

# South Balkan II as a Romani dialect branch: Bugurdži, Drindari, and Kalajdži

NORBERT BORETZKY

Various suggestions have been made on how to classify the large number of Romani dialects, but the proposed classifications should be put on firmer ground by drawing on as many details as possible. It is very difficult to find features exclusively characteristic of the individual groups. The three dialects subsumed here under 'South Balkan II' undoubtedly form a densely coherent group, mainly due to a number of typical sound changes (palatalisation of dentals, velars, and laterals), but also because of some morphological traits rarely found outside this group (specific copula forms, zero tense marker, and *-iz-* for the adaptation of loanverbs). There are interrelations with some of the South Balkan I dialects, but no specific similarities with other groups can be found.

## 1. General

The three dialects spoken in Bulgaria, Macedonia and Kosovo (see Maps 1 and 2<sup>1</sup>), differ from the South Balkan I group (see Boretzky 1999c) in a number of phonological traits, mainly processes of palatalisation. It is the aim of this paper to give as complete an account as possible of the dialects thus contributing to a future dialectology of Romani. A more special aim is to find out if the changes shared by the three dialects have a common source, that is if they took place in a common ancestor dialect, or if at least some of them started after the ancestor dialect assumed here had fallen

Norbert Boretzky is Professor of Comparative Linguistics, Institut für Sprachwissenschaft, Ruhr-Universität Bochum, 44780 Bochum, Germany. Email: boretzky@linguistics.ruhr-uni-bochum.de

*Romani Studies* 5, Vol. 10, No. 2 (2000), 105–183. ISSN 1528–0748

apart. Further we want to clarify whether the ancestor dialect displayed roughly the same characteristics as South Balkan I,<sup>2</sup> before palatalisations and other changes produced a group of its own.

The tribes belonging to our group are known by other names as well: some Bugurdžides ‘drill-makers’ (from Turkish *burgucu*) call themselves Rabadžides ‘carters’ (from Turkish *arabacı*), and in Macedonia they are known as Kovači ‘blacksmiths’ (from Slavic *kovač*). The Drindarja or Drəndarja (from Bulgarian *drāndar* ‘carder’) are called by other Roms Kitkadžides (from *kītka* ‘hither’) because of this peculiar form of the adverb, or Čalgîdžides ‘musicians’ (from Turkish *çalgıcı*). According to Kenrick (ms. 1969: 13), the self-denomination of the group residing in Kotel is Rōmaicja, apparently a derivation from *romano* (with Greek -ιτης?). The Kalajdžides ‘tanners’ (from Turkish *kalaycı*) are called by others Pavazja or Pavazes (etymology unknown). As can easily be seen, most groups owe their names to traditional professions. Therefore, it is not astonishing that there are groups bearing identical denominations but speaking different dialects. In Bulgaria, there are Bugurdžides (also called Pārpulja) speaking a dialect similar to Erli<sup>3</sup> (cf. the data given in Gilliat-Smith 1915–16, part I, 5f.), and in Hungary there are Kalajdži(des) belonging to the Vlach group. Kalajdži(des) speaking a Vlach dialect are reported also from Bulgaria (p.c. Iglā). The Kalajdži dialect belonging to our group appears to be extinct now, after most of the speakers adopted Turkish as a first language. Gilliat-Smith reports that when he made his investigations many subgroups were about to abandon their dialects or had already done so.<sup>4</sup>

Since we have no common name for the three dialects, I propose to call them South Balkan II, in distinction to South Balkan I comprising such dialects as Arli, Erli, Sepeči, Paspatian as well as some others (see Boretzky 1999c). I prefer to refer to all these dialects as ‘South Balkan’ rather than ‘Balkan’, because most Vlach dialects are also spoken in the Balkan area (in the linguistic sense of the word).

The information available is rather scanty, especially regarding the morphological data of Drindari and Kalajdži. What is more, we lack information about linguistic variation between the subdialects. The only published monograph is on Kosovo Bugurdži (Boretzky 1993), but even for this dialect we have nearly no data about the varieties spoken in Macedonia. As for Drindari, we have some articles by Gilliat-Smith (1914, 1925, 1931, 1962) and Kenrick (1967) as well as some data on the subdialect of Zlatarica east of Tārnovo (p.c. Iglā). From Kenrick I had at my disposal the unpublished

monograph on the dialect of Kotel (ms. 1969), which however does not give all morphological details necessary for comparison, and what is more, the Drindari words are written in a kind of phonological deep structure that is difficult to interpret in phonetic terms.<sup>5</sup> Information on Kalajdži is restricted to what can be found in the articles of Gilliat-Smith (1935, 1940, 1944), and to a translation of the Gospel of Matthew by Atanasakieff (1932; Tatar Pazardžik). As already shown by Gilliat-Smith, the linguistic forms of the gospel translation are not fully reliable. In my view, it is Kalajdži with an admixture of Erli word forms, together with hypercorrect morphological forms. The translator appears to have had certain preconceived ideas about how to write the dialect.<sup>6</sup>

Judging even from the insufficient sources available for Drindari we can state that at the time around 1900 no uniform Drindari existed, and the same can be said for Kalajdži. Therefore, we cannot exclude the possibility that there existed transitional dialects between typical Drindari and typical Kalajdži. Especially, we have to be cautious with texts displaying poetic or metrical language (ballads, songs) since, as already pointed out by Gilliat-Smith, in this type of text dialect mixture comes about very easily.

The dialects under examination preserved a considerable number of Greek words, in all likelihood more than any other Romani dialects outside of Greece, the reason being that contact with speakers of Greek was interrupted only a few centuries ago.<sup>7</sup> Unfortunately, we do not know whether our dialects are also spoken in Greece even now. At least for the Kosovo Bugurdžides it is evident that their ancestors immigrated from the south (Macedonia) or from Bulgaria: the phonetic make-up of some Slavic words is Macedonian-Bulgarian rather than Serbian.

## 2. Sound changes

As will be seen in the following there are many sound changes not only common to the three dialects but unknown outside this group.

### 2.1 *Palatalisations*

Processes of palatalisation that can be said to be typical of our group, exceed those known from other dialects, for instance from South Balkan I (Arli or Erli), both in number and degree, although our dialects did not undergo precisely identical changes. On the other hand, it cannot be claimed that one of the dialects has gone farther in this respect than the other two. Some

outcomes of palatalisation can be viewed as sound reductions, and there are other reductions independent of palatalisations, for instance reductions of intervocalic consonants. All these processes cooperated in making the words shorter. This tendency holds least for Bugurdži, it is stronger in Kalajdži, and most characteristic of Drindari, and it is therefore the case that Drindari is less understandable for speakers of other dialects and less transparent for linguistics than other dialects of Romani.

The older state can be identified in most cases by comparing South Balkan II words with those of the more conservative South Balkan I or Vlach dialects. A feature shared by all three dialects is that palatalisation takes place before [i] and [j], but never before [e] (as for instance in the Gurbet-type dialects of Vlach).

2.1.1 Palatalisation of velars (see Map 8). The velars [k, kh, g] changed to [ç, ch, (d)z] before [i] and [j], which is exceptional in Romani, because normally it goes no further than [č, čh, dž] or [č, čh, dž]. This also means that the outcome of this late palatalisation did not merge with the affricates [č, čh, dž] of Old Indo-Aryan, which is only understandable if we assume [č, čh, dž] to be the intermediate stage of the process. Examples from Bug. are *cin-* ‘to buy’ < *kin-*, *chil* ‘butter’ < *khil*, *vozi* ‘soul’ < *vogi*; *locola* ‘relieves’ < *lokjola*, *phadzilo* ‘broke’ < *phagilo*. In a number of other cases the palatalisation has been leveled owing to analogy, as in Fem. *loki* ‘easy’ instead of \**loci* (Masc. *loko*), and *kašukjar-* ‘to deafen’ from *kašuko*, or did not occur because the following [i] is not original, as in *kid-* ‘to collect’ (cf. *ked-* in other dialects), and *kibor* ‘how much’ (cf. *kebor* or *kabor/kobor* in other dialects).

In Kalajdži, the same state has been reached throughout; cf. *cinne* ‘they bought’ < *kinde*, *kercipe* ‘bitterness’ < *kerkipe*, *pacona* ‘they believe’ < *pakjona*, *ozi* ‘soul’ < (v)*ogi*, *manzin* ‘wealth’ < *mangin*, *zeb-* ‘to sing’ < *giljab-*, as well as *ćiro* ‘time’ < Greek *καρός*, and *ci* ‘and’ < *καί*. Exceptions are *kid-* ‘to collect’, *kibor* ‘how much’, *kidja* ‘thus’ and *biken-* ‘sell’, precisely as in Bugurdži.

Drindari displays the same results; cf. *cina* ‘I buy’ < *kinava*, *cirus* ‘time’ < *καρός*, *cisi* ‘bag’ < *kisi*, *machi* ‘fly’ < *makhi*, *zab-* ‘to sing’ < *giljab-*, *sizila* ‘accelerates’ < *sigjola*, and additionally *loci* ‘easy’ < *loki*, *banzi* ‘crooked’ < *bangi*, that is the process has not been blocked by analogy as in Bugurdži. Exceptions are *khid-* ‘to collect’ (Bug. and Kal. *kid-*), *kide* ‘thus’ (Bug. and Kal. *kidja*), *kikesai* ‘such’ < *kikisavi*, *kikid-* ‘to squeeze’, and *kici* ‘how much’

(perhaps *keti* originally). Totally different is *alapilu* ‘was found’ < *arakhilo* (which gave *arapcilo* in Bugurdži).

The change of the velars to [c] and [dz] or the transition of [č] and [dž] to [c] and [dz] is not known from Balkan-Slavic, but it does occur in Aromanian. However, it is unlikely for other reasons that this language had an influence on the development of the dialects of our group. Since average Greek has only [c] and [dz] and no [č] and [dž], it cannot be excluded that it furthered the transition from [č] and [dž] to [c] and [dz] in South Balkan II, but apparently not in a direct way.

2.1.2 Palatalisation of velar + [l] clusters (see Map 10). In Bugurdži, the clusters [kl, khl, gl] became [č, čh, dž]<sup>8</sup> before [i] and [j]. The resulting sounds thereby merged with the Old Indo-Aryan affricates. We must assume that this change occurred later than the one discussed under 2.1.1 because [lʰ] first had to be changed into [j], which in turn would affect the velar. Examples from Bugurdži are *dičhom* ‘I saw’ < *dikhljom*, *ičava* ‘I go out’ < *ikljava*, *kandži* ‘comb’ < *kangli*. There are exceptions to this change even among the few words containing the clusters in question: *ikicilo* ‘came out’ < *iklistilo*, very likely in order to avoid \**ičicilo* or similar; *čirikli* ‘bird’, apparently in analogy to *čiriklo*; further the strange passive forms *dišola* ‘is visible’<sup>9</sup> instead of \**dičhola*, and *arapcola* ‘is found’ instead of \**aračhola*, with an additional *p*. It should be stressed here that both passives have meanings deviating from those of the respective active forms, *dikhla* ‘sees’ and *arakhla* ‘finds’.

The other two dialects did not develop along the same lines. In Kalajdži, [klj] resulted in [k], that is a depalatalisation of [k(h)lj] via [k(h)j] to [k(h)] must have taken place; cf. *mukom* ‘I left’ < *mukljom*, *dikhan* ‘you saw’ < *dikhljan*, and *ikola* ‘comes out’ < *ikljola* (no examples available for the development of [glj]). Kalajdži *ikist-* ‘came out’ < *iklist-* reminds us of Bugurdži *ikicilo*, but we cannot be sure whether the forms of the two dialects go back to a common ancestor form.

In Drindari, the clusters under discussion gave [ki] and [kj] respectively, that is at this stage the velar resisted palatalisation (as in Modern Greek and Bulgarian, where [kʰ] is not further changed). In some inflectional forms [j] disappears through vowel assimilation; cf. *dikhjas* ‘saw’, but 1st Sg. *dikhim* < *dikhljom*; *inkjel* ‘comes out’ < *inkljel*, but *inkila* id. < *inkljola*; *raki* ‘girl’ < *rakli*, *inkistu* ‘came out’ < *inklisto* (Žeravna near Kotel; Gilliat-Smith 1931); *nakhistu* ‘passed’ < \**nakhlisto* (instead of *nakhlo*).

The examples cited demonstrate that Drindari is the most conservative of the three dialects, its forms can be seen as pre-forms for both Bugurdži and Kalajdži, which developed in different directions: in Kalajdži the palatal element was lost, whereas Bugurdži underwent further palatalisation. There are some apparent exceptions in Drindari: *phazi* ‘broken’ < *phagli*, and *phuci* ‘swollen’ < *phukli*, but these forms might go back to *phag-ol-i* and *phuk-ol-i*, and if this were the case, the sound change would have been regular. I was not able to find any ‘regular’ reflexes of [glj].

The surrounding languages did not undergo similar changes. In Aromanian, [klʲ] and [glʲ] have been preserved; in the Albanian dialect of Kosovo former [kl] and [gl] gave mere non-palatal [k] and [g], which is not in congruence with the changes observed in Bugurdži (cf. above); in Bulgarian and Greek, [kli] and [gli] survived. In Macedonian, *kl’u-* became *klu-*, that is *l* has been preserved even in this special case.

2.1.3 Palatalisation of dentals (see Map 9). In Bugurdži, [ti, thi, di] became [ci, chi, (d)zi] respectively, that is the results are the same as for the velars. Palatalisation of dentals can be found in Vlach, it is also known from Northern Central dialects,<sup>10</sup> but almost unattested for South Balkan I dialects. In our dialects, this change is the most irregular of all processes discussed so far, it did not include all pertinent words; it did not occur in *dikh-* ‘to see’ anywhere (cf. Modern Indic *dekh-*, Lomavren *lekh-* with [e], which does not trigger palatalisation in our group), but it did affect *tikno* ‘little’ even in dialects showing no palatalisation of this type in general (generalised emphatic change, supported by forms used in baby talk?). Examples of this change are *buci* ‘work’ < *buti*, *phurca* ‘bridges’ < *phurtja*, *gozi* ‘brain, sense’ < *godi*. There exist, alongside the regular preterite forms *kerdzom* ‘I did’, *phendzom* ‘I said’, forms like *kerom*, *phenom*, without any preterite marker. These forms can be reconstructed as *kerom* < \**kerjom* < *kerdjom*, but we have no explanation as to why it is only the *d* preterites that underwent total reduction. Palatalisation affected even loan words that must have entered the dialect in more recent times; cf. *vrizla pes* ‘turns around’ < Slavic *vrti se*, or *sahaci* ‘hour’ < Turkish *sahat* (via \**sahati*). The preterite *rodjom* ‘I sought’ as well as the passive *cidjola* ‘is drawn’ appear to be aberrant, but as can be gathered from the 3rd Pl. *rodle*, they must be reconstructed as \**rodljom* and \**cidljola*. This secondary [j] did not provoke further palatalisation. On the other hand, the preterite *čičom* ‘I put’ < \**čhitljom* (cf. 3rd Pl. *čhitle*) does not agree with *rodjom*, since here [j] did palatalise the dental.

Kalajdži is at the same stage as Bugurdži: we have *buci* ‘work’ < *buti*, *bucate* < *butjate*, *uxci-* ‘to jump’ < *uxti-*, *morcende* ‘skin’ < *morthjende*, *pacarde* ‘they wrapped’ < \**patjarde*, *šunze* ‘heard’ < *šundja*, *gozever* ‘clever’ < *godjaver*, *benzipe* ‘birth’ < *biandipe*, *zis* ‘day’ < *dives*, *avzies* ‘today’ < *avdives* (against *do diesendar* ‘of two days’ and *-disengo*), and in loan words *vakæci* ‘time’ < Turkish *vakit*, *paranzer-* ‘command’ < \**porændjar-* (from Slavic *po-ročiti* ‘to order’, with the nasal vowel preserved?). Verb forms are not treated homogeneously. On the one hand, we have depalatalised forms like *beštom*, *beštas* ‘sat’, *bičhaldas* ‘sent’, *antomas* (for *andjomas*) ‘brought’, *sastav* ‘I recover’ < *sastijav*, on the other hand forms without any preterite marker like *kerom*, ‘I made’, *sasterom* ‘I recovered’, *ašarom* ‘I praised’ (cf. Gilliat-Smith 1935: 28, but in other texts only *kerdom*, *kerdan*), *kere* < *kerdja*, *kerem* < *kerdjam* (cf. Bugurdži), but alongside regular *šunze* ‘heard’ < *šundja*, and *anzen* ‘you brought’ < *andjan*.

Drindari shows the same results; cf. *cikořo/cikuřu* ‘tiny’ < *tiknoro*, *macisa* ‘you get drunk’ < *matjosa*, *sahcil* ‘he recovers’ < *sastjol*, *taci* ‘warm’ < *tati*, *kici* ‘how much’ < *keti*, *buceti* ‘work’ < *butjate*; in Greek -ίτικο: *altæjciki* ‘golden’ < *altîn-itiki*; in turcisms: *sahaci* ‘hour’ < *sa(h)at*, *vakæci* ‘time’ < *vakit*. For [di] > [zi] we can quote *zis/zes* ‘day’ < *dives*, *zivisejlu* ‘day broke’ < *divesajlo*, *xurzjaras* ‘shortened’ < *xurdjaras*, *xæizer-* ‘to feed’ < \**xadjar-* (causative of *xala*), *cizjas* ‘collected’ < *cidjas*, *vazini* ‘they lifted’ < *vazdine*, and with oppression of one sibilant *čujzina* < \**čuzizina* ‘they surprize’ < *čudizena*, *khisejli* < \**khizisejli* ‘gathered’ < *khidisajle*.

The *d*-preterites have different forms in the various sub-dialects; in Razgrad (Kenrick 1967) *kerim* ‘I made’ < *kerdjom*, comparable to Kalajdži and Bugurdži *kerom*, *čhim* ‘I put’ < \**čhivdjom*, *gaim* ‘I took out’ < (*in*)*galdjom*, *pheras* ‘filled’ < *pherdjas*, but the stems ending in *d* show other reflexes: *roziim* ‘I sought’ < \**rodijom*, *phuzjas* ‘blow’ < *phurdjas*. Exceptions with [dj] preserved are *thodjas* ‘put’, *odija* ‘that’ < *odoja*, *kidje* ‘thus’ < \**kidija*. Of these, only *odija* is explainable (because of secondary *di*). The preterites of the *ker-Ø-om* type that lack a marker are present in all three dialects, which is a very important congruence, because no convincing explanation for the dropping of [dj] can be put forth. Therefore, we must assume that the process began in the common ancestor dialect, and was continued in the individual dialects that emerged after the split-up. On the other hand, we do not know why the regular reflex of [dj], [(d)z], has been preserved in a number of cases. Note the parallel development of *ker-d-jom* ‘I made’ > *ker-Ø-om* and *muk-l-jom* ‘I left’ > *muk-Ø-om* in Kalajdži.

Among the surrounding languages only Aromanian changed [ti] and [di] to [ci] and [dz], the results being the same as with the dentals.

2.1.4 Palatalisation of [st] and [št]. In Bugurdži, [sti] changed to [ci] and [šti] to [či], but the few cases do not behave homogeneously; cf. *anrucu* ‘ring’ < *anrusti*, *kucik* ‘belt’ < *kustik*, *uči!* ‘get up!’ < *ušti!*, but *paracuj* ‘friday’ < *paraštuj* and *sascola* ‘recovers’ < *sastjola* (apparently influenced by *sasto* ‘healthy’. The intermediate stages should have been \*[sci] and \*[šči] respectively.

From Kalajdži we have *naši* ‘cannot’ < *našti*, but *uči* ‘get up!’ < *ušti*, and *učias* ‘got up’ < *uštjas*. There are too few examples to set up rules for regular changes.

Drindari shows different reflexes, apparently depending on the morphological category involved; cf. (*a*)*ši* < *ašti* ‘can’ and *naši* < *našti* ‘cannot’, but *našči/nahči* < *našti* ‘she fled’; *uša* ‘I get up’ < *uštjava*, *uši!* < *ušti!* ‘get up!’ (but in a song old *ušti!*; p.c. Kenrick), *uxici* ‘went down’ < *uxistili*, *nakici* ‘passed’ < *nakistili*; on the other hand *beščim* < *beštjom* and *bešči* ‘sat’ < *bešti*. Former *anrucu* ‘ring’ has given *anrucu* although there was no related grammatical form to enable a restitution of [sci] for [ci]. Zlatarica (p.c. Igla) has *inkiscim* ‘I came out’ < *inklistjom* and *nakhiscim* ‘I passed’ < *\*nakhlistjom*.

As can be seen from the data, Kalajdži and Drindari go together against Bugurdži. The complex process may have started in the common ancestor dialect, but the development followed a different path in Bugurdži: after original [šti] became \*[šči], the cluster was simplified to [či] in Bugurdži, but to [ši] in the other dialects. Forms like *sascola* in Bugurdži or *bešči* in Drindari can be explained by the influence of related non-palatalised grammatical forms like *sasto* and *bešto*. Generally, analogical restoration of forms is possible at any stage of development. To the best of my knowledge there are no neighbouring languages displaying an identical change.

2.1.5 Palatalisation of [l] (see Map 11). In the dialects of our group, [l] has become [i] and [j] before [i] and [j], respectively; that is, the lateral was lost just like in the clusters [kli] and [gli] (cf. 2.1.2). While the process [lj] > [j] is known from other dialects of Romani as well as from other languages, [li] > [i] word-initially can be classified as an exceptional sound change bearing witness to the high degree of uniformity within our group. Examples from Bugurdži are *il* ‘paper’ < *lil*, *ikha* ‘nits’ < *likha*, *momej* ‘candle’ < *momeli*,

*hojame* ‘enraged’ < *xoljame* (Greek), *imori* ‘grave’ < *limori* (Greek), *bešjom* ‘sat’ < *bešljom*, *gjom* ‘went’ < *geljom*, *muj* ‘died’ < \**muí* < *muli*, but *ljom* ‘I took’ and *livazi* ‘meadow’ < *livadi* (Greek) remained unchanged.

In Kalajdži, the results of the process are the same, except for minor details; word-initially we have *il* ‘paper’ < *lil*, *indralate* ‘sleep’ < *lindra-*, but again unchanged *liom* ‘I took’, *line* ‘they took’, word-internally *muí* ‘died’ < *muli*, *ainč* ‘field’ < *alindž*, *zoraí* ‘strong’ < *zoralí*, *misiz-* ‘to think’ < *misliz-* (Slavic; against *misliz-* in Bug.), *žuí* ‘woman’ < *džuvlí*, *gíne* ‘went’ < *gelíne* and *gijas* < *geljas*, *aen* ‘you came’ < *aljan*, *xajom* ‘I ate’ < *xaljom*, *pašiar-* ‘to put to sleep’ < *pašljár-*, *kojona* ‘they become tender’ < *kovljona*, and *deino* ‘crazy’ < *delino* (against Bug. *dinilo*). In agreement with other formations of the preterite depalatalisation has taken place; cf. *barilom* ‘I grew up’ < *bariljom* (roughly parallel to *kerdom* ‘made’ and *dikhom* ‘I saw’, *čitas* ‘put’ < *čhitljas* (3rd pl. *čitle*), diminutive *moloři* ‘wine’ < *moljoři*. Matth has no examples for depalatalisation: *giom* ‘I went’ < *geljom*, *bokajom* ‘I got hungry’ < *bokhaliljom*, and even *trušaom* ‘I got thirsty’ < *trušaljom* (authentic form?) for intervocalic [lj].

Drindari has carried out the same changes, for instance *il* ‘paper’ < *lil*, *im* ‘snot’ < *lim*, *indra* ‘sleep’ < *lindra*, *imori* ‘grave’ < *limori*. This palatalisation is often accompanied by umlaut and contraction creating word forms considerably deviating from those of the non-palatalised ones; cf. *ai* ‘came’ < *ali*, *píi* ‘drank’ < *píli*, *alakhim* ‘I found’ < *arakhljom*, *alápíi* ‘was found’ < \**arakhlíli*, *melái* ‘dirty’ < *melalí*, *bokí* ‘white bread’ < *bokolí*; in loanwords: *možila*/*možela* ‘bid’ < *molizela*, *kubica* ‘small hut’ < *kolibica* (both from Slavic), *xalaj* ‘well-being’ < *xalalí* (Turkish). Examples for [lj] > [j/i] are: *bujeti* ‘bottom’ < *buljate*, *šileim* ‘I got cold’ < *šilaliljom*, *zabála* ‘sings’ < *giljabala*, *pašá* ‘I lie down’ < *pašljovava*. This sound change must have been without exception in earlier times, but it did not comprise later instances of [li] or [lj]; cf. *khangíli* ‘church’ < *khangeri*.

Aromanian, the language sometimes cited to explain sound changes in our dialects, is more conservative, the lateral being preserved as [lʲ] in all cases.

2.1.6 Palatalisation of [n] (see Map 12). Only in this process do the three dialects differ significantly. Bugurdži is conservative in preserving the nasal throughout; cf. *pani* ‘water’, *khoni* ‘suet, fat’ not even with slight palatalisation ([nʲ]) before [i], but Pl. *phenja* ‘sisters’, denominal *sanjarava* ‘I make thin’ with [nʲ]. There is a problem with the preterite forms *ljom*

'I took' and *đjom* 'I gave' for which in other dialects *liljom/linjom* and *dinjom* can be found, and in Bugurdži 3rd Pl. *line* and *dine*, that is forms containing an [n]. If *ljom* and *đjom* are regular outcomes they should be derived from *liljom* and \**diljom*, forms attested in some of the South Balkan I dialects. The problem arising from this reconstruction is that we would have to accept non-homogeneous stem formation: *lin-* for the 3rd Pl., but *lilj-* for the other persons. Since 'mixed' paradigms are attested for Bugurdži in other cases, the explanation given here looks acceptable.

Kalajdži and Drindari reduce [ni; n'] to [i; j] respectively. Kalajdži has *kui* 'elbow' < *kuni*, *phejatar* 'sister' < *phenjatar*, *die* 'gave' < *dija* < *dinja*, *guri* 'cow' < *guruvni*, and in loans from Turkish *bostai* 'garden' < *bostani*, *zəndai* 'jail' < *zəndani*. A few cases of depalatalisation can be presented too: *romnatar* 'Gypsy woman' (apparently because of the cluster *mn*), and diminutive *phenorije!* 'little sister!'.

Drindari too lost the nasal consonant throughout, but in some varieties the vowel originally following [n] has been nasalised (see Gilliat-Smith 1914: 271f.), that is a phoneme is transformed into a distinctive feature of another phoneme. Undoubtedly, this must have been the older stage, but some varieties have no nasal vowels (anymore). Examples are *pureĩ* 'old' < *purani*, *ruĩ* < *runi* 'wept', *beĩpi* 'birth' < *bianipe*, *peĩnda* 'fifty' < *peninda*, *eja* 'nine' < *enja* (after vowels); *kahĩ* 'hen' < *kaxni*, *řomi* 'Gypsy woman' < *řomni*, *parĩ* 'white' < *parni* (after consonants). It is possible that even *sias* 'was' derives from \**sinjas*, a form based on *sine*.

