayapezava* ‘I love’. Note that the markers are productive, and are assigned to Greek stems irrespective of the current or original Greek inflection class; cf. Greek *γράφω* ‘I write’, *βοιθώ* ‘I help’, *παράκαλό* ‘I thank’, *αγαπώ* ‘I love’. In the past tense, the marker is dropped, and the default Romani perfective marker in *-d-* appears: *yradom* ‘I wrote’, *vojθidom* ‘I helped’ etc. The predominance of *-Vz-*, which is normally found in the Romani dialects of the Black Sea coast (Crimean Romani, Ursari, Rumelian Romani, and dialects of the Drindari ‘type’ and their outposts throughout northern Bulgaria and into Macedonia), is at first glance somewhat surprising, and reinforces the impression of the relative isolation of PR. In the more well-known dialects of the western or southernmost areas of the southern Balkans, such as Sepeçides, Erli, and Arli, the marker is *-Vn-* or *-in-* (cf. Boretzky 1999). This too is found in PR, but rather sporadically, and only in the past tense: *rotindjom* ‘I asked’, *rotizava* ‘I ask’, cf. Greek *ρώτῶ* ‘I ask’; *sinanindom* ‘I met’, *sinandizo* ‘I meet’, cf. Greek *σινανδῶ* ‘I meet’; *arçinda te phenel* ‘he started to say’.

Another pattern coexists with the one just described. Here, Greek verbs are used with their Greek person-inflection in the present tense, while in the past tense the perfective formation *-isájl-* is employed. The latter can be reconstructed as an ER Greek-derived aorist marker *-is-* followed by the intransitive valency-marking integration morph, and appears to have been used originally only with mediopassives or intransitives. The two patterns may alternate, and it appears that speakers have individual preferences. Thus we find *petázo* ‘I throw’ (cf. standard Greek *πέτο*), past *petasájljom*, alongside *petazava* ‘I throw’, past *petadjom*; *arçindas* ‘he began’ alongside *arçinisájllo*. The perfective in *-isájl-* still predominates however with intransitives and mediopassives: *nevrijazo* alongside *nevrijazava* ‘I become angry’, (cf. Greek *nevriázo*), past *nevrijasájljom*; *pandrévo* alongside *pandrezava* ‘I am getting married’, past *pandresájljom*.

It seems therefore that the patterns of loan verb adaptation in PR are still undergoing levelling and re-organisation. For a state in the not-so-distant past we might reconstruct the following PR markers: present-tense *-Vz-*,

transitive past *-in-* (assigned to the perfective class in *-d-*, hence *-in-d-*), and intransitive past *-isájl-*. This can be inferred from the current distribution, and would coincide with the typical pattern of simplification in Romani dialects, where the valency distinction is generally more likely to be maintained in the past tense, whereas typically just one form is selected for the present (cf. Matras 2002: 131). The ongoing shift is now showing (1) a reduction of the transitive past-tense marker *-in-d-* to just *-d-*, (2) a tendency for *-isájl-* to infiltrate transitive verbs, and (3) a tendency to use Greek inflection with Greek-derived verbs.

The presence of the latter option, available through contemporary contact with Greek, scrambles the resources available for loan verb adaptation in PR significantly. Whatever system existed so far in the dialect is now competing with a license to adopt Greek inflection marking wholesale, in a fashion similar to the integration of Turkish verbs in some of the Romani dialects of Bulgaria, or Russian verbs in some of the North Russian Romani dialects. The process obviously begins with modal verbs, which are conjugated as in Greek throughout, and then infiltrates spontaneous lexical loans as well: *na boro te diavazo soske prepri te vojtzav me daja* 'I cannot study because I have to help my mother', *boris te doljevis* 'you can work', *xrijazome me malen* 'I need my friends', *kon bori te xtizi nje kher xoris karfja* 'who can build a house without nails?'

7.3. Present inflection classes and present concord

Like all other Romani dialects, PR retains the ER split in present stems between consonantal (*ker-* 'to do') and vocalic classes (*asa-* 'to 'laugh'), the latter showing partial assimilation of subject concord markers to the final, vocalic segment of the stem (Table 6).

TABLE 6. Short forms of the present conjugation

	Consonant	Vowel
1SG	-av	-v
2SG	-es	-s
3SG	-el	-l
1PL	-as	-s
2PL	-en	-n
3PL	-en	-n

Long forms of the present conjugation are followed by *-a*. In PR there are usually no contractions in the long forms, with the exception of occasional alternation between /s/ and /h/ in the 2SG and 1PL: *kamesa* alongside *kameha* ‘you.SG want’, *kamasa* alongside *kamaha* ‘we want’. As in many other dialects, the verb *kam-* ‘to want, to love’ shows a 1SG ending in *-am*: *kamam*, *kamama* ‘I want’. With the verb *džan-* ‘to know’, short forms are often used in the indicative: *na džanav* ‘I don’t know’.

In the 3SG, the marker *-i*, borrowed from Greek into ER, appears with Greek-derived modals, as well as with some recent Greek loans. However, unlike other dialects, no clear case can be argued for its continuation from ER into PR, since PR tends to show complete Greek conjugations both with Greek-derived modals and with recent loans (see 7.2, ‘Loan verb adaptation’).

The mediopassive marker is *-(j)ov-* throughout. There is no contraction in any of the persons, and the mediopassives behave like other consonantal present stems: *barjovava* ‘I grow’, *bar(j)ovena* ‘they grow’, etc.

The imperative is generally formed by using the plain present stem, with no additions in the singular, and the regular 2PL conjugation marker in the plural: *xa!* ‘eat’, *xan!* ‘eat!’ (PL). This includes derivatives in *-d-* such as *phand!* ‘shut!’ and *kid!* ‘gather’, mediopassives, e.g. *na marjoven!* ‘do not quarrel!’ (PL), as well as loans, e.g. *petoz o bar!* ‘throw the stone’, *pandreviz sigo* ‘get married quickly!’, and analogous formations such as *čumiz!* ‘kiss!’.

7.4. Perfective inflection classes and perfective concord

Person concord in the perfective shows endings inherited from ER, except for the 2PL, where we find the pattern of partial analogy to the 3PL shared by all Romani dialects of the southern Balkans (both Balkan and southern Vlax), *-en* < **-an* (Table 7).

The personal forms—all except 3SG *-o/-i* and 3PL *-e*—are generally jotted, though occasional de-jotation is encountered as a variant. As elsewhere in

TABLE 7. Perfective subject concord markers

1SG	-om
2SG	-an
3SG	-as/-o, -i
1PL	-am
2PL	-en
3PL	-e

the southeastern European dialects of Romani, the adjectival endings 3SG M. -o F. -i appear with a series of intransitive verbs. These include verbs indicating movement, such as *alo* ‘came’, *gelo* ‘went’, state, such as *bešto* ‘sat’, *rundo* ‘cried’, *darándilo* ‘feared’, mediopassives, such as *bárido* ‘grew’, and loan verbs that are integrated with the intransitive (past-tense) extension ER *-a(vi)l- > -ájl-, e.g. *arčisájlo* ‘began’. The person-inflected form of the 3SG, -as, is sometimes aspirated to -ah, especially when followed by a word with an initial vowel (see Section 5, ‘Phonology’).

PR shows a rather conservative formation of perfective inflection classes, matching those found in quite a number of other dialects of the southern Balkans. The class of verb stems in -r, -l and -n retains the perfective marker -d- (jotated where appropriate): *kerdjom* ‘I did’, *kerde* ‘they did’, *kheldjom* ‘I played’ *khelde* ‘they played’, *čhindjom* ‘I cut’ *čhinde* ‘they cut’. The irregular stems *mer-* ‘to die’ and *per-* ‘to fall’, and the perfective of ‘to go’, belong, as elsewhere, to the class in -l-, having in ER already continued MIA vocalic perfective stems: *mulo* ‘he died’, *pelo* ‘he fell’, *gelo* ‘he went’.

Verbs in -v are generally a diverse and somewhat volatile class in Romani. PR is conservative in maintaining the affiliation of ‘straightforward’ phonological stems in -v to the class of the other voiced consonants, namely those in -d-: *čhivdjom* ‘I put’. This class also comprises transitive derivations with the stem extension -av-: *garavdjom* ‘I hid away’, *daravdjom* ‘I frightened’. The verb *dživ-* ‘to live’ behaves in a similar fashion, though as an intransitive verbs of state and motion, it is then assigned the perfective extension -il-: *dživdiljom* ‘I lived’. Inherited, irregular perfective formations appears for the verbs *rov-* ‘to cry’ (*rundjom* ‘I cried’) and *sov-* ‘to sleep’ (*sutjom* ‘I slept’). As in many other dialects, the perfective of *av-* ‘to come’ is contracted from the ER regular intransitive formation of motion verbs **avil-*, to *aljom* ‘I came’.

