0. INTRODUCTION

The Yugoslav lands are the most ethnically diverse in the Balkans and have a recent history of the most pluralistic language policies.1 Within Yugoslavia as it emerged from World War Two, Macedonia was the only republic in which the official language underwent complete shift from the inter-war period. Until 1944, the language of the majority of the population — Macedonian — was proscribed (see Kloss 1968) not only in Yugoslavia but also in every other country where it was spoken (Albania, Bulgaria, and Greece). The recognition and establishment of Literary Macedonian occurred in 1944, although activity leading to its codification dates from the nineteenth century (see Friedman 1975). The language of the next largest ethnic group in Macedonia — Albanian — was likewise not permitted in Yugoslavia during the interwar period (although it developed as a literary language in neighboring Albania), and after the War the official form of Albanian in Yugoslavia remained a Cicg-based standard related to the one developed in Albania before the War. It was not until 1968 that the post-war, Tosh-based standard in use in Albania since 1952 was officially adopted for use in Yugoslavia.2 Macedonia also has the largest numbers and percentages of Turkish speakers, as well as the largest percentage of Romani-speakers in Yugoslavia. Moreover, the Roma of Macedonia are in the forefront of the current international movement to establish a broadly based Romani literary language. The Macedonian ethnic and linguistic picture is further complicated by a significant Arumanian (Vlah) minority of over 6,000.3 In the context

---

1 This article was originally written in 1991 for a volume entitled Variation and Between in the Languages of Yugoslavia, ed. by Olga Mikesa-Tomík, to have been published by Benjamins, Amsterdam.

2 Since 1981, however, there has been an erosion of legal linguistic pluralism.

3 Albanian is also spoken in parts of Montenegro bordering on Albania, and it is the majority language in Kosovo (Serbo-Croatian Kosov, Albanian Kosova/Kosovë, in the spirit of linguistic neutrality that is receiving increasing recognition among linguists sensitive to these issues I have adopted a form in keeping with the principles of English toponymy).

4 These are not to be confused with the Wallachian population of Eastern Serbia, who in the Yugoslav census are also called Vlah but who speak a dialect of Daco-Romanian. Arumanian (and Megleno-Rumanian) are separate languages that split off from Daco-Romanian about a thousand years ago. Arumanian is also spoken in Albania, Greece, and Bulgaria while Megleno-Rumanian is limited to seven villages in southeastern Vardar (Yugoslav) Macedonia and adjacent parts of Aegean (Greek) Macedonia (Atanasov 1991).
of this complexity, the relationship of dialectal variation to standardization and codification takes on increased significance. The choices and compromises that have been made and that are being made in the creation of linguistic norms in Macedonia reflect a multicultural diversity that is a source of fascinating richness but also of potential serious conflict. I shall examine here similarities and differences in the relation of dialectal variation to standardization in Macedonian, Albanian, and Romani by looking at a selection of salient features in phonology, morphology, syntax, and lexicon that are comparable among these languages. While a complete account of variation in any one of these languages is a topic for a monograph, it is my intention here to emphasize comparable features. Thus this account should by no means be taken as exhaustive, but rather as highlighting some of the comparable points among these languages.

In each language, there have been two major dialect groups involved in the codification process. For Macedonian, the main bundle of isoglosses follows the course of the river Vardar southeast to its junction with the Crna and then follows the Crna southwest to to Aegean (Greck) Macedonia, dividing the language into Western and Eastern dialect regions. The two major dialect groups of Albanian are Gheg (north) and Tosk (south). A significant bundle of isoglosses dividing the two dialect regions runs through the area south of the river Shkumbi in Albania and then along the north of shore of lake Ohrid in Macedonia. For Romani, the main dialectal division is between the so-called Vlax and Non-Vlax dialects. Due to Romani nomadism, the distinction is not so much based on geographically defined isoglosses as on specific features, in particular a large Romanian-based vocabulary in the Vlax dialect, whence their name. In the context of the Balkans, Non-Vlax dialects are characteristic of groups whose ancestors probably did not leave the confines of the Ottoman Empire.

The West-Central dialects of Macedonian spoken in the Veles-Prilep-Bitola-Kiza regions serve as the basis of Literary Macedonian. The overwhelming majority of Albanian speakers in Yugoslavia speak one or another of the varieties of Gheg, the only exception being the Tosk speakers of the Ohrid-Prespa region in southwestern Macedonia. Of the many dialects of Romani spoken in Macedonia, the two main dialects of Skopje, viz. Arlija and Dzambaz, have been most involved in codification efforts. Arlija is a Non-Vlax dialect, whereas Dzambaz appears to be a Vlax dialect, or a Vlax-based Non-Vlax dialect.

In terms of the interaction between dialect and standard, three different types of patterns emerge from the histories of codification efforts in these three languages. Macedonian and Albanian are sharply differentiated from one another in this respect. For Macedonian, the process of codification involved establishing the separateness of Macedonian from two closely related languages (Servo-Croatian and Bulgarian) and resisting influence from a third, more distantly related language (Russian; see Friedman Forthcoming), but the dialectal base of standard Macedonian has remained the West-Central dialects since the earliest efforts at codification (Misirkov 1903: 145). Albanian codifiers had to concentrate on overcoming deep internal dialectal division and moreover experienced a major shift in the dialectal base (from Gheg to Tosk). The situation in Romani is still in flux. In Skopje, the Arlija dialect is the one used in the television news broadcasts began in 1991, but speakers of other dialects complain that the broadcasts are difficult to understand, and some public notices show a mixture of features. Aside from competition between Arlija and Dzambaz (Arlija as the older dialect has more prestige), there is competition between Non-Vlax Balkan dialects like Arlija and Vlax dialects from other parts of Europe, especially Kalderash.

1. PHONOLOGY

1.1. Stress

The stress systems of the three languages show variations that have three different types of importance: in Macedonian East/West stress variation is of major significance in literary codification, in Albanian, Gheg/Tosk variation is limited to a few classes of lexical items, whereas in Romani stress variation is a feature that opposes the dialects of Macedonia (and elsewhere in the Balkans) to those spoken elsewhere. Macedonian dialects show considerable complexity in their stress patterns (see Alexander in this volume), and this relationship of dialects to the standard is changing. The codified standard prescribes the West-Central dialect pattern of fixed antepenultimate stress. Exceptions are lexical, usually unadapted loan words or suffixes, e.g. epërë completë (Turkish loan), and phrasal, usually from an expanded word boundary, that is, one stress (usually antepenultimate) within a phrase (accidental) unit, e.g. gë so mëne 'with me' (shift onto monosyllabic prepositions with personal pronouns). Many accentual units prescribed in normative grammars, however, are now considered

---

Footnotes:

1. The Turkish dialects of Macedonia were marginalized in the formation of modern Literary Turkish, although language reforms were also attempted in Macedonia during the Ottoman period (Hasan 1991). It is only since the end of 1990, with the appearance of newspapers in Arumanian and the beginning of an Arumanian news program for twenty minutes once a week on TV Skopje, that current efforts potentially leading to a codified Arumanian have begun.