We have no convincing explanation as to why in Bugurdži the nasal is more constant than in the other dialects, though cross-linguistic evidence proves that the lateral is accessible to palatalisation more easily than the nasal (cf. dialects of Croatian, Hungarian, and Spanish). This would mean that in Bugurdži the nasal simply lags behind (although the surrounding Kosovo Albanian changed [n'] to [j]). But how is it that the dental plosives [t] and [d] did undergo palatalisation, while the dental nasal was preserved?

It is evident, then, that the three dialects display a very similar behaviour with regard to palatalisation. The differences are restricted to the following phenomena: in Bugurdži, [ni] and [n'] have been preserved, perhaps since this dialect was separated from the common ancestor before the nasal underwent palatalisation. On the other hand, Kalajdži and to a lesser degree also Drindari show depalatalisation (dejotisation) processes not known in Bugurdži. In this, Bugurdži is more similar to Arli and other dialects of the South Balkan I group. If we follow the assumption that all palatalisation

processes started as early as in the ancestor dialect, the split into individual dialects must have taken place at the stage of [k'/t'] < [kli, klj].

The languages of the world abound with palatalisation processes, which are considered natural because they can be interpreted as processes of assimilation in most cases. Nevertheless, we have no explanation for the fact that our dialects have been exposed to such processes to a much greater extent than any other group. In most cases, contact with Balkan languages could not have triggered changes like these, since similarities between palatalisation processes in our dialects and in the neighbouring languages are only superficial. Note especially that in the Balkan languages the velars are weakly palatalised before both [i] and [e], whereas in our dialects [e] does/did not have a palatalising effect. Aromanian is similar to our group in that the palatalisation of both velars and dentals gave [c] and [dz], but it did not serve as a model for the treatment of [kli], [sti], [li-] and [ni, n']. We have no clues for determining the time of these changes, and therefore we cannot exclude that some of them started as early as in the pre-European period. The palatalisations make our dialects clearly distinct from the more conservative South Balkan I dialects (see Boretzky 1999c).

## 2.2 *Various other changes*

2.2.1 *Vowel changes.* The vowel system has not been changed much; the familiar five basic vowels /i e a o u/ survived as phonemes, but were changed by assimilation and contraction in individual words (see below).

In Bulgaria, the new vowel [ə] (see Map 3) was introduced into the inherited lexicon, apparently under the influence of Bulgarian dialects having changed [a] to [ə]. In varieties of Drindari [ə], and sometimes also [i], arose under conditions that are unclear; cf. *kər̥la* 'make' < *kerela*, *p̥iʁo* 'foot' < *piʁo*, *ž̥inla* 'know' < *džanela*, *nənəj* 'is not' < *nanaj*. According to Kenrick (1967: 71), an essential characteristic of Drindari is the 'instability of vowels and mouillierung of consonants'. (Examples of the optional change of *a*, *e*, *i*, and *u* are given on p. 71, *op. cit.*<sup>11</sup>)

In Kälajdži, we find [ə] and other vowels, mostly from [e], especially in the environment of [ř] and [r]: *saor̥u* 'all' < *savoře*, *avru* 'other' < *avre*, *prənžanesaa* 'knew' < \**prindžandja(sa)*, *č̥rxan* 'star' < *č̥erxan*, *e tagarə* 'the king' < *e thagares*, *mařəja* 'bread' < *mařesa* (a process also documented for Kalderaš; see Boretzky 1991). The central vowel occurs quite often in the vicinity of sibilants as well; cf. *brešəndar* 'year' < *brešendar*, *pašə* 'near' <

*paše, šužu-vozenge* ‘of clear conscience’ < *šuže-vogjenge, zisənde* ‘day’s < *divesende, perəseja* ‘word’ < *pherjasesa, žəne* ‘persons’ < *džene, čača* ‘true’ < *čače*, and sometimes even in the 3rd Sg. -əl < -el.

The assumption of late Bulgarian influence is corroborated by the fact that Bugurdži did not develop centralised vowels, but that Rhodope Erli (South Balkan I; see Iгла 1997: 150) does have them. It has to be stressed that Bulgarian dialects did not offer a simple model for changing other vowels to [ə] since normally only unstressed [a] has been reduced to [ə] (see Stojkov 1993: 97). One may imagine that there was a pressure on the part of the second language spoken so frequently in everyday life to introduce the centralised vowel in the first language as well, under whatever conditions. Reduction of unstressed [a] is documented in Bulgarian from the thirteenth century on (cf. Mirčev 1963: 129), but this does not imply that this process was immediately transferred to Romani. Therefore, we do not obtain by this distinction a terminus ante quem for the separation of Bugurdži from the other dialects.

In Kalajdži, an unusual fluctuation of vowels can be observed, at least outside the word roots. Often the vowels appear to have been harmonised, but by no means consistently; cf. from Matth *manušunge* ‘men’ < *manušenge* (but conservative *manušen*), *akušuna* ‘they hear’ < *akušena, šoro* ‘head’ against *šəre* ‘heads’ (an alternation also found in Gurbet and Kalderaš), *počun!* ‘ask!’ < *phučen!*, and *oxujela* ‘goes down’ instead of *uxjela*. No such changes are known from Drindari.

Another change, characteristic of some varieties of Drindari, is [e] > [i] (see Map 4), again without being clearly conditioned by the phonetic environment. At any rate, stress does not play a decisive role (as for instance in a similar reductive process of Bulgarian), that is [i] does not appear in unstressed syllables only. This new [i] does not cause palatalisation, which allows for the conclusion that the palatalisation processes were accomplished before [e] became [i]; cf. *dikhisa* ‘you see’ < *dikhesa* and *dikhim* ‘I saw’ < *dikhjom*, but exceptional *khəila* ‘plays’ < *khelela* and *dəiske* ‘father’ < *dadeske*. This process has been extended to late loanwords by way of analogy.

Umlaut processes took place in Drindari and Kalajdži but, again, not in Bugurdži:

Kal. [ja] > [e]: *asev* ‘mill’ < *asjav, bev* ‘marriage’ < *biav, zeb-* ‘to sing’ < *ziab-*, *poze* ‘lap, menstruation’ < *podja, dike* ‘saw’ < *dikja* < *dikhlja* (but in Matth *pav* ‘I drink’ < *piav, pel* < *piel*);

Drind. [ai/aj] > [ei/ej] or [əj]; [ja] > [je/e]; [jo] > [i]: *pai* 'water' < *pani*, *zarej* 'hairy' < *zarali*, *xojějlu* 'got angry' < *xoljajlo*, *purej* 'old' < *purani*, *kizěja* 'child' < *kizanja*, *čalejna* 'they hit' < *čalavena*, *zejěki* 'saddle' < *zenjaki*; in preterite and passive: *dikhim* 'I saw' < *dikhljom*, *roz(i)im* 'I sought' < *rodljom*, *kouřila* 'becomes blind' < *kořijola*; *turje* 'around' < *trujal* or \**truje?* In most cases [j] is secondary, the outcome of palatalisation of *n* or *l*.

Undoubtedly, the process of [ja] > [e] has to do with the alternation between [ʼa] and [e] in Bulgarian, it occurs in the Rhodope Erli (see Iglá 1997: 150f, and Boretzky 1999c: 26) and a dialect spoken in the vicinity of Serres (Northern Greece) as well, but apparently not in the Vlach dialects of Bulgaria (not indigenous in this country), nor is it documented for dialects spoken outside the Bulgarian linguistic area. On the contrary, in Vlach dialects of Romania [ja] or [ʼa] has been preserved in Romanian loans even where Romanian itself changed to [e]. The circumstances of the transfer of this change to our dialects are not at all clear. We would expect it to have been adopted somewhere in western Bulgaria, in an area where, in the Bulgarian dialects, a general change from [ʼa] to [e] occurred in more recent times, probably under the influence of the western-most dialects than as an internal sound change. In the northeast [ja] as a reflex of Old Bulgarian /ě/ has been preserved before back vowels in the following syllable, but in the Rhodope Bulgarian /ě/ has been replaced by an open *e* in all positions (see Stojkov 1993: 127). This explains the change in Rhodope Erli, but not immediately in Drindari.

2.2.2 Changes of consonants. Of the inherited consonants, nearly all have been preserved in most of the varieties.

The aspirated plosives are absent from the Kalajdži variety of Varna (see Gilliat-Smith 1935: 27). In Matth, they are marked in only a few words, by doubling the consonant; cf. *ppuv* 'earth' for [phuv]. The loss of the aspirates must be viewed as a late process characteristic of individual varieties. In the Drindari of Zlatarica, *cil* 'butter' and *maci* 'fly' with loss of aspiration can be heard (p.c. Iglá).

The strong, cerebral [ř] has not been preserved in the whole group (see Map 5): it occurs in Drindari and in Kalajdži, but not in Bugurdži; cf. Drind. *řom*, *kouřu* 'blind' < *kořo*, *kouř-ila* 'becomes blind' < *koř-jola*, *cikořo* 'little', Kal. *mořo* 'my', diminutive *mol-oři* 'wine', *mařo* 'bread'. In Bugurdži, I was not able to find any traces of it. One is led to assume that this loss took place some time ago, before contact with Albanian in Kosovo was es-

established, for otherwise the contrast between /r/ and /r̥/, present in Gegue (northern) Albanian in general and in most of the Kosovo local dialects, might have contributed to the preservation of two *r*-phonemes rather than to a merger.<sup>12</sup>

Word-final and ‘grammatical’ *s* (see Map 6) is preserved throughout in Bugurdži, and has not been lost in any of the grammatical morphemes known to have an inherited *s*, that is Obl.Sg.Masc. *-es*, Adv. *-es*, Instr.Sg. *-esa* and *-asa*, 2nd Sg.Pres. and 1st Pl.Pres. *-es* and *-as*, 3rd Sg. Pret. *-as*, and in loanwords Masc. Sg. *-os*, Masc.Pl. *-ides*, Fem.Pl. *-es*. Since the same holds for Kalajdži and Drindari as represented in the publications of Gilliat-Smith and Kenrick, the group would be very homogeneous (contrary to South Balkan I and Vlach), if there were not Matth Kalajdži, where *s* has been abandoned throughout; cf. Obl.Sg. *prale* ‘brother’, reflexive *pe*, Instr.Sg. *leja* ‘with him’, *jagaja* ‘fire’, Adv. *čača* ‘right’, 2nd Sg.Pres. *tove* ‘wash’, 1st Pl.Pres. *poklonizaa* ‘present’, 3rd Sg.Pret. *benze* ‘give birth’ < *biandjas*, *isie* ‘was’ < *isjas*, and in loans like *proroko* ‘prophet’, Pl. in *-ide*. To be sure, there are pluperfect forms like *isiesaa* ‘they were’ or *zebiesaa* ‘they sang’, but those we can qualify as non-forms coined by the translator, as was already done by Gilliat-Smith (see 1935: 31–32). It remains an open problem why *s* was lost only in this variety. The possibility that the translator created an artifact can be ruled out, since to be able to do so he would have been a sophisticated linguist with knowledge of many other dialects, which however was not the case. We can also rule out foreign influence. The varieties of Erli known to me did preserve *s*, and the Vlach dialects spoken in Bulgaria (roughly of the Gurbet type) retained it as well (p.c. Igla).

The affricate [dž] is stable in Bugurdži, but not in the two dialects spoken in Bulgaria. Here [dž] has become [ž] but, as can be read from the texts, not consistently. A tendency towards this change can be observed elsewhere in Bulgaria, in recent varieties of Erli and in Vlach dialects of the Kalderaš type, but not in those of the Gurbet type (p.c. Igla). In the older Erli as described by Gilliat-Smith (see Boretzky 1998: 126) as well as in the Rhodope Erli of today (cf. the texts collected by Nounev 1994: 48–61) [dž] did not change. This change can be characterised as a natural simplification which first took place in the marked member of a contrasting pair, that is voiceless [č] is not simplified to [š] in the same process. It is not clear whether the change under scrutiny has to do with the fact that, in Bulgarian, /č/ is a phoneme with normal distribution, while its voiced counterpart is largely restricted to Turkish loans.

In Kalajdži, [c] has often been changed to [č] (see Map 7), apparently under the influence of Turkish, which itself substitutes [č] for the [c] of loanwords. Examples are *čikono* ‘little’ (Drindari *cikoŕo*), *alapčon* ‘are found’ (Bugurdži *arapcona*), *piičajola* ‘gets married’ < \**pilitjavola* (Drind. *pilceila*), *čide!* ‘collect!’ (Drind. *cide!*), *čima* ‘a little’ < \**cima* (Drind. *thrima* < θριμα, and elsewhere also *frima*). I believe that this change is a late one, initiated in groups that were about to adopt Turkish as their first language. This assumption takes into account that, today, Kalajdži seems to be extinct.

An aberrant change occurred in *pašjar-* ‘to turn over’ and *našjar-* ‘to make run’, which gave *pasjar-* and *nasjar-* respectively.

All in all, then, our group does not display, in a reconstructed stage, sounds or phonemes or distribution rules that would constitute pre-European differences between South Balkan I and South Balkan II dialects.

### 2.3 Other syntagmatic changes: elision, vowel assimilation, and contraction

Since changes of this type often affected grammatical morphemes, they frequently resulted in a blurring of grammatical forms, to a minor degree in Bugurdži, but culminating in Drindari, where palatalisation of *l* and *n*, elision of intervocalic *v*, and vowel contraction enabled by the processes just mentioned co-occurred.

2.3.1 Elision of [e] in present and imperfect (see Map 13). This process has been carried out in Bugurdži with only a few exceptions; cf. for ‘speak’:

2nd Sg. vaker-ésa > vakér-sa	2nd Pl. vaker-éna > vakér-na
3rd Sg. vaker-éla > vakér-la	3rd Pl. vaker-éna > vakér-na

Elision did not take place for phonetic reasons in some verbs, for instance in *desa* ‘give’, *dela*, *dena*; *phutresa* ‘open’, *phutrela*, *phutrena*.

For Kalajdži it is more difficult to determine under which conditions [e] was dropped; we have *čhinna* ‘they cut’ < *čhinela*, *khella* ‘plays’ < *khelela*, *inkalla* ‘takes out’ < *inkalela*, *džanlas* ‘knew’ < *džanelas*, *vakerna* ‘they speak’ < *vakerena* on the one hand, but *o(v)esa* ‘become’, *ojela* < *ovela*, *aela* ‘comes’ < *avela*, *roela* ‘weeps’ < *rovela*, *resela* ‘arrives’ on the other hand. It seems that when *-v-* or *-j-* were already elided, [e] was preserved, but this does not explain why *resela* has not been changed (preservation of *e* after non-sonorant consonants?). More material is needed to ascertain which instances of the Cl/Cn-cluster are tolerated and which are not.

Drindari shows elision to a varying degree, apparently depending on the sub-dialect. Elision is possible where clusters containing *r* or *n* as the first constituent emerge, but not with other combinations; for instance *karsa*, *karla* ‘make’ and *žinsa*, *žinla* ‘know’, and even *bikinna* ‘they sell’ < *bikinena* with a long consonant, but unchanged *marela pe* ‘beats oneself’, *genila* ‘counts’ < *genela*, and *makela* ‘smears’. No elision occurred in other environments: *ušila* ‘gets up’ < *ušijela*, *khelisa* ‘you play’ along with *kheisa* < *khelesa*, *aina* ‘they come’ < *avena*, and loan verbs like *mizisena* ‘they think’, *čujizena* ‘they wonder’ (Slavic *misl-*, *čud-*).

In all likelihood, this was a phonetic-morphological process beginning with individual verbs, perhaps the most frequent ones. It reached a ‘natural’ end in Bugurdži, but stopped somewhere in the dialects of Bulgaria for reasons yet unknown.

2.3.2 Loss of [v] (see Map 14). In Bugurdži, intervocalic [v] is stable in most cases, grammatical suffixes included. It is lacking in the lexemes *del* ‘God’ < *devel* and *zis* ‘day’ < *dives*, in the verb *čhjava* < *čhivava*, *čhila* < *čhivela* ‘put’, and in the preterite of *av-* ‘to come’ *alo* < *avilo* (*alo* also in most South Balkan I dialects). The passive present forms are contracted throughout: *-josa* < *-jovesa*, *-jola* < *-jovela*, *-jona* < *-jovena*, which is quite usual in many other groups as well. The plural marker for loans in *-os*, *-ov-ja* (from Bulgarian/Macedonian *-ove* or *-ovi* + indigenous *-a*) has been reduced to *-oja*. A rare allegro form for 1st Sg.Pres. *-ava* is *-a*, but this form would never be used for quoting verbs. Some verbs containing the transitivity suffix *-av-* drop *-v-* and contract the rest, but I have heard such forms from a few aged speakers only, most probably allegro forms again; cf. *hohes* ‘you cheat’ < *hohaves*, *pharel* ‘tears’ < *pharavel*, *urjel* ‘flies’ < *urjavel*, and *t’ el* ‘becomes’ < *te ovel*.

We cannot exclude the possibility of the full forms having been restituted under the influence of Arli and Vlach after the Bugurdžides immigrated to Macedonia and Kosovo, but we have no direct proof for this assumption. It is not known how far elision has gone before the contact between Bugurdži and the dialects of Bulgaria was interrupted.

Elision is found in Kalajdži to a greater extent, but it did not blur distinctions between words or grammatical forms. Loss of *v* occurred in verbal roots, for instance in *tho-av* ‘I wash’, *či-en* ‘they put’, *a-ela* ‘comes’, *o-ela* ‘becomes’, *ro-ela* ‘weeps’, *ro-indos* ‘weeping’, *xoxa-ibe* ‘fraud’. The hiatus arising from this may be bridged by *-j-*, as in *ojava*, *ojela*, *ajasa*. As in Bu-

gurdži, the plural *-ovja* has been reduced to *-oja*, and the passive forms are reduced and contracted too. More reduction is found in *gurujořen* ‘cows’ < *guruvnjořen* and *kojona* ‘they become tender’ < *kovljona*. There was no change in the 1st Sg. *-ava* and the demonstrative pronouns *kodova* ‘that’, *kadava* ‘this’.

In Drindari, *v* is generally lost except for the word-initial position, that is even two *v*'s per word may be dropped thus opening the way for far-reaching contractions. 1st Sg. *-ava* appears as *-a* or *-aa*; cf. *dikhá* ‘see’, *ušaa* ‘get up’ < *uštjava*, *zába* ‘sing’ < \**zijabaa* < *giljabava*. The monosyllabic stems *l-ava* ‘I take’ and *d-ava* ‘I give’ assume the irregular forms *ləá* and *dəá*, with non-syllabic [e] (perhaps in analogy to passive forms like *barəá* ‘become big’ < *barjovava*).

In causative and passive forms *v* is dropped throughout; cf. the paradigms for ‘lie down’ and ‘understand’:

pašá < pašljovava	axája < axáljovava
pašísa < pašljovesa	axóisa < axáljovesa
pašíla < pašljovela	axóila < axáljovela

Through these changes the 1st Sg. has become isolated within the paradigm, but even the less reduced forms give room for confusion, especially because now [i] represents not only the passive (from *-jove-*), but also active *e*; cf. *dikhíla* ‘see’ < *dikhéla*.

Other noteworthy forms are: *sái* ‘which’ < *save*, *čhais* ‘boy’ < *čhaves*, *mothejla* ‘tells’ < *mothávola*, *loíza* ‘I wait for’ < *lov-iz-ava* (Bulg.), *karazi* ‘crayfish’ < *karavidi* (Greek), *žuel* ‘woman’ < *džuvel* (instead of *džuvli*). Internal *-v-* has been preserved in *everat* ‘the day before yesterday’ < *aver-rat*, *avzin* ‘honey’, *zivise* ‘in daytime’ (alongside *zis*, *zes*), and mostly in loanwords: *tavani* ‘pan’ (Turk.), *dəržava* ‘state’ (Bulg.), *sivamlo* ‘engaged’ < \**simavlo* (from Greek σημάδι ‘sign’, a loan translation mirroring the semantic relation between Turkish *nişanlı* ‘engaged’ and *nişan* ‘sign’).

2.3.3 Contraction. A precondition for contractions is the loss of consonants, mainly of primary *-j-* and *-v-*, then of secondary *-j-* arising through the palatalisation of [n'] and [l'] as well as a hiatusilger (see above). In Bugurdži there are only slightly more instances of contraction than generally found in Romani, for instance *šošes* ‘hare’ < *šošojes*, *phaba* ‘apples’ < *phabaja*, and *-o-* < *-ove-* in the present passive. Secondary *-j-* < *-li-* often leads to the loss of a syllable; cf. *zi* ‘song’ < \**ziji* < *gilí*, *gej* ‘went’ < \**gei* < *gelí*,

*kojba* ‘hut’ < *koliba* (Slavic). On the other hand, *-j-* derived from *-lj-* has not been dropped; cf. *hojame* ‘enraged’ < *xoljame*, *ajum* ‘I came’ < *aljom*.

Judging from the spelling, in Kalajdži prevocalic [i] has not been changed to [j], that is the number of syllables has not been reduced; cf. *čhien* ‘they put’ < *čhiven*, *zies* ‘day’ < *dives*; *liom* ‘I took’ < *liljom*, *gi(j)as* ‘went’ < *geljas*, *aine* ‘they came’ < *aline*, *deino* ‘crazy’ < *delino*. The same holds for [j] from [ni/n’]; cf. *bostái* ‘garden’ < *bostani*, *sias* ‘was’ < *sinjas*. However, we cannot be sure whether this ‘i’ is pronounced as a non-syllabic [j] rather than an [i]. New recordings are necessary to solve this problem.

Drindari is clearly distinct from the two other dialects in displaying more contraction, but even here no syllable is lost where only one consonant is dropped, at least in most cases. As already mentioned, even nasality survives in the shape of a distinctive feature; cf. *peí* ‘fell’ < *peľi*, *alápii* ‘was found’ < \**arakhľili*, *zareí* ‘hairy’ < *zaralı*; *kizēja* ‘children’ < *kízanja*, *grasteĩ* ‘horse-’ < *grastani*; *roín* ‘they weep’ < *roven*, *áina* ‘they come’ < *avena*. A syllable has been lost in *zábala* ‘sings’ < *giljábala*, *karazi* ‘crayfish’ < *karavidi*, *kubica* ‘hut’ < *kolibica*, *bokí* ‘white bread’ < *bokoli*, *čalejna* ‘they hit’ < *čalavena*. In other cases there have been two elisions per word, but contraction is perhaps due to only one of them; cf. *xojějlu* ‘got angry’ < \**xojájlo* < *xoljanilo*, and *urejla* ‘flies’ < \**urjeela* < *urjavela*.

Normally, two elisions per word automatically give rise to contraction. Spelling suggests that *aáa* ‘I come’ < *avava* contains three syllables, but the phonetic correctness of this spelling can be doubted. Contraction has been carried out in a small group of verb forms, as for instance *barena* ‘they get big’ < \**bariena* < *barjovena*, *dišá* ‘I am seen’ < *dišľjovava*, *axája* ‘I understand’ < *axaljovava*, *axáim* < *axaliljom*, *bilájim* ‘I (was) melted’ < *bilaniljom*, and *deĩ* ‘crazy’ < *delini*.

These processes, too, have led to a greater differentiation of our group from the South Balkan I dialects.

#### 2.4 Stress

Stress shift has not altered the outward appearance of our dialects decisively, that is the basic rules valid for South Balkan I and Vlach hold for our group too. In Bugurdži, however, there is much variation in the distribution of stress even with inherited words, which are often heard as bearing the stress on the penultimate syllable, but this may be a pragmatic phenomenon because informants, when asked to pronounce a word in isolation, give the conservative form. It is difficult to find an explanation for this variation. As

a rule, the root syllable is stressed in verbs ending in *-d*, for instance *ćidel* ‘draws’, *phándel* ‘binds’ although practice is far from consistent. On the other hand, stress does not appear to fall on the antepenultimate syllable, that is *kérdjovel* and *kérdilo* ‘make’,<sup>13</sup> as must have been the case in old Romani. Rather, the stress is on the passive formant for monosyllabic roots, while for dissyllabic roots the older accentuation has been preserved. Thus we have younger *kerdzóla* ‘is made’ and *kerdzílo*, *phandjóra* ‘is bound’ and *phandílo*, *pecóra* ‘is baked’ and *pecílo* in contrast to *astárdzola* ‘is seized’ and *astárdzilo*, *biándzola* ‘is born’ and *biándzilo*. Stress was generally shifted in present forms after the ‘thematic’<sup>14</sup> vowel [e] was elided, as for instance in *džánla* ‘knows’ < *džanéla* and *džánna* < *džanéna*. Worth of noting are *ziábava* ‘I sing’, *másek* ‘month’ and *kókalos* ‘bone’ (as in Greek κόκκαλο).

We have no direct information if there is stress vacillation also in Drindari. As in Bugurdži, we have *zába* ‘I sing’ and *másək* ‘month’. In the publications no accentuation rules are given for the *d*-verbs, but according to Kenrick (p.c.) the regular stress is on the root syllable in *vádela* ‘lifts’ although *vadíla* can also be heard. In verbs no longer perceived as compounds but originally formed with *de-* ‘to give’ stress has been shifted to the formant, hence *phudíla* ‘blows’ and *kidíla* ‘collects’. This subclass, then, has been adjusted to the general pattern. If I understand Kenrick correctly, the passive forms preserved the old accentuation even with monosyllabic roots, that is we have *phérila* ‘is filled’ < *phérdjola*, *phánzola* ‘is locked’, *lócila* ‘is relieved’. In the preterite, *bénzim*<sup>15</sup> ‘I was born’ should be distinguished from *benzim* ‘I gave birth to’ by stress exclusively, but according to Kenrick (p.c.) this is hardly maintained systematically. These rules are in contrast with what can be found in Gilliat-Smith (1914: 275f.); according to this author we have oxytonic forms with monosyllabic roots, but paroxytonic forms with dissyllabic roots, for instance *baríla* ‘gets big’ and *barilú* (older than *barílo* in Bugurdži), *sikíla* ‘learns’, but *axóila* ‘understands’ and *axáílu*.

In Kalajdži, the old accentuation has been preserved with *d*-verbs; cf. *ćidela* ‘draws’, *kídela* ‘collects’, and *čándel* ‘calls’. The scanty information available for the passive points to shift of stress; cf. *ikóla* ‘is enough’ < \**ikljola*, *čivziól* ‘is placed’, *sastílo* ‘recovered’, and *barilóm* ‘I became big’.

### 3. Morphology

Many characteristic forms of nominal and verbal inflection have come about through the sound changes outlined in the preceding sections. We

will mention them in the following sections, but the emphasis will be on the typical forms by which our group is/was distinguished from other groups.

### 3.1 *The definite article (see Map 15)*

All three dialects have the same original forms, which are also identical with those reconstructed for the South Balkan I group (see Table 1).

Table 1. The definite article

	Sg.Masc.	Sg.Fem.	Pl.
Nom.	o	i	o
Obl.	e	e	e

Bugurdži uses only these forms, but in Drindari *o* and *u* are found for the Masc.Sg. as well as for the Nom.Pl., and *e* and *i* for the Oblique; cf. Pl.Fem. *o riza* ‘the vineyards’. The variants *u* and *i* prevailing in most of the subdialects are not the product of regular sound change. According to Kenrick (p.c.) *o* and *u* are in free variation, which does not look like a sound change not yet accomplished. Instead of oblique *e* and *i* the nominative forms *o* and *u* can also be found, for instance in *o rozibnastar* ‘from seeking’ and *u Štalenca* ‘with the Turks’. Substitution of *o/u* for the Sg.Fem., which is normal in Rhodope Erli, could not be observed.