The inherited class in vocalic stems shows continuation of ER *-l-, as in *piljom* ‘I drank’ (*pi-* ‘to drink’) and *xaljom* ‘I ate’ (*xa-* ‘to eat’). With psych verbs in -a, however, we find the common, ER pattern of extension involving both the adjectival-participial marker -n- and the intransitive -il-, with stress falling on the root, rather than the conjugation marker: *asándiljom* ‘I laughed’, *darándiljom* ‘I feared’.

Velar and affricate stems are assimilated into the class in -l-, a process that apparently began already in ER: *mangljom* ‘I demanded’ (*mang-* ‘to demand’), *mukhljom* ‘I left’ (*mukh-* ‘to leave’), *dikhljom* ‘I saw’ (*dikh-* ‘to see’), *pučljom* ‘I asked’ (*puč-* ‘to ask’). However, on the final position of the hierarchy of class re-assignment (cf. Matras 2002: 139) we find an archaism in the retention

of **-t-* with stems in sibilants: *beštjom* ‘I sat’ (*beš-* ‘to sit’), *arestjom* ‘I reached’ (*ares-* ‘to reach’).

Monoconsonantal stems show traces of the adjectival-participial extension in *-in-*, which is found sporadically in Romani, and appears to have been at least one of the norms, perhaps a variant, in ER: The verb *d-* ‘to give’ has a straightforward extension in *-in-* (*dinom* ‘I gave’). The verb *l-* ‘to take’ shows an analogy to stems in *-n*, evidently based on the original extension, giving *lindjom* ‘I took’. In the copula, the extension is found in the past tense of the third person, *isine-* ‘was’. The extension *-in-* also appears, as mentioned, with psych verbs (*dara-* ‘to fear’, *darándiljom* ‘I feared’), as well as with the verb *ušt-* ‘to stand up’ (*uštíndiljom* ‘I stood up’).

Mediopassives form, as in ER, a separate class, showing, irrespective of stem phonology, the intransitive ending *-il-*: *báriljom* ‘I grew’, *phúriljom* ‘I aged’. The class also assimilates the intransitive verbs *uštíndiljom* ‘I stood up’ and *dživdiljom* ‘I lived’, as well as psych verbs in *-a*, *darándiljom* ‘I feared’. Irregular perfective stem formations are *nikhivava* ‘I go out’, *nikhistjom* ‘I went out’, and *čumizava* ‘I kiss’, *čumidjom* ‘I kissed’, the latter behaving like a loan verb.

7.5. Copula inflection

The PR copula appears to be in a state of transition and shows considerable variation (Table 8). The initial segment *i-* is volatile, and optional (*išom* : *šom* ‘I am’). Jotation has led to a change in the stem sibilant, to *-š-* (*šom* ‘I am’). There is a tendency to distinguish singular from plural forms through the stem vowel, by dropping jotation in favour of umlaut *-e-* in the first and second persons plural (e.g. *išem* ‘we are’). This partly mirrors the state of affairs

TABLE 8. Forms of the copula

	Pronoun	Present	Past
1SG	me	(i)šom	(i)šomas
2SG	tu	(i)šan	(i)šanas
3SG.M	ov	(i)si/išos	isine(s)/isindo(s)
3SG.F	oj	(i)si/išoj	isine(s)/isindi(s)
1PL	amen	(i)šam/išem	(i)šamas/ (i)šemas
2PL	tumen	(i)šan/išen	(i)šanas/ (i)šenas
3PL	ol	(i)si/išole	isine(s)/ isinas/ isinde(s)

in the perfective conjugation of lexical verbs, where 2PL *-en* is distinguished from 2SG *-an*. The stem extension *-in-* is preserved only in the past tense of the third person, which in turn tends to adopt final *-s* by analogy to the other persons (*isines* ‘he/she was’). The third-person forms may also adopt gender–number distinctions. In the present tense, the palatal stem is extended to the third person, followed by pronominal clitics resembling the enclitic object pronouns discussed above (M. *-os*, F. *-oj*, PL. *-ole*; cf. *išole* ‘they were’). In the past tense, there is an optional regular perfective extension of the copula stem in *-d-*, followed by a regular gender–number agreement marker as found in past participles and adjectives (cf. *isindo* ‘he was’).

A curious feature is the formation of the future tense of the copula in a regular way, with *ka* followed by the present copula: *ka (i)šom* ‘I will be’. Transitions of state—‘to become’—are expressed by *av-*: *ama pjava thud but k'avav but zurali* ‘if I drink a lot of milk I will be very strong’. The presence of *av-* in this function—hitherto identified mainly for Romani dialects in central, western and northern Europe (but see Boretzky 1999 for this area)—confirms that *av-* was, alongside *ov-*, a variant of the transition and non-indicative copula already in ER.

7.6. *Tenses and moods*

The tense formation shows a combination of conservative features, and regional innovations that are typical of many of the Balkan dialects of Romani. A carry-over from ER, the long forms of the present conjugation in *-a* are used in the present tense (*džava* ‘I go/am going’) while the short forms are used in the subjunctive (*boró te džav* ‘I can go’). The remoteness marker *-as* attaches to the short present conjugation to form the imperfect, as in *bešavas* ‘I was sitting’, and to the past conjugation to form the remote or completed past, as in *šundomas* ‘I heard’, or *aljomas* ‘I have become’. In the third person following adjectival-participial endings, it takes the form *-sas*: *sutosas, suti-sas, sutesas* ‘he/she/they slept’.

An innovation that is typical of the dialects of the southern Balkans (including Balkan and southern Vlax), is in the formation of the future, which features a (late) Balkanism, the future particle *ka* (from *kam-* ‘to want’), followed by the subjunctive: *ka kerav* ‘I shall do’. In combination with the remoteness marker, the future particle *ka* can be used to form a conditional mood: *ama theravas xandi love ka davas tu* ‘if I had any money I would give it to you’. The imperfect is used also for the past subjunctive: *ov kamelas*

te arakhelas njekesa 'he wanted to meet with somebody'. The imperfect may also function as a politeness form: *kamamas te džavas* 'I would like to go'.

A further innovation that is limited to the dialects of Greece is the emergence of a perfect tense, employing the verb *ther-* 'to have' (itself an innovation in this region, from *ther-* 'to hold'), in conjunction with the past participle: *ov therel našto* 'he has left'. The periphrastic and synthetic past tenses appear to be interchangeable, cf. *oj na therel aresti akoma* 'she hasn't yet arrived', alongside *oj n'arestjas akoma*. Inflected for remoteness, the auxiliary *ther-* used in a similar format expresses the pluperfect, as in *ov therelas našto* 'he had left', or the counterfactual, as in *ama therelas arakhlo o kher* 'if he had found the house'.

No present participles are attested in PR. Past participles are formed in the usual way, by adding adjectival endings to perfective stems (cf. 7.4). Past participles of loan verbs (so-called athematic participles) show ER Greek-derived endings in *-ime* and *-ome*, which do not inflect, and which have lost the final segment **-n*: *amborome* 'ill', *pandrim* 'married'.

7.7. Verb negation

The negation particle is the inherited *na*: *na džanav* 'I don't know', *tu na dikhljan ma* 'you didn't see me'. The negated future shows the same negation particle, but no future particle: *ama pudela aeras, na nikhavav avri* 'if the wind is blowing I will not go out', *tajça na išom khere* 'tomorrow I will not be at home'. The negation of the imperative is with the inherited particle *ma*: *ma mar man* 'don't hit me'. Greek verbs with Greek inflection take Romani negation: *na boró te diavázo* 'I cannot study'.

7.8. Modals

With the exception of 'want', which is expressed by the inherited *kam-*, modal verbs are borrowed from Greek. The modals *prep-* 'must' and *bor-* 'can' retain Greek tense and where relevant person inflection: *prepi te džas ti poli* 'we must go to town', *éprepe te džakerasas* 'we had to wait'; *na borume te džas palal* 'we cannot go back', *boris te doljevis* 'you can work', *na borusa te phiravavas e vudar* 'I couldn't open the door'. The verbs *arç-* 'start', *stamat-* 'stop', *arez-* 'like' are integrated into Romani inflection: *oj arçinisajli te phinel nje istoria* 'she started to tell a story', *stamatisajlo te del biršindo* 'it has stopped raining', *arezela ma te therav je findžani kafes to sabahi* 'I like to have a cup of coffee in the morning'.