2. For example, Byron (1768), which is just such a treatment of Albanian.

3. Outside the context of Romani, Vlax or Vlak refers to Balkan Romance speakers, cf. note 8 above.

4. For the most part, the dialects of Montenegro belong to Northwestern Gheg, those of Kosovo to Northeastern Gheg, and those of Macedonia to Central Gheg. The dialects of Central Albania (southern Gheg and Transitional) are spoken in some villages between Debar and Struga (Gjinari 1989: 53–55).

5. The classification of Romani dialects cannot be considered as definitively settled (cf. Hancock 1988).

6. Other Vlax dialects include Gurbet and Malcovo in southeastern Europe, Lovari and Çarari in northeastern Europe. The Vlax dialects have the most speakers and are spoken in the most countries, including the Western Hemisphere and Australia (Hancock 1991).
localisms or dialectisms by educated Macedonians, especially in the younger generations. This is due to the influence of the Skopje dialect (and Serbo-Croatian) on the literary language combined with apparent resistance on the part of speakers from outside the Western area to adopt these specifically Western types of pronunciation. Thus, prescribed pronunciations as in (1):

(1) *Eve ti go (as opposed to Eve ti go.)
here you-DAT him/it-ACC
*Here he/it is for you.*

are now considered local Western rather than literary pronunciations. Even in those regions where accentual units are native, it seems that the educated younger generation tends to avoid many of them.

The stress patterns of Albanian dialects do not show significant variation (Gjinari 1970: 26). In general, Albanian stresses the final syllable of the stem, although there also exist certain classes of exceptions (see Newmark, Hubbard and Pfitz 1982: 15–18). One major difference between Tosk and Geg is the treatment of Turkish loans in original final stressed -a or -e. In Tosk, Turkish loans are treated like other Albanian words with a final stressed vowel, hence *babë 'father', babdi 'the father', teneqë 'tin', teneqjia 'the tin', like native vëllë 'brother', vëldi 'the brother'. In Geg, however, the stress is shifted back and Turkish loans are adapted to the pattern of nouns in final schwa or original short -e, hence halvë 'sweetmeat', halvha 'the sweetmeat', pënsërë 'window', pënsërëja 'the window', like native çapë 'girl', çapa 'the girl', dëve 'ewe', dëjia 'the ewe'. In the standard language, the solution has been to adapt a mixture of these. Thus both Tosk babë 'father' and Geg babë 'father' are acceptable, but Geg tenqjia 'tin' and Tosk pënsërëja 'window' and halvë 'sweetmeat' are all excluded (Kostallari 1976).11

Although the stress patterns of a number of Romani dialects tend to conform to those of the chief non-Romani contact language, e. g. some Romani dialects of Germany, Hungary, and Czechoslovakia tend to have initial stress, those of Poland tend to have penultimate stress (Ventcel' and Čerenkov 1976: 297), the Balkan Romani dialects have generally converged on what must have been the original oxytone stress pattern. Exceptions are affixes of clitic origin (e. g., cases, the imperfect marker -sk), some final syllables in polysemous inflectional endings (e. g., final the -a of the present tense), and foreign loans. Within the context of Romani in Macedonia, stress is not an issue of standardization.

11 It should be noted that given the stress pattern in the Turkish source, it is Tes that uses stress shift and not Tosk, pace Pipa (1989: 9). There are a few other lexical items where Tosk oxytone corresponds to Tes paroxytone, e. g. Tes vëll Toks vëll 'north', etc. In such words, Southern Tes (Central Albania) patterns with Tosk. In clusters of two vowels such as [hai/iai, jei], some northern Tosk dialects shift the stress on to the second vowel and treat the sequence as a rising diphthong (see Gjinari 1975: 91–95). In all these instances, Standard Albanian follows Tosk (or southern Tosk).

12 Romani dialects in contact with languages with vowels other than these often have the additional vowels of the contact language in borrowed words.

13 Macedonian Turkish dialects, too, do not differ much in vocabulary from the eight phonemes of Standard Turkish, although the distribution does.

14 This is the opposite of the situation in Bulgarian, where lower, more open /e/, or are characteristic of women's speech.

15 The dialects of Dëbar (Albanian Dibër) in Macedonia and Ulqin (Albanian Ulqin) in Montenegro are unique among Romany in their lack of nasality.

16 The Southernmost Tosk dialects of Labëri in Albania and Çarëc in Greece have phonemic length. All Albanian dialects except those without phonemic high front rounded /ï/ (Albanian orthographic [i]) have these seven vowels as their unstressed inventory.
should be noted, however, that intellectuals of Kosovia and Macedonia insist
that phonemic length is a legitimate feature of the Kosovian variant of
Standard Albanian, i.e. this is one of those features that is a subject of
variation within Standard Albanian, at least for speakers living outside

As indicated above, Geg lacks stressed schwa (Albanian orthographic
[ø]). One area of important variation in this regard is the correspondence
of Geg nasal /n/ (orthographic [nj]) to Tosk /n/ or /n/, the distribution of which is
complex, with southern Tosk being most consistent in the use of /n/ (cf.
Byron 1976: 102). Pre-War Standard Tosk had /n/, but the post-War Stand-
ard has codified /n/, e.g. zemër 'hart' vs the dialectal variants zemër and
zemër. Most of the Albanian dialects of Macedonia have fourteen vowels, viz.
/a/, e, i, o, u, long, short, and nasal except nasal /n/, which does not occur in
any Albanian dialect. The /ə/ of other dialects here corresponds to /ə/. The
Central Geg dialects of Albania also lack /ə/ but add to this inventory mid
front rounded /æ/ and /ʌ/ thus having the same number but not the same
inventory of phonemes as Northeastern Geg including Kosovia and Southern
Geg including the villages between Struga and Debar (Albanian Dilër) in
Macedonia, which all have /u/ and nasal /n/ (Gjinari 1989: 102–106).\(^{17}\)

1.3. Consonants

The examination of consonantal phenomena will focus on those parts of
the sound systems in which a type of variation is shared between one or more of
the languages under consideration. From this point of view, palatals, liquids,
velar/post-velar fricatives and glides, the opposition voiced/voiceless, and the
elimination of certain consonants in intervocalic position (Albanian /n/ ~ /l/,
Macedonian /r/ ~ o, Romani /n/, ~ o) are relevant to this discussion.

Standard Albanian, like Standard Macedonian, has a phonemic contrast
between strident palatal affricates /tʃ, ʃ/ and mellow palatal stops /k, ɡ/
(Albanian orthographic [ç, ð] and [g, q], IPA [ʃ, ʒ, c, ŋ]). Most dialects of
Kosovia (also parts of Macedonia such as Kumanovo and Debar), however,
merge the two into palatal affricates (Gjinari 1989: 156). Some Macedonian
dialects, particularly urban ones such as Prilep, exhibit the same phoneme-
non. The Romani dialects of Macedonia do not have a phonemic contrast
between mellow and strident palatals, but both occur at the phonetic level.
In Džambaz, original /h, d/ is backed to /k, ɡ/ (realized as [k, ɡ] due to automatic
fronting before front vowels) before stressed /l/, e.g. bukë 'work' from the

\(^{17}\) The Albanian dialects of Macedonia and adjacent Central Geg dialects of Albania also
have a characteristic diphthongization of stressed /u/ and /n/ to [a], [ɛ, ɔ] and [a, ɔ] in many
environments, e.g. in Debar [sɔp] for standard shër seli 'IMP, [mog] for standard suq 'corn'.
(Basha 1989: 148–50). On the other hand, Geg in general tends to monophthongize original
diphthongs, e.g. /ou > o, i > ɛ, y > ɔ. As mentioned above, in Tosk there is variation between
treating these as diphthongs, which is characteristic of northern Tosk on vowel sequences,
which is southern Tosk and standard (Gjinari 1989: 201). In general, Geg has /n/ and Tosk /n/
from original initial *o, e, g. voar 'water' hearth' and heor where Tosk has /a/, e.g. mua vs
mua 'me'.

older form buqi.\(^{18}\) In Arlja, morphophonemic jotation in the aorist stem results in the same phoneme,
e.g. kerdu, doncu [PART], kerdumu = [kergum] I did' (cf. Džambaz kerdem, I did').