In Kalajdži there is much instability in the use of the article forms. For Sg.Obl. we also find *o* along with *e*; cf. *o grasteske* ‘to the horse’, *o Dželi-laastar* ‘from Dželil Aga’. The same holds for the Pl.Obl., for instance *o romen* ‘the Gypsies’, *o ilvalen* ‘the scribes’, *o čejenge* ‘to the girls’, *k-o sikavdenge* ‘to the teachers’ (Matth; incorrect). Also, the feminine shows a tendency towards *o* instead of older *i*; cf. both *o čərɣan* and *i čərɣan* ‘the star’, *o Güla kadına* ‘Ms. Gül’ and *o Gülasa* ‘with Ms. Gül’. In this, Kalajdži behaves like varieties of Erli.

### 3.2 *Case forms (see Map 16)*

The forms of Bugurdži are those valid for non-Vlach dialects in general: the genitive has long *-koro*, *-kiri*, *-kere*, sometimes reduced to *-kro*, *-kri*, *-kre*, but according to my impression short *-ko*, *-ki*, *-ke* can also be heard. Fem.Pl.Obl. has both *-jen* and *-en*, but never *-jan* as in Vlach. All this is congruent with what we know from South Balkan I. Noteworthy is the

jotation in Pl. *čhibja* ‘tongues’ and the singular form *phurc* ‘bridge’ along with *phurt* (influenced by the plural *phurca*?).

In Drindari, the ablative ends in *-ta/-da* but also in *-tar/-dar*, the latter occurring in Zlatarica and in the dialect of Kotel as described by Kenrick (1969). Contrary to Bugurdži the genitive has only short forms. In Kotel, the Voc.Pl. is on *-aler!* instead of *-alen!* (= Bugurdži and South Balkan I). There are some instances of Ø-plural as in *trin phal* ‘three brothers’ and *pe brăș* ‘his years’, but we have no data on the frequency of such forms. Kenrick (p.c.) claims that both *phrala* and *phral* are in use for the plural, and that Fem. *rat* ‘night’ has both singular and plural value. Other deviations from the average Romani inventory have been caused by sound change: Masc.Obl. *-is/-əs* < *-es*, Fem.Obl. *-e* < *-ja*, Masc.Pl. *-i/-ə* < *-e*, Pl.Obl. *-in/-ən* < *-en*, and Gen.Masc. *-ku* < *-ko*. In some details Drindari has forms identical with those of Bugurdži, for instance Sg. *phurc* and *phruc* along with *purt* and *prut* ‘bridge’, Dat. *čhibjaki* ‘tong’ with jotation, but Pl. *čhiba* in (both in Razgrad; see Kenrick 1967: 72).

Kalajdži parallels Drindari in having only short genitives, but the Voc.Pl. is *-ale!*; cf. *barale!* ‘you great ones!’, *romale!* Zero plurals can be found after numbers with *dies* ‘days’, *brăș* ‘years’ and Fem. *rat* (cf. *trin rat* ‘three nights’), and further in *kici drom* ‘how many times’; with non-temporal expressions in *dešuduj phal* ‘12 brothers’ and even in the ambiguous *o rom* ‘the Gypsies’. Since these forms are normal for Kalderaš and other Vlach dialects and since they frequently occur in Paspatis (1870), we may assume that they were the oldest forms for masculine plural with nouns ending in a consonant. In *amari ozi* ‘our soul’ *ozi* is treated as a feminine noun, apparently triggered by the isolated masculine ending *-i* interpreted here as a feminine marker (Drindari *ozi* and Bugurdži *vozi* remained masculines).

As for the genitive morphemes, we have no clues for deciding which of the forms is the older one. Provided *-ko* did not arise through mere phonetic reduction, this difference may go back to pre-European times. If we wanted to set up *-ko* as the common pre-form for our group Bugurdži might be assumed to have introduced *-koro* under the influence of Arli, but if *-koro* is to be viewed as the older form, Drindari and Kalajdži might have adopted the short variant from Vlach dialects. However, both solutions are not convincing. We have to take into account that Drindari and Kalajdži differ from Erli, the clearly dominating and firmly established dialect of Bulgaria, whereas Vlach is younger and less well established in this country. Perhaps *-koro* and *-ko* continue cases forms with originally different func-

tions. There is some evidence from Gurbet for the assumption that the long forms were reserved for marked uses (nominalised genitives, and stressed genitives following the head noun).<sup>16</sup>

### 3.3 *Nominalisations* (see Map 17)

The nouns formed from verbs and adjectives differ largely in the three dialects, but the weight of the differences is diminished by the fact that competing forms exist even within one dialect.

Bugurdži has *-iba*, but *-ipa* and *-ipe* occur as well. There are clues allowing us to assume that once *p*-forms were derived from adjectives, and *b*-forms from verbs, but now there are no clear-cut rules, most probably because the category is in decay. The oblique form is in *-ipnas* (for *ha-* ‘eat’ in *ha-mnas-*), but no plurals could be found.

According to Kenrick, Drindari has *-ipi* with Obl. *-ipnas* for denominal and *-ibi* with *-imnas* for deverbal formations, again with exceptions. Forms found in other texts diverge from this distribution; cf. *čhazipi* from *čhad-* ‘to vomit’, *pureïbnastar* from *purano* ‘old’, *beïbnaski* < *bianibnaski* from *bian-* ‘to give birth to’. The strange forms in *-imatar* found in Colocci (1889), *tchumimatar* ‘kissing’, *kidjimatar* ‘collecting’, for which Gilliat-Smith (1914: 281) had no explanation are most likely deformed plurals in *-imata*.

Kalajdži is even more heterogeneous: There are forms in *-ibé*, but also in *-ibó*, with Obl. *-imnas* for denominal formations, and in *-ipél-ipó*, Obl. *-ipnas* for deverbal formations, again without clear distribution; cf. *keripó* and *keribó* from *ker-* ‘to make’. In Matth plural forms in *-imata* like *roimata* ‘crying’ and *sivamata* ‘engagement’ can be found. The exceptional *-ó* in *-ibó* is most probably an adaptation to the masculine inherited nouns in *-ó* of the type *rakl-ó*, it does not seem to occur in all subdialects in all likelihood.

Loan verbs are nominalised in Romani by adding *-imos* to the stem, mainly in Vlach, but also in Anglo-Romani (see Smart and Crofton 1875: 22f.) and Welsh Romani,<sup>17</sup> and traces of this formant can be found in Russian Romani and perhaps in Caló as well (see Sampson 1926: 77), which suggests that this formant was once more widespread than in recent times. This formation has been replaced in our dialects as well as in South Balkan I and the Central dialects by *-ibel-ipe* and similar forms. The single *xamus* ‘food’ of Drindari beside *xabí* is classified by Kenrick (1969: 53) as a Vlach loan. This may be true in view of the fact that *-mos/-mus* would be expected with loan verbs in the first line.

The discrepancies between our three dialects look grave, but we have to

keep in mind that they do not mirror old differences and that the individual dialects are not homogeneous in themselves. Also, we cannot infer old differences between our group and South Balkan I from the forms attested.<sup>18</sup>

### 3.4 Borrowed nouns

The types of adaptation in Table 2 are attested for our dialects (see also Map 18).

In Drindari, the *-a* of Pl. *-uja* can be lacking; cf. *mastur-uja* ‘masters’, but *čjar-uj* ‘medicine’. Greek πέταλο ‘horseshoe’ appears in Bugurdži and Drindari as *pétalos*, but the plural is *pétala* as in Greek. Bugurdži has also *kókalos* ‘bone’ and Pl. *kókala* as in Greek, and Matth Instr.Pl. *kokal-anca* points to the same inflection. Note also the strange Masc. *o mámos*, Pl. *o mámoja* for ‘grandmother’!

Beside *-os* we have *-i* for masculine nouns in all dialects, which appears to prevail after *r* and *l*, but occurs after *k*, *t* and other consonants as well. What our group has in common with South Balkan I is that *-i* is not limited to a few final consonants, as is the case with Vlach. In Bugurdži, some nouns form an oblique in *-is* instead of *-es*, mainly but not exclusively those denoting persons. Thus we have animate *argac-es* ‘servant’ but *oficir-is*, and inanimate *inac-es-ke* ‘in defiance of’, but *šandan-isa* ‘candlestick’. The speakers are not always certain how to form the oblique, but *-is* seems to be the more productive form. For Drindari, it has become difficult to establish if there existed old *-is* alongside *-es* since both forms had to merge into *-is* (see 2.2). Most texts and descriptions available for Kalajdži do not provide us with examples, but Matth has *xambar-is-te* ‘store’, *ses-ija* ‘voice’ and *dušman-ija* ‘enemy’ (for *-isa*).

A number of words ending in *-i*, *-a* and *-e*, mainly of Turkish origin, add *-s* in the Nom.Sg.; cf. Bug. *mušteri-s* < *müşteri* ‘customer’, *báhča-s* < *bahçe*

Table 2. Adaptation of borrowed nouns

Bugurdži	Kalajdži	Drindari
-os, Pl. -oja	-os, Pl. -oja	-us, Pl. -uj(a)
-i, Pl. -ja	-i, Pl. -ja	-i, Pl. -ja
Obl. -es/-is	Obl. (-is or -ij ?)	Obl. -is
-is; -as, -es	-is; -as; -es	-is; (-as?); -es
Pl. -ides; -ades; -edes	Pl. -ide(s); (-ades ?); -ede(s)	Pl. -ides; (-ades ?); -edes

(*bahça*) ‘garden’, *kafé-s* < *kahve* ‘coffee’; Drind. *čifci-s* < *çiftçi* ‘peasant’, *gjuvé-s* < *güve* ‘moth’; Kal. *bostandži-s* < *bostancı* ‘gardiner’, *lulá-s* < *lüle* ‘pipe’, *xazné-s* < *hazine* ‘treasure’. The respective vowel also occurs in the oblique cases. In the plural, these words adopted the Greek formant *-des*, for instance Bug. *Bugurdži-des*, *hebé-des* < *heybe* ‘saddlebag’, *pašá-des* < *paša* ‘pasha’, Drind. *čifci-des*, *gjuvé-des*;<sup>19</sup> Kal. *kesidži-des* < *kesici* ‘slaughterer’, *kjušé-de* < *köşe* ‘corner’ (Matth without *-s!*). In accordance with the Turkish accentuation rules stress should be on the last syllable in singular forms, but sometimes it is placed on the penultimate syllable. In Bugurdži, *-ides* can be replaced by *-ida*, a contamination between *-ides* and indigenous Pl. *-a* being productive now. Oblique forms are rare, but we have Drind. *gjuved-uin-ca*, apparently with *-uin-* from *-ojen-* formed in analogy to Pl. *-oja* for loans in *-os*, and *parad-uin-ca* from Pl. *pár-es/par-és* ‘money’. Though Greek has Pl. *παππούδες*, Romani *papus* ‘grandfather’ did not adopt this inflection; cf. Bug. *pápus*, Pl. *pápuja*, and Drind. *pápus*, obl. *papós* (no plural attested).

The formant *-des* also occurs in South Balkan I (Paspatian, Erli, Crimean, Sepeči, but no traces in the westernmost Arli and Prilep), but it is lacking in Vlach, Central dialects, Russian and Baltic dialects. Very likely, then, *-des* has been taken over from Greek in a later contact phase, for otherwise we would find at least traces of it in other groups.

Feminine nouns ending in *-a* take the plural in Greek-derived *-es*, again as in South Balkan I. For the Kalajdži variety of Matth recall that *-s* has been dropped throughout.

### 3.5 Prepositions (see Map 19)

Deviations from the general inventory and behaviour can be noted in only a few cases. For ‘at, to’ old *k-o/k-i* has been preserved, but Kenrick gives *k-o* and Fem. *kaj*. Independent *kaj* exists as well; cf. Bug. *kaj mo phral* ‘to my brother’, Drind. *kaj mo papus* ‘to my grandfather’, Kal. *kaj po romoro* ‘to her husband’, *kaj hamami* (Masc.!) ‘to the bath’. For ‘around’ Bug. has the widespread *trujal*, but Drind. *turja/turje* and Kal. *turjal* (older form or metathesis). The preposition ‘for, because of’ governing the dative is present, but no common form can be reconstructed; cf. Bug. *ača* (< \**ašt,a?*), Drind. *astar* and Kal. *aštal*. Strange as the differences might look, we should not draw from this single case too far-reaching conclusions. ‘over’ is expressed by *opral* (*opra*), not by *opre*. Except for *bi* ‘without’ and *astar* etc. all prepositions govern the locative or the nominative (of nouns), but in Kalajdži (Matth) the oblique can also be found, for instance *palal man* ‘behind me’

instead of normal *palal mande*. This reminds us of the syntactic behaviour of Erli and of some Vlach varieties (see Boretzky 1998: 131f).

In a Drindari variety from Sliven (Marušiakova and Popov 1997: 27–32) even *te/ti* is to be found, that is a preposition known from Paspatian only; cf. *ti tu datoru* ‘to your daddy’, *te pə kəzani* ‘to her child’ (without palatalisation of *n* typical of Drindari!).

In Kalajdži, adpositions are postposed in rare cases, apparently under Turkish influence; cf. *panže kurkendar sora* ‘after five weeks’ with the Turkish adverb *sonra* in postpositional function, and *mandar tele* ‘under me’ with an indigenous element.

With regard to the prepositions our group displays no conspicuous traits nor does it differ much from other groups.

### 3.6 Pronouns

This is a section of grammar where Romani dialects differ considerably, hence it is important for historical classification. Therefore, it is essential to find out whether our group has largely identical pronominal forms or not.

3.6.1 Personal pronouns (see Map 20). In Bugurdži, the 3rd Pl. is *on*; the 1st and 2nd Pl. show the non-reduced forms *amen*, *tumen* for both nominative and oblique; the reflexive is *pes-* throughout, and there are no special plural forms such as *pen-* or *pumen-*. Drindari has precisely the same original forms: *on*; *amin*, *tumin* for both cases; *pes-* or *pis-*. For Kalajdži not all the necessary data are available, but there seems to be a differentiation between Nom. *ame*, *tume* and Acc. *amen*, *tumen*. The 3rd Pl. is *olla* in Matth, reminding us of Erli *ol/ola*.<sup>20</sup> Despite the deviating forms found in Matth, the inherited inventory for the plural must have been *amen*, *tumen*, and *on*.

In Drindari, there are the reinforced nominatives *min-da* and *tun-da* instead of *me-da* and *tu-da*.

3.6.2 Possessive pronouns (see Map 21). All dialects have differentiated short from long forms for the 1st and 2nd Sg. as well as for the reflexive.

The forms containing *e* in the feminine and the oblique, *meři*, *meře* etc., appear to be the older ones, while *mři*, *tři* can be explained by assimilation. The long forms of Drindari are likely to be *pořul/puřu* and *peři*, but they are not mentioned in Kenrick (1969). Also, *mo* is used instead of *mə* and *mi*; cf. *mo thanistar* ‘in my place’, *mo pejăki šapka* ‘my sister’s cap’, *mo k^izăja* ‘my children’—an indication of the considerable confusion of vowels rather

Table 3. Possessive pronouns

Bugurdži	Drindari	Kalajdži
moro/mo miri (meri)/mi mere/me	mořu/mu meři/mi məřə/mə	mořo/mo miři/mi miře/me
toro/to tiri (teri)/ti tere/te	tuřu (tořu)/*tu təři/ti təřə/tə	tořo/to tiři (teři)/ti tiřə (teřə)/te
poro/po piri/pi pere/pe	*pořu/pu *pəři/pi	pořo/po piři (peři)/pi peře/pe

than a tendency to generalise the masculine nominative. A special plural for the reflexive, *pomare*, is only found in Kalajdži. Since such forms have been developed in South Balkan I (among them Erli) they may have been taken over from there, especially as there are reasons to assume an influence from Erli with other phenomena too. The existence of the personal pronouns *pumen* (*pomen*) could not be verified.

3.6.3 Demonstrative pronouns (see Map 22). Romani shows a great number of formations for demonstratives, but since individual dialects possess only a section of all the forms that have become known the demonstratives are very important for recognising inter-dialectal relations. All forms described to-date are based on the following constituent parts: *ka-/ko-* and *da-/do-*, which are combined to *kada-/kodo-*, *kaka-/kuka-*, *daka-/doka-*, and extended by *-va*. Further, a particle *-ka* can be added giving series such as *akav-ka* and *adav-ka*, and an element *a-/o-* can be preposed, as in *a-kava* or *a-kada*. For our dialects at least three combinations are characteristic (see Table 4).

In Bugurdži, the other inflectional forms for *kava* are: Fem. *kaja*, Pl. *kala*, Obl. *kale*, for *kaka*: *kakja*, *kakala*, *kakale*, and for *kada*: *kadja*, *kadla*, *kadle*, that is vowel elision has affected the various paradigms to a differing degree.

Table 4. Demonstrative pronouns, masculine singular

Bugurdži	Drindari	Kalajdži
kava/kova (rare)	(kava/kova ?) <sup>21</sup>	(a)kava/(o)kova
kaka/kuka	kaka (kəka)/kuka	kakava/kokova (Matth)
kada/koda (frequent)	kada (?), kəda/kuda (odova <i>in Sliven</i> )	kadava (?)/kodova odova

As can be seen from table 4 we do not have access to all data for the two other dialects. In Drindari, there are *kəle*, *kəla* and *kole* (*kuli*), identified by Kenrick as inflectional forms of *kəka* and *kuka* respectively, but since there exist forms containing a second *k* such as Fem. *kəke* (< *kəkja*), *kəkala*, *kəkale* and *kukala*, I would prefer to assign *kəle* etc. to the nominatives *kava* and *kova*, at least in some varieties. Also, for *kuda* (*koda*) the non-reduced *kodola* is attested. Nevertheless, we cannot exclude the possibility that two paradigms, those of *kava* and *kaka*, have been contaminated. The paradigm with initial *d-* is fragmentarily documented, and seems to be of low frequency (only Fem. *odija* in Drindari and Kalajdži).<sup>22</sup>

The forms attested for Kalajdži are: *kava*, *kaja*, *kale*, *kole*. There are also *kalla* and *kanna* (see Gilliat-Smith 1915/16, II: 80 ff.), either through spontaneous lengthening or by assimilation of *kadla*, less probably of *kakla*. An indication of the existence of the third series is *kodoleske* ‘therefore’. In Matth, isolated *alkale* and *kalka* (*peresa*) ‘these (words)’ can be found, probably borrowed from Erli *akavka*, Pl. *akalka*.

For ‘such’ (see Map 23) our dialects have forms containing *i* instead of *a* or *o* in other dialects; cf. Bug. *kikiso* and *kidiso* (with distal contrast!), Drindari *kikesəu* and *kisavu/kisalu* (according to Kenrick the latter is well-attested). In Kalajdži no other form but non-inflected *kidja*, originally ‘thus’, could be found, but in Matth *ikisavo* is used, a form closely related to Drindari *kisavu*. ‘So much, so many’ is expressed by *kidibor* in Bug., by *kidebor* in Kal., but by *kikici*, *kikeci* and even *keici* (for the latter see Gilliat-Smith) in Drind., again all with the characteristic *i* (see Map 24).

This inventory is quite homogeneous and characteristic of our group, especially if we take into account that the demonstrative series *adava*, *adavka* and *akavka*, which are widespread in the South Balkan I dialects, do not occur in our group (except for *akavka* in Matth), whereas *kaka* is absent from South Balkan I. These differences are highly important since South Balkan I shows a rather low degree of group coherence. Also, *kava* or *kova* do not function as interrogatives as is partially the case with South Balkan I (again with the expectation of Fem. *kaja* ‘which?’ in Matth). There is also a clear distance between our group and the Central dialects showing *ada(va)* throughout but lacking the *kaka(va)* series.

3.6.4 Interrogative and other pronouns. The interrogative forms are those normally found in Romani: *kon* (*ko*) ‘who’, *so* ‘what’, *kaj* ‘where’, *katar* ‘from where’, *kana* ‘when’, *sar* ‘how’ and *savo* ‘which’, the latter being reduced in

Drindari to Masc. *su* and Fem. *sai* according to dialect-specific phonetic rules (see Map 25). The difference between Bug. *ko* and older Drind./Kal. *kon* is easily explained through combined Serbian and Arli influence on the dialect spoken in Kosovo. Adverbs such as *kate/kote* in the interrogative function of ‘where?’ and pronouns such as *kava/kova* as interrogatives for ‘which, who?’ that are typical of South Balkan I are only found in Matth Kalajdži, once more a candidate for Erli influence on Kalajdži.

It must be stressed here that *kate* and *kote* are lacking in our group even in their traditional meaning ‘here’ and ‘there’; Bugurdži has *katka*, *kot(h)ka/kutka*, Drindari *kətka*, *kutka*, and Kalajdži *kaka* < *katka* and *kokha* < *kot(h)ka* instead, again identical formations (see Map 26).

The expressions for ‘how much, how many?’ are much less homogeneous, but this does not mean much in view of the fact that ‘how much?’ is the element with the greatest inter-dialectal variation in Romani; Bug. *kibor*, Kal. *kibor/kebor*, but Drind. *kici* (see Map 24).<sup>23</sup> Of course, there must be a reason for this ‘wild’ distribution of so many forms in Romani. There might have been differences in meaning, for instance *kabor* ‘of what size’ against *keti* ‘how much’, but this does not explain all of the forms.

Negative pronouns (see Map 27) make use of Slavic *ni-*; cf. Bug. *nikho* ‘nobody’ (with emphatic aspiration as in Obl. *nikhas*), *nijek* ‘no one’, but *ništo* ‘nothing’ from Serbian; Drind. *nikon*, and the strange formation *nikhanžuřu*, probably for ‘nothing’;<sup>24</sup> with the same meaning also *nisu* and *nikhač*; Kal. *khonik* and diminutive *khonikořo* instead of *niko*, further *nijek* and again *nikač*, and in Matth *na i kači* (an artificial form?). The formations *khonik* and *khanči* are typical of Vlach, which makes me believe that the translator borrowed *khonik* from a Vlach variety deliberately, while *nikač* differs so much from Vlach *khanči* that it must be indigenous. Bugurdži may have had a similar form later replaced by Slavic *ništo*.

Indefinite pronouns (see Map 28) are not formed according to homogeneous rules; cf. Bug. *haj* and *hajek* ‘some’, *haj-či* and *hajči-pota* ‘something’, Drind. *kajek* < *kaj-jekh*, *vare-so* and *čipa* ‘something, anything’.<sup>25</sup>

### 3.7 Numbers

The numbers have more or less identical expressions in our dialects, except where they have been replaced by later loans. For ‘nine’ we have *enja* (not *inja*), for 30 and 40 *trianda* and *saranda*, and for 50 *penda* in Bug. and *peinda* in Drind. Tens and units are tied by *-u-*, not only with 11–19, but also with higher numbers, for instance *biš-u-duj* 22. This is an innovation shared

by South Balkan I, Southern Central dialects and our group.<sup>26</sup> While in Bugurdži for 60–90 the conservative formations *šovardeš* etc. have been preserved, they are replaced by Turkish or Bulgarian loans in Drindari. For the ordinal ‘first’ *jek-to* can be found, but more frequent is old *avgol/avgu* in Kalajdži and Drindari, and Slavic *prvo* in Bugurdži. For ‘half’ Bugurdži has *epaš* + article, but Drindari *ek-kaš* < *ek-paš* and *ek-varš* (see Map 29).

### 3.8 Verb

3.8.1 The Copula inflection (see Maps 30–32). The three dialects have in common that the present stem contains a [j], and that the suppletive form is (was) conservative *ovel* (Pres.) and *ulo* (Pret.), but there are important differences in the 2nd Pl. Pres. and the 3rd Sg. Pret.

Table 5. The Present paradigm (see Map 30)

	Bugurdži	Kalajdži	Drindari
Sg. 1st	sjom (sijom)	(i)sjom	sim
Sg. 2nd	sjan (sijan)	(i)sjen	sjan
Sg. Pl. 3rd	si; Neg. nane	(i)si), (i)sine; Neg. nanaj	si (isi); Neg. nanəj
Pl. 1st	sjam (sijam)	(i)sjem	sjam (sijam)
Pl. 2nd	sjen (sijen)	(i)sjen (u. isine!)	sjan(əs)

Table 6. The Imperfect paradigm (see Map 31)

	Bugurdži	Kalajdži	Drindari
Sg. 1st	sjomas	(i)sjoma(s)	simas
Sg. 2nd	sjanas	(i)sjanas	sjanas
Sg. Pl. 3rd	sine; Neg. nasine (nanas)	(i)sija(s), isje	sjas (ses)
Pl. 1st	sjamas (sijamas)	(i)sjemas	sjamas (semas)
Pl. 2nd	sjenas (sijenias)	(i)sjenas <sup>27</sup>	sjanas (senas)

The forms *ses*, *semas*, *senas* are used in Zlatarica (p.c. Igla).

Table 7. The Subjunctive

	Bugurdži	Kalajdži	Drindari
Sg. 1st	te ovav	te ojav	te oja
Sg. 2nd	te oves	te o(j)es	te o(j)es/ oís
Sg. 3rd	te ovel	te o(j)el	te o(j)el/ oíl
Pl. 1st	te ovas	te ojas	te ojas
Pl. 2nd	te oven	te o(j)en	te o(j)en/ oín
Pl. 3rd	te oven	te o(j)en	te o(j)en/oín

Table 8. The Preterite of 'become' (see Map 32)

	Bugurdži	Kalajdži	Drindari
Sg. 1st	ujom	ojom/ujom	(uím)
Sg. 2nd	ujan	ojan/ujan	(ujan)
Sg. 3rd	ulo; Fem. uj	ojas/olo/ulo	(ulu; Fem. uí)
Pl. 1st	ujam	ojam/ujam	(ujam)
Pl. 2nd	ujen	ojan/ujan	(ujan)
Pl. 3rd	ule	ole/ule	(uli)

A number of forms are easily explained by sound changes observed elsewhere too: Matth *isje* 'was' instead of *isjas* by 'umlaut' and loss of -s in this Kalajdži variety; Drin. *sim* < *sjom* as in the preterite of other verbs; and in *ovel* loss of intervocalic -v-. As can be seen from the paradigms, different varieties of Kalajdži underwent different changes. If the translator did not commit an error, the 2nd Pl. *isine* in Matth can be explained by analogy to the present paradigm of verbs (2nd and 3rd Pl. in -en).