8. Syntax

8.1. Adverbs

Location adverbs continue either the ER stative/directive forms in *-e* or the ablative forms in *-al*, but the forms are polyfunctional, and a distinction between stative, directive and ablative is normally not preserved. We find *andre* 'inside', *avri* 'outside', *paše* 'nearby', *maškare* 'in between, in the middle', and *opral* 'above', *telal* 'below', *anglal* 'in the front', *palal* 'in the back', *perdal* 'away', *dur(al)* 'afar'. The only distinct ablative form appears to be *andar* 'from inside'. Other location adverbs are Greek loans, e.g. *apénadi* 'opposite'. Basic lexical adverbs are derived from adjectives through *-es*: *lačhes* 'well' (cf. *lačo* 'good'). Deictic adverbs are *akate* 'here' and *othe* 'there' (ablative *akatar* 'from here' and *othar* 'from there'), and *aboka* 'so much'. Reversal (spatial) is expressed by *palal* 'back', while repetition (temporal) is expressed by *pále* 'again', an ER Greek-derived loan. Other adverbs tend to be borrowed from Greek. This includes the phasal adverbs, *akóma* 'still, yet' and *pja* 'anymore', and temporal and modal sentential adverbs such as *sixná* 'often', *sígura* 'certainly', *ksafniká* 'suddenly', *fáre* 'entirely', or *kanoniká* 'originally, generally'.

8.2. Case representation

One of the most distinctive typological features of PR and neighbouring dialects of Greece is the emergence of a possessive verb, *ther-* (from *ther-* 'to hold'). The possessive construction thus has the possessor in the subject role, and the possessed in the direct object role: *i džuvél therel trin čhaven* 'the woman has three children'. The external possessor is in the direct object case: *mo kočja dukhana ma* 'my knees hurt', *oleski por dukhalos* 'his belly is aching'. The possessor in local-existential constructions can appear either in the instrumental case, as in *ka o love? olesa* 'where is the money? with him', or with a preposition, in the locative, as in *na therava abuka but love pal mande akana* 'I don't so much money on me now'. The experiencer can remain with no overt expression: *prepi te džas ti poli* 'we must go to town', *therav kati buti ja te kerav* 'I have some work to do'. It can also be in the direct object case: *arezela ma te therav je findžani kafes to sabahi* 'I like to have a cup of coffee in the morning'. The predicate of a promotion to a state takes the subject case: *kamama te avav demosiyráfos* 'I want to become a journalist'.

The direct object shows the typical animacy split. The animate direct object appears in the independent or unmodified oblique (acting as a quasi-

accusative): *geljom te dikhav me dajá* 'I went to see my mother.OBL', *dikh-lom je muršes te phirel to drom* 'I saw a man.OBL walking down the street'. The inanimate direct object takes the same (nominative) case as the subject: *dikh-lom o kher ka vakeresas* 'I saw the house that you spoke [about]', *dikh-lom nje suno darano* 'I saw a frightening dream', *oj džanelas e gilja lačhes* 'she knew the songs well'. Animals are treated as animates: *dikhljem jek grastes* 'I saw a horse', *ol marde e ruves idžara* 'they killed the wolf yesterday', *ov mardas i kahnja njeke čhurjasa* 'he killed the chicken with a knife'.

The object of comparison is expressed by the synthetic (Layer II) ablative case: *oj serdas o polimos mo lačhes sarenendar* 'she remembered the war better than everyone', *olesko kher si mo baro mi kherestar* 'his house is bigger than my house'. The object of equation appears in the nominative: *oj si abuka bari sar išjan tu* 'she is as big as you are'. The object of reference is generally in the dative: *isine nje murš ka pučhelas tuke* 'there was a man who was asking about you', *i istoria isi njekeske mindeske dženeske* 'the story is about a relative of mine'. The recipient is split among different cases, depending on the verb. The recipient of 'to give' appears in the quasi-accusative or unmodified oblique: *boró te dav oles e romaciles xande cigara* 'I can give that man.OBL some cigarettes'. Other recipients, such as those of 'to tell' and 'to send', take the dative: *ama te dikhes kati phin mange* 'if you see something tell me', *oj bičhavdas mange lil* 'she sent me a letter'. The benefactive is generally expressed alternately by the dative or locative: *ov džala to kafenijas te kinel xandi guldibe pe čhavenge* 'he is going to the café to buy some sweets for his children', *pidžardom ola e džula sarinende to gav* 'I introduced that woman to everybody in the village', and *anava olestesupa* 'I am bringing him soup', alongside *anava oleskesupa*. The recipient of 'to give' may also appear in the locative or dative, if it is also regarded as a benefactive of the action: *kindomas nje polikano vraxoli ka ka davos mi dajate* 'I bought a golden bracelet which I shall give [it] to my mother'.

The case of the goal is the dative: *ol marjovena e panjeske* 'they are fighting over water', *nje phuri geli to veš kaštenge* 'an old woman went to the forest for wood'. Dative also expresses the object of reason and cause: *dinas ov abuka but love nje aftokindoske?* 'did he spend so much money on a car?', *ov nevr-ijasájlo manza solaveske* 'he was angry with me because of something'. The source object of the verb 'to ask' appears in the unmodified oblique or accusative: *pučhljom e daskalós ama k'avel to bjav* 'I asked the teacher whether he is coming to the wedding'. Material source is expressed by the prepositional combination *ka + t-*: *ola o guldibe isi kerde kato varo, kati zaxari, kato thud eli*

vande 'this cake is made of flour, sugar, milk and eggs'. Origin is expressed by the ablative: *pirdom ola luludja mi phenjatar* 'I received those flowers from my sister'.

Associative relations are expressed by the Layer II instrumental case, both for the instrumental proper, *tho te vast tatesa panjesa* 'wash your hands with hot water', and for the comitative, *khelava me phralesa* 'I am playing with my brother'. The privative object shows inherited *bi-* 'without' only with pronouns, where it attaches to the genitive: *išom kokori bitukoro* 'I am lonely without you'. With full nouns, the Greek preposition *xoris* is used: *xoris amaksi* 'without a car'. Exemption is similarly expressed by Greek *extós apo*: *extós apo ti phuri konjek to gav na džanel te phinel i mira* 'except for the old woman nobody in the village knows how to tell the story', or *ėksan*, accompanied by *kato* and the ablative case: *o káthe dženo eksan kato me papostar uštindile našte* 'every person except my grandfather stood up and left'.

8.3. Prepositions and local relations

The expression of local relations in PR is characterised by the use, typically, of location adverbs in combination with a prototype preposition, with nouns usually appearing in the nominative. The basic inventory of prepositions consists of just two: The simplex *to* (*ti*) indicates contact or the possibility of contact, either stative or as a result of movement: *išomas to kher* 'I was at home', *čumidas e džuvlja to vošt* 'he kissed the woman on the lips', *o lil isi opral to trapezi* 'the letter is on the table', *džava to dukjani* 'I am going to the shop'. The complex *kato* (*kati*) indicates separation or separateness: *i šiše peli kato trapezi* 'the bottle fell off the table', *oj našti kato gav* 'she went past the village', *isi njek ikona opral kato kravati* 'there is a picture above the bed'. In the absence of full grammaticalisation of adpositions from local relations adverbs, the system relies on these two prototype prepositions, both mirroring Layer II case ending forms, locative *te* and genitive *k-*, to categorise the noun. This categorisation is crude, and is structurally separated from the semantic specification, which is either inferred from the verb, or expressed by the adverb.

In actual grammaticalisation of prepositions from adverbs, the new prepositions carry both the categorising function, and the semantic specification. Such prepositions are represented sporadically in the corpus. There are a few tokens of *ando* 'in'. Syncopated adverbials *pal(al)* 'behind, on, with' and *paš(e)* 'next to' appear with pronouns, which are marked in the locative: *na therava abuka but love pal mande* 'I don't have so much money on

me', *džakeravas paše tute* 'I was standing next to you'. The preposition *pa* appears in the expression *pa ada* 'therefore'. A number of prepositions are borrowed from Greek, but they too, like Romani location adverbs, tend to be followed by one of the two 'prototype' adpositions: *oj phirdas jira kato kher* 'she walked around the house', *éksan kato me papostar* 'except for my grandfather'.

Alongside this predominant format of adverb+prototype preposition indicating either 'contact' or 'separateness/separation', there is some, rather marginal, use of synthetic (Layer II) nominal case for similar functions. In the synthetic format, the general scheme of categorisation of the noun is rather similar: The locative case *-te//de* is used to indicate contact: *tu aljan me khereste* 'you came to my house', *isi nje bari júrti amare gaveste* 'there is a big celebration in our village'. The ablative case *-tar/-dar* is used to indicate separation or separateness: *ov alo vaverestar gavestar* 'he came from another village', *o čikoro čhavo garadilo pala njeke rukhestar* 'the little boy hid behind a tree'.

Location and movement toward a target resulting in contact, irrespective of containment, are expressed by *to*: *mo phral dživela ti rusija* 'my brother lives in Russia', *gelo to dukjani* 'he went to the shop'. Explicit containment (incorporation) is expressed by *andre to*: *oj isi andre to kher* 'she is inside the house'. Movement away from a target is classified as separation, and expressed by *kato*: *oleski phen irizi kato pazari* 'his sister is returning from the market'. Explicit incorporation can be expressed by the adverbs *avri* or *andar*: *oj nikhisli avri kato kher* 'she came out of the house', *šundomas bašaipe andar kato kher* 'I heard music [coming] out of the house'. Similarly, proximate location is categorised as showing contact, and so it takes *to*, while movement with reference to an object does not entail contact, and so here we find *kato*: *o pazari si paše to gav* 'the market is near the village', but *oj našti kato gav* 'she went past the village'.