Albanian and Macedonian have a phonemic opposition between clear /l/
and velar /l/, whereas these two are in automatic alternation in Romani
(clear [l]) before front vowels and /y, velar [l] elsewhere.\(^{19}\) In Macedonian
there is a tendency to shift clear [l] to palatal [l] under the influence of
Serbo-Croatian (and the Skopje dialect), or, among some younger speakers,
to have velar /l/ in all positions.

The presence of the velar fricative /ʃ/ is characteristic of the Eastern
Macedonian dialects, whereas in Western Macedonian it is lost initially and
intervocally (except before /u/) and replaced by /o/ (devocalized to /o/ finally
and before voiceless consonants) elsewhere. In this instance, Literary
Macedonian follows the Western dialects except under a limited number of
circumstances, viz. Church Slavonicisms (dus 'spirit'), recent loan words
(hotel 'hotel'), neologisms (dohod 'income'), toponyms (Ohrid) and cases
where otherwise there could be ambiguity in which case an Eastern form is
used (hrana 'food' vs cena 'cost'). Standard Albanian, like most of Tosk,
has a phonemic laryngeal glide /h/, but northern Geg and the dialect around
Debar eliminate /h/ either by a shift to /o/, e.g. shoq ~ shok 'see', foqot ~
foqi, cold', or by dropping it altogether, e.g. ekur ~ ekur 'iron'.\(^{20}\) The
elimination of /h/ also occurs in the north Tosk dialects of Berat, Korçë,
and Devoll (Gjinari 1989: 155, 163). All Romani dialects have the glide /h/,
but most contrast in with velar and/or uvular /h/ or /h/ (e.g. Džambaz sugas~
I cough, nasav 'laugh'). In Arlja, however, these phonemes merge and are
in free variation. Although both sounds were prescribed at the Romanian
World Congresses (Kenrick 1981, Balci et al. 1991), some codifiers in
Macedonia reject adopting the distinction into the Arlja-based standard (cf.
Jusuf and Kepeski 1980).

Each of the three languages also has a basic dialectal variation that results
from an earlier consonant shift or loss in intervocalic position. In
Albanian, original /n/ became /l/ in Tosk but remained in Geg, e.g. Džambaz
Shqiptar 'Albania' ~ Geg Shqipini 'Albania'. In Macedonian original inter-
vocalic /l/ was lost in the West but not in the East, e.g. Western voloj 'oxen' ~
Eastern volovi 'oxen'. In Romani intervocalic /l/ was eliminated in certain
inflectional endings in Arlja but not in Džambaz and not in Arlja, e.g. Arlja
devela ~ Džambaz devlesa 'with god', Arlja pani ~ Džambaz pani 'water'. In
Standard Albanian, rhotacized forms are prescribed with a few lexical
exceptions (e.g. dashnor 'lover', cf. dashuri 'love'). In Macedonian

\(^{18}\) This is also a position of variation in other Romani dialects, e.g. Burguči buki, Burgub
buki 'work'.

\(^{19}\) Macedonian has a similar distribution, except that clear /l/ occurs before back vowels
and word-finally in words of Turkish origin and from original sequences of *"l* or *"l*
before back vowels (e.g. bela, white' FEM, bela 'troubled'.

\(^{20}\) Loss of /h/ is also characteristic of the Turkish dialects of Macedonia and Kosovia (cf.
Friedman 1982).
intervocalic /v/ was retained, although the plural ending -oj was permitted as a variant until 1950. In Literary Romani usage as it is developing in Macedonia, Arilja variants such as those described above are generally favored over Džambaz variants. In Romani dialects spoken outside of Macedonia, however, retention of intervocalic /v/ is characteristic of the majority of dialects. Moreover, loss of /l/ before stressed /l/ occurs in Vlax dialects, which are also being used for literary purposes by some Romans in Europe and the United States.

Final devoicing is automatic throughout Macedonian. In Albanian it is characteristic of Northern Tosk and transitional Southern Gëg but not of the Standard, while in Romani it occurs sporadically in some dialects, presumably under foreign influence but is not part of the developing standard.

2. MORPHOLOGY

Morphology is a rich source of dialectal variation in all three languages. Among the comparable phenomena of variation in verbs are shapes of inflectional endings, choice of base for the formation of paradigms (e.g. present vs aorist stem), types of analytic constructions (e.g. choice of auxiliary), and grammatical categories expressed (e.g. presence vs absence of an aorist/imperfect distinction in some paradigms). In the nominal systems, sources of variation include the treatment of definiteness, case morphology or the expression of case relations, the shape of pronouns, and the expression of reflexive possession.

2.1. Verbs

Among the most characteristic variations in the shapes of inflectional endings in Macedonian are the generalization of -am for all verbs in the first singular present, in the West and East Central dialect groups (as opposed to a variety of inflectional endings, e.g. -mi, -mi, -em, -em elsewhere), the presence of final -t in the third singular present in the West and its absence in the East, and phonological variation in the marker of all but the second/third singular of the synthetic pasts (aorist and imperfect), which is /t/ in the East and /l/ or /l/ in the West.

In Albanian, there is a host of variants in verbal morphology. Thus, for example, in Northern Tosk the ending of the first singular present and subjunctive in vocalic stem verbs contains /l/ else it contains /j/, e.g. Northern Tosk punoi, 'I work' (PRES) and (SUBJ) ~ other Tosk puno, 'I work' (PRES) and (SUBJ). The morphologically distinct subjunctive, which

occurs only in the second and third persons singular, is lost through generalization in Kosova and Macedonia. Thus, for example, Standard Albanian (and the dialect of Korça) opposes puno, 'work' (2SG PRES INDIC) to punoi, 'work' (2SG PRES SUBJ) and puno, 'work' (3SG PRES SUBJ). In Kosova, however, the suffix -n is generalized for both persons in both moods while in some dialects of Macedonia, e.g. Debar, the form in -n is generalized for the third singular in both moods and the suffix -nh is used for both second singulars (v. Ajeti 1978: 11–17, Barsha 1989: 186–190). The presence of /j/ in the imperfect is characteristic of Gëg (except the southeastern Gëg of Albania), whereas its absence is characteristic of Tosk, r. g. Gëg puno, 'work' (1SG IMPF = 'Tosk puno, 'work') 1SG IMPF ~ 'Tosk puno, 'work' 1SG IMPF. The shape of the standard variant, -ja, e.g. puno, 'work' 1SG IMPF ~ is characteristic only of a small region just north of Southern Tosk, although the suffix /j/ is followed by schwa, /e/, or /o/ occurs throughout Southern Tosk as well as in Southeastern Gëg (Gjini 1989: 248).