The original copula stem cannot be reconstructed for our group with absolute certainty. Drind. *sjan* and Kal. *isjen* may be derived from \**isinjan*, that is a form with an additional *n* as found in many South Balkan I dialects (though not in the three easternmost dialects, Paspatian, Ursaritiko, and Crimean, preserving *som*), mainly because there is a 3rd Pres. *isine* in Kalajdži, but Bug. *sjan* (older *sijan*) appears to run counter to this reconstruction, because in this dialect [ni, n'] has not been reduced even in grammatical 'endings', for instance in Pl. *phenja* 'sisters'. This argument, however, is not fully convincing since there is reason to assume that exceptional sound changes often start in very frequent and unmarked categories; cf. Arli *sijum*, *hijum*, along with *sinjum*, *hijnum*, as a parallel process. Thus, two possibilities of deriving the actual forms remain:

1. Our group goes together with the majority of the South Balkan I dialects that developed the extended *sinjom* from (*i*)*som*. This derivation is supported by Kal. Pres. *isine* and Bug. Pret. *sine*, but not necessarily by the negated by-form *nanas* (less frequent than *nasine!*), which may be older than *sine*, provided even *nane* and *nanas* must be derived from \**na-hine* and \**na-hinas* respectively.
2. Our group inserted *-j-* into the copula for one of two reasons: either by analogy to the preterite of other verbs (*šun-d-j-om*, *nakh-l-j-om*), or by re-shaping the other persons of the copula on the basis of the 3rd person *si*, that is *sijom* from \**si-* + *om*, and so on. One factor may have reinforced the other. The second possibility must also be considered for explaining the copula of the dialect of Wales (for the paradigm cf. Sampson 1926: 208), and perhaps the Iberian forms as well (see Boretzky 1995: 19). All in all, more speaks in favour of the second possibility, but then we must explain Pret. *sine* as influenced by Arli and Erli respectively.

The same problem arises with *djom* and *ljom* (Drind. *dim* and *lim*), the preterite forms of *del* 'give' and *lel* 'take': since Bugurdži and Kalajdži have 3rd Pl. *dine* and *line*, it might seem most natural to trace *djom* back to *dinjom* but, as can be gathered from Bug. 3rd Sg.Masc. *uhcilo* in contrast to 3rd Sg.Fem. *uhcini*, *l-*forms and *n-*forms do co-exist in one paradigm, that is we cannot know for certain which of them is the original one and which has been introduced via analogy.

An important difference turns out between 2nd Pl. *sjen* of Bugurdži, but *sjan* in the two other dialects.<sup>28</sup> If *sjan* is the older form (cf. *-an* in Paspatian, Northern Vlach, Central dialects, Sinti etc.; Boretzky 1995: 17ff.), Bugurdži must have been influenced by Arli.

In a Drindari variety the copula form *isi-tar* was found, but we do not know whether this is a usual formation or a single form borrowed from a Vlach dialect, for to my knowledge *-tar* is hardly found outside Vlach.

Attention must be drawn to the 3rd Pres. (!) *sine* of Kalajdži, since elsewhere this form is only known as a preterite, for instance in a part of the South Balkan I dialects and the Southern Central dialects. Here *-ne* can hardly have been a preterite marker. Rather it was a morpheme added to neutral *si* in order to create an emphatic variant. Only later *sine* must have been re-interpreted as a preterite form (see Boretzky 1995: 26f.). As already mentioned, Bug. Pret. *sine* may have been taken over from Arli, but another

explanation is also possible: *sine* was in use as an emphatic present, but later it was re-interpreted as a preterite under Arli influence.

The suppletive *ovel* and *ulo* have not been re-shaped much. According to Kenrick (1969: 96, 146), they are in retreat in Drindari giving way to *ačh-* ‘to remain’, but some instances can be found in the texts. There is no reason to believe that *ovel* and *ulo* have secondarily been borrowed into Drindari from other dialects. Thus, our group agrees with South Balkan I and Southern Central, whereas in other groups *ovel* and *ulo* have been replaced mainly by *avel* and *avilo* ‘come’, that is suppletion has been maintained.

In the following the tenses of the verb are discussed.

3.8.2 The Present Active. The forms in *-a* are the unmarked present, but the short form is in use for the present as well. The paradigms have been changed by the loss of *e* (see 2.3.1) but to a differing degree in the individual dialects.

Table 9. The Present Active

	Bugurdži	Drindari	Kalajdži
Sg. 1st	ker-áva	kər-á (-a:)	ker-áva
Sg. 2nd	kér-sa	kər-sa	kér-ca
Sg. 3rd	kér-la	kər-la	kér-la
Pl. 1st	ker-ása	kər-ása	ker-ása
Pl. 2nd	kér-na	kər-na	*kér-na
Pl. 3rd	kér-na	kər-na	*kér-na; (sér-na)

In Bugurdži, elision of *e* is compulsory, also in the copula *ovla*; there are only a few exceptions (for instance *dela* ‘gives’, *lela* ‘takes’, *phutrela* ‘opens’, *učela* ‘gets up’, *uhcela* ‘jumps’) conditioned by phonetic factors. Drindari drops *e* regularly only after *n*, *l* and *r* but keeps it in cases such as *dikh-íla* ‘sees’, *žínk-íla* ‘comes here’, *a-íla* ‘comes’, *uš-íla* ‘gets up’, and in *khe-íla* ‘plays’ where the first *l* is dropped. As for Kalajdži, we are not able to establish clear-cut rules for vowel elision; so far we can only quote stems ending in *n*, *r* and *l* where *e* is elided; cf. *annas* < *anelas* ‘brought’, *čhinna* < *čhinela* ‘cuts’, *phenla* < *phenela* ‘says’, *khella* < *khelela* ‘plays’, and even *kan-čhinlas* < *kan-čhinenas* ‘would cut’ with confusion of *n* and *l*. For the 2nd Sg. the only form attested is *kerca* ‘you make’ from Montana (NW Bulgaria). The elision of *e* has no parallel in South Balkan I but occurs sporadically in other groups too, for instance in Northern Vlach.

1st Sg. *-ava* is shortened throughout in Drindari, but never in the two other dialects. The long form in *-a* is usually used for the present tense only, but in Kalajdži it also appears after *te* in the so-called subjunctive; cf. *te vakerava* ‘that I tell’, *te džava* ‘that I go’ alongside *te džav(v)*; further after *naši* ‘cannot’ in *naši kerava* ‘I cannot do’. This also holds for Drindari (p.c. Kenrick), but never for Bugurdži. More material is required in order to clarify if the two constructions render an aspectual distinction.

In Drindari, *džala* ‘goes’ has been combined with *inkja* ‘hither’ forming a new verb, in all likelihood via the imperative *dža inkja!* ‘come here!’, inflected *žənkja*, *žənkisa*, *žənkila* etc., in Zlatica *žainkja*, *žainkisa*, *žainkila* etc. without contraction, in other varieties *džankjava* etc. A similar case is *lamáa*, *lamáisa*, *lamáila* etc. ‘begin’, probably accreted from *lava ma* ‘id.’, literally ‘I take myself’. In Kalajdži *an-ava* ‘I bring’ has been re-shaped into *antava* etc., apparently on the basis of the imperative *an-ta!* (with an emphatic particle).<sup>29</sup> In Drindari, *dava* ‘I give’ and *lava* ‘I take’ gave *dəa* and *ləa* respectively, although *\*da(a)*, *la(a)* should have resulted after the loss of *-v-*. A causative derivation from *uxtjel* ‘jumps’ (Bug. *uhcel*) is *uxtáa* ‘pick, gather’ < *uxtjavava* in Drindari. *uxla-* ‘draw down’ appears to be a causative from *uxj-el* ‘goes down’.

The verbs for ‘get up’, ‘jump’, and ‘go down’ (see Map 33) have been transformed by sound change considerably, and it may therefore be useful to give an overview for the three dialects; see Table 10.

Table 10. The 3rd Sg.Pres. of ‘get up’, ‘jump’, and ‘go down’

	uštjela	uxtjela (xutjela)	uxljela (xuljela)
Bug.	učela	uhcela	uhjela
Kal.	učela	?	uxela
Drind.	ušila	uxcila	uxila <sup>30</sup>

A verb typical of Kalajdži is *muri-ker-* or *mudi-ker-* ‘to kill’, in contrast to *mudarel* in the two other dialects. The stem *ter-* (elsewhere also *ther-*), originally ‘to hold’ and ‘to stand’ (in the passive), assumed the meaning ‘to remain’ and ‘to stop (at a place)’ in Drindari, but it is absent from Bugurdži. *dex-el* ‘loves’ has only been found in Matth, very likely a deliberate borrowing of this rare verb from Erli.

Loan verbs are adapted by adding *-iz-* or *-az-* to the root in all dialects of

our group, for instance *ir-iz-el* ‘returns’. In varieties of Drindari *-inz-* is used; cf. *kla-inz-* ‘to pray’, *živ-inz-* ‘to live’ (Kenrick 1969: 64 for Sliven, but *živ-iz-* for Kotel). Apparently, this formant has come into being as a contamination of *-iz-* and South Balkan I *-in-*. The translator of Matth added *-in-* even to inherited verbs; cf. *laža-in-ker-* ‘to ashamed’ instead of *ladž-av-*, *vaz-in-ker-* ‘to lift’, *muk-in-zjola* passive of ‘to let, leave’, and passive *vaz-in-zjol*. This ‘superfluous’ element must be interpreted as a sign of weak language competence.

3.8.3 The Preterite Active. The preterite has been changed mainly by palatalisation processes that led to a differentiation of the dialects in some cases; cf. for *ker-* ‘to make’, *šun-* ‘to hear’, and *vaker-* ‘to speak’; see Table 11.

Table 11. The Preterite Active with *d-* formant (see Map 34)

	Bugurdži	Drindari	Kalajdži
Sg. 1st	ker-dz-om /-Ø-om	kər-Ø-im šun-z-im	ker-d-om ker-Ø-om
Sg. 2nd	ker-dz-an /-Ø-an	kər-Ø-ean šun-z-ean	ker-d-an ker-Ø-an
Sg. 3rd	ker-dz-as /-Ø-as	kər-Ø-eas šun-zeas	ker-d-a(s) vaker-Ø-e
Pl. 1st	ker-dz-am /-Ø-am	kər-Ø-eam šun-z-eam	ker-d-am ker-Ø-em
Pl. 2nd	ker-dz-en /-Ø-en	kər-Ø-ean(as) šun-z-ean	ker-d-an
Pl. 3rd	ker-d -e	kər-d-i šun-d-i	ker-d-e

The difference in the 2nd Pl. (*-en* in Bugurdži, *-an* in the other dialects) has already been mentioned and interpreted with the inflection of the copula form *ulo*.

The palatal element *-j-* to be presumed for all our dialects must have palatalised *-d-* in a first step, but the ensuing development resulted in differing paradigms even within one and the same dialect. On the one hand, *-dj-* gave *-dz-* in Bugurdži, on the other hand the formant disappeared altogether, perhaps via *-dj-* > *-j-* > Ø. The same processes must be assumed for Drindari, but distributed over different verbs, that is we will not find *kərdzeas* along with *kər-Ø-eas*, at least not in the same sub-variety. The problem is that we do not have sufficient knowledge about the behaviour

of the sub-dialects, but both *phenzas* ‘said’ and *phejās* are attested. The syllabic *ɟ* in *-ɟan* etc. may be viewed as a reflex of the preterite marker. In another variety of Drindari [ea/ja] becomes [e], for instance *kæres*, *kærem*, that is only an indirect trace of the marker survives. The stems ending in *-v* have zero marker too; cf. *phiřeim* < *phiřavdjom* ‘I opened’, *nařaim* < *nařavdjom* ‘I made run’, *lamaim* < *lamavdjom* ‘I began’ (but *čhungazim* which, however, belongs to *čhungadela* ‘spit’). In Kalajdži, the palatal element has been dropped in the course of a non-automatic sound change thus giving *kerdom*, but there are also varieties with zero marker, much like in Bugurdži and Drindari;<sup>31</sup> cf. *kerom* (Gilliat-Smith 1935: 28). From *vakere* (Matth) ‘spoke’ we can infer that there existed a pre-form \**vaker-j-a*, and *benze* ‘give birth to’ (Matth) displays the ‘normal’ reflex of [dj] > [dzʹ] > [z], that is there were no less than three different developments of original [dj]. It is most difficult to understand how and why *-dj-* has been reduced to *-d-*. Note that Kalajdži is not a single case, some Vlach dialects show forms such as *kerdem*, *kerdan* etc. as well. The differentiation of the dialects’ paradigms need not be old, it may have taken place in the last centuries.

In addition to the *d*-preterites *t*-formations are attested, which must be the original formants with verb roots ending in a sibilant (see Map 35); cf. for Bugurdži *beř-č-oml-anl-ilo* etc. ‘sat’, *doř-t-e* ‘milked’, *nař-č-i* ‘fled’, sometimes after *k* too, as in *muk-t-e* ‘they let’, *arak-t-e* ‘they found’, *pek-t-e* ‘they baked’, for Drindari *beh-č-im*, *beř-t-u*, *beř-č-i* etc., *res-c-im* 3rd Pl. *res-t-i* ‘arrived’, *nař-č-im*, for Kalajdži *erəs-t-o* (in Matth for *araslja*). This formant is in retreat: it is not applied in all instances where it might be expected according to the rule.

It is most likely that *-l-* has been palatalised in all three dialects, first giving [j]. This *j* has been preserved in Drindari and in Bugurdži (here with further palatalisation), while in Kalajdži it disappeared without leaving a trace. However, a form such as *dik-em* in Matth presupposes \**dikh-j-am*.

Table 12. The Preterite Active with *l*-formant (see Map 36)

	Bugurdži		Drindari		Kalajdži
Sg. 1st	res-j-om	dičh-om	dikh-Ø-im		dikh-Ø-om
Sg. 2nd	res-j-an	dičh-an	dikh-j-an		dikh-Ø-an
Sg. 3rd	res-j-as	dičh-as	dikh-j-as		dikh-Ø-as
Pl. 1st	res-j-am	dičh-am	dikh-j-am		dikh-Ø-am (dik-em)
Pl. 2nd	res-j-en	dičh-en	dikh-j-an(əs)		dikh-Ø-an
Pl. 3rd	res-l-e (-t-e)	dikh-l-e	dikh-l-i		dikh-l-e

Details. The verb ‘to find’ appears in Drindari as *alakh-ava*, Pret. *alakh-im*, the two other dialects have regular Romani *arakh-*, with Pret. *aračhum* in Bug. and *arakom* in Kal., but even here the passive stems are irregular; cf. Bug. *arapcola*, Drind. *alapila*, and Kal. *arakjola*, and in Varna Kal. *alapčola* with *l* as in Drindari (see Map 43). It is most likely that the forms containing [p] have come about through contamination of the passive forms with the reflexive *arakh-el pes*. Innovated *ant-* ‘bring’ for *an-* in Kalajdži had an *l*-preterite too; cf. *ant-om*, but 3rd Pl. *ant-l-e*. Original *čhiv-* ‘put, throw’ has been expanded in Bugurdži and Kalajdži by *-t-*; cf. Bug. *čhičom* < \**čhit-l-j-om*, 3rd Pl. *čhitle*; Kal. *čitas pes*, *čitle*. More conservative is Drind. *čhim*, *čas*, 3rd Pl. *čhidi*, to be reconstructed as \**čhiv-d-j-om*, *čhivdjas*, *čhivde*.

Table 13. The Preterite Active of verb roots ending in *-d*  
(see Map 37)

	Bugurdži	Drindari	Kalajdži
Sg. 1st	rod-j-om	roz-(i)-im	
Sg. 2nd	rod-j-an	roz-e-an	
Sg. 3rd	rod-j-as	roz-e-as	cid-ie (< *cidjas)
Pl. 1st	rod-j-am	roz-e-am	
Pl. 2nd	rod-j-en	roz-e-an	
Pl. 3rd	rod-l-e	rod-i (roz-in-i)	kid-in-e

In Bugurdži and presumably in Drindari as well an *l*-formant has to be reconstructed, contrary to conservative Vlach dialects such as Kalderaš showing no formant at all; cf. *rod-em*, *rod-an*, 3rd Pl. *rod-e* ‘seek’. Whereas in Bugurdži [j] from [lʰ] had no longer a palatalising effect, Drindari [z] is a product of further palatalisation. The 3rd Pl. *rod-i* makes us doubt that there once has been an *l* throughout the paradigm, because then we would expect \**rod-l-i* < \**rod-l-e*. Do we have here a mixed paradigm, or must *rozeas* be derived from \**rodjas* instead of \**rodljas*? If the latter were correct *rodi* might be a residue of the oldest inflection fully preserved in Vlach.<sup>32</sup> Drind. *cidinu* ‘to draw’ is a participle rather than an active preterite with additional *-in-*.

We have seen that while South Balkan I dialects roughly preserved the original past forms for most verb-classes (cf. *kerdjom/kerdjom*, *makhljom*), in the dialects of our group the forms underwent modifications to a considerable degree. This does not hold, however, for the verbs ending in *-d*, since here South Balkan I expanded the verb stem by *-inj-* or *-indj-*, whereas in

our group *-in-* invaded the paradigm only sporadically. In summary, the two Balkanic groups diverged in nearly all paradigms, but the original inflection must have been very similar.

The *a*-verbs (see Map 38) except *džala* ‘goes’ and *xala* ‘eats’ show a tendency to add *-an-* in the preterite; according to Kenrick (1969) Drindari has *asan-* ‘to laugh’, *dukhan-*, ‘to hurt’, *lažan-* ‘to ashamed’, *pacan-* ‘to believe’, *xasan-* ‘to cough’ and even *zaban-* ‘to sing’, but these appear to be the abstract stems that do not tell us whether the real inflectional forms display an [n] or not. A real 3rd Sg. is *trašilu-tar* ‘fear’, which is aberrant anyway since it does not even preserve the *a* of Pres. *traš-a-la*. B. Iгла (p.c.) was given the following paradigm in Zlatarica: *asaim*, *asajan*, *asalu/asai*; *asajamas*, *asajan(as)*, *asale (asali)*, which looks much more like an original *l*-preterite. For Kalajdži, no data is available except for the participle/adjective *bi-ladžando* ‘shameless’ from Matth., containing a formant *-and-*.<sup>33</sup> Bugurdži does not present a clear picture; on the one hand there is *-an-ilo* in the 3rd Sg., for instance *prastanilo* ‘ran’, *trašanilo* ‘feared’, *tromanilo* ‘dared’, that is a passive form to an active present, but on the other hand the other persons have *-ajom*, *-ajan* etc. against a single *trom-anj-um*. As shown in Section 2.1.6, *-nj-* does not normally change to *-j-*, and therefore *-aj-* should go back to *-alj-* rather than to *-anj-*. The problem could be solved with the aid of Kalajdži, since here *-aljom* must have given *-alom*, but I was not able to find preterite forms of this type. Thus, there is no solution to this problem with the material available so far, but it cannot be excluded that an irregular change of [n'] to [j] has taken place. The latter would allow for a reconstruction of *djom*, 3rd Pl. *dine* as *djom* < *dijom* < *dinjom* ‘to give’ opening the possibility to identify all these forms with *-anjom* etc. and with *dinjom* of the older layer of South Balkan I.

In what follows a synopsis of the irregular *l*-preterites is given (original *avilo* ‘came’, *gelo* ‘went’ (see Map 39), *pelo* ‘fell’, *xaljas* ‘ate’, *piljas* ‘drank’, *ačhilo* ‘remained’, and *uštילו* ‘got up’); see Table 14.

Additional details. The verb *mer-el*, Pret. *mulo* ‘to die’ is inflected like *perel*, *pelo*. In Kalajdži, forms containing *n* must have been in progress, apparently favoured by the fact that both [l'] and [n'] gave [j] and that the participle could be expanded by *-in-*. Thus *ale* and *gele* were replaced by *aline* and *geline*, which in turn were regularly changed to *aine* and *gine*. This model was then transferred on *xal* and *piel*. Bug. *učinilo* contains even two formants, perhaps an indicator for ongoing morphological change. Drind. *genkim* etc. developed a variety of forms; cf. 3rd Sg. *genku/genklu/gelko*, and

Table 14. Irregular *l*-preterites (1st Sg., 3rd Sg. Masc., 3rd Pl.)

	av-el	dža-l	per-el	xa-l	pi-el	ačh-el	uštj-el
Bug.	ajum	g(e)jum	pejum	hajum	pijum	ačhjum	učjom
	alo	gelo	pelo	hajas	pijas	ačhílo	uč(in)ilo
	ale	gele	pele	hale	pile	ačhíle	uč(in)ile
Drind.	aim	genkim	peim		piim	ačhíim	ušiim
	alu	genku	pelu		pijas	ačhilú	ušinu
	ali		pele		pili	ačhilí	ušini
Kal.	ajom	giom		xajom			
	alo	gias		xajas		ačhjas	učias/uči-lo/
	aine	gele/gine		xane	pine	ačhine	-no

in Zlatarica (p.c. Igla) conservative Masc. *gelo*, but Fem. *gengi*. The segment *-nk-* (*-ng-*) goes back to *inkja* (as in Pres. *žanjjava* from *dža inkja!* ‘come here!’).

The verbs for ‘to give’ and ‘to take’ (see Map 40) are conspicuous in that the initial consonants *d-* and *l-* have not been palatalised even in these extensively palatalising dialects, that is there is no *\*zijom*, *\*ijom* or similar forms (Table 15).

The 3rd Pl. has *n*-expansion in all three dialects, but this does not mean that the other persons underwent the same change. The preservation of *d-* and *l-* would be understandable if the pre-forms had been *\*denjom* and *\*leljom* (*\*lenjom*) with non-palatalising [e].

The two or three verbs adding the formant *-ist-*, *inklisto* ‘got out’, *xulisto* ‘went down’ and perhaps *uklisto* ‘climbed’ too were expanded in Bugurdži, but maintained this allomorph in the two other dialects; cf. Bug. *ikicilo* (younger *icilo*) < *\*ikl-ist-ilo*, and *uhicilo* (younger *uhinilo*) < *\*uhl-ist-ilo*,

Table 15. Preterite forms for *del* ‘give’ and *lel* ‘take’

	Bugurdži		Drindari		Kalajdži	
Sg. 1st	djom	ljom	dim	lim	djom	ljom
Sg. 2nd	djan	ljan	djan	ljan	djan	ljan
Sg. 3rd	djas	ljas	djas	ljas	djas (die)	ljas (lie)
Pl. 1st	djam	ljam	djam	ljam	djam	ljam
Pl. 2nd	djen	ljen	djan	ljan	djan	ljan
Pl. 3rd	dene/dine	lene/line	dini	lini	dine	line

but Drind. *inkistu/inkici* and *uxistu/uxici*, and Kal. *ikistas* (Matth *ink-ist-ilo*) and *uxustas*. From this picture we can draw the conclusion that Vlach *-isto* represents the oldest Romani form. The formant *-ist-* has been extended to *nakh-* ‘to pass’ in Drindari; cf. *nakistu* (Kenrick and Gilliat-Smith), *nakiscim* (Zlatarica).

3rd Sg. gender-inflected *-o/-i* must be viewed as the oldest formants for the intransitive motion verbs and a few other verbs. This state has been maintained in Bugurdži and Drindari, while in Kalajdži a change must have been going on at the beginning of our century replacing *-o/-i* by *-as*; thus we have *alo* ‘came’, *ačhilo* ‘remained’, *učilo* ‘got up’ on the one hand, and *gjas* ‘went’, *učjas* ‘got up’, *ikistas* ‘came out’ on the other hand.<sup>34</sup> The gender-inflected forms have also been preserved in South Balkan I (and moreover in the Vend-group of the Southern Central dialects), that is the two groups agree in this feature too.

The element *-tar*, typical of motion verbs in Vlach, occurs in Drindari only; cf. *thoim-tar* ‘I washed’, *asani-tar* ‘I laughed’, passive *alapiim-tar* ‘I was found’, *umblazilu-tar* ‘hung’, according to Kenrick (1969: 58) limited to preterite forms, but there is also the copula form *isi-tar* (see 3.8.1). Here, again the question arises if these are the residues of an old element once widespread in Romani.

Loan verbs adopted by *-iz* form their preterite on *-d-* in Bugurdži and perhaps in the two other dialects as well (no reliable data available), the *-d-* showing up in the 3rd Pl. only; cf. *ir-iz-om*, *ir-iz-an* etc., but *ir-iz-de* ‘return’. In Zlatarica Drindari the forms are *piš-iz-im*, *piš-iz-jas* etc., but *piš-i-di* ‘write’ without *z* (p.c. Iglă). The formant *-iz-* can be said to be characteristic of our group, although it occurs in Paspatian and in a few Erli-like varieties of Bulgaria as well. Contrary to this, South Balkan I has *-in-* instead.

3.8.4 The Present Passive. The forms to be met in our dialects have been altered, mainly by sound change, but they clearly go back to the inherited paradigm. In a first step they have been simplified from *-jovava* to *-java*, *-jovesa* to *-josa* etc., and this is the state found in Bugurdži and Kalajdži, whereas in Drindari further sound change has taken place (see Map 41); cf. the forms for *chinjola* ‘gets tired’ and *pilcjola* ‘gets married’ in Table 16.

Old, non-contracted forms are not attested, nor do other reduced forms occur, such as 1st Sg. *-ivav* and 1st Pl. *-ivas* (as for instance in some Kalderaš varieties). The palatal element may be dropped as in Bug. *phuterdzola* ‘opens’, Kal. *ikolal/inkola* ‘is enough’ (but *benzjola* ‘is born’), and Drind. 1st

Table 16. The Present Passive

	Bugurdži	Kalajdži	Drindari
Sg. 1st	chin-jáva	pilc-java	pilc-ǵá
Sg. 2nd	chin-jósa	pilc-josa	pilc-isa
Sg. 3rd	chin-jóla	pilc-jola	pilc-ila
Pl. 1st	chin-jása	pilc-jasa	pilc-ǵasa
Pl. 2nd	chin-jóna	pilc-jona	pilc-ina
Pl. 3rd	chin-jóna	pilc-jona	pilc-ina

Sg. *pašá*. ‘lie down’. The extreme reduction in Drindari in the 1st Sg. draws on the loss of intervocalic *v*. The emergence of *-isa*, *-ila* and *-ina* instead of *-josa*, *-jola* and *-jona* is remarkable, first because of the irregular sound change, and second because by this change the passive paradigm merged with the active paradigm except in the 1st persons; cf. active *dikh-á*, *dikh-isa*, *dikh-ila* ‘see’ with passive *phiraz-ǵá* < \**phiravdjava*, *phiraz-isa* < \**phiravdjosa* etc. ‘open, be opened’. A number of verbs are reduced to such a degree as to become historically transparent only at a second glance, for instance *najá* < \**najovava* ‘I bathe’, *šileá* < \**šilaljovava* ‘I feel cold’, *axája* < \**axaljovava* ‘I understand’, *alápǵa* < \**alaljkovava* ‘I am found’ (see Map 43). There are also passives to *xala* ‘to eat’ and *piela* ‘to drink’, *xaizila* ‘it is eatable’ and *piizila* ‘it is drinkable’, that cannot be traced back to regular pre-forms but must have been modelled in analogy to other verbs displaying a sequence *-zila*, or else they are formed with the aid of *-iz-* (normally only with borrowed verbs). As far as I know *xala* ‘to eat’ and *piela* ‘to drink’ do not have passive forms in Bugurdži, and neither Bugurdži nor Drindari have developed a passive for *dela* and *lela*. Therefore, the passive *dinzjol* ‘is given’ in Matth could be an artificial creation. The stem *čhiv-* ‘to put, throw’ has no passive in Bugurdži, but in Kalajdži (Matth) two different formations are used; cf. *čivzjola* < \**čhiv-dj-ola* and *čičjola* < \**čhi-tlj-ola* (?). It is very likely that the translator was not familiar with the passive of this verb. In using the uncontracted forms *-jovesa*, *-jovela* etc. along with the newer forms many of the South Balkan I dialects represent a more conservative state.