Horizontal relations imply absence of contact: *olesko kher isi palal kato kafenes* 'his house is behind the café', *isi nje rukh anglal kato pendžeri* 'there is a tree in front of the window'. With vertical relations, there is a distinction, with 'on' implying contact—*o lil isi opral to trapezi* 'the letter is on the table'—while other relations imply separation or separateness: *i šiše peli kato trapezi* 'the bottle fell off the table', *isi nje ikona opral kato kravati* 'there is a picture above the bed', *o lil isi telal kato trapezi* 'the letter is under the table'. Location of an object in physical overlap with another object or point of reference is regarded as entailing contact: *isi maškare to gav* 'it is in the middle of

the village', *isi but plusi to romacel olate ti poli* 'there is a lot of wealth among the Roma in this town'. Other categorisations of location or movement with reference to points lacking physical contact with the object of categorisation, appear with *kato*: *me čhave dikhenas kati pendžeri* 'my children were looking through the window', *apénadi kati kangerin isi njek skoljos* 'opposite the church there is a school', *alem jıra kati kangerin* 'I went around the church', *o káthe dženo éksan kato me papostar uštindile našte* 'all the people except my grandfather stood up and left'.

8.4. Temporal relations

Deictic expressions of day-distances are *adives* 'today', *tačja* 'tomorrow', which phonologically matches the word for 'tomorrow' in the regional Greek dialects, and appears in a more conservative form than cognate expressions in other Romani dialects (such as *taiša*, *taha-ra*), *idžara* 'yesterday', *proxsara* 'the day before yesterday' and *tačarovre* 'the day after tomorrow', as well as *o aver o dives* 'the next day'. The expression of hours through prepositions is variable. We find *to deš* 'at ten o'clock' alongside *uštindilom kato šov* 'I got up at six', while Greek numerals often attract Greek articles, without a preposition, as in Greek: *uštjava tis epta to sabahi* 'I get up at seven in the morning'.

Times of the day show demonstratives for *ova sabahi* 'this morning' and *avarát* 'tonight', the second having been fused into one word with word-final stress; and the preposition *to* for general reference: *to sabahi* < Turkish 'in the morning', *to apójama* < Greek 'in the afternoon', *ti rati* alongside *rati* 'at night'. Days of the week and months of the year are generally Greek, and accompanied either by Greek oblique definite articles, such as *tin pempti* 'on Thursday', *ton julio* 'in July', *to sabatokirjaka* 'on weekends', or by Romani articles, as in *i paraskevi* 'on Friday', or prepositions: *to kurko* 'on Sunday'. This irregularity is found also with seasons: *o behari* < Turkish 'in spring', but *to vend* 'in winter'. Deictic expressions of year-distances are *beršeske* 'next year', *persi* < Greek 'last year', and *ova berš* 'this year'.

Sequential-durative relations are expressed by a demonstrative, if the reference point is deictic, as in *therasas but buti akala duj berš* 'we had much work during these past two years', or by a complex preposition, for a lexically specified point of reference, as in *ov dživela akate kato junjos* 'he has been living here since June'. Distance ahead is expressed by the locative case: *nje beršeste* 'in a year's time', *štar divesende* 'in four days'. Distance backwards is expressed by Greek-derived *prin*, as in *prin jek berš* 'one year ago'. Durational extent is expressed by the dative: *štar divesenge na geljom avri* 'I haven't gone

out for four days', *dživdilom to gav pandž beršenge akana* 'I have lived in the village for five years now'.

8.5. *Embeddings and relative clauses*

As in other Romani dialects, interrogatives are employed as complementisers introducing embeddings: *ov phučhljas kon gelo ti poli tin triti* 'he asked who went to town on Tuesday'. The conditional embedding (indirect conditional) is introduced by the conditional conjunction *ama*: *pučhljom e daskalós ama k'avel to bjav* 'I asked the teacher whether he would be coming to the wedding'.

Relative clauses are introduced by *ka*: *o murš ka alo to bjav therelas nevo aftokindos* 'the man who came to the wedding had a new car'. Pronominal resumption of the head noun is not obligatory, even when its case-role within the relative clause is low on the thematicity (accessibility) hierarchy: *dikhlom o kher ka vakeresas* 'I saw the house that you spoke [about]', *ka si i čhuri ka phiravdan o grama?* 'where is the knife that you opened the letter [with]?', *o murš ka dikhlom idžara isi akate pale* 'the man that I saw yesterday is here again'. Resumptive pronouns may however occur, in which case they appear as enclitic object pronouns that accompany the verb of the relative clause, as in *kindomas nje polikano vraxoli ka ka davos mi dajate* 'I bought a golden bracelet which I shall give [it] to my mother', or else as a possessive pronoun, as in *ol si o džene ka olenge kherimata phabile* 'they are the people whose [= who their] houses burned down'.

8.6. *Complementation*

Complementation is split, as elsewhere in Romani, between factual and non-factual clauses. Non-factual complements in PR are always finite, the verb appearing in the subjunctive and agreeing with the subject of the complement clause, and are introduced by the inherited complementiser *te*: *oj na kamela te džal ti poli* 'she didn't want to go to town', *prepi te džas ti poli* 'we must go to town', *na borusa te phiravavas i vudar* 'I couldn't open the door'. Optional omission of the complementiser is attested only for 'can': *boris aves manza?* 'can you come with me?'. With manipulation clauses (modal complements in which the subject of the complement clause differs from that of the main clause), the complementiser is likewise always *te*, and the embedded subject or manipulee is expressed as an antecedent, as in *phindom olake te kinel vande* 'I told her to buy eggs', *olesko dad kerdasos te bičhavel o lil* 'his father made him send the letter', or as antecedent with resumption, as in

ov manglja mandar te davos love 'he asked me to give him money', or else as an exposed subject of the complement clause, in a position external to the complementiser: *kamama ov te našel* 'I want him to go away'.

Factual or epistemic complements are introduced by the Greek-derived complementiser *oti*, as in many other Romani dialects, in which the epistemic, but not the non-factual or subjunctive complementiser, is borrowed. The complement clause shows independent tense: *dikhljom oti oj na sines to kher* 'I saw that he was not at home', *phindom oleske oti o dikani phiravela to deš* 'I told him that the store opens at ten', *nomizava oti ov dživela akate* 'I think he lives here'.

Purpose clauses behave in Romani much like non-factual complements: they are introduced by a non-factual complementiser, and the verb appears in the subjunctive. In PR, simple purpose clauses, in which the semantic connection between main and subordinated clause is straightforward, are introduced by *te*: *geljom te dikhav me dajá* 'I went to see my mother', *dinom e gadžes xande love te kinel kafes* 'I gave the man some money to buy a coffee'. Complex purpose clauses are introduced by *ka te*. Here, there is either intensified planning and effort and so 'reflective intent' on the part of the actor to achieve the goal, as in *sigo ka therav but love ka te kinav mindo mo aftokindos* 'soon I shall have enough money in order to buy my own car', or else the connection between the two propositions needs to be reinforced, as in *ov kindas neve cavala ka te džal ti poli* 'he bought new clothes in order to go to town'.

8.7. Adverbial clauses

PR relies exclusively on the linking of finite clauses to express adverbial propositions (propositions that modify the main proposition). Individual semantic relations between the two clauses are expressed by the choice of adverbial subordinator (conjunction). Simultaneity can be expressed in several different ways. Subject complements in *te* indicate state or action of the object of the main clause: *dikhlom je muršes te phirel tele to drom* 'I saw a man walking down the street', *dikhljjas tu konjek te džas othe?* 'did anybody see you going there?'. In the absence of gerundial constructions, predications that describe the actions of the subject of the main clause are arranged in a paratactic way: *ov alo manza garavelas kati palal kato pike* 'he came towards me hiding [= he was hiding] something behind his back'. Temporal overlap is expressed by *kana* 'when': *oj darándili kana dikhljjasos* 'she was frightened when she saw him'. Durative relations are expressed by Greek *oso*: *gilavdam gilja oso kerahas buti* 'we sang songs while we were working'.