Throughout European Romani there is variation in the present conjugation between the presence versus absence of final /l/ in all persons, e.g. kerava, 'I do' ~ kerav, 'I do', kera(s), 'you do' (SG) ~ ker(s), 'you do' (SG). In some dialects, grammatical significance is assigned to final /l/ in others pragmatic, i.e. in some dialects this /l/ marks a grammatical category, in others it is a non-obligatory signal that tends to occur in certain contexts. Thus, for example, in Lovari, Kadërani and some other dialects of Northwestern European, this /l/ is added to the present in order to make it a future. In some of these same dialects, however, this /l/ can also be added to the present tense to mark oratorical style without implying future meaning (Hancock 1991: 31, 42). In Arilja it is a maker of the progressive whereas in Džambaz it does not appear to have a grammaticalized function, although it is generally avoided in subordinate clauses, especially after the subjunctive marker te (Jusuf and Kepeshi 1980: 104–105). Arilja and Džambaz are also differentiated in their use of jetating and rounded versus non-jetating unrounded first person singular aorist markers, e.g. Arilja /kerdum/, 'I did' (AOR) ~ Džambaz /kerdem/, 'I did' (AOR), which reflect the variation in the first person singular present of, 'be', whence they are descended: sijum ~ sem, 'I am'.

In Albanian, there is significant variation in the formation of pluperfects and in the shape of the participle, which is essential for a variety of analytic constructions. Gëg has a series of compound perfects (perfects composed of a perfect or pluperfect auxiliary plus the participle) which are used to express what I have called preterior taxis (Friedman 1981), i.e. a past resultative event prior to another past resultative event, as in kam pasë skhuke literally, 'I have having gone', kisha pasë skhuke literally, 'I had having gone'; both

21 The desinence of the first person plural synthetic aorist/imperfect is subject to additional variation since original /xme/ gives /xme/ in the West, which becomes /xme/ in most dialects due to analogical pressure from the first singular /-i/ and second singular /-i/. In Bitola-Ressen there is additional dissimilation to /-i/ in scattered parts of the West /-me/ is preserved.

22 This is part of a larger phenomenon of Gëg /-i/ Tosk /j/ (cf. Byron 1976: 99 – 102).

23 Southern Tosk also has /j/ for the third singular subjunctive but it has /j/ for the second singular subjunctive (Gjini 1970: 64–66).

24 There is also considerable variation in the shape of the final syllable /ja, je, te, te, tam, ta ~ ja, je, te, te, tam, ta/ (Gjini 1989: 248).

25 In the dialects of the Balkans, the future is formed by means of a particle, usually ka or kum (< kame, want) plus the present.
translatable as 'I had gone' but with nuances of great distance in time. The following example illustrates this concept:

(2) Eshë vetinja brengë ... shqiptor (AOR) më qartë ai, pasi kishte folar (PLUP) ... pas një vajë ... me të cilën e kishin pastë fevuar (CPD PLUP) prindet qesh në femijët.

It is the only trouble he said (AOR) more clearly, after he had spoken (PLUP) about a girl ... to whom his parents had engaged him (CPD PLUP) in childhood. (Friedman 1981: 278)

Tosk would simply use a pluperfect in both instances. In the participle, Tosk has a single, non-inflected form as in shkuar 'going.' Also, in dialects where the aorists of the auxiliary verbs 'be' and 'have' are lost, the pluperfects formed with those aorist forms in the standard language and other dialects are replaced by the imperfect (Basha 1989).

In Macedonian the old perfect in 'be' plus the verbal-form competes with the new perfect in 'have' plus the newer verbal adjective, e.g. sum dosto 'I came/I have come' - ivam dojdeno 'I have come.'26 The nuances of meaning attached to the old perfect vary from perfect in the East, where the new perfect is lacking, to pure nonconfirmative (reported, admisive, dubitative) in the Southwest (for details see Friedman 1988 and also Miščeska-Tomík 1988).

In Romani, there is competition between synthetic imperfect and pluperfect in Džambaz and analytic equivalents in Arlja using the present and aorist followed by the third person preterite of 'be,' e.g. keravas = kerava sine 'I was doing', kerđamas = kerđam sine 'I had done.' There is also variation in the expression of aspect. In Albanian, as noted above, the aorists of the verbs 'be' and 'have' are excluded from some dialects, which in turn affects the formation of pluperfects. In Macedonian the imperfective aorist forms showing degrees of marginality, which in turn affects the contexts in which the imperfect can appear: the more restricted the imperfective aorist, the broader the contexts in which the imperfective imperfect occurs. In general, loss of the imperfective aorist is more common than geographic. Although it is less common in some western dialects, it is especially uncommon or totally absent for speakers of the youngest generation. Those dialects without the imperfective aorist will also not use the imperfective aorist when they do not form the I-participle. Moreover, the free-standing perfective non-past is characteristic of the East as opposed to the West. There is considerable variation in the grammaticalization of aspect in Romani as a whole - e.g. in the use of Slavic or Germanic type prefixation - but within the context of Macedonia the main source of aspectual variation is in the use of the final -a of the present tense noted above.27

26 For intransitives, there is also a perfect construction using 'be' plus agreeing verbal adjective, e.g. dojden sum 'I have arrived,' but this is a synthetic collocation and not paradigmatic (cf. Friedman 1977: 16-17).

27 The Džambaz marker is also descended from the that dialect's third singular imperfect of 'be' sur, but the marker is now inchoational.

28 Prefixes do occur in Arlja, e. g. dotërdë 'arrive' [3 SG AOR] < do-...up to' (Slavic) + dol 'go.'

Morpho-syntactic differences between Gëg and Tosk in analytic verbal constructions include the following: Gëg infinitive of the type me shkuar (= 'with' + long participle) Tosk për të shkuar (= 'for' + subordinator + participle) to go;29 the Gëg use of 'be' as the auxiliary of the perfect of intransitive active verbs as opposed to the Tosk generalization of 'have' for all active verbs (both dialects use 'be' for the perfect of medio-passive verbs), e. g. Gëg jam shkuar 'I am gone' but Tosk kam shkuar 'I have gone'; the Gëg future of the type Gëg këm me shkuar ('have' + infinitive) 'I will go,' as opposed to the Tosk type do të shkoj (particle based on 'want' + subordinator + present subjunctive verb) 'I will go,' the Gëg progressive of the type jam këh shkoj (present tense of 'be' + 'from' + present tense verb) 'I am going' as opposed to the Tosk type po shkoj (progressive marker + present indicative verb) 'I am going' or jam duke shkuar (present indicative of 'be' + 'while' + participle) 'I am going,' and the formation of gerunds Gëg ëmë shkuar 'while going' Tosk duke shkuar 'while going.'