Formations of the type *-ávola*, for instance *divesavola* ‘day breaks’ in other dialects, are not in use; instead we have shorter forms such as Kal. *kan-disjol* (note that *d-* has been preserved) and Drind. *zivesila*. As in other dialects, most passives are formed from a preterite stem, but passives formed on the basis of the present stem occur as well (e.g. all deadjectival

formations). Passives from loan verbs in *-iz-* are not attested, and it is possible that they never existed. In conclusion we may state that there is hardly a property that is characteristic only of our group of dialects.

3.8.5 The Preterite Passive (see Map 42). The basis for all forms is *-il-* giving *-ij-* with subsequent *-j-*.

Table 17. The Preterite Passive

	Bugurdži	Drindari	Kalajdži
Sg. 1st	bar-(i)jóm phutér-dzjom	bar-íim bén-ziim	bar-il-óm ben-zil-óm
Sg. 2nd	bar-(i)ján phutér-dzjan	bar-iján bén-zijan	bar-il-án ben-zil-án
Sg. 3rd	bar-íl/-i phutér-dzilo	bar-ilú/-íí bén-zilu/-ii	bar-iló/-í ben-zil-ó/-í
Pl. 1st	bar-(i)jám phutér-dzjam	bar-ijám bén-zijam	bar-il-ám ben-zil-ám
Pl. 2nd	bar-(i)jén phutér-dzjen	bar-iján bén-zijan	bar-il-án ben-zil-án
Pl. 3rd	bar-íl-e phutér-dzile	bar-ilí bén-zili	bar-il-é ben-zil-é

In Bugurdži, passives formed from *d-*stems are rare, apparently because they are distinguished from the respective active forms only by means of stress (shaky in this dialect) and because the palatal element [j] is often not audible; cf. *phuterdzóm* 'I opened' alongside *phutérdzjom* 'I was opened'. Of course, the distinction is easier in the 3rd Sg. *-ilo* and the 3rd Pl. *-ile*, and perhaps it is not by chance that these persons are more frequent. Slightly more differentiated are active *pečóm* 'I baked' and passive *pécjom*; cf. *pécjom bukhatar* 'I nearly starved of hunger'. Kalajdži has undergone depalatalisation, hence the preservation of *-il-* in all persons. According to Gilliat-Smith (1915/1916, II: 95f.) the 3rd Sg. should be on *-il-as*, but in Matth the older forms are also found. As with the preterite active, the forms that should contain a *d-*formant occur in two varieties in Bugurdži: *ker-dz-ilo* and reduced *ker-ilo* (as in active *kerdzom* and *kerom*). Roots ending in *-d* display only *-ilo*, for instance *phand-ilo* 'closed, bound' and *cid-ilo* 'drew'.

Contrary to the *d*-preterites, the formant *-l-* has not been extended to the passive preterite; cf. Bug. *pecilo* < \**pek-ilo* (not \*\**pek-l-ilo*) ‘baked’, *phadzilo* ‘broke’, *mucilo* ‘let, left’. Passive \*\**pekliljom* should have resulted in \**peč(i)jom*, which is very close to active *pečom*.

Drindari has [z] from [dj], for instance *astarzilu* ‘caught’. Kenrick (1969: 41) points out that active and passive forms are distinguishable mainly by stress, that is we are faced with the same relations as in Bugurdži; cf. active *benzim* < \**biandjom* ‘I gave birth to’, but passive *bénzim* < \**biándiljom* ‘I was born’ (reconstruction by the author) - rather idealised than normally used forms. As with the present passive, some of the forms, especially denominal formations, are extremely reduced; cf. *deřim* < \**deniljom* (*denilo*) ‘I went crazy’, *bilájim* < \**bilaniljom* ‘I melted’. We have no data about the passive of active *l*-preterites. It is likely that Kalajdži had the same parallel formations as Bugurdži; cf. *benzilo* with a reflex of *d*, but *pherilo* ‘was filled’ (Matth) < \**pherdilo* without *z*. Here too sufficient data are lacking.

Loan verbs adapted through *-iz-* (see Map 44) in the present active have preterite forms ending in *-is-ajl-*, a complex formant present in most dialect groups, partially even in those displaying *-in-* in the present active. From this we can conclude that *-is-* was the most important and perhaps also the oldest formant borrowed from Greek. The distribution of *-iz-* and *-is-* is not the same as in Greek, since in our group the preterite active has *-iz-* instead of *-is-*; cf. Greek pres. γυρίζω ‘turn’, but Pret. γύρισα.

In Drindari, [aj] has regularly been changed to [ej]. For Kalajdži *-ij-lo* is given, but it is not clear if this formant occurred in all varieties. According to Gilliat-Smith (1914: 276) the preterite passive is formed with the aid of *-sar-*, but it is not specified what the concrete forms are like. Kenrick (p.c.) did not come across such forms.

Table 18. The Preterite Passive of loan verbs

	Bugurdzi	Drindari	Kalajdži
Sg. 1st	ir-is-aj-um	ir-is-ej-im	*xoj-as-il-um (?)
Sg. 2nd	ir-is-aj-an	ir-is-ej-an	*xoj-as-il-an (?)
Sg. 3rd	ir-is-aj-lo/-aj	ir-is-ej-lu/-eji	xoj-as-ij-lo
Pl. 2nd	ir-is-aj-am	ir-is-ej-am	
Pl. 2nd	ir-is-aj-en	ir-is-ej-an	
Pl. 3rd	ir-is-aj-le	ir-is-ej-li	

3.8.6 Transitive and causative formations. The two procedures for forming transitives in Romani, that is first with *-av-* (unstressed) and second with *-(j)ar-*, are also known from the dialects of our group; additionally, a formant *-ker-* derived from ‘to make’ is in use in Drindari and in Kalajdži. The latter formant occurs as well in South Balkan I, in Central dialects, and in Baltic dialects.

Bugurdži has forms with *-av-* with a number of old, widespread formations; cf. *phir-av-* ‘to lead’, *traš-av-* ‘to frighten’, *ačh-av-* ‘to stop’. Derivations employing *-jar-* are not very frequent. Apparently, it must be reconstructed as *-ljar-*, the element *l* being imported from the preterite stem; cf. *beš-jar-* ‘to put down’ < \**bešljar-*, *rov-jar-* ‘to make cry’. Original *-ar-* or *-jar-* is regularly used in verbs derived from adjectives; cf. *čuč-ar-* ‘to empty’, *lačh-ar-* ‘to settle’, and *kov-jar-* ‘to soften’ < *kovljar-* (*kovlo*). Sporadically inherited *-al* is found in *ik-al-* ‘to take out’, *baš-al-* ‘to play (instrument)’ and *naš-al-* ‘to drive away’.

In Kalajdži, we have *phir-a(v)-ela* ‘opens’, *ušt-av-el* ‘steps’, *traš-av-zilo* (passive) ‘was frightened’, *bolavela* ‘baptizes’, and *dišavela* ‘shows’ from \**dišola* ‘seems, appears’; *šuž-a-əl* ‘dries’ from *šužo* ‘dry’ is irregular since what we would expect is \**šuž-ar-*. Formations with *-jar-* (partially from *-ljar-*) are *pas-jar-* ‘to turn over’ (sic!) < \**pašljar-*, *beš-ar-* ‘to place’, *pen-z-er-* ‘to make speak’ < \**phen-d-jar-*, *ikanzar-* ‘to take out’ instead of \**ink-al(-d)-*. Innovations formed with the aid of *-ker-* are applied to a stem expanded by *-i*, for instance *beši-ker-* ‘to put down’, *lonzi-ker-* ‘to salt’, and *muri-ker-* ‘to kill’ instead of older *murdar-*. From loan verbs some forms in *-is-ker-* are attested; cf. *poklon-is-ker-* ‘to bow, to bend’ from the base *poklon-iz-*.

Drindari does not differ much from what is found in Kalajdži. Formations in *-av-* are *phiraá* ‘I carry’ < *phir-av-ava*, *našaá* ‘I lose’ (innovation for *naš-al-*), *traš-aá* ‘I frighten’; in *-ar-* *diš-ar-* ‘I show’ alongside *dišaá*, *ker-z-jar-* ‘to cause to make’, *čor-z-jar-* ‘to cause to be stolen’. Formations with *-ker-* show stem extension in *-i*, precisely as in Kalajdži; cf. *asai-ker-* ‘to make laugh’, *mothazi-ker-* ‘to make tell’, *rozi-ker-* ‘to make seek’. This *-i* might be an irregular reduction of *-ja* in Drindari, because normally [ja] gives [je] or [e], but not in Kalajdži. Therefore, it is more likely to be of unknown origin. Another form that is difficult to explain is *čum-is-ker-* ‘to kiss’.

In South Balkan I *-ker-* occurs as well, but the full formant is *-av-ker-* (Sepeči) or reduced *-a-ker-* (most other dialects), a reinforcement of old *-av-*. It is not clear how *-i-ker-* and *-av-ker-* are historically related. This late *-ker-* can be viewed as a parallel development to inherited *-ar-l-jar-*, which

in turn presumably goes back to Indic *\*kar-*. Since *-jar-* must be derived from *\*-i-kar-*, both average Romani *-jar-* and *-i-ker-* of Kalajdži and Drindari turn out to be made up of the same elements, but *-i-ker-* is a later formation.

3.8.7 Future (see Map 45). There is no synthetic future form in the South Balkan dialects in general, since the extended form in *-av-a*, *-es-a* etc., a future in the Central dialects, has present value. The future marker goes back to *kam-* ‘wish, love’, a loan translation from the Balkan languages: *kam-/kan-* in Kalajdži and *ka-* in Bugurdži. Drindari has a particle *mə-* (with the variants *ma-*, *mo-*, *mu-*), probably a derivation from *mang-* ‘ask, demand, beg’, and ‘want’, but it is possible that the pre-form of *mə-* is the same element as in all other Balkan dialects, come about by reduction of the variant *kama-* attested for southern Gurbet.

The future negator is inherited *na-*, but there are competing forms apparently triggered by the Bulgarian/Macedonian model: Kal. *nanaj te* ‘is not that’ approximately corresponding to Bulg. *няма да* ‘has not that’ (see Boretzky 1999a: 172), and Drind. *njama-ma-*, a direct borrowing from Bulgarian combined with indigenous *ma-*. In Bugurdži, positive *si te* is occasionally utilised as a future marker, but I was not able to find instances of negative *nane te* with this function. As mentioned above, even the long forms in *-ava*, *-esa* etc. can be combined with future markers, but we cannot say if by this an aspectual distinction is expressed.

3.8.8 Non-finite verbal categories. Since the past participle provides the basis for the formation of the preterite stem, the two stems are identical in most cases, the formants being *-d-*, *-t-*, and *-l-*. In a few cases, however, the participle preserves an older form whereas in the preterite another formant was added; cf. regular Bug. *naklo* and *načhum* from *nakh-l-* ‘to pass’, Drind. *beš-t-u* and *beš-č-im* from *beš-t-* ‘to sit’, but Bug. *beš-t-u* against 3rd Sg. *beš-(č)-il-o*.

Loan verbs form their participles with the aid of the non-inflecting *-ime*, *-ame*, *-ome*, Drind. *-imi* etc., borrowed from Greek *-ημένος*; cf. Bug. *sprem-ime* from *sprem-iz-* ‘to prepare’, *prand-ome* from *prand-oz-* ‘to engage’, *hoj-ame* from *hoj-az-* ‘to get angry’; Kal. *piš-ime* from *piš-iz-* ‘to write’; Drind. *piš-imi*. This formant is also used in South Balkan I and in Vlach, but in the latter group the distinction between inherited and borrowed verbs has not been upheld.

As is the case with most modern Romani dialects the gerund (see Map 46) is only sporadically used, but it does exist. Its formant is *-indos* (Matth *-indo* because of the general loss of final *-s*), which is uniform in the three dialects: Bug. *mar-indos* ‘killing’, Kal. *naš-indos* ‘running away’, Matth *vakir-indo* ‘speaking’, Drind. *ro(v)-indus* ‘crying’. Occasionally one finds *-indoj* in Bugurdži, which must be attributed to Vlach influence. In Kosovo Bugurdži this formation is being replaced by Albanian (dialectal) *tuj* + present. There is a sporadic reduplicative formation of the type *phir-i phir-i* ‘walking’ in Kalajdži, attested only for a couple of verbs, according to Gilliat-Smith (1915/1916, II: 95) reminiscent of the Turkish formation *gül-gül-e*, but it is unlikely that it has been taken over directly from that language. Both *-indos* and the reduplicated construction are attested for at least some of the South Balkan I dialects.<sup>35</sup>

3.8.9 Modal verbs and related auxiliaries. There is an inventory of a few inherited elements, along with a certain number of borrowings. Bugurdži is the most conservative dialect in preserving all the old auxiliaries: *kam-el* (in retreat in most other dialects of the southern Balkans) and *mang-* ‘to wish, want’, *šaj* ‘can’ and *nači* ‘cannot’, and *si te* ‘must, have to’, a calque on Slavic *ima da* (along with the direct loans *valjazla* and *mora* from Slavic).

In Kalajdži, *mang-* is used for ‘wish’, cf. *mangel t’ oel* ‘wants to become’, but I was not able to find instances of *kam-el*. Positive *šaj* is lacking as in the majority of South Balkan I, but negative *naši* is regularly used; cf. *naši kerava* ‘I cannot do’, and combined with a future *naši kan-kerav* ‘I shall not be able to do’. Matth has *našti* and, derived from it, *ašti*, but both forms appear to be loans from Erli, since [šti] cannot be indigenous. We have no information about the existence of ‘can’ in normal Kalajdži. For *si te* ‘must’ cf. *isine man te džav* ‘I must go’.

For Drindari, Kenrick (1969) gives *mang-* and *kam-*, and it is possible that both are used as auxiliaries; cf. *mangaa tə motha tuki* ‘I want to tell you’ (Kenrick), and *jek na kamila ti inkalil la* ‘nobody wants to remove it (sc. the ring)’ (see Marušiakova and Popov 1997: 27ff.). To judge from the use of the independent verbs there is no actual semantic difference between them; cf. *me manga čha* ‘I want a girl’ with *na k(h)amila čha* ‘he does not want a girl’. As is generally the case, negative *naši* has been preserved, while for positive ‘can’ a new form *aši* has developed in a process of reinterpretation and decomposition. It is noteworthy that neighbouring dialects belonging to different groups went through the same process. Therefore, it is most likely that

the process started in one of the dialects extending over time to others. The phonological shape of Drindari *aši* makes it impossible to take it as a direct loan from Erli *ašti*. For ‘must’ Bulg. *trjab-iz-la* has been borrowed, while no instances of *si te* could be found (see Maps 47–49).

As can be seen from the list of elements given above our dialects do not go together in the important details, Bugurdži being the most diverging one (no traces of *ašti*, but preservation of *šaj*).

#### 4. Evaluation of the relations between the dialects

The three dialects treated here display a strong cohesion, mainly due to the results of shared sound changes, but also due to some important morphological traits. Most conspicuous are the palatalisation processes producing sound types that differ from those observed in other dialect groups (cf. 2.1). The morphological traits characteristic of, though not totally exclusive for, our group are (a) plural forms in *-oja* (*-uja*) to borrowed nouns in *-os* (*-us*) (cf. 3.4); (b) a typical set of demonstrative pronouns (*kada*, *kaka*, *kava*, cf. 3.6.3); (c) the characteristic vowel [i] in ‘such’, ‘so much’ and ‘so’ instead of *a* and *o* elsewhere (*kidiso*, *kikesu*, *kidja*, *kidibor* etc.; cf. 3.6.3); (d) the absence of the local adverbs *kate*, *kote*, otherwise widespread in Romani (cf. 3.6.4); (e) a present copula in the form *sjom* (*sijom*), *sjan* (*sijan*) etc. (cf. 3.8.1); (f) systematic elision of *e* in the present paradigm (cf. 3.8.2), occurring rather sporadically in other dialect groups; (g) past forms lacking a tense marker, known from varieties of Sinti as well (cf. 3.8.3); and (h) loan verbs adapted by *-iz-* (cf. 3.8.2; in Paspatian too)—except for (b) all of them shared innovations. Through these features South Balkan II differs from most of the South Balkan I dialects. Borrowed nouns in *-is* form their plurals in *-ides*, which is widespread in South Balkan I as well, but does not occur in its westernmost dialects neighbouring with Bugurdži (cf. 3.4). The verb stems ending in *d* of the type *rod-* form their past with the *l*-marker (cf. 3.8.3), a feature encountered in Kosovo Gurbet as well, but not elsewhere in the Southeast. In view of the fact that it is very difficult to find features that are exclusively characteristic of one Romani dialect group, the list of features given above appears more than sufficient to separate our three dialects from the bulk of a common South Balkan group and to establish a group of its own.

Little evidence can be adduced from the lexicon. First, we do not know if there was a lexicon shared by all Romani dialects, or if the dialects (or groups) differed slightly even before their immigration to the Balkans.

Second, the lexicon of our three dialects has not been documented completely, and data are lacking especially for Kalajdži. Third, we cannot be sure whether actual lexical differences between the dialects are due to rather late losses. Characteristic of our group is *trašala* ‘fears’ (instead of *darala*), quite rare in Romani, the Greek loan *rahciri/rahsiri* ‘smithy, smith’s tools’ in Bug. and Kal., as well as the peculiar passive formations for *dikhela* ‘sees’ and *arakhela* ‘finds’, *dišola* ‘is visible, it seems’ and *arapcola/arakcola/alapila* ‘is found, is there’, but hardly anything else. It may well be that our dialects preserved more Greek words than any other dialects. Some might indeed be found only here, though the lack of data for Kalajdži makes it impossible to prove this assumption.

Most of the grammatical features can be explained as the result of changes that took place after the common ancestor dialect split up into individual dialects, but there remain some serious problems, especially concerning the formation of the genitive (long *-koro* in Bugurdži, but short *-ko* in the other dialects) and the formant for the 2nd Pl. Pret. (*-en* in Bugurdži, but *-an* in the other dialects). However, the solutions offered for the two problems (cf. 3.2 and 3.8.3) seem acceptable.

## Notes

1. The dialect areas marked on the maps must be taken as an approximate indication of the settlement patterns of the respective groups. The region around Tatar Pazardžik is marked as Kalajdži for the only reason that the translator of the gospel of Matthew lived in this town. We have no other sources for this subdialect.
2. A comparison of the South Balkan I dialects known to us (Paspatian I and II, Erli, Ursaritiko, Crimean, Arli, Prilep, Sepeči as well as the dialects spoken in Iran) has shown that this tentatively established group is quite heterogeneous and does not allow for subgroupings. Therefore, the dialects examined here cannot be expected to be more closely related to the whole South Balkan I group but at most to a few of its members.
3. There are word forms like *našti* ‘cannot’, *geljom/gelom* ‘I went’, *ukljava* ‘I climb’, *uštjava* ‘I get up’, *pani* ‘water’ (op.cit. part II, 77 f.) without palatalisation, and *nanaj* ‘is not’ is different from Kosovo Bugurdži *nane*. Gilliat-Smith groups it with Paspatian, but this is not very helpful in view of the fact that Paspati (1870) encompasses a variety of dialects without specifying where the material stems from.

4. The Demirdžides of Kazanlık reported in Gilliat-Smith (1915/1916, II: 8off.) do not belong in our group. Their speech did not undergo the palatalisations typical of our group, but has verbal abstracts in *-iben* and participles from loan verbs in *-imen*, which is exceptional for the Balkan area.
5. In many cases this abstract form is identical with the historical predecessor of the word in question, for instance /kirmu/ 'worm' instead of [cirmu], in distinction to /kIkId-/ 'squeeze' for [kikid-] with a fictitious non-palatalising phoneme /I/, but sometimes the 'deep' forms have been established on morphological grounds, as for instance in /tutřu/ for [tuřu] 'your' Sg. in analogy to oblique *tut*, or /konIskI/ for [kaski] 'whose' requiring a detailed knowledge of morphology for correct reading. Therefore, I am not sure if I have 'reconstructed' the right phonetic form in all cases. An edition using a phonetic-based spelling is being prepared. The author provided me with a lot of additional information which enabled me to clarify most of the open questions. I wish to thank him for his readiness to help me.
6. Gilliat-Smith (1935: 26) characterizes this dialect as 'grammatically so well preserved that it must rank with those recorded by Paspati.' However, a careful study of the texts reveals much uncertainty, doublets, non-forms as well as the use of many loanwords where other dialects did preserve 'good' Romani words, which makes us doubt the translator's competence. Gilliat-Smith himself lists the following irregularities: a strange pluperfect instead of the normal preterite, a superfluous reflexive *pe* with already passive verbal forms, a periphrastic perfect imitating the Bulgarian perfect, oscillation in gender, case, and in the use of prepositions. We want to add here: lacking mastery of the genitive, and word formations of a doubtful quality. We cannot say whether this was the state of the dialect in 1935, but Gilliat-Smith's assessment is too positive. Many elements appear to be adopted from Erli, but it is not clear if they have been borrowed into this variety or if it was the translator who adopted them with this special opportunity.
7. Sporadic contact with Greek may have been maintained until the beginning of twentieth century, when large numbers of Greek and Slavic speakers were exchanged between Greece and Bulgaria.
8. In Boretzky (1993) I used the spelling 'č' and 'dž' for the products of late palatalisation in order to distinguish them from the inherited Old Indo-Aryan affricates, although in the varieties known to me there is no phonetic difference between the two series.

9. Gilliat-Smith (1914) convincingly suggests to derive this form from *drś-* rather than from \**drkh-* (giving *dikh-el*).
10. In the Northern Central dialects some instances of *ki-* > *ci-* can be found, but the parallel *gi-* > *dzi-* is not attested. The dental has been affected in *cikno* 'little' < *tikno*, *keci* 'how much' < *keti*, and *ajci* 'so much' < \**ajti* (see Boretzky 1999b: 256, 264, 266, and Elšík et al. 1999: 288 ff.).
11. It is not clear whether 'ĭ' denotes Bulgarian [ə] or Turkish [ɪ].
12. In recent times northern Albanian is undergoing a change by which the two phonemes collapse in a quality of *r* that resembles the *r* sound spoken in most varieties of American English, but this process reached only some places in western Kosovo.
13. Accentuation of the root syllable has come about, in all likelihood, by the contraction of two originally independent words and the subsequent loss of the stressed syllable: \**kerđi ovel* 'is made' > \**kerđiovel* > *kérđjovel*. This process is paralleled by what has happened in the contracted present forms of our dialects: *kérla* < *keréla*.
14. I do not use the terms 'thematic' and 'athematic' in the sense proposed by Ian Hancock, but in the sense they traditionally have in comparative (Indo-European) linguistics: roots are expanded by stem-forming elements, among them thematic vowels.
15. According to Kenrick (1969) *bian-* exists beside *ben-*; cf. also *bianimaski dukh* 'labour pains'.
16. In Bosnian Gurbet constructions like *e dasnja gole mulehćira* 'the wife of that dead man' (see Uhlik 1941) are quite frequent, and in Kosovo Gurbet forms like *lesiri* and *leseren* 'his' can be found (Krasnići 1981). This is a striking parallel to the double forms found for the possessives in a number of dialects (cf. *mrno/mo* 'my' and *ćiro/ćo* 'your' in Gurbet).
17. The old distribution, that is *-ipe/-ibe* for inherited elements and *-imos* for loanverbs, has been blurred even in Vlach.
18. The old distribution rule between *-ibe* and *-ipe* holds more or less for Arli, Erli and Prilep, while Crimean and Sepeči generalised *-ipe*, probably under the influence of Vlach.
19. D. Kenrick informed me that he did not come across formations in *-as*, Pl. *-ades*.
20. The forms containing an *l* have been shaped, in all likelihood, according to the model of Bulgarian where earlier *on, ona, ono, oni* were replaced by *toj, tja, tova, te*, original demonstrative pronouns (suggestion

by B. Iglá, p.c.). The immediate model forms in Romani were the demonstrative plural forms *kola*, *kodola* and similar. This process affected some of the South Balkan I dialects and, perhaps indirectly via Erli, Kalajdži too.

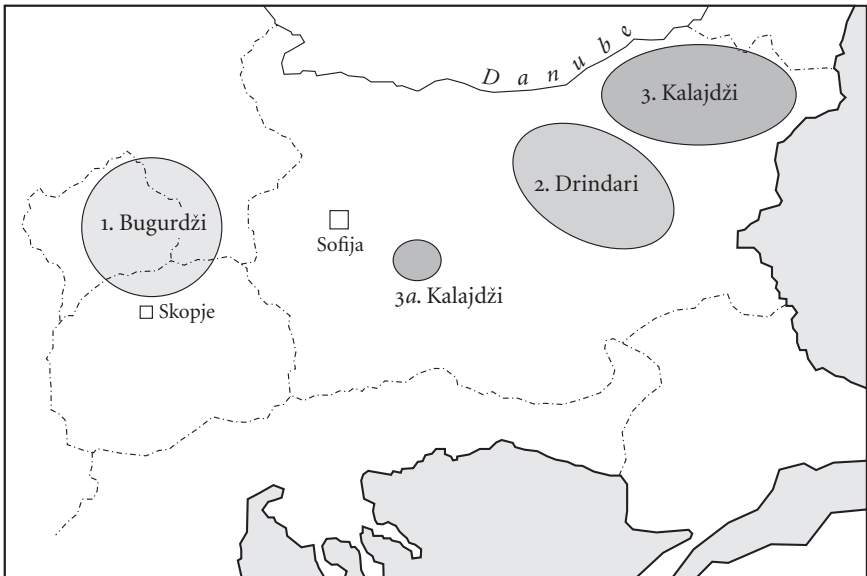
21. In the dialect described by Kenrick *kava/kova* does not exist.
22. The *adava*-series is present in nearly all South Balkan I dialects except Erli. Therefore it is less probable that the series was borrowed into Drindari and Kalajdži from other dialects.
23. In Erli, *kibor* and *kiti* are found alongside, and in Arli *kobor* and *kozom*.
24. Kenrick (p.c.) is not sure about the meaning of this obsolete form.
25. *Vareso* is a form also found in Vlach, with the element *vare-* supposed to be Romanian *oare-*; *hajči-pota* looks much like Greek τίποτα ‘something’, but *či-pa* ought to have another source.
26. Since the Southern Central dialects share a number of traits with the South Balkan I dialects, one is led to assume a common, more or less simultaneous innovation. On the other hand, Northern Central dialects such as East Slovak appear to have expanded the use of *-u-* in an internal process: *-u-* is the connector for 11–19 and for 21–29, but from 31 on conservative *-the-* has been preserved.
27. Not all forms could be found in the sources available, some of them have been constructed on the basis of the attested forms.
28. In distinction to Bug. *sjen*, Kal. *sjen* owes its form to umlaut from *sjan*; cf. also Kal. *-an* in the 2nd Pl.Pret.
29. These processes have parallels in other Romani dialects; cf. *kampel* < *kamel pes* in the Northern Central dialects, and *pekal* < *pe kamel* in Southern Central dialects, both meaning ‘must’; further *pelinen* ‘they quarrel’ < *pes line/line pes*, literally ‘they took themselves’. Even Sinti *humte/homte* ‘must’ must have come about by the generalisation of one inflectional form, in this case the 1st Sg.
30. The second meaning is ‘arrive’. Contrary to this semantic change, the Vlach *teljarel*, derived from *tele* ‘down’ has acquired the opposite meaning ‘depart, set out’.
31. Preterite forms without any formant are also known from Sinti, but it is by no means clear if they came about in the same way. Examples are found in Knobloch (1950, 1977) and in Holzinger (1993), but no clear-cut rules can be formulated. Forms containing *-d-* are *šundies* ‘heard’, *čingerdien* ‘quarrelled’, *kedien* ‘collected’, *rakedjum* ‘talked’, *taperdjum* ‘lost’ (Knobloch 1950), *xačdoməs* ‘burned’, *kerdom* < *keravdom* ‘cooked’

(Knobloch 1977), *d*-less forms are *phenom*, *phenies* along with *phendiās* ‘spoke’, *anies* ‘brought’ (Knobloch 1950), *čingæres*, 3rd Pl. *phenan* instead of *phende* (Knobloch 1977). Forms lacking a preterite marker are to be found with original *l*-preterites as well, for instance *dikom* ‘I saw’, *mukom* ‘I left’, *našom* ‘I fled’, *kamom* ‘I loved’ and *kimies*, *bešien* ‘sat’ and analogically re-shaped *sovom* ‘I slept’ for *sutom*. It is very likely that dissyllabic roots, among them loans, do not appear without *-d-*. If the data given in Finck (1903) is correct, older Sinti was more conservative: *-d-* is preserved while *-lj-* has been reduced to *-j-*; cf. *dikhjom*, *vejom* ‘I came’ < *aviljom*, and even *phučejom* ‘I asked’. Dialect-internal transition from *dikhjom* to *dikhom* is plausible, but how does one get from *kerdom* to *kerom*? It seems more plausible to reconstruct the following sequence: *kerdjom* > *kerjom* > *kerom*.