A series of adverbial relations are marked by borrowed Greek subordinators, followed by the non-factual complementiser *te*. Anteriority is expressed by *prin*: *haj te džas avri prin te del biršindo pale* ‘let’s go out before it rains again’, *prin te avav te dživav akate džividom but dur akatar* ‘before I came to live here I lived far away from here’. Addition is marked by *éxthos*: *éxthos te bešela to kafenijos, na kerela but praymata* ‘apart from sitting in the café, he doesn’t do much work’. Negative addition is marked by *xoris*: *oj phirdas jıra kato kher xoris te arakhel phiravdi e vudar* ‘she walked around the house without finding the door open’. Here, the irrealis reading of the adverbial modification triggers on the one hand the use of a tense-less, subjunctive verb, introduced by the non-factual subordinator, which is stable, and not prone to borrowing. On the other hand, irrealis and exceptionality are semantic relations that are susceptible to borrowing, and we find here, as in the parallel semantic domain of local relations (*xoris* ‘without’, *éksan* ‘except’), an abundance of Greek items.

Cause and result are expressed by *soste*: *ka džav te sovav akana soste šom čindi* ‘I shall go to sleep now because I am tired’. Conditional clauses are introduced by *ama*, etymologically an adversative conjunction in Turkish and Albanian, possibly introduced into its current function in PR via its role as an interjection of amazement in Albanian (but cf. its original Arabic meaning, ‘as for’). Occasionally, *ama* is followed by inherited *te*, which immediately precedes the verb. In potential conditional constructions, the verb appears in the present or future tense in the conditional clause, and in the future tense or imperative in the main clause: *ama pjava thud but k’avav but zurali* ‘if I drink a lot of milk I will become very strong’, *ama k’aves, ka dikhav tu* ‘if you will come, I shall see you’, *ama te dikhes kati phin mange!* ‘if you see something, tell me!’. In realis constructions, the verb of the conditional clause appears in the imperfect, and that of the main clause takes the conditional mood (*ka* + imperfect): *ama theravas xandi love ka davas tu* ‘if I had some money I would give you’, *na pučhavas tu ama džanavas ka isines ov* ‘I wouldn’t ask you if I knew where he was’ (note that the future particle *ka* is deleted in the negative). In irrealis constructions, the verb of the conditional clause is in the counterfactual (pluperfect or remote past), and that of the main clause in the remote conditional (*ka* + counterfactual): *ama tu t’avehas idžara ka thesasi dikhlo* ‘if you had come yesterday you would have seen her’. Concessive conditionals are expressed by Greek-derived *eθ an*: *prepi te džah othe eθ an na kameha te džas* ‘you have to go there, even if you don’t want to’.

8.8. *Word order*

In the noun phrase, indefinite articles precede adjectival attributes, which precede the noun, as is generally the case in Romani: *nje baro džukel* 'a big dog'. Exceptionally, adjectival modifiers may follow the noun: *dikhljom nje suno darano* 'I saw a frightening dream'. Other modifiers, such as possessives, definite articles, and demonstratives, also precede the noun: *mo čhavo* 'my son', *o kher* 'the house', *ova kher* 'this house'. The definite article may appear at the beginning of a chain of modifiers: *o čukore čhaja isi sixná ladžane* 'young girls are often shy'. Alternatively, it may be repeated to introduce each lexical member of the noun phrase: *o terne o rakle theren but θáros* 'young men are very courageous'. Quantifiers appear outside the scope of the definite article: *sare o kherimata si parne* 'all the houses are white'. The demonstrative precedes adjectives: *kamamas te džavas ti poli olenza e trinenza e muršenza* 'I would like to go to town with those three men'. Unlike other dialects of Romani, under Greek influence the PR demonstrative precedes prepositions, which are immediately adjacent to the noun: *štar romacel dživena olesti to kher* 'four Roma live in this house'. As in Greek, and unlike the norm in other dialects of Romani, the overwhelming tendency is for the genitive to follow the head noun: *o kher e muršeskoro isine but purano* 'the man's house was very old'. Adjectives describing the genitive noun agree with it for case, and precede it: *o kher e dujengoro e phralengoro isi but čikoro* 'the two brothers' house is very small'.

In the verb phrase, the object usually follows the verb, as in *ov dikhljias e phures* 'he saw the old man', and the direct object usually follows the indirect object, as in *oj sikhavela pi čhajake i nevi angrusti* 'she is showing the new ring to her daughter'. In regard to the positions of subject and verb, there are two principal word-order formats: categorical andthetic (for the terms see Sasse 1987). In categorical word order, where an attributive statement is being made about the subject, the subject precedes the verb: *o čhavo sovela* 'the boy is sleeping', *o čhave mi bibjakere isi plusi* 'my aunt's children are rich'. Thetic word order highlights the event, and has the verb in initial position, preceding the subject. Typical subjects inthetic constructions are new and unknown topics, indefinite subjects, quantified subjects, and contrastive subjects, and typical predicates inthetic constructions are existentials and verbs of appearance: *gele but džene to pazari* 'many people went to the market', *isine nje murš ka pučhelas tuke* 'there was a man who was asking about you', *ov si etimos kana šan tuja etimi* 'he is ready when you are ready'.

In questions,thetic word order prevails wherever both subject and verb are expressed, as is to be expected, since subjects are largely presupposed: *so dinas tu ov?* ‘what did he give you?’, *kaj sas ov saro dives?* ‘where was he all day?’, *soske kerdasos oj ova?* ‘why did she do this?’, *kaj geli oj?* ‘where did she go?’. Categorical word order is also possible, however: *sar ov arestjas othe?* ‘how did he get there?’.

In subordinations, similar rules seem to apply. Thetic order is preferred with existential predicates and those of appearance, and with subjects that are unknown: *na pučhavas tu ama džanavas ka isines ov* ‘I wouldn’t ask you if I knew where he is’, *šundomas oti dživena aver romacel akate* ‘I heard that other Roma are living here’. Categorical order is preferred when a statement is made about a known topical entity: *phindom oleske oti o dikani phiravela to deš* ‘I told him that the shop opens at ten’, *nomizava oti ov dživela akate* ‘I think that he is living here’. In non-factual complements, the non-identical subject (manipulee), if expressed overtly by a subject pronoun, precedes the verbal complement: *kamama ov te našel* ‘I want him to go away’. The pronominal object is attached to the verb in enclitic position: *olako dad na mukhelos te phandreveli* ‘her father won’t let him marry her’.

Pronominal object doubling is optional, but frequent. The pronoun then tends to precede the full nominal or pronominal-demonstrative object: *d-os nje kotor guldibe oles e čhaves!* ‘give[him] a piece of chocolate to this boy!’, *kon kerdasos ova?* ‘who did[it] this?’, *na borusa te phinavasi olake te avel pale manza* ‘I couldn’t tell[her] her to come back with me’.

8.9. Connectors and discourse markers

Connectors are a mixture of items from Albanian and Greek, the recent and current contact languages. From Albanian we find *éli, éle* ‘and’, *pastánja* ‘then, after’, and *po* ‘but’. From Greek we find *alá* ‘but’ and *ómos* ‘however, nevertheless’. Connective particles are also of these two sources. Albanian gives *ele . . . ele* ‘both . . . and’ and *as . . . as* ‘neither . . . nor’: *ele mo dad ele mo phral* ‘both my father and my brother’, *as ov as olesko phral* ‘neither he nor his brother’. Greek gives *i . . . i* ‘either . . . or’ and *ute . . . ute* ‘neither . . . nor’: *to vend i te dela biršindo i te avela eras i o duj* ‘in winter it either rains, or it is windy, or both’, *ute mi phen ute me* ‘neither my sister nor I’. Focus particles are the Albanian-derived *ja, eli* ‘too’, and inherited *kokori* ‘only’, and Greek *iđjon* ‘same’. Typically, then, the more additive-continuative markers are more conservative, deriving from the Recent L2 Albanian, while the more contrastive–restrictive markers are prone to earlier replacive borrowing, stemming from the Current L2 Greek (cf. Matras 1998).

9. The historical and regional position of Romacilikanes

PR belongs to the type of Romani dialects that is most prevalent among the Balkan group of southeastern Europe: it has retained an adjectival past tense (active participle) with the 3SG form of intransitive verbs of motion and state, and mediopassives; it shows the retreat of subject clitics to copula predications, an analytic future in *ka*, partial analogy of the 2PL perfective concord marker to the 3PL, giving *-en*, reduced reflexive possessive in *po* and reduced possessives *mo* and *to*. The challenge is to determine its more specific regional position, in relation to neighbouring dialects.