Romani shows variation in the forms of 'be,' e. g. Arlja sijm - Džambaz sem 'I am.' Likewise in the preterite Arlja sine 'I was' (IMPF). - Džambaz sa' 'I was' (IMPF). There is also tremendous variation in shape of the participle and aorist. The standard rule is stems in -i, -a, -v, take -'e, those in -d take -in, and others take -'e, e. g. ker- 'do' > kerd 'done', čumid- 'kiss' > čumidino 'kissed', be- 'sit' > bëlo 'seated'. Among the many sources of variation is the fact that the dental stop can be generalized, however, and thus forms such as čumidino 'kissed,' and bešo 'seated' (with progressive voicing assimilation) also occur.

In Macedonian, aside from the variation in present tense first singular endings mentioned above, there is significant variation in stem classes. In the dialects of Polog, Skopje, Veles, Mačësce, and in parts of Kičevo and Poleče there are the three thematic conjugations characteristic of the literary language: -a, -i, -e, e. g. gëna 'book,' nos 'carry,' kate 'tell' (all 3 SG PRES). Elsewhere there are only two thematic conjugations, -e and -i, and conjugations merge in a variety of patterns (see Elkon 1983).

2.2. Nominals

Both Macedonian and Romani show variation in the treatment of animate or inanimate objects. Macedonian and Albanian have variation in the shapes of their case endings. Macedonian and Romani show variation in the shape of oblique personal pronouns (as does Albanian, but for purely phonological reasons), and all three languages show some variation in their reflexive possessives. All three also have variation in the shape of the definite article.

In Macedonian, there is variation in the facultative oblique forms of personal nouns and in analytic as opposed to synthetic dative pronouns. The oblique form is a Westernism accepted into Literary Macedonian. It is always facultative and is limited to masculine proper and family names, kinship
terms ending in a consonant, -i, -o, -el, and the nouns čovek 'person', bog 'god', gavel 'devil', and gospod 'lord'. Oblique forms for nouns denoting domestic animals are now dialectal. Nouns in a consonant, -el, or -i take -al, nouns in -eI add -t: brat 'brother', bratija 'brother' (OBL), tatkce 'father', tatak 'father' (OBL), Gorgi 'George', Gorgija 'George' (OBL), Blaže 'Blaže', Blažeta 'Blaž' (OBL). These forms can occur wherever an oblique pronoun would occur:

(3) Mu rokov na Ivanu pred Blažeto. 
     him-DAT said-1SG-AOR to Ivan-OBL in-front-of Blaž-OBL.
     'I said to Ivan in front of Blaže.'

(4) Ee go Gorgija. 
     Behold-there him-ACC George-OBL.
     'There's George.'

The Western dialects are characterized by the use of synthetic dative pronouns, e.g. nam 'to us' while the Eastern dialects — like Bulgarian — use the indirect object marking preposition na 'to' plus the accusative, e.g. na nas 'to us'.

In Albanian the remnants of the locative in -i are dialectal feature in Northern Tosk and Southern Gëg (Gjini 1970: 77, Byrom 1976b: 112).

In Romani there is variation in the use of the animate accusative:

(5) a. dikkja argijia. (Jusuf and Kepeski 1981: 33) 
     saw-3SG-AOR field-hand-NOM-PL.
     'He saw some field hands.'

(5) b. dikkja argi žiajia. (Jusuf and Kepeski 1981: 32) 
     saw-3SG-AOR field-hand-ACC-PL.
     'He saw some field hands.'

In most dialects, the accusative is only used for animate objects and pronouns while inanimate substantives are left in the nominative. However animate objects in the nominative and inanimate objects in the accusative also occur:

(6) Dikela jekh but dàndio gau. (Punon 1981: 79) 
     saw-3SG-IMPF one-NOM very knowing non-Rom-NOM.
     'He saw a wise non-Rom.'

(7) Gele jekh vàdke nelt. (Gillett-Smith 1912: 85) 
     went-3PL-AOR into one-OBL forest-ACC.
     'They went into a forest.'

In terms of the category of definiteness, all three languages show variation in the shape of their respective definite articles. In Macedonian, there is variation between the three-way definite opposition in the West (and in the literary language) and the single article of the East. The tripartite distinction mirrors the same distinction in demonstratives, viz. unmarked /-i/-proximate /-v/-, distal /-n/-, e.g. decata 'the children' (NEUTRAL), decava 'the children' (PROXIMATE), decava 'the children' (DISTANT). In the codification of Literary Macedonian, the eastern shape of the masculine definite article (/o*/u/) in the West it has the form /-a/ was integrated with the three-member system of the western dialects, so that the literary language has a new compromise system without the introduction of artificial grammatical distinctions. 30

In Albanian the main variation is the generalization of /i/ as the definite article for all masculine nouns in Kosovo (or those in /h/) but not in /-ki/-gi/ in Macedonia (normally masculines in /h, k, g/ take /i/) (Gjini 1991: 254—56), e.g. bilbil/bibili ñightingale/the nightingale" but zog'zo gu bird/the bird", krai/krahi, wing, side/the wing side", but in Kosovo zog 'the bird", krai 'the wing, side' and in Macedonia zogu 'the bird', krahi 'the wing, side'. Albanian also has a difference in the indefinite article (Gjeg na 'one, a', Tosc një 'one, a'), which although phonologically based, was the subject of a potential, artificial morphological distinction. 31

Romani dialects, too, have a variety of definite article systems, the two major distinctions being between the presence or absence of initial /l/ in oblique forms on the one hand and the degree of syncretism on the other. These problems are not yet solved, e.g. Jusuf and Kepeski (1980) mix the Arlja and Dzambaz systems. Four representative systems (based on Jusuf 1990) are given in Table One.

<table>
<thead>
<tr>
<th>Maced</th>
<th>Roman</th>
<th>Dzambaz</th>
<th>Arlja</th>
<th>Gurbet</th>
<th>Kalderas</th>
</tr>
</thead>
<tbody>
<tr>
<td>masc nom sg</td>
<td>fem nom sg</td>
<td>masc nom sg</td>
<td>fem nom sg</td>
<td>masc nom sg</td>
<td>fem nom sg</td>
</tr>
<tr>
<td>1E</td>
<td>O</td>
<td>O</td>
<td>I</td>
<td>E</td>
<td>LA</td>
</tr>
<tr>
<td>all other</td>
<td>all other</td>
<td>all other</td>
<td>all other</td>
<td>all other</td>
<td>all other</td>
</tr>
<tr>
<td>masc obl sg</td>
<td>fem obl sg</td>
<td>masc obl sg</td>
<td>fem obl sg</td>
<td>masc obl sg</td>
<td>fem obl sg</td>
</tr>
</tbody>
</table>

In the expression of reflexive possession, Macedonian and Albanian show interesting parallel variation. Tosk lacks a morphologically distinct reflexive possessive pronoun, whereas Gëg has i véet, 'one's own' for the third person, and this form has been adopted into the standard language. Macedonian 30

30 The Bulgarian situation is relevant to this discussion and is described in note 31 below.
31 Folk Kônica proposed that Gëg një be used for feminine nouns and Tosk një for masculine, but his proposal was not accepted (Byrom 1976a). In the codification of Literary Bulgarian, which is older than the codified standards of the languages of Macedonia, the one-member article system was chosen in connection with the northeastern dialectal base (three-member systems are restricted to the Rhodopian dialects and a small pocket around Trnëv near the Serbain border), but an entirely artificial distinction was created in order to incorporate both shapes of the masculine definite article (/o*/u/ and /-a/), which have a complex dialectal distribution. It was declared that the form in a consonant would be used in nominative functions and that in a vowel in oblique functions, despite the fact that no such distinction occurs in any Bulgarian dialect.
onian has a reflexive possessive svoj 'one's own', which can refer to any person. Romani has a variety of forms based on the stem p(e)-, 'one's own' which are also limited to third persons. The use of i već, 'one's own' in some Tosc writers for non-third persons appears to be a hypercorrection (Byron 1976b: 118). In Macedonian, svoj 'one's own' has taken on nuances of emphatic rather than purely reflexive, and so unlike elsewhere in Slavic, where svoj is obligatory in reference to a third person subject, Macedonian svoj is actually avoided in contexts where it might appear too emphatic (cf. English 'He came to the party with his own wife'), which is closer to the Tosc model. In the Macedonian dialects of Romani the reflexive possessive appears to be used regularly as in Gag, although elsewhere this may not be the case (cf. Paspati 1870: 70–71).