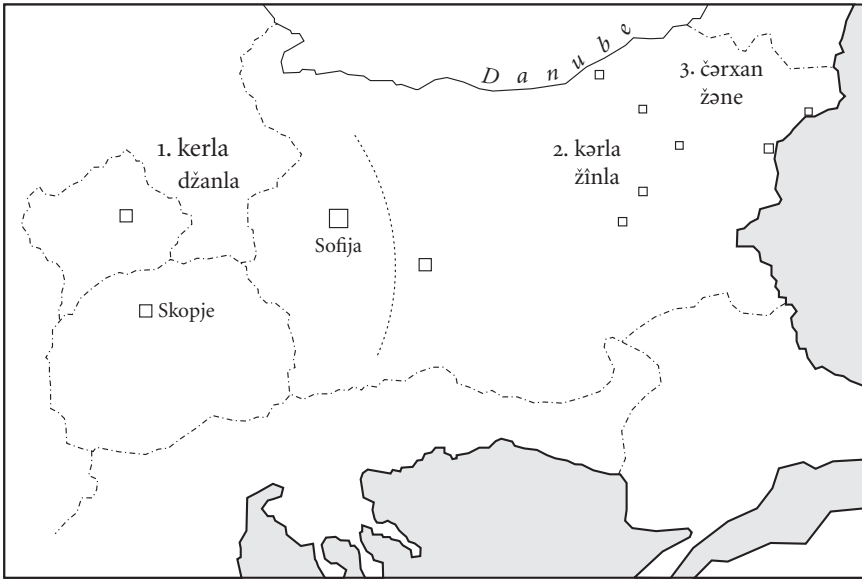
32. Kalderaš varies between *rod-a* and *rod-j-a* in various sub-dialects, which raises the question whether the *j*-preterite or the *j*-less one is older. We have also to keep in mind that, very often, the inflection of the *d*-verbs has been adjusted to the inflection of *d-el* ‘gives’.
33. This formant is attested from some of the South Balkan I dialects (see Boretzky 1999c: 87, Map 46).
34. Gilliat-Smith (1915/1916: 95f.) claims that *-ol-i* disappeared altogether, but this is an over-generalisation.
35. This formant also occurs in Paspatian and Sepeči (both South Balkan I), and in the North-Russian dialect, which makes a Turkish origin unlikely (cf. Boretzky 1999c: 116).

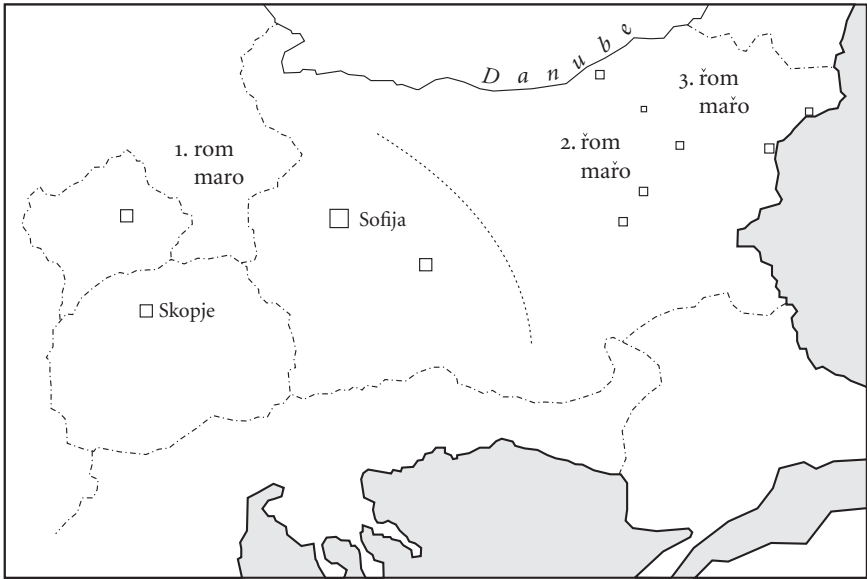


Map 1. Political overview

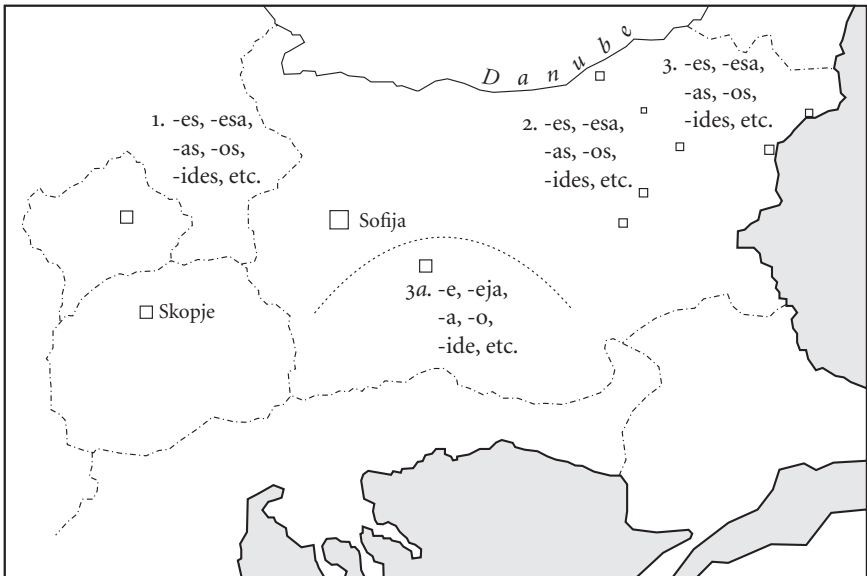


Map 2. Distribution of the dialects

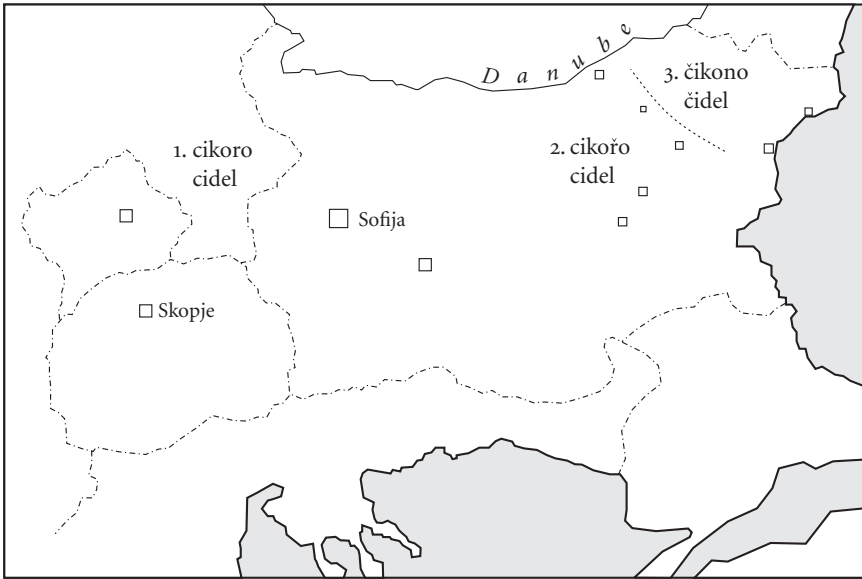
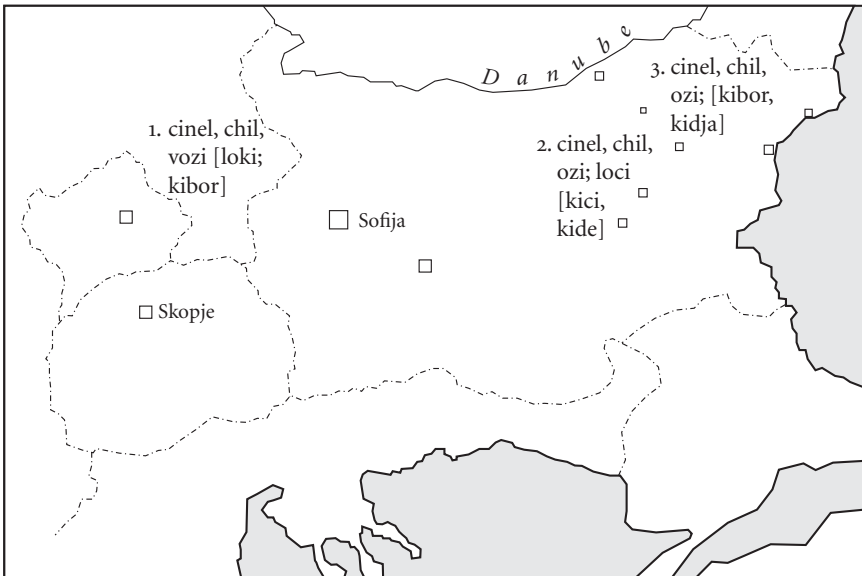
Map 3. Centralisation of *e* and *a*Map 4. *e* > *i* in inflectional morphemes

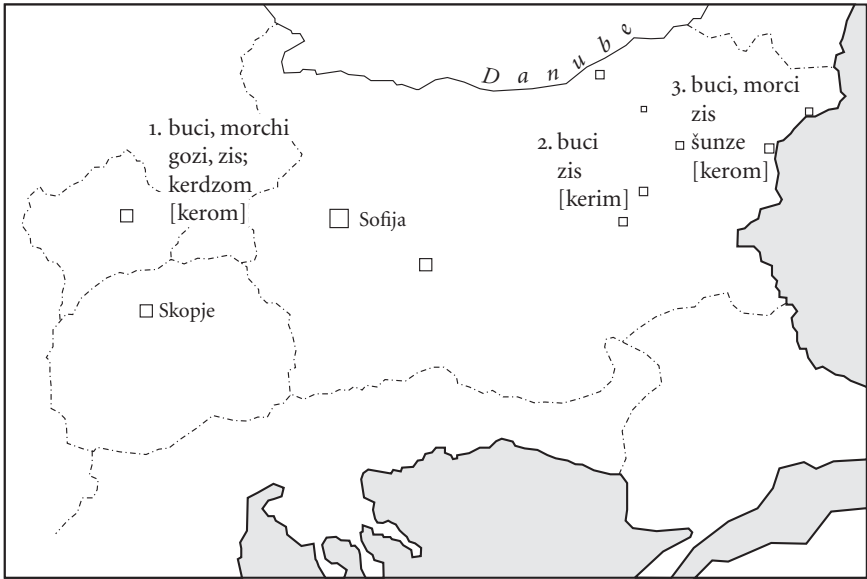


Map 5. Preservation of ř

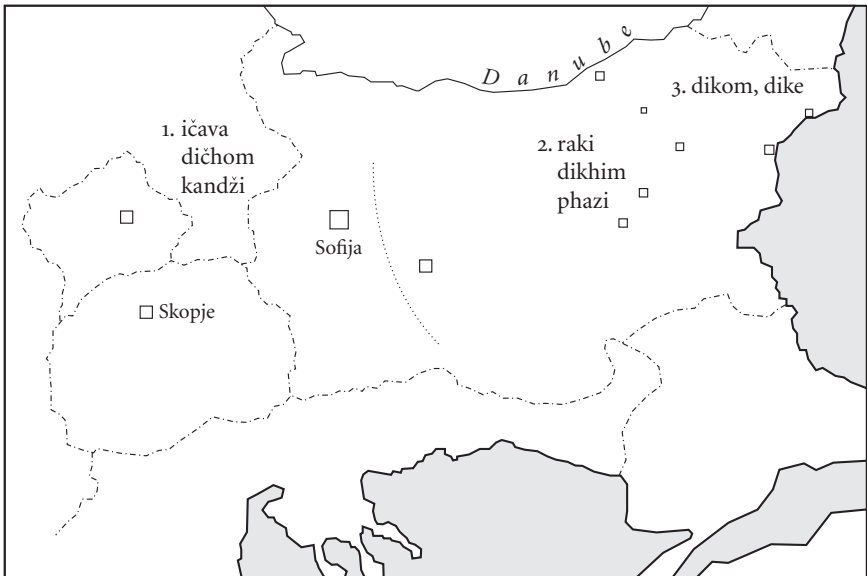


Map 6. Preservation of grammatical -s

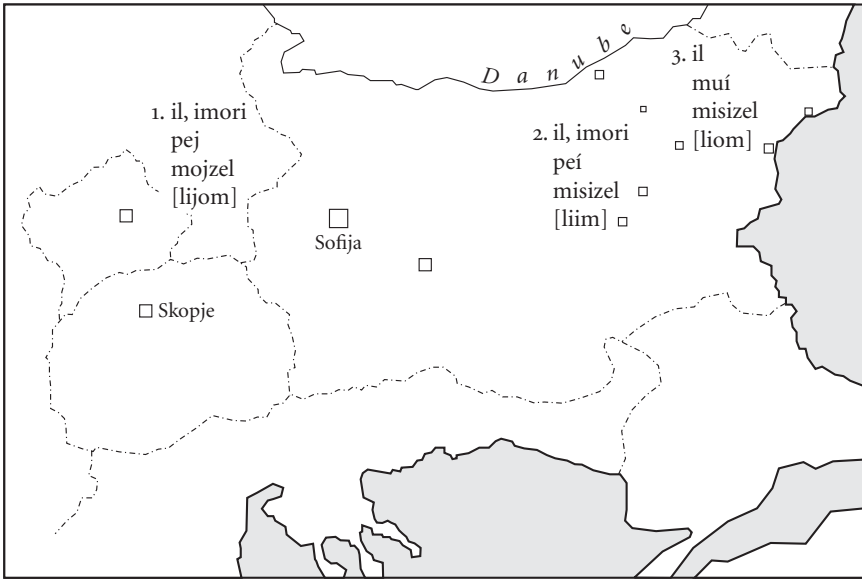
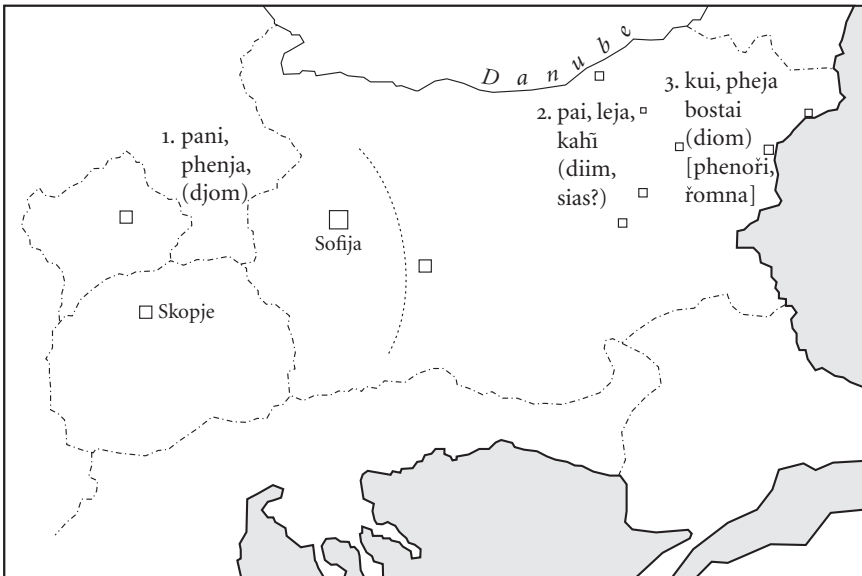
Map 7. Change of *c* > *č*Map 8. Palatalisation of velars (*k*, *kh*, *g*)

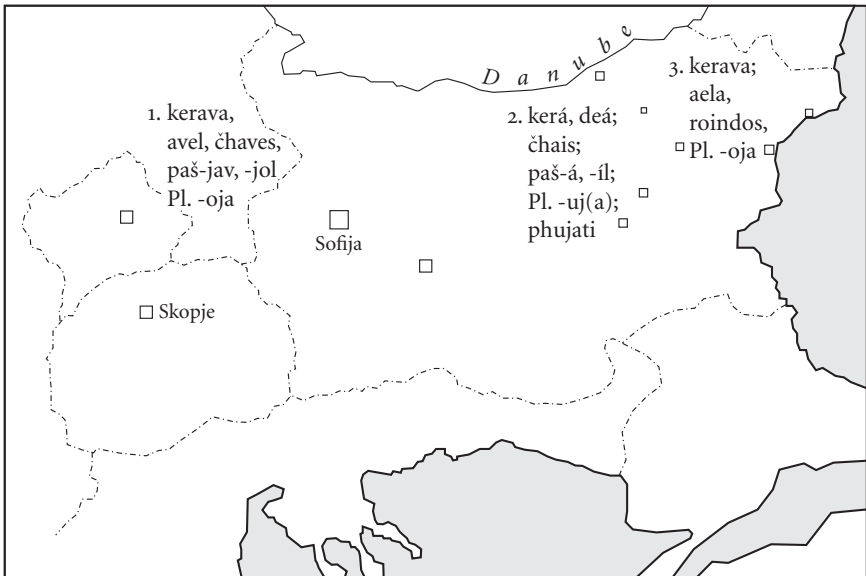


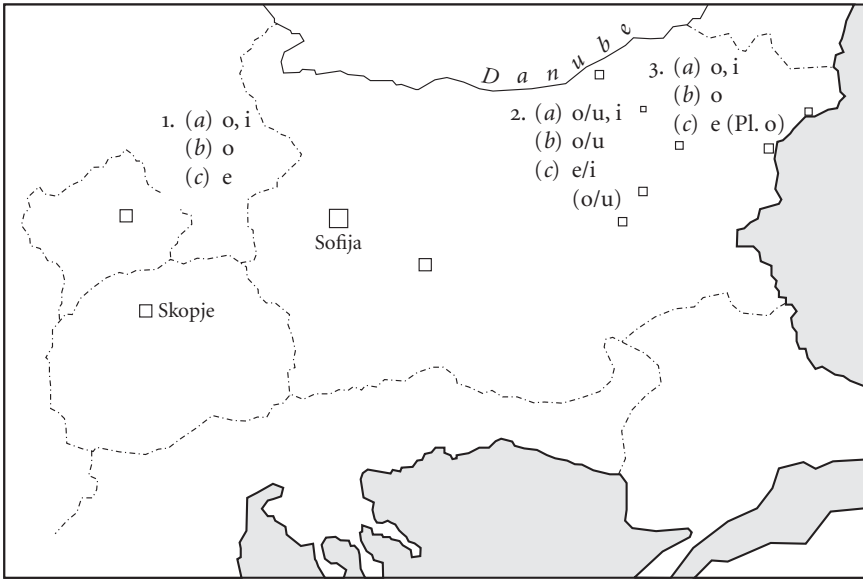
Map 9. Palatalisation of dentals (*t, th, d*)



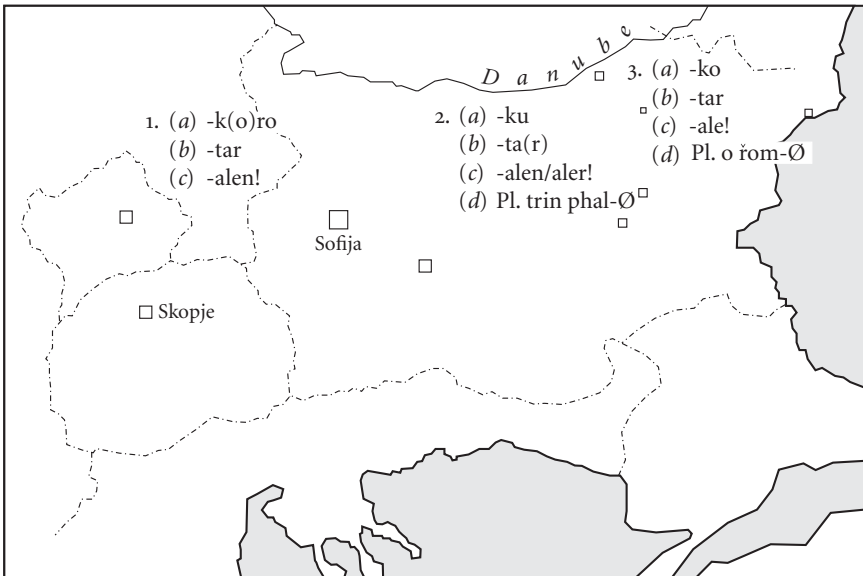
Map 10. Palatalisation of *kli, khli, gli*

Map 11. Palatalisation of *li/lj*Map 12. Palatalisation of *ni/nj*

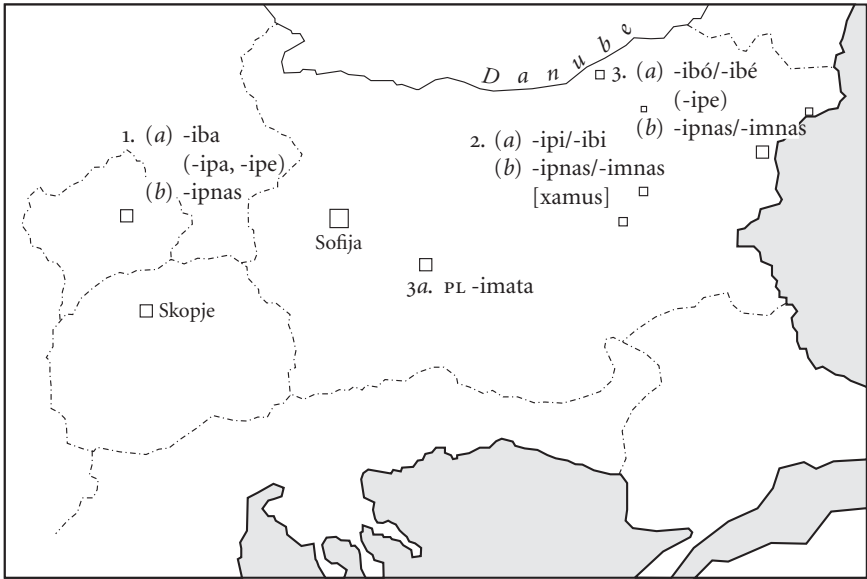
Map 13. Elision of *e* in the Present paradigmMap 14. Elision of *-v-* (and contraction)



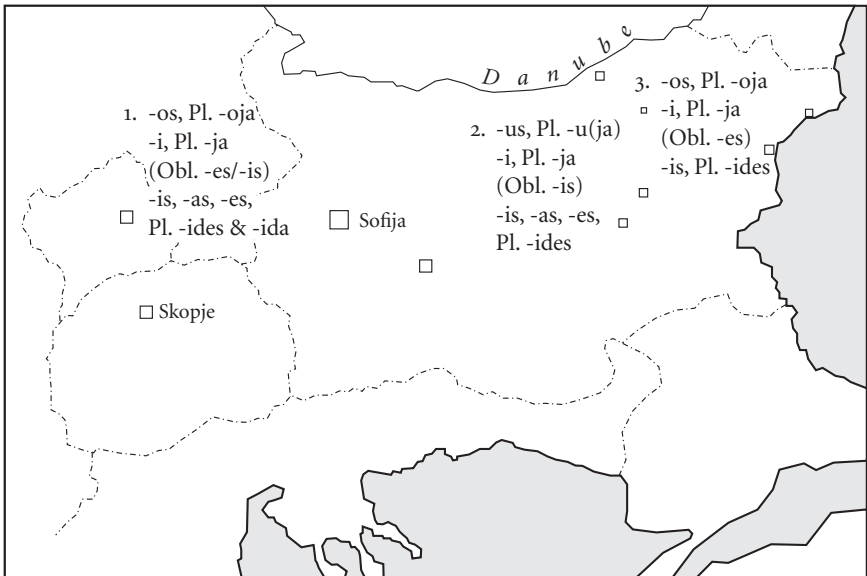
Map 15. The definite article: (a) Nom.Sg., Masc., Fem; (b) Nom.Pl.; (c) Obl.