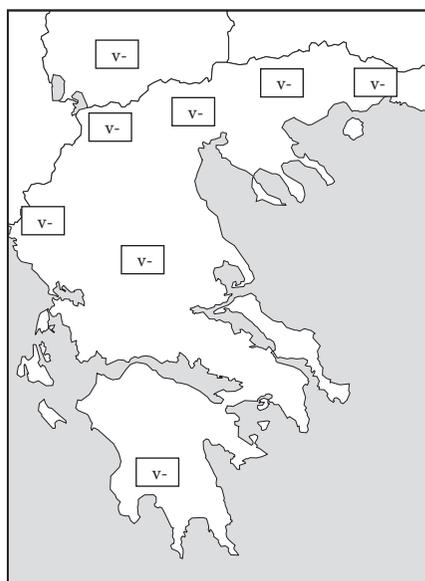
As mentioned above (Section 1), Boretzky (1999), in a pioneer discussion of the southern Balkan dialects of Romani, plots features of several documented dialects of the area onto maps, and draws a number of isoglosses that appear, at first glance, to separate the dialects of Greece from those of adjacent regions to the north—Bulgaria and Macedonia. Boretzky's only published source for the area is Paspatis (1870) for the so-called Rume-*lian* sedentary dialect. His unpublished sources include fieldnotes on the Sepečides dialect, which in the meantime have appeared in print (Cech and Heinschink 1999), as well as fieldnotes from Prilep in Macedonia, and fragmented fieldnotes from Serres in northeastern Greece.

Through recent documentation work, more data on the dialects of Greece has now become available, as of yet unpublished, but contained in the RMS database. These include detailed documentation of the Prilep Arli variety, edited and collected by Petra Cech based on fieldwork carried out by Mozes Heinschink; elicitations of the RMS-based Romani Dialectological Questionnaire—from Florina on the Greek–Macedonian border (Arli dialect), and from Kalamata on the southern edge of Pelopennese, recorded by Irene Sechidou, and from Karditsa in central Greece, recorded by Veronica Schulman. In addition, Sechidou's (2002) presentation on the Serres dialect (apparently a different variety from that considered by Boretzky) has already seen limited distribution, and covers some basic structural information on that dialect. In the following, I discuss some key features of PR in relation to these dialects (see Map 1 for locations).

The dialects of the region share a number of features that are not found to the north of the area. These include the rather consistent *v*-prothesis, covering at least the words for 'egg' (*vandřo*), 'flour' (*vařo*) and 'lip' (*vuřt*), and usually more. This clustering of *v*-prothesis might be considered the most outstanding feature of the 'Greek' group of Romani dialects, as it is not paralleled in any other region across the Romani-speaking landscape (cf. map



MAP 1. Location of dialects

MAP 2. Prothetic *v-* in 'egg', 'flour', 'lip'

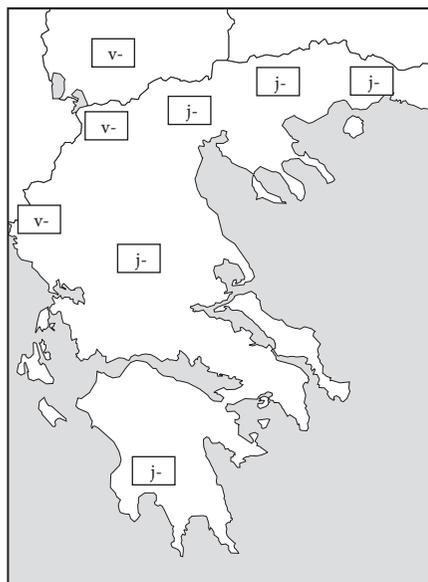
in Matras 2002: 216). Moreover, it could well represent the historical centre of diffusion of ER *v-*, giving rise to pan-Romani *v-ast* 'hand'. Another innovation that is typical of the region, though it is triggered by contact and therefore somewhat less diagnostic, is the emergence of *ther-* in the meaning 'to have', which also appears in the role of a perfect auxiliary, and the formation of the periphrastic perfect and pluperfect tenses that draw on it. Greek loans are of course also shared by these dialects, but they also appear in the Vlax dialects of the area, and are, for this reason too, not diagnostic of the 'Greek' group of Balkan Romani dialects.

Maps 3 and 4 show a sub-division within the region. In the northwestern sub-group, to which PR belongs, *v-* prothesis is carried further to include the adjectival *v-aver* 'other' (Map 3). The remainder of the area shows prothesis of *j-* here: *j-aver*. The split between nouns and the adjective is understandable: the original prothesis derives from the attachment of the old definite article, M. **ov*, F. **oj*. Following the model of the masculine noun **ov-ast* 'hand', the dialects of the region generalised the article incorporation with other masculine nouns: **ov-ařo* 'flour', **ov-andřo* 'egg', etc. With the adjectival **aver*, the choice of prothetic segment is not as obvious, and in principle each definite article, masculine and feminine, stand an equal chance of being selected. Each region shows a different selection from what were in all

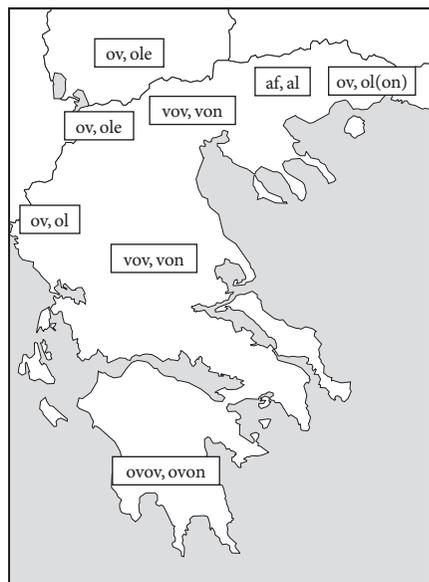
likelihood two ER variants, **ov-aver* and **oj-aver*, alongside the non-prothetic **aver*. The geographical division is evidence that preferences were influenced by two distinct networks of contacts.

From Map 4 we get partial confirmation of a coherence among the north-western subgroup—Prilep, Florina, and Epirus (historically, in all likelihood prior to settlement in Parakalamos itself). The situation here is somewhat more complex: The northwestern corner shows absence of *v-* prothesis in the 3SG.M pronoun ‘he’, and selection of a 3PL form in *ol-*, from the available ER variation **ol/on*. The more central region, comprising Sepeči, Karditsa, and Kalamata, show *v-* prothesis (in Kalamata, the rather aberrant reduplication *ovov*, a kind of literal analogy to the original prothesis of nouns, or alternatively an attempt at a reconstructive retention of the initial vowel segment despite prothesis). While Rumelian retains the conservative ER pattern, Serres selects proximate demonstrative forms in *a-*. Taking a somewhat broader view, then, we might say that there is a division between a northern and a southern zone—the first, northern zone, with no prothesis, and showing continuation of *-l* forms in the plural; the second, southern zone, with prothesis and continuation of *-n* only.

We find this pattern of a north-south division confirmed in Map 5, which shows the distribution of forms of the verb ‘to become’. Here, however, Epirus

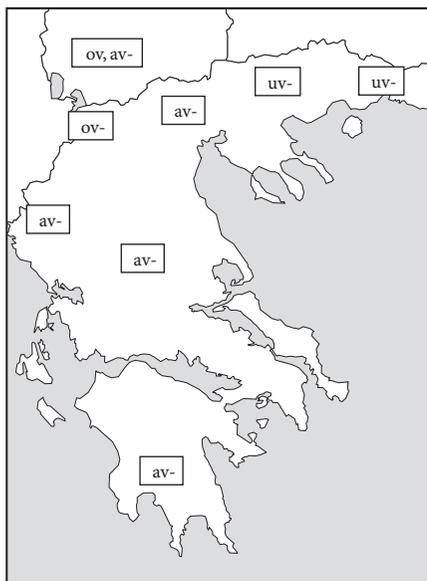


MAP 3. Prothesis *v-/j-* in ‘other’

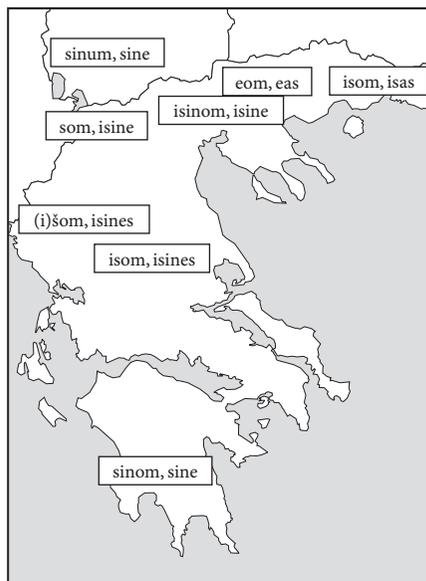


MAP 4. Anaphoric pronouns ‘he’, ‘they’

(PR) goes with the southern group, rather than with the northern one. In fact, the picture here is complex, too. The northeastern corner, Rumelian and Serres, actually show a variation on ER *ov-*, namely *uv-*. Florina retains *ov-*. Both forms continue MIA *bhuv-* ‘to become’, and are cognate with the grammaticalised mediopassive marker *-jov-*. In the Romani dialects north of the southern Balkans—such as Vlax, the Central dialects, and elsewhere—the form for ‘become’ is *av-*, a functionalisation of the verb of motion ‘to arrive’ (promotion to place > promotion to state). Elsewhere I suggested, based on the presence of *av-* in loan verb adaptation markers in the south, and the occasional though isolated presence of *ov-* in the north, that *av-* was already an ER Romani innovation, and that ER had both variants (cf. Matras 2002:138). The generalisation of *av-* to the south of the zone in which *ov-* prevails (and which in turn is south to the main *av-* zone in central and northern Europe), confirms that hypothesis: In this region, a different ER variant was selected and generalised than in the adjoining regions to the north of the southern Balkans. From the swaying affinities of PR—with the north on Map 4, with the south on Map 5—a differentiated picture emerges of a gradual diffusion of innovations (or, in both these cases, ‘option selection’ from among ER variants) across regional space, with varying degrees of progression. By necessi-