3. SYNTAX

Of the many types of syntactic variation, two that are comparable among the three languages being considered here are word order in genitive phrases and object reduplication, i.e. the repetition of a direct and/or indirect object by means of a clitic pronoun agreeing with it in gender, number and case:32

(8) a. Mu ja davam na momčeno knjiga.
him-DAT it-ACC-FEM give-1SG-PRES to boy-DEF book-DEF-FEM
'I gave the boy the book.' (Macedonian)

(8) b. Ja dhashë djali libarin.
him-DAT it-ACC give-1SG-AOR boy-DEF-DEF book-DEF-ACC
'I gave the boy the book.' (Albanian)

3.1. Word Order in Genitive Phrases

With regard to genitive word order, both Albanian and Macedonian have the inherited type head + genitive, whereas Romani has the type

genitive + head. Apparently under Turkish influence the Albanian dialects of Thrace (as well as their descendants spoken on the Black Sea littoral) display a unique genitive-head word order (Dcrzavin 1934, Sokolova 1983):

(9) a. jani jana she plaka mima
    Jani Janaki-GEN old mother
    'Jani Janaki's old mother'

(9) b. mima e plake e
    mother FEM-NOM old
    FEM-GEN Jani Janaki-GEN
    'Jani Janaki's old mother'

Macedonian has both orders, and Golab (1960: 41) identifies the genitive + head order as a calque on Turkish:33

(10) a. majka mu na carot
    mother him-DAT to king-DEF
    'the mother of the king' (Macedonian)

(10) b. na carot majka mu
    to king-DEF mother him-DAT
    'the king's mother' (Macedonian)

(10) c. sultanin anes
    king-GEN mother-3SG-POSS
    'the king's mother' (Turkish)

In keeping with its postpositional origin, the Romani genitive normally precedes its head, as is the case in Turkish as well as the languages of India. South Balkan Romani dialects, however, permit considerable variation, quite probably due to the influence of the other Indo-European Balkan langua-

32 Another source of variation in the syntax of the languages of Macedonia is the influence of Serbo-Croatian. Macedonian clitic order differs from that of Serbo-Croatian in, among other things, the fact that Macedonian does not allow the verbal subordinator da to be separated from the finite verb except by a small group of clitics, whereas Serbo-Croatian da can be separated by clitic phrases. In Macedonian technical writing, however, such Serbo-Croatian influenced constructions are common (Koceni 1991). Albanian has a paradigmatic medio-passive that is lacking in Slavic, which often uses a construction of the type he + verbal adjective for passive constructions. In the Albanian print media of Kosovia and Macedonia, constructions of the type he + participle are often used as present passives under Serbo-Croatian influence, when they are used only as the medio-passive perfect. According to Korubin (1951) Serbo-Croatian has similarly influenced Macedonian, which is more likely to use the intransitive marker se + finite verb and limit he + verbal adjective to syntactic perfects. Cf. also the concern with Slavic influence expressed in an article on syntactic errors in the Albanian-language press due to the influence of du ca gje do pe some neighboring languages' (Flaka e vellaztimit 14: XII: 86–17).

33 On the other hand, the Turkish dialects of south western Macedonia show the opposite order due to Slavic and Albanian influence:

(i) 
    familiati
    family-3SG-POSS man-GEN
    the family of the man

This Slavicized syntax is also to be found in the Gagauz dialects at the other end of the Balkan Turkic speech area in Dobrudja and Moldavia:

(ii) a. gečnist
    past-PERF-PART life-3SG-POSS-ACC person-PL-GEN
    past-perfect part of life of a person

b. insanlari
gennist
    person-PL-GEN
    the past-life of people

Such dialectal variation demonstrates, among other things, that the processes of Balkan linguistic contact move in various directions, that they are motivated by diverse sociolinguistic, demographic and structural factors, and that they can give diametrically opposed results in different dialects. Popov (1979: 248) also attributes the genitive-head construction in Bulgarian to Turkish influence although his examples are all from proverbs or poetry:

(iii) Na mališ xobat ne se gleda
    Of man-DEF beauty not REFL watch
    One doesn't pay attention to the beauty of a man
DIALECT VARIATION AND QUESTIONS OF STANDARDIZATION IN MACEDONIA:...

In some Balkan dialects, however, such genitives are possible, e. g. Mohadžer from Prishița (Ramadani 1989: 350):

(14) Isto iš an-o chitá ē Indiae
same is in-the languages-NOM the-OBL India-GEN
'It is the same in the languages of India.'

One of the goals of standardization is precisely the comparison of choices and compromises in dialectal variation, which have differing degrees of significance. In the foregoing example of variation in genitive word order, the processes of standardization reflect dialectal distribution. Thus the order genitive-head is completely excluded from Standard Albanian (just as the reverse order – head + genitive – is rejected by the Literary Turkish), i. e. in both norms variation in word order in genitive phrases is entirely unacceptable because of the marginality of the dialects in which variation occurs. On the other hand, the Macedonian (and Bulgarian) literary standards allow both possibilities, although the calqued word-order is more characteristic of colloquial style. In Roman, the influence of the Slavic head + genitive construction has resulted in variation from the inherited genitive + head construction so that now both are possible. It remains to be seen how these two constructions will be integrated in to Literary Roman, but at present it appears that the borrowed word order (head + genitive) is more characteristic of literary style.

3.2. Reduplication

The use of anaphoric or cataphoric clitic pronouns triggered by syntactic objects is another feature subject to variation in all three languages. In Macedonian it is completely grammaticalized, i. e. obligatory, in the case of definite direct and most indirect objects. As one moves further east, however, grammaticalization weakens, and in Bulgarian is reduced to pragmatics. In other words, while reduplication is required by Macedonian grammar in certain contexts, Bulgarian grammar never requires the use of this syntactic device but rather permits its use to signal discourse functions such as focus. The phenomenon is also highly restricted in neighboring Serbian dialects, where it is limited to expressing emphasis with personal pronouns (Bečić 1908: 293). Reduplication in Macedonian is also triggered by personal pronouns when emphasis is desired. The numeral eden ‘one’ functioning as an indefinite article denoting specificity can also trigger object reduplication, especially colloquially (cf. Kazazis and Pchentchev 1975):

(15) a. Barav edna marka no ne najdov seek-1SG-IMPF one stamp but not find-1SG-AOR
'I was looking for a stamp but didn't find one.'