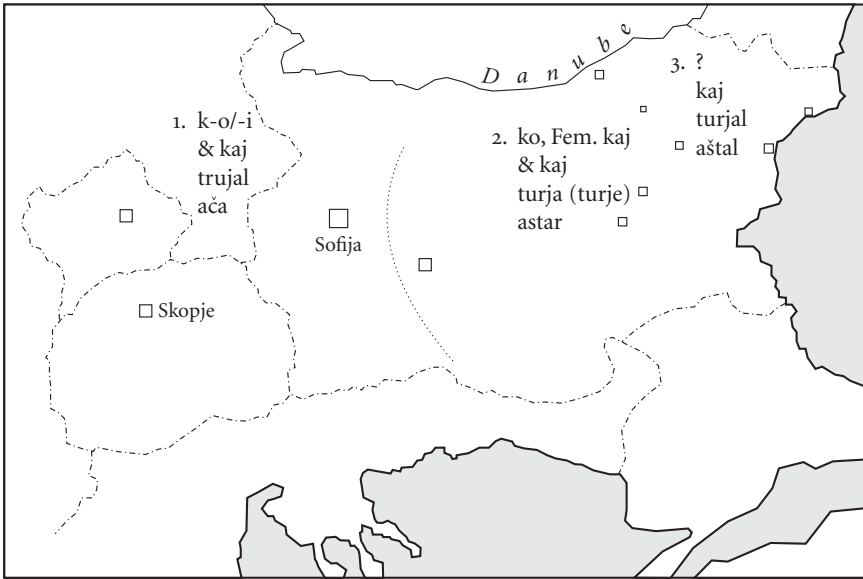
Map 16. Case forms: (a) Gen.; (b) Abl.; (c) Voc.Pl.; (d) Nom.Pl. in -Ø



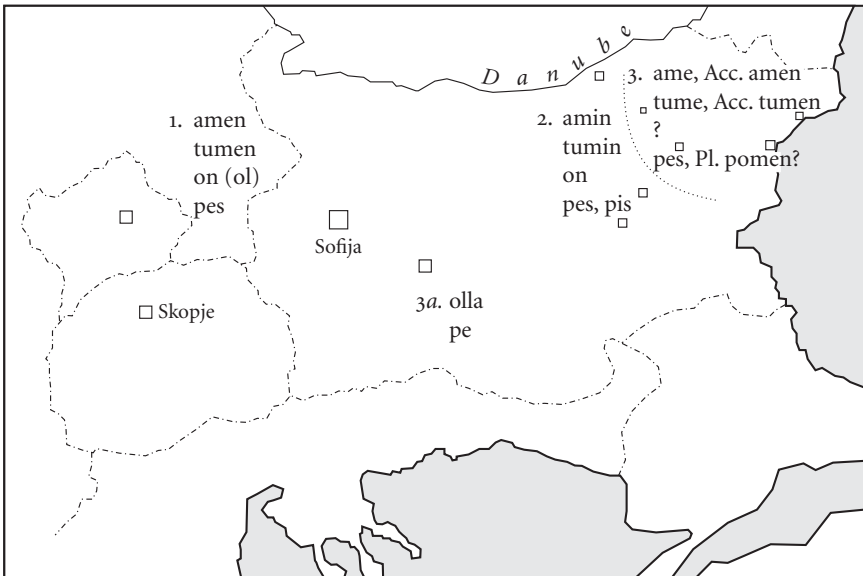
Map 17. Deverbal and deadjectival nouns: (a) Nom.Sg.; (b) Obl.



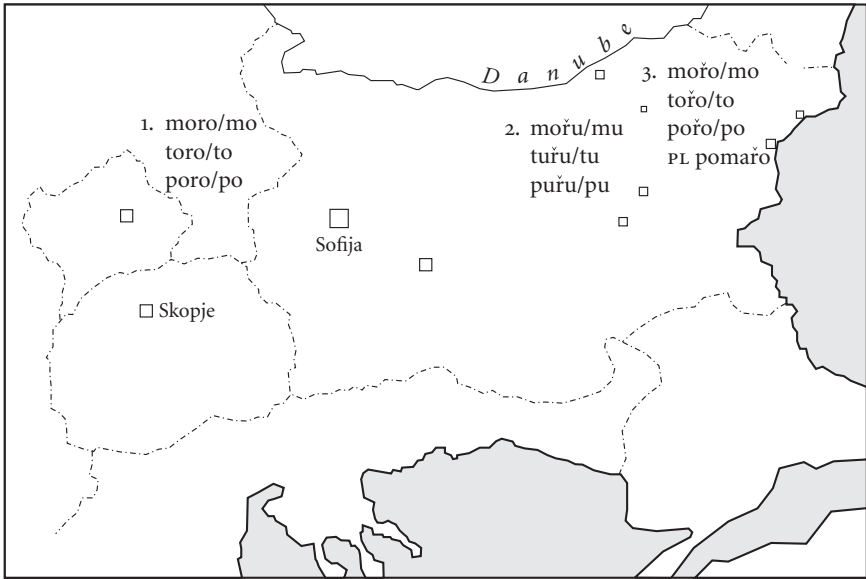
Map 18. Adaptation of borrowed nouns



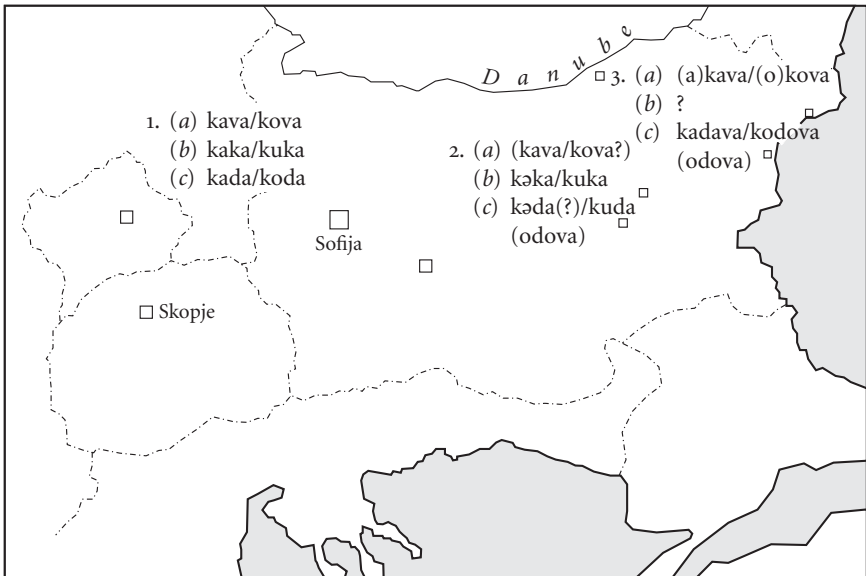
Map 19. Prepositions: *ke* and *kaj*, *trujal*, *\*vaš*



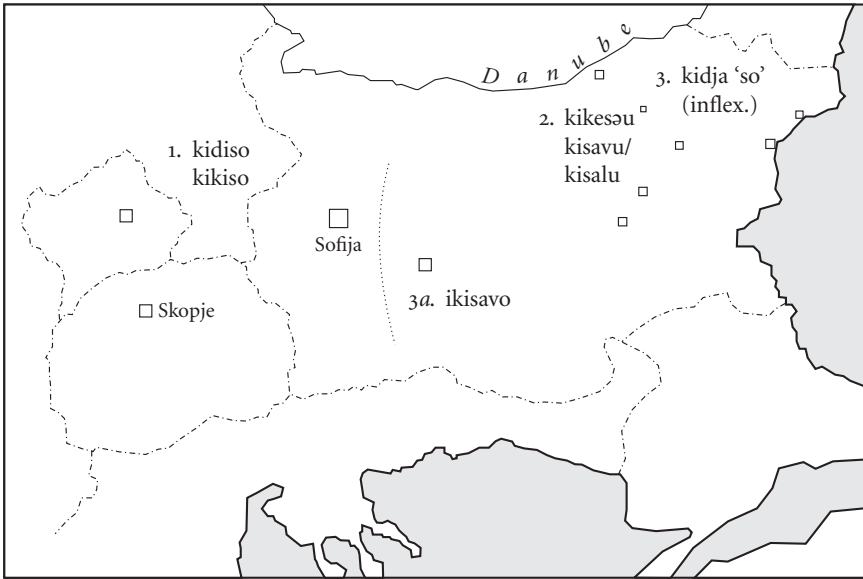
Map 20. Personal pronouns: 1, 2, 3 Pl.; Reflexive



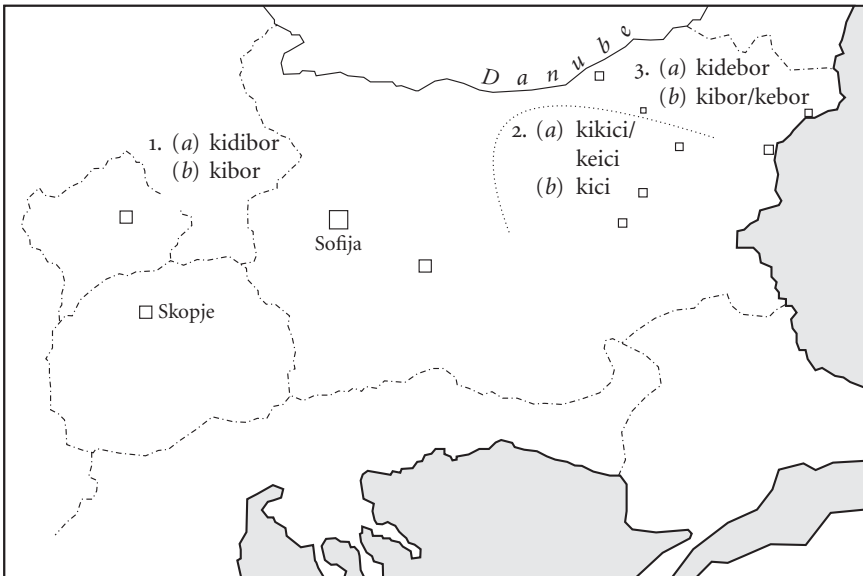
Map 21. Possessive pronouns: 1, 2 Sg; Reflexive



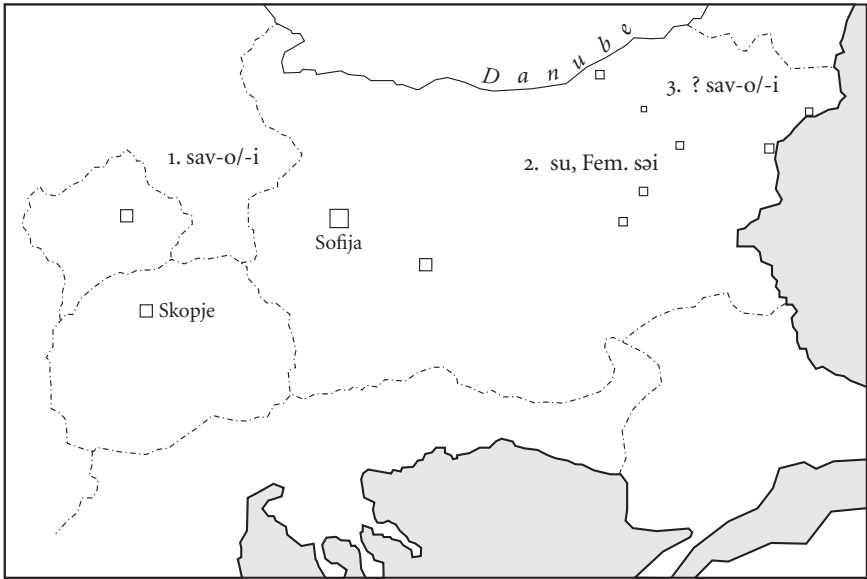
Map 22. Demonstrative pronouns



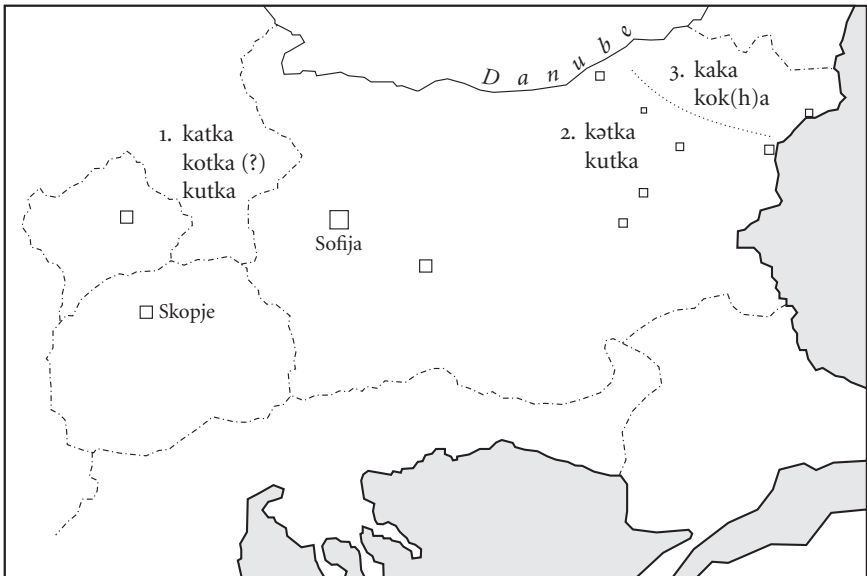
Map 23. Forms for 'such' (Masc.)



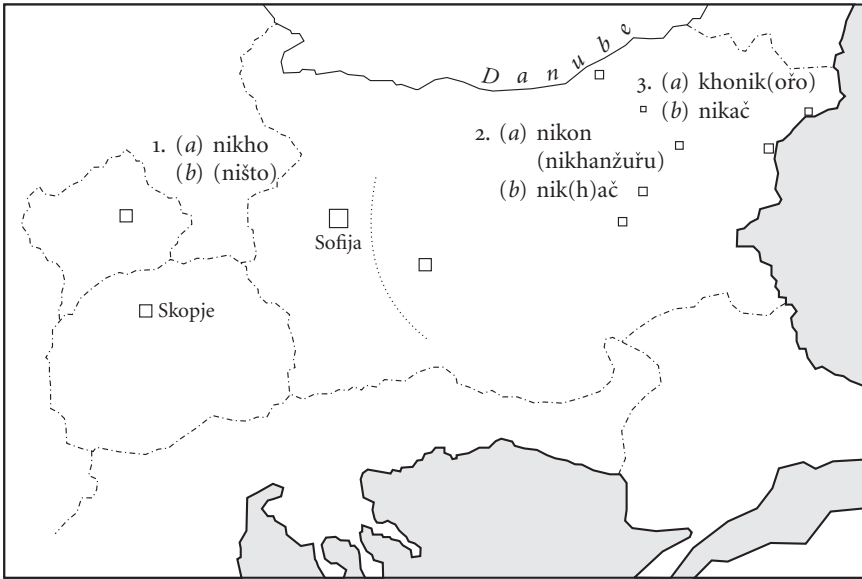
Map 24. Forms for (a) 'so much'; (b) 'how much'



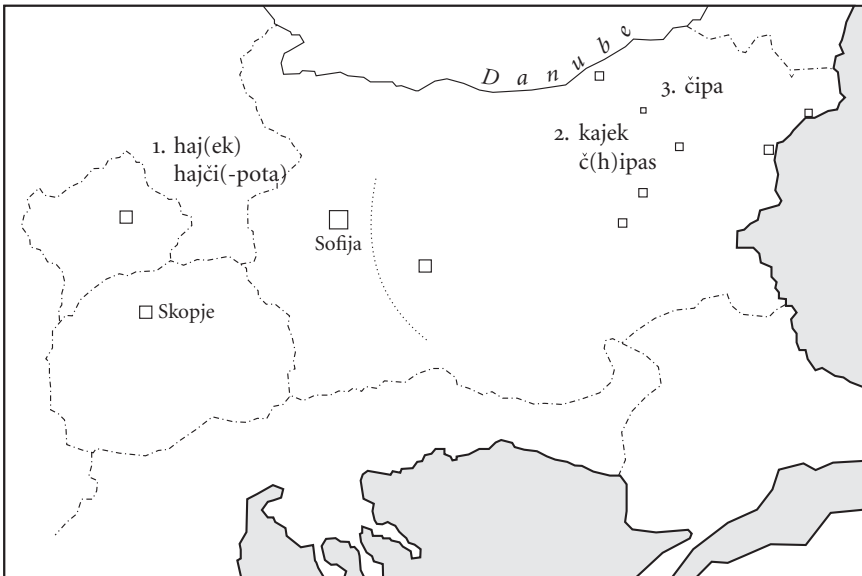
Map 25. Forms for 'which'



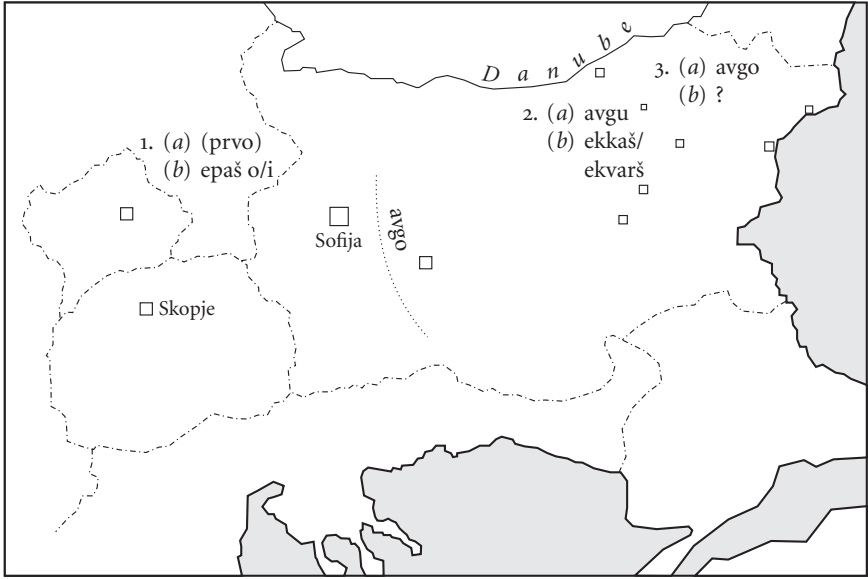
Map 26. Forms for 'here' and 'there'



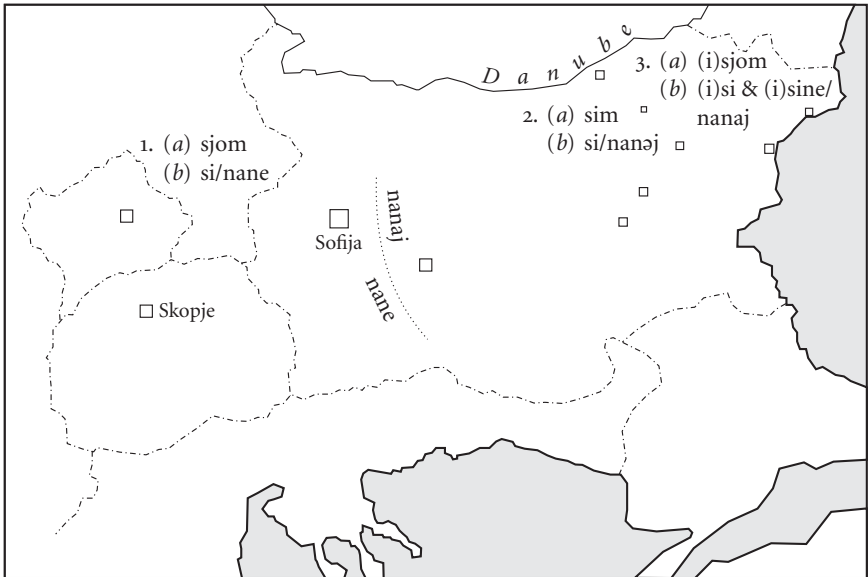
Map 27. Forms for (a) 'nobody'; (b) 'nothing'



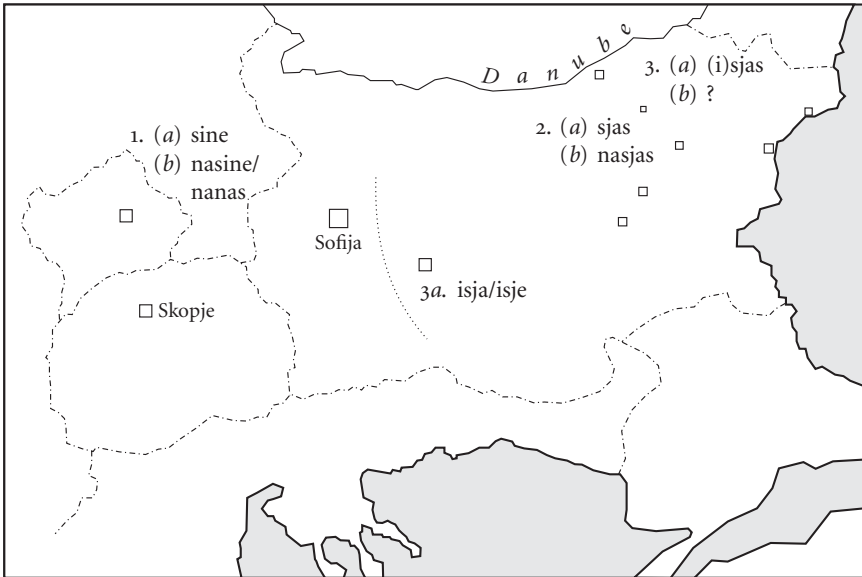
Map 28. Forms for 'something'



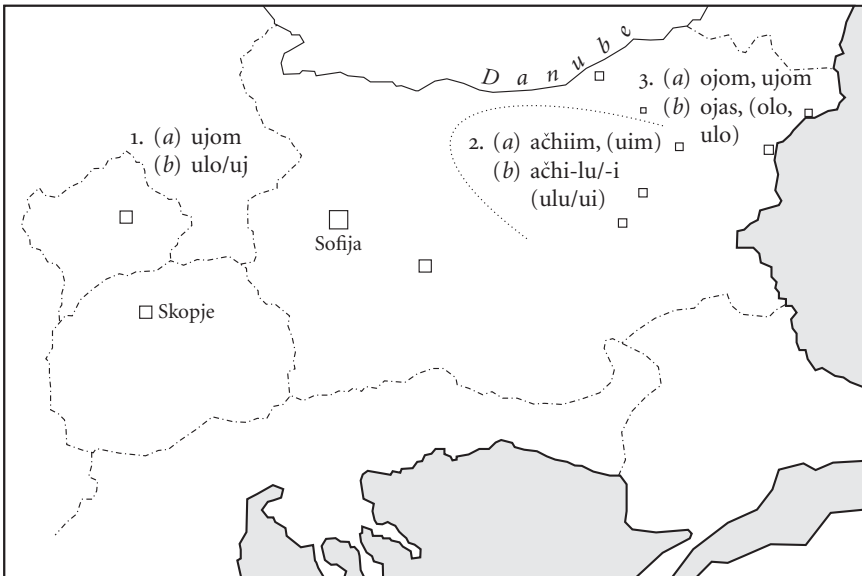
Map 29. Forms for (a) 'first'; (b) 'half'



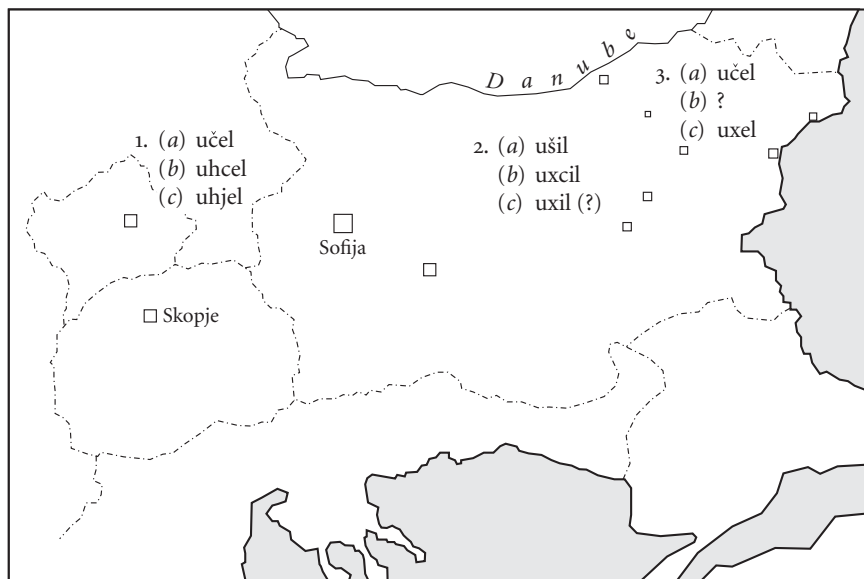
Map 30. Copula Present: (a) 1st Sg; (b) 3rd Sg. Pos./Neg.



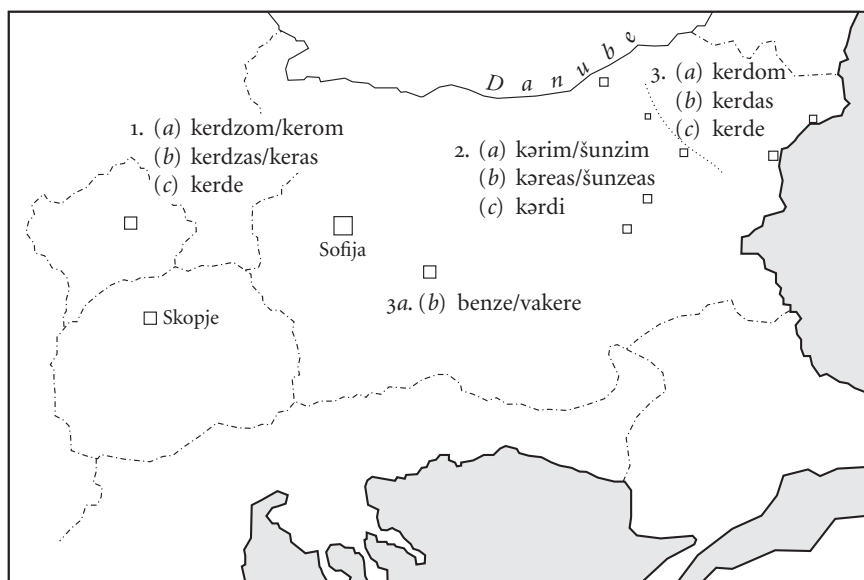
Map 31. Copula Preterite: (a) 3rd person positive; (b) 3rd person negative



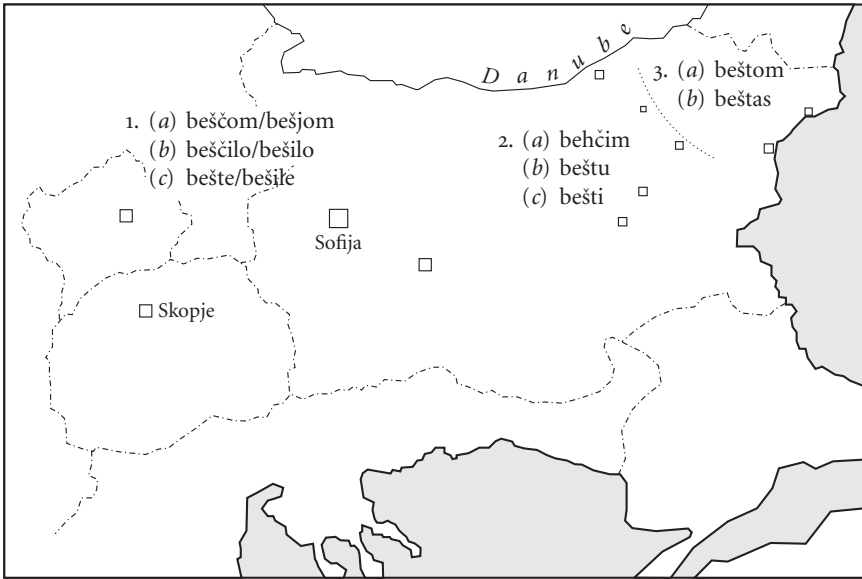
Map 32. Suppletive copula preterite: (a) 1st Sg.; (b) 3rd Sg. Masc./Fem.