MAP 5. ‘To become’



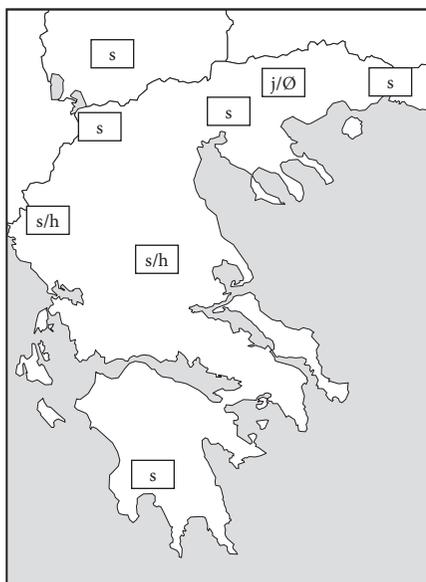
MAP 6. Copula, ‘I am’, ‘he/she was’

ty this implies contact between the community in Epirus and neighbouring Romani-speaking communities in both directions.

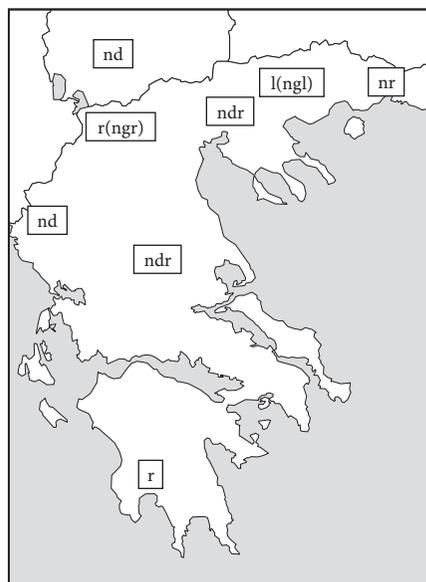
Another good illustration of these shifting orientation targets, indicating contacts in both directions, is provided by the pattern of geographical distribution in the region of copula forms in the 1SG present and 3SG past. A number of dialects show idiosyncratic behaviour in the treatment of the 1SG present 'I am'. The intrusion *-in-* (in all likelihood an ER option) is retained in Prilep, Kalamata, and Sepeči, with no geographical connection. The Serres dialect selects forms of the copula in **h > Ø*. Initial *i-* is lost in the extreme north, and in the extreme south, leaving a belt of *i-* retention in between. In the 3SG past 'he/she was', Rumelian and Serres share a similar formation pattern, albeit with different stems, while PR and Karditsa share the suffixation of *-s* to the form in *-ine-*. On the whole, the PR forms are closest to Karditsa. However, PR partly goes with Florina to the north in dropping initial *i-*, and it shows an idiosyncratic palatalisation of the stem consonant in *šom* 'I am'.

The emerging pattern is that of strongest contacts to immediately adjacent dialects, both to the south and to the north. In both cases, PR appears at the edge of a spread zone, suggesting that it was a recipient of forms diffusing from elsewhere, but not necessarily a promoter of forms that served as a model that was copied by others. In addition, PR shows idiosyncratic behaviour, which suggests some degree, at some historical period, of social isolation.

In Map 7 we see once again a link to the central Greek dialects, in the distribution of *s/h* forms in intervocalic affixes—the variation of the type *kere-sa/kereha* 'you will do'. Here, we find variation both in PR and in Karditsa, historically *h-* forms only in Serres, and *s-* forms generalised elsewhere. The *s-* dialects of Florina and Prilep appear to separate PR from the principal zone in which *h-* is generalised, which lies further to the north, affecting the dialects of Macedonia and Kosovo, including immigrant southern Vlach dialects in that area, and extending north, through Serbia, the Central dialects of the Hungarian contact area, and the Sinti dialects of Germany. It is possible of course that PR adopted (or retained) *s/h*-variation through contact with Romani dialects of Albania. This scenario would imply that PR was the target of imitation for the more southern varieties, represented here by Karditsa, and so a prestige variety in its own right. Another possibility, however, is that the variation we find in PR and to the south of it is an old ER legacy of alternation in these positions, inherited into ER from MIA, and preserved in some of the dialects of Greece.



MAP 7. *s/h* alternation in intervocalic affixes

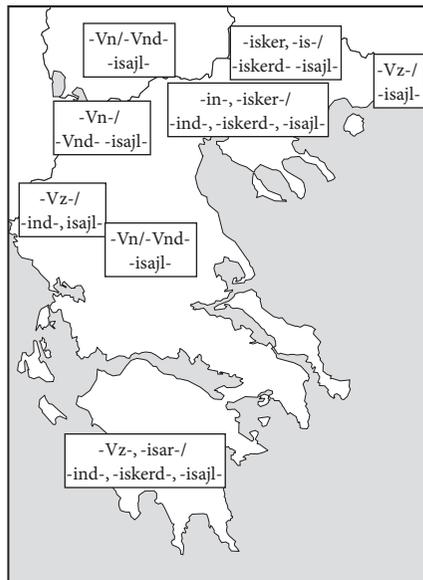


MAP 8. Cluster *-ndř-*

Map 8 represents the reflexes of the ER cluster **-ndř-* (MIA *-nd-*) in words such as **mandřo* ‘bread’, **pindřo* ‘foot’. It is well known that Romani dialects of southeastern Europe tend to preserve a cluster, whereas those of northern and western Europe tend toward cluster simplification to *-r-*. The patterns of cluster re-structuring in the Balkans however are idiosyncratic, and lead to simplifications here too. Map 8 illustrates this rather nicely: apart from the zone created by the contiguity of Sepeči and Karditsa, there are no coherent zones for the formation of the cluster. There could be several reasons why the pattern of cluster re-structuring differs so radically from the diffusion zones illustrated by some of the earlier maps: We could be dealing with a very recent development, which might have occurred after settlement in more contained regions and loss of contact between the groups. This is supported indirectly by the fact that in some areas of the southern Balkans, retroflex sounds can still be heard in the relevant words, suggesting that ER *-ř-* was, and was continued as, a retroflex until a rather late period. In both Serres and Florina, there are exceptions to the general development, namely the words *kanglo/kangro* ‘thorn’ and *vanglo/vandro* ‘egg’ (otherwise *-l-* and *-r-* respectively). This too might suggest a development that has not yet been completed. Alternatively, the presence of a salient cluster in a small

group of frequently used vocabulary items might have served as a token of group-identity and might therefore have been more resistant toward change and accommodation to neighbouring dialects.

A somewhat comparable picture, though more complex, is portrayed by Map 9, which shows loan verb adaptation markers. Regional patterns here are limited. The same forms appear in Prilep and Florina, which they, in turn, share with the bulk of Balkan dialects of Romani, i.e. those of the Arli and Erli types. Indeed, this pattern is to some extent—in *-in-* rather than *-Vn-*—shared by the Central as well as Northeastern Romani dialects. Prilep and Florina thus form the southernmost edge of a vast continuum, stretching all the way to the Baltic sea. Karditsa happens to show the same pattern, apparently a result of its own, idiosyncratic development. Serres and Sepeči, on the other hand, share the form *-isker-*, which is otherwise encountered further to the north, along the Black Sea coast. In Sepeči, it is rather specialised (for transitive derivations and causatives), and coexists with *-in-*. The Rume-
lian dialect shows forms that are common along the Black Sea area, while Kalamata shows a strikingly wide range of forms, shared individually with Vlax, Iberian, and Azerbaijanian Romani (*-isar-*), with the Black Sea coast dialects (*-iz-*), with the Balkan, Central and Baltic dialects (*-in-*), and the pan-Romani *-isajl-*.



MAP 9. Loan verb adaptation markers:
present/past

Only one conclusion can be drawn from the remarkable diversity of markers in this comparatively small area, and that is, that ER had the full range of markers, which was then inherited wholesale by the dialects. These, in turn, simplified the system, selecting and generalising few markers at the expense of the others (cf. Matras 2002: 128–34). In the area under consideration, the older, more diverse system was available until a rather late stage. Its collapse appears to have coincided with the breakdown in the network of contacts between the dialects, and so to have occurred relatively recently. The patterns are therefore idiosyncratic, or shared at best with immediately neighbouring dialects. PR's behaviour in this respect testifies to considerable isolation in the relevant period, for it is the only dialect in its area to have selected *-Vz-* as a primary marker. The coexistence of *-ind-* and *-isájl-* as past-tense markers on the other hand is quite common, not just in this region, but throughout the central zones of Europe. Evidently, simplification resulting in the present arrangement for the past tense occurred first, still in the context of a regional network of contacts, while levelling in the present tense continued well into the period of comparative isolation.