(15) b. JA barav edna marka no ne ja najdov it-ACC seek-1SG-IMPF one stamp but not it-ACC find-1SG-AOR
'I was looking for (a specific) stamp but didn't find it.'

There is also variation in the shape of the Romani genitive. The basic shape of the suffix is -ku, where k represents either k, š, or š depending on the head of the genitive (i.e. for MASC NOM SG, š for FEM NOM SG, š for elsewhere, although š is generalized for all nominative plurals in Arlja; cf. above on the definite article), e. g. raskleskari daj ‘the boy’s father’ NOM SG, e raskleskari daj ‘the boy’s mother’ NOM SG, e raskleskari umudi ‘the boy’s friend(s),’ umali, amulan, ‘the boy’s friend(s)’ ACC SG, NOM PL, ACC PL. The chief variant is without the š/ and its preceding vowel, e. g. namesko eski, eski ‘the Rom’s of the Rom.’ Other variants drop only the vowel that precedes ši or še precede ši with invariant ši. The variant without šiš is standard in Vlah dialects but also occurs in the Non-Vlah dialects of the Balkans. The long genitive is favored in Arlja, whereas Džambaz favors the short genitive but can use both. The form and order of the genitive are two of the many questions that have not yet been definitively decided in the process of codification, as can be seen in Jusuf and Kepeski (1983), which utilizes many orthographic, phonological, morphological and syntactic variants without choosing those that would be unique in a fully codified standard (for details see Friedman 1985).

Note that since ‘house’ is in the ablative — an alternative form of expression possession (probably calqued on west Macedonian possessives using the preposition of from) — the genitive can be either short or long.

24 There is also variation in the shape of the Romani genitive. The basic shape of the suffix is -ku, where k represents either k, š, or š depending on the head of the genitive (i.e. for MASC NOM SG, š for FEM NOM SG, š for elsewhere, although š is generalized for all nominative plurals in Arlja; cf. above on the definite article), e. g. raskleskari daj ‘the boy’s father’ NOM SG, e raskleskari daj ‘the boy’s mother’ NOM SG, e raskleskari umudi ‘the boy’s friend(s),’ umali, amulan, ‘the boy’s friend(s)’ ACC SG, NOM PL, ACC PL. The chief variant is without the š/ and its preceding vowel, e. g. namesko eski, eski ‘the Rom’s of the Rom.’ Other variants drop only the vowel that precedes ši or še precede ši with invariant ši. The variant without šiš is standard in Vlah dialects but also occurs in the Non-Vlah dialects of the Balkans. The long genitive is favored in Arlja, whereas Džambaz favors the short genitive but can use both. The form and order of the genitive are two of the many questions that have not yet been definitively decided in the process of codification, as can be seen in Jusuf and Kepeski (1983), which utilizes many orthographic, phonological, morphological and syntactic variants without choosing those that would be unique in a fully codified standard (for details see Friedman 1985).

25 There is also variation in the shape of the Romani genitive. The basic shape of the suffix is -ku, where k represents either k, š, or š depending on the head of the genitive (i.e. for MASC NOM SG, š for FEM NOM SG, š for elsewhere, although š is generalized for all nominative plurals in Arlja; cf. above on the definite article), e. g. raskleskari daj ‘the boy’s father’ NOM SG, e raskleskari daj ‘the boy’s mother’ NOM SG, e raskleskari umudi ‘the boy’s friend(s),’ umali, amulan, ‘the boy’s friend(s)’ ACC SG, NOM PL, ACC PL. The chief variant is without the š/ and its preceding vowel, e. g. namesko eski, eski ‘the Rom’s of the Rom.’ Other variants drop only the vowel that precedes ši or še precede ši with invariant ši. The variant without šiš is standard in Vlah dialects but also occurs in the Non-Vlah dialects of the Balkans. The long genitive is favored in Arlja, whereas Džambaz favors the short genitive but can use both. The form and order of the genitive are two of the many questions that have not yet been definitively decided in the process of codification, as can be seen in Jusuf and Kepeski (1983), which utilizes many orthographic, phonological, morphological and syntactic variants without choosing those that would be unique in a fully codified standard (for details see Friedman 1985).
Thus emphasis in pronouns and in marking specific indefinite direct objects are subject to variation in terms of signaling by reduplication. There is also variation due to the omission of the reduplicated pronoun in imperatives and in narratives with significant historical distancing (Ugrinova-Skalovska 1960–61).  

(16) a. Odi, kati na majka si  

go-IMP tell-IMP to mother self-DAT  

'Go tell your mother.'  

(16) b. Odi, kati i na majka si  

go-IMP tell-IMP her-DAT to mother self-DAT  

'Go tell your mother.'  

(17) a. Otsod, i katal na majka si  

went-3SG-PERF and told-3SG-PERF to mother self-DAT  

'He went and told his mother.'  

(17) b. Otsod, i katal na majka si  

went-3SG-PERF and told-3SG-PERF to mother self-DAT  

'He went and told his mother.'  

The variants without reduplication are especially characteristic of folk speech and are not encountered in modern literature even in contexts where folk speech is being imitated.

The rules in Albanian are similar to Macedonian, but Albanian permits more variation in the use of the anaphoric pronoun. Thus, for example, it is possible to emphasize a subordinate clause by reduplication (Buchholz and Fiedler 1987: 442), which is not possible in Macedonian:  

(18) a. Unë e dija se do so vonohet  

I-it-ACC knew-1SG-IMP that FUT SUBJUNCT be-late-3SG  

'I knew that he would be late.' (Albanian)  

(18) b. *Go znae... deka ke se sadoci  

I-it-ACC knew-1SG-IMP that FUT INTR be-late-3SG  

'I knew that he would be late.' (Macedonian)  

On the other hand, Albanian can omit the pronoun if there is lexical focus (Buchholz and Fiedler 1987: 443), whereas Macedonian does not:  

(19) a. Në mbledhje kritikuan, sidomos drejtoria  

At meeting criticized-3PL-AOR especially director-DEF-ACC  

'At the meeting they especially criticized the director.' (Albanian)  

(19) b. *Na sotanokos kritikuan osobeno direktorot  

at meeting-DEF criticized-3PL-IMP especially director-DEF  

'At the meeting they especially criticized the director.' (Macedonian)  

In Standard Albanian reduplication is used for emphasis with personal pronouns, grammatically for definite or quantified direct and all indirect objects (Ejntrej 1982: 62–63). Specific direct objects can also trigger reduplication (Buchholz and Fiedler 1987: 442), especially in Kosovia and Macedonia, but this is not the norm (Gjinari 1970: 77, cf. Ajeti 1978: 275):  

(20) Athërrë po e dreth  

then PROG it-ACC roll-3SG-PRES 1 one cigarette  

cigare 'Then I'm rolling a cigarette, too.'  

Albanian also has so-called pleonastic reduplication where the clitic does not refer to an object — with a limited class of verbs that take indirect objects, although in some cases the pleonastic reduplication is facultative:  

(21) a. Ia hëni  

him-DAT it-ACC mount-3SG-AOR horse DEF-ACC  

'He mounted the horse.'  