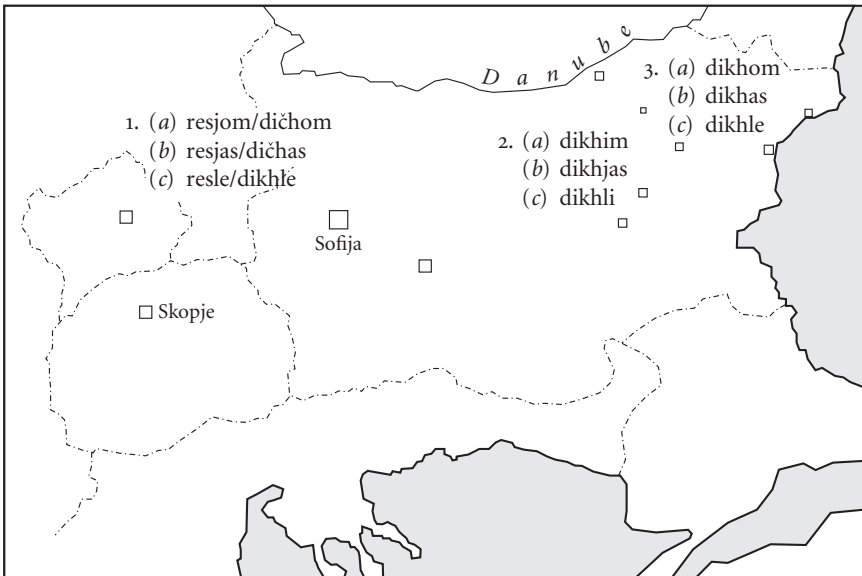
Map 33. (a) 'get up'; (b) 'jump'; (c) 'go down'



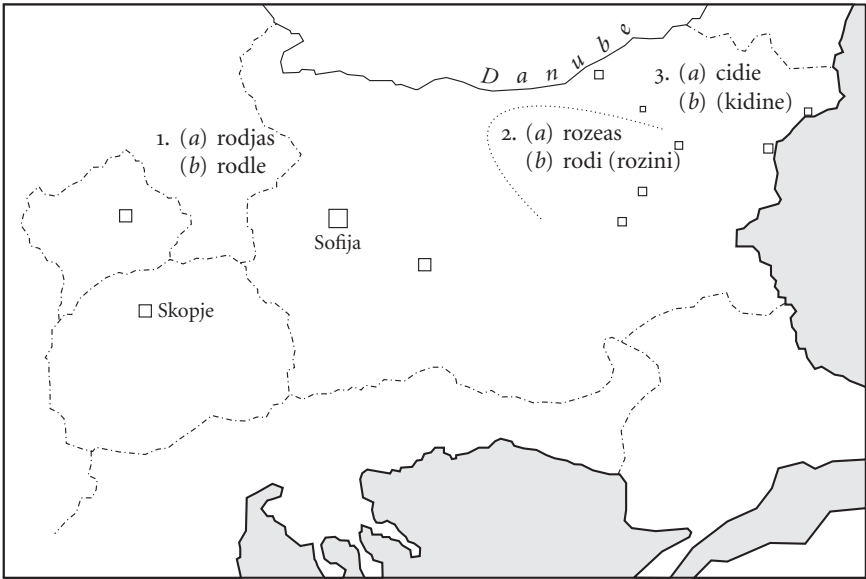
Map 34. *d*-Preterites: (a) 1st Sg.; (b) 3rd Sg.; (c) 3rd Pl.



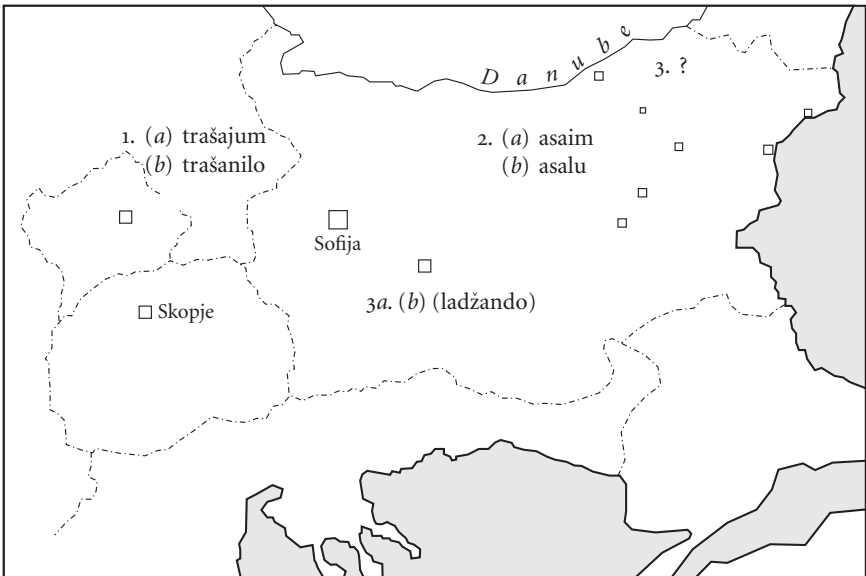
Map 35. *t*-Preterites: (a) 1st Sg.; (b) 3rd Sg.; (c) 3rd Pl.



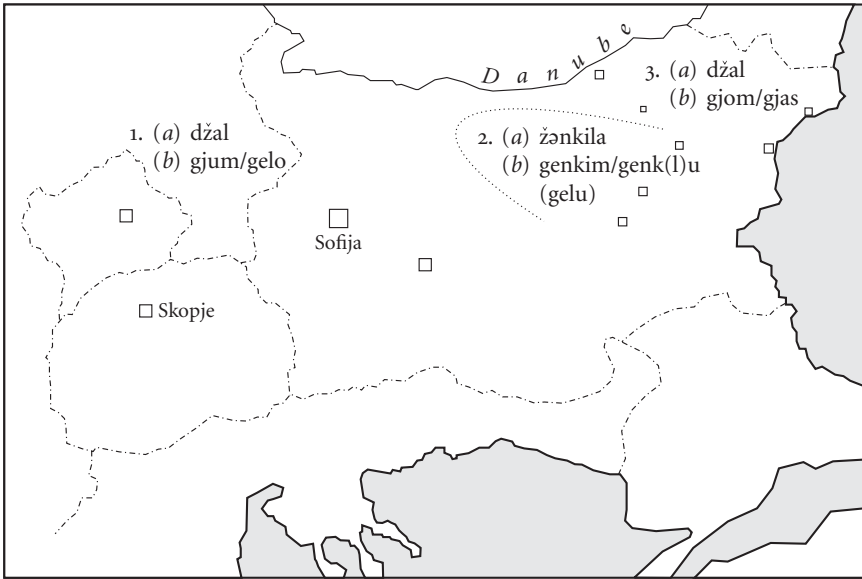
Map 36. *l*-Preterites: (a) 1st Sg.; (b) 3rd Sg.; (c) 3rd Pl.



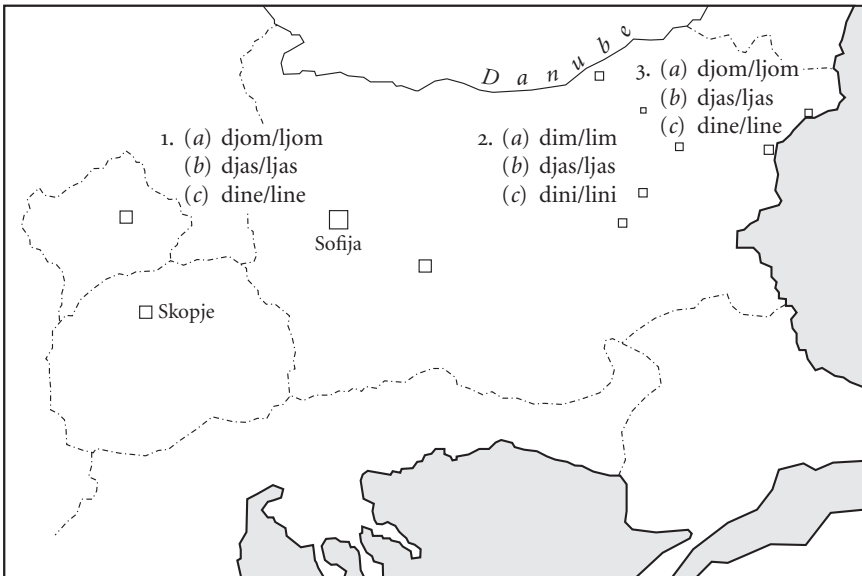
Map 37. Preterite of *d*-stems: (a) 3rd Sg.; (b) 3rd Pl.



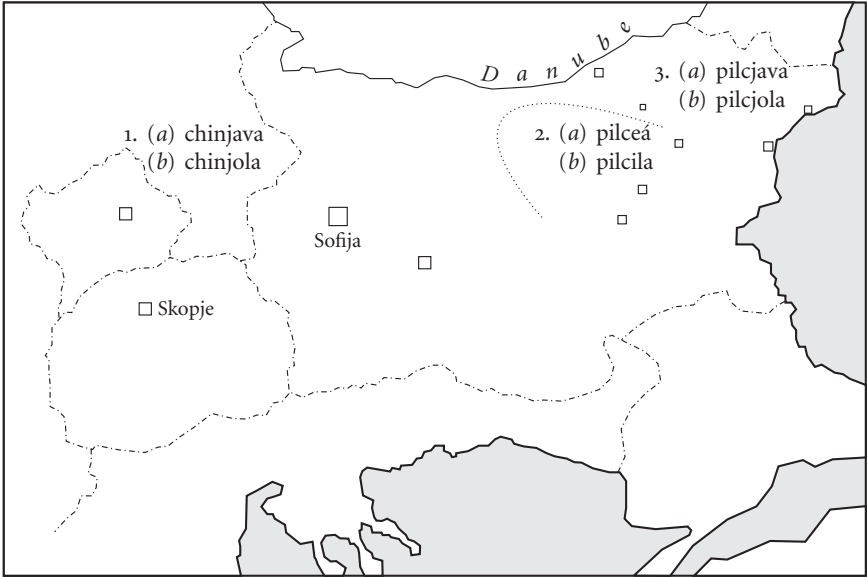
Map 38. Preterite of *a*-stems: (a) 1st Sg.; (b) 3rd Sg.



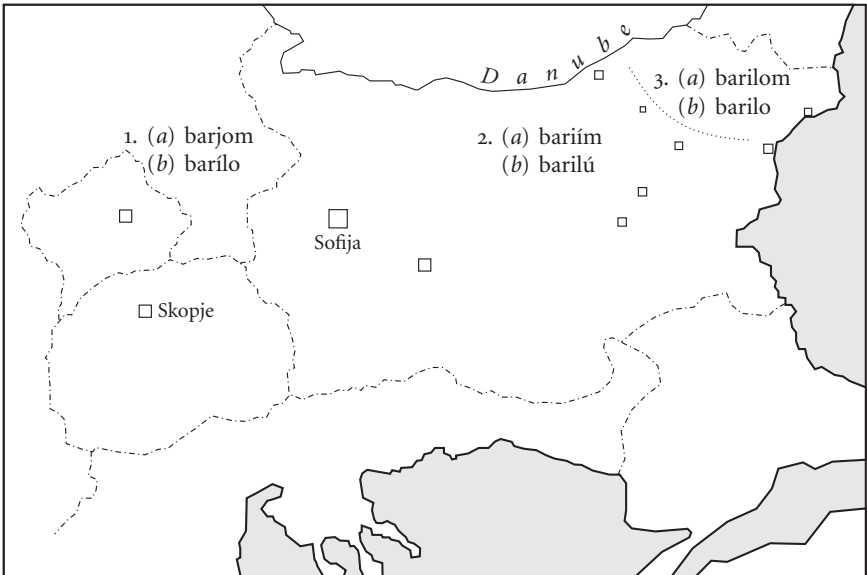
Map 39. 'go': (a) Present 3rd Sg.; (b) Pret. 1st Sg./3rd Sg.



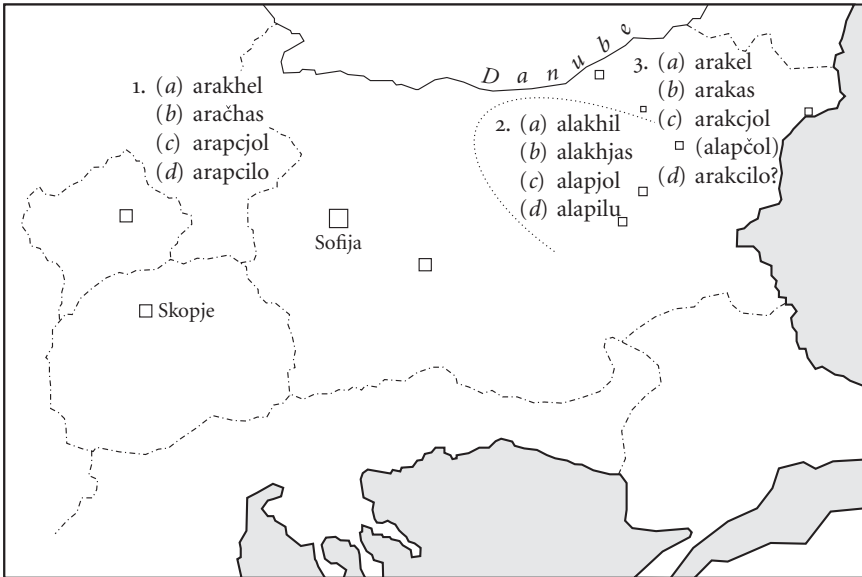
Map 40. 'give' and 'take': (a) Pret. 1st Sg.; (b) Pret. 3rd Sg.; (c) Pret. 3rd Pl.



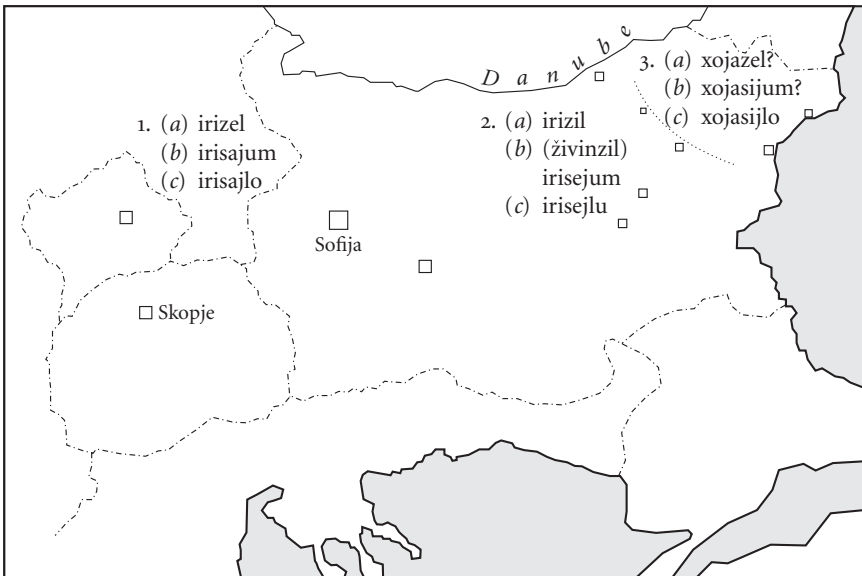
Map 41. Present passive: (a) 1st Sg.; (b) 3rd Sg.



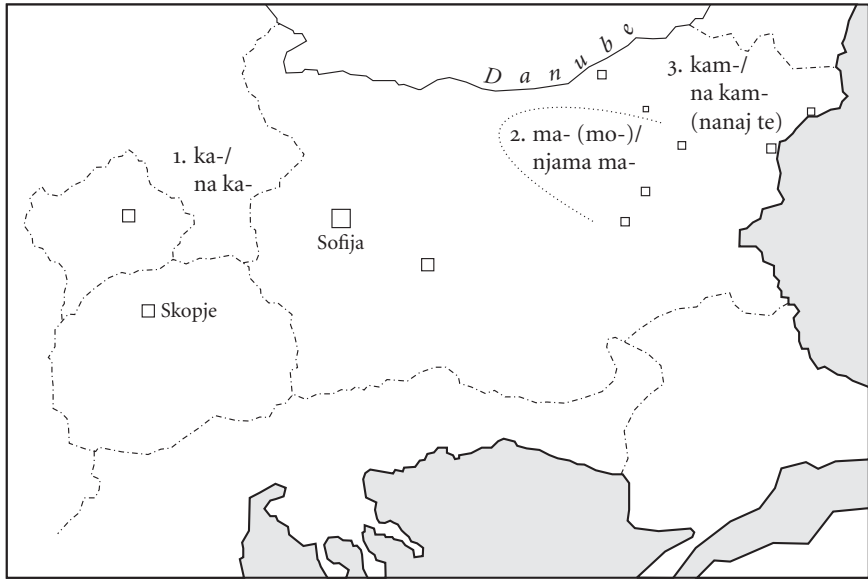
Map 42. Present passive: (a) 1st Sg.; (c) 3rd Sg.



Map 43. 'Find' 3 Sg.: (a) Pres. act.; (b) Pret. act.; (c) Pres. pass.; (d) Pret. act.



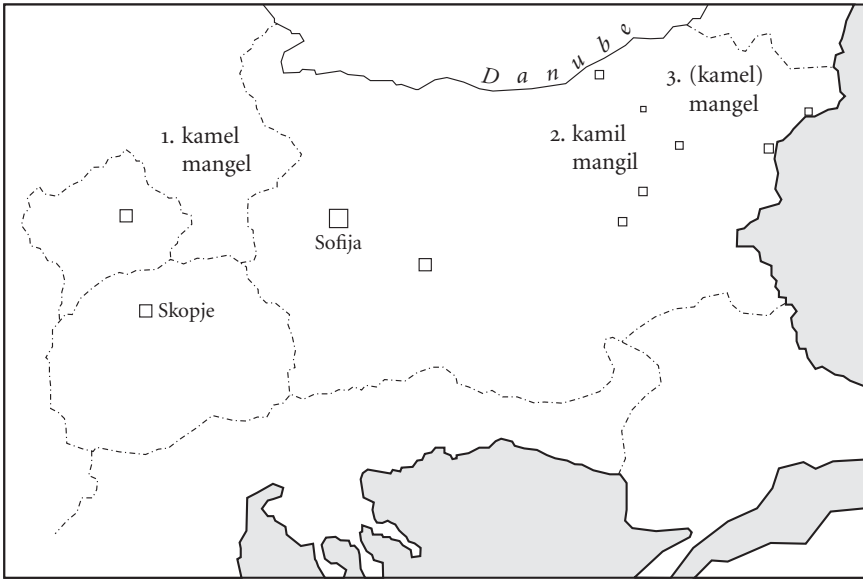
Map 44. Passive of borrowed verbs: (a) Pres. 3rd Sg.; (b) Pret. 1st Sg.; (c) Pret. 3rd Sg.



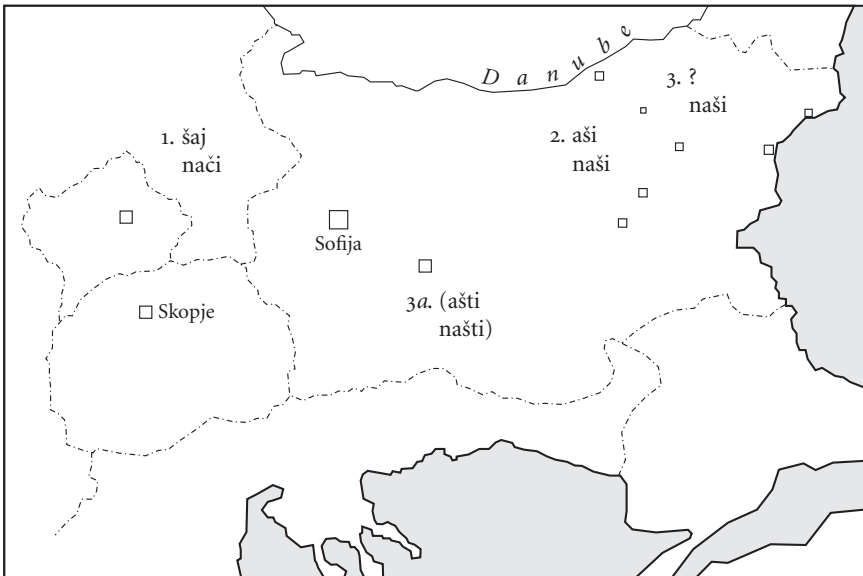
Map 45. Future marker: Pos./Neg.



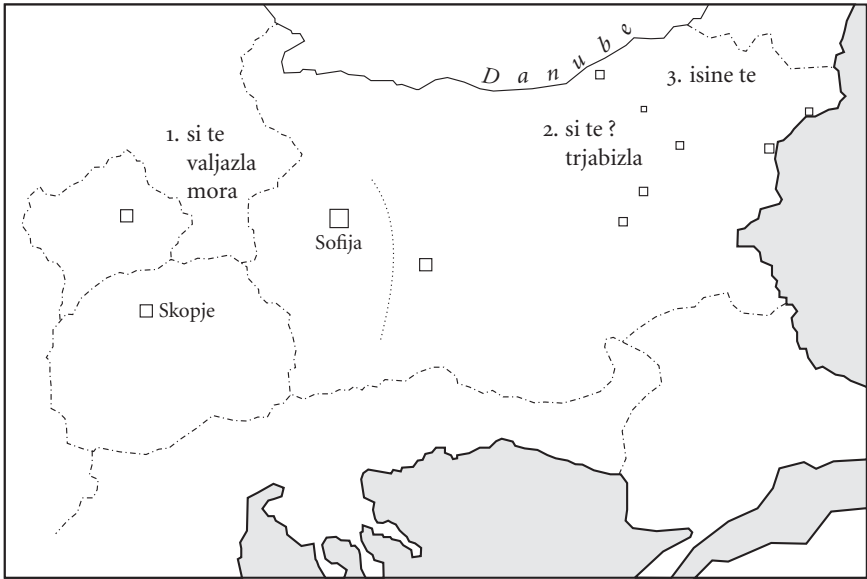
Map 46. Gerund



Map 47. Modal verb 'want': *kamel* ('love'); *mangel* ('demand')



Map 48. Modal element 'be able': 'can' and 'cannot'



Map 49. Modal element 'must', 'be necessary'

## References

- Atanasakieff, A. 1932. Somnal evangelie (il) Matejatar. Sofija: American and British Bible Society.
- Boretzky, Norbert. 1991. Contact-induced sound change. *Diachronica* VIII (1): 1–15.
- 1993. Bugurdži. Deskriptiver und historischer Abriß eines Romani-Dialekts. Wiesbaden: Harrassowitz.
- 1995. Die Entwicklung der Kopula im Romani. *Grazer Linguistische Studien* 43: 1–50.
- 1998. Erli. Eine Bestandsaufnahme nach den Texten von Gilliat-Smith. *Studii Romani* 5–6: 122–60.
- 1999a. Grammatical interference in Romani: Loan formations for foreign categories. *Acta Linguistica Hungarica* 46 (3–4): 169–200.
- 1999b. Die Gliederung der Zentralen Dialekte und die Beziehungen zwischen Südlichen Zentralen Dialekten und Südbalkanischen Romani-Dialekten. In *Die Sprache der Roma. Perspektiven der Romani-Forschung in Österreich im interdisziplinären und internationalen Kontext*. D. Halwachs and F. Menz, eds. Pp. 210–76. Klagenfurt: Drava-Verlag.
- 1999c. Die Verwandtschaftsbeziehungen zwischen den Südbalkanischen Romani-Dialekten. [Studien zur Tsiganologie und Folkloristik 27]. Frankfurt: Peter Lang Verlag.
- Colocci, Adriano. 1889. *Gli zingari*. Torino: Loescher.
- Elšík, Viktor, Milena Hübschmannová, and Hana Šebková. 1999. The southern central (*ahi-* imperfect) Romani dialects of Slovakia and northern Hungary. In *Die Sprache der Roma. Perspektiven der Romani-Forschung in Österreich im interdisziplinären und internationalen Kontext*. D. Halwachs and F. Menz, eds. Pp. 277–90. Klagenfurt: Drava-Verlag.
- Finck, Franz Nikolaus. 1903. *Lehrbuch des Dialekts der deutschen Zigeuner*. Marburg: Elwert.
- Gilliat-Smith, Bernard J. 1914. The dialect of the Drindaris. *Journal of the Gypsy Lore Society. New Series*. 7: 260–98.
- 1915/16. Report upon the Gypsy tribes of North-East Bulgaria. *Journal of the Gypsy Lore Society. New Series*. 9: 1–54, 65–109.
- 1925. The Song of the Bridge. *Journal of the Gypsy Lore Society. Third Series*. 4: 103–14.
- 1931. A Drindari folk-tale. *Journal of the Gypsy Lore Society. Third Series*. 10: 76–86.

- Gilliat-Smith, Bernard J. 1935. The dialect of the Moslem Kalajdzis (tanners) of the Tatar Pazardžik district (being a review of Somnal Evangelje Matejatar by A. Atanasakieff, 1932). *Journal of the Gypsy Lore Society*. Third Series. 14: 25–43.
- 1940. The ballad of Lady Rose. *Journal of the Gypsy Lore Society*. Third Series. 9: 146–52.
- 1944. A Bulgarian gypsy tale. Ali Üstas. *Journal of the Gypsy Lore Society*. Third Series. 23: 14–21.
- 1962. A new version of the Song of the Bridge. Communicated by L.N. Čerenkov. *Journal of the Gypsy Lore Society*. Third Series. 41: 124–33.
- Igla, Birgit. 1997. The Romani dialect of the Rhodopes. In *The typology and dialectology of Romani*. Yaron Matras, Peter Bakker, and Hristo Kyuchukov, eds. Pp. 147–58. Amsterdam: John Benjamins.
- Kenrick, Donald S. 1967. The Romani dialect of a musician from Razgrad. *Balkansko Ezikozanie* 11–12: 71–8.
- 1969. Morphology and lexicon of the Romani dialect of Kotel (Bulgaria). Unpublished Ph.D. thesis, London.
- Knobloch, Johann. 1950. Volkskundliche Sinti-Texte. *Anthropos* 45: 223–40.
- 1977. Berichte und Erzählungen württembergischer Zigeuner. In *Zigeunerkundliche Forschungen, I* [Innsbrucker Beiträge zur Kulturwissenschaft. Sonderheft 42]. J. Knobloch and I. Sudbrack, eds. Pp. 43–57. Institut für Sprachwissenschaft der Universität Innsbruck
- Krasnići, Ali. 1981. Čergarendje jaga. Čergarske vatre. Priština: Jedinstvo.
- Marušakova, E. and Veselin Popov, eds. 1997. The Song of the Bridge [Studii Romani 3–4]. Sofia: Litavra Publishing House.
- Matth *see* Atanasakieff
- Mirčev, Kiril. 1963. Istoričeska gramatika na bälgarskija ezik. Sofia: Nauka i Izkustvo.
- Nounev, Yosif. 1994. Legends. *Studii Romani, I*: 48–61. Sofia: Club '90 Publishers.
- Paspati, Georgios A. 1870. Études sur les Tchinghianès ou Bohémiens de l'Empire Ottomane. Constantinople: Koromela.
- Sampson, John. 1926/1968. The dialect of the Gypsies of Wales. Oxford: Clarendon Press.
- Smart, Bath C., and Henry Th. Crofton. 1875<sup>2</sup>. The dialect of the English Gypsies. London: Asher and Co.
- Stojkov, Stojko. 1993. Bälgarska dialektologija. Sofia: Bälgarska Akademija na Naukite.

Uhlik, Rade. 1941. Serbo–Bosnian Gypsy folk-tales, No 6. *Journal of the Gypsy Lore Society. Third Series.* 20: 25–36.