Though the spatial distribution of numerous other forms might be considered, Maps 10 and 11 help us round up the picture inasmuch as they represent the patterns of isolated developments, out of the inherited stock of shared ER legacy. The southern Balkan dialects of Romani tend to preserve the ER inventory of demonstratives (Map 10), and we find this pattern throughout our region as well, though individual dialects tend to add one or two forms to the general inventory. The Rumelian dialect stands out in its innovation of forms that show suffixed reduplication in *-ka*, a pattern that extends farther north, into the Erli varieties of Bulgaria. PR is aberrant, in showing just one of the two common series, that in *akava/okova*, while on the other hand preserving the functionality of what is clearly a very archaic set, *ava/ova*. The latter preceded reinforcement of the demonstratives through local deixis *akaj/adaj* 'here' etc., a development that is attested everywhere in Romani and so must have been an ER development. The archaic character of the set is further confirmed by its tendency to survive primarily in compositions of the type *av-dives* 'today', or in contracted form, in the anaphoric pronoun set *ov* etc. (Serres *af*). From its prevalence in PR we must conclude that we are dealing with a dialect that has resisted some trend toward changes that dominated the Romani-speaking landscape elsewhere. The impression is confirmed, finally, by Map 11, which shows the assimilation of the perfective marker of sibilant stems in most of the region's dialects

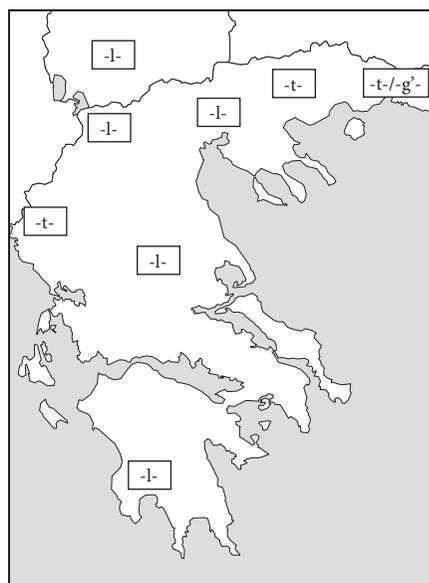
to the pattern of (originally) vocalic stems in *-l-*. This development too is rather typical of dialects of the southern Balkans, though conservative pockets can be found, for instance in Serres. PR too retains the conservative formation.

These latter few idiosyncratic traits of PR join a somewhat longer list: The dialect is conservative in retaining vocalic forms of the oblique pronouns (*oles-*, *ola-*, *olen-*), and in relying mainly on adverbs rather than prepositions for semantic differentiation of local relations. Like some of the other dialects in the Balkans it retains some of the older indefinite forms, based on the formants *kaj-*, *-ni-* and *-muni* (e.g. *kajnijekas-* ‘nobody’, *katemuni* ‘nowhere’). The retention of *romacel* is no doubt also a conservatism, while on the other hand its generalisation as the only self-appellation, and the construction of a label for the language that is derived from it, *romacilikanes*, is an innovation that serves as a token of the group’s self-perception as distinct and separate.

This separateness is reflected linguistically in a number of idiosyncratic innovations. They include the replacement of plural *-a* in inherited masculine nouns ending in consonants by borrowed *-imata*, re-assignment of some pre-European masculines in consonants to the class of European



MAP 10. Demonstratives

MAP 11. Perfective marker of *beš-* ‘to sit’

loans in *-i*, full Layer II case inflection in adjectives, the emergence of pronominal object clitics, the emergence of a new use of *nje(k)* as indefinite article, apparently from *ni-jekh* ‘any-one’, use of indicative copula forms in the future tense (*ka šom* ‘I will be’), palatalisation of copula forms in the first and second person: *išom*, *išan*, and a number of contact phenomena in the syntax of the noun phrase, notably doubling of determiners (definite articles and possessive pronouns) with adjectives, postposition of the genitive noun, prepositioning of the demonstrative before the adposition.

List of Abbreviations

ER	Early Romani	NOM	nominative
F	feminine	OBL	oblique
INSTR	instrumental	PL	plural
LOC	locative	PR	Parakalamos Romani
M	masculine	SG	singular
MIA	Midde Indo-Aryan		

References

- Bakker, Peter. 1997. Athematic morphology in Romani: The borrowing of a borrowing pattern. *In* The typology and dialectology of Romani. Yaron Matras, Peter Bakker, and Hristo Kyuchukov, eds. Pp. 1–21. Amsterdam: John Benjamins.
- 1999. The Northern branch of Romani: Mixed and non-mixed varieties. *In* Die Sprache der Roma: Perspektiven der Romani-Forschung in Österreich im interdisziplinären und internationalen Kontext. Dieter W. Halwachs, and Florian Menz, eds. Pp. 172–209. Klagenfurt: Drava.
- Boretzky, Norbert. 1999. Die Verwandtschaftsbeziehungen zwischen den Südbalkanischen Romani-Dialekten. Mit einem Kartenanhang. Frankfurt am Main: Peter Lang.
- Cech, Petra, and Mozes F. Heinschink. 1999. Sepečides-Romani: Grammatik, Texte und Glossar eines türkischen Romani-Dialekts. Wiesbaden: Harrassowitz.
- 2001. A dialect with seven names. *Romani Studies. Fifth Series.* 11: 137–84.
- Elšík, Viktor. 2000a. Romani nominal paradigms: Their structure, diversity, and development. *In* Grammatical relations in Romani: The noun phrase. Viktor Elšík and Yaron Matras, eds. Pp. Amsterdam: Benjamins. 9–30.
- 2000b. Inherited indefinites in Romani. Paper presented at the Fifth International Conference on Romani Linguistics, Sofia, 14–17 September 2000.
- Friedman, Victor A. and Robert Dankoff. 1991. The earliest known text in Balkan (Rumelian) Romani: A passage from Evliya Celebis Seyâhat-nâme. *Journal of the Gypsy Lore Society. Fifth Series.* 1: 1–20.

- Hancock, Ian F. 1995. *A handbook of Vlax Romani*. Columbus: Slavica.
- Igla, Birgit. 1996. *Das Romani von Ajia Varvara. Deskriptive und historisch-vergleichende Darstellung eines Zigeunerndialekts*. Wiesbaden: Harrassowitz.
- Marushiakova, Elene and Vesselin Popov. Forthcoming. Segmentation vs. consolidation: The example of four Gypsy groups in CIS. *Romani Studies*.
- Matras, Yaron. 1998. Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36: 281–331.
- 1999a. Johann Rüdiger and the study of Romani in eighteenth-century Germany. *Journal of the Gypsy Lore Society. Fifth Series*. 9: 89–106.
- 1999b. The state of present-day Domari in Jerusalem. *Mediterranean Language Review* 11: 1–58.
- 2002. *Romani: A linguistic introduction*. Cambridge: Cambridge University Press.
- Viktor Elšik, Katrin Hietam, Christa Schubert, Barbara Schrammel, and Irene Sechidou. 2001. Romani dialectological questionnaire. University of Manchester: Romani Morphosyntactic (RMS) Database Project.
- Messing, Gordon. 1987. *A glossary of Greek Romani. As spoken in Agia Varvara (Athens)*. Columbus: Slavica.
- 1991. Some sample texts in Greek Romani. *Journal of the Gypsy Lore Society. Fifth Series*. 1: 67–83.
- Miklosich, Franz. 1872–1880. *Über die Mundarten und Wanderungen der Zigeuner Europas X–XII*. Wien: Karl Gerold's Sohn.
- Paspati, Alexandre G. 1870 [reprint 1973]. *Études sur les Tchinghianés ou Bohémians de l'Empire Ottoman*. Osnabrück: Biblio.
- Sasse, Hans-Jürgen. 1987. Thethetic/categorical distinction revisited. *Linguistics* 25: 511–80.
- Sechidou, Irene. 2002. The Romani dialect of Serres. Paper presented at the 6th international conference on Romani linguistics, Graz, September 2002.
- Tenser, Anton. 2003. *A description of Lithuanian Romani*. Unpublished M.A. thesis, University of Manchester.
- Theodosiou, Aspasia. 2004. 'Be-longing' in a 'doubly occupied space'. *The Parakalamos Gypsy musicians. Romani Studies*. 14: 25–58.
- Wentzel, Tatjana. 1980. *Die Zigeunersprache. Nordrussischer Dialekt*. Leipzig: Enzyklopädie.