(21) b. I hëni kati  

him-DAT mount-3SG-Aor horse-DEF-ACC  

'He mounted the horse.'  

Pleonastic reduplication can also be used for lexical differentiation (Buchholz and Fiedler 1987: 445–46):  

(22) a. la dha gazat  

him-DAT it-ACC give-3SG-AOR smile-DEF-DAT  

'He burst into laughter.'  

(22) b. la dha besten  

him-DAT it-ACC give-3SG-AOR faith-DEF-ACC  

'He gave him his word.'  

(22) c. I dha best  

him-DAT give-3SG-AOR faith-INDEF-ACC  

'He believed him.'  

In Romani, reduplication of verbal objects in general is facultative (Jusuf and Kepaöki 1980: 10):  

(23) a. E Romans dende lën po  

the-OBL-PL Romans-ACC give-3PL-AOR them-ACC on-the  

anav Egyptians  

'In France they called the Romans Egyptians...'  

(23) b. Kì Franceja e Romans dende po  

in-the France the-OBL-PL Romans-ACC give-3PL-AOR on-the  

anav Boimeni  

'In France they called the Romans Slovenians.'  

This is not a feature that has been specifically addressed in language codification documents.

---

36 Cf. also (Hancovc 1992: 14) which shows that Romani allows both reduplicated and unreduplicated pronouns, and if unreduplicated both short and long (clitic) forms as in the following examples with dijkhet ‘see’ (3 SG PRES), man ‘me’ (ACC) and man ‘me’ (ACC CL): man(n) dijkhet; dijkhet man(n); dijkhet man man; man dijkhet man all meaning ‘he sees me’.
4. LEXICON

In all three languages, the major sources of variation on the lexical level are Turkish, Slavic, neologisms together with new borrowings, and competition between the standard and the dialectal base of the aspeaker. In this section, I shall consider the roles of Slavic and Turkish in contributing to lexical variation.37

Unlike Albanian and Romani, Macedonian is itself a Slavic language, as is Serbo-Croatian. Thus, in the context of Macedonian, what is meant by a Slavic source of lexical variation is essentially competition from Serbo-Croatian. This competition is seen as a serious problem in Macedonian because the two languages are so closely related, and because Serbo-Croatian has been the dominant language or so-called lingua communis of Yugoslavia (the official language of communication in the army, predominant in the mass media, etc.). Also, the prestige of the dialect of the Macedonian capital, Skopje, which is located at the border of the northern Macedonian dialects that are transitional to Serbian, has added to the effect of Serbo-Croatian competition.

In the Romani dialects of Macedonia, Macedonian represents a source language for borrowings and for glosses on neologisms. Thus, for example in a poster urging Romans to vote in the referendum on Macedonian independence of 8 September 1991, the Indo-neologisms nijamija 'right', sansara 'peace', sansaralo 'peaceful', and rasita 'state' as well as the use of tromalipa 'acts of daring/courage' to mean 'freedoms' (a Vlax Romani term of Greek origin) were all glossed in parentheses with their Macedonian equivalents (prava, mir, mirno, drzava, slobođi, respectively).38 Slavic is also the source of calques, e.g. a vazi 'vote' (Turkish avaz, voice, shout), < Persian owatz 'voice') based on Slavic glas meaning both 'voice' and 'vote'. Although exposure to Hindi through popular films available on video cassette may hasten the process of adoption, at present Slavic must still be used to gloss neologisms and it remains to be seen how successful the new vocabulary will prove.

In Albanian, recent Slavic borrowings are especially characteristic of colloquial style, e.g. the exclamation of a frustrated shopkeeper in discussing new tax laws:

The word for 'state' is based on Macedonian (and Serbo-Croatian) država 'state' rather than Standard Albanian shtet 'state'. Even in Macedonia, however, recent Albanian borrowings from Slavic often reflect Serbo-Croatian rather than Macedonian, e.g. a headline in the Macedonian Albanian-language tri-weekly Flaka e vellacërimit (17: VI; 90–9): Baku i komunizmit 'The bogeyman of communism' (Serbo-Croatian bauk, Macedonian bauè, Standard Albanian bogol). Serbo-Croatian is also the source of calques in colloquial style, e.g. jëfikisht 'intensely' (< jëfik 'life-long' < jëf 'life') based on Serbo-Croatian živo 'vital', intensely', Standard Albanian me gjithë zemër 'with all heart'.39 The reasons for this are twofold. Firstly, Albanian has been under the same pressure from Serbo-Croatian as Macedonian in terms of the status of Serbo-Croatian in Yugoslavia. Secondly, the majority of Albanian speakers in Yugoslav lands live in Kosovo, whereas many Albanians have emigrated to Macedonia, and in Kosovo the dominant Slavic language is Serbian. Thus the Albanian of Kosovo is influenced by Serb and in turn influences the Albanian of Macedonia. To this can be added the fact that the largest concentration of Albanians in Macedonia is in the north, where the Macedonian dialects are transitional to Serbian (cf. Ismaili 1988: 18).

Another source of variation in the Albanian of Macedonia centers around the use of toponyms. Most toponyms in the Balkans have different names in each of the languages of the peoples that have occasion to refer to them. In some cases the differences are rooted in the phonological history and structure of the respective languages, e.g. Macedonian Skopje, Serbisk Skopje, Albanian Shkup, Turkish Üsküp, Arumanian Scopia, all from Latin Scupi. In other cases the name is quite different due to translation or calquing as in the case of Slavic Crna Gora, Albanian Mal i Zi, Turkish Karadağ 'Black Mountain' (English Montenegro), similarly Macedonian Bitola (from Slavic obelite, monastery) but Turkish and Albanian Manastır, from Greek Monastirion. Sometimes the toponyms have separate histories as in Macedonian Tetovo and Albanian Teto, but Turkish Kalkandelen. In the belief that choice of toponym is connected with claims of sovereignty and suffering from a sense of territorial insecurity, Macedonian authorities have attempted to require the use of Slavic toponyms in non-Slavic language textbooks and sometimes in other printed media since 1982 (cf. Flaka e vellacërimit 14: IV; 91–13, 16: IX; 90–13, 14: VI; 89–12).40 Albanians

---

37 Lexical items from dialects not serving as the base of the standard are incorporated for purposes of vocabulary expansion in all three languages. Both Romani and Albanian show significantly greater degrees of dialectal lexical differentiation than Macedonian. The coining of neologisms and borrowing of so-called international words is a significant process in all three languages. Macedonian and Romani are differentiated from Albanian in that the former have attested relatives (other Slavic and Indo languages, respectively) that can also serve as sources of vocabulary, whereas Albanian has no close living relatives and the ancient languages of the Balkans to which Albanian is related are too poorly attested to be of any use in lexic expansion.

38 A comparable phenomenon was the use of Turkish to gloss Slavic terms in Macedonian texts of the nineteenth century, when Turkish was the state language and Macedonian was struggling for literary status. It is noteworthy that Turkish even penetrated the realm of Christian religious terminology, which, given the identification of Turkish with Islam, should have been the most impervious to such influence, e.g. Kurilija Savor (Turk. kurut, 'sue') for Slavic svastil (Jasmaris Nasteva 1970).

39 Excluded from consideration here are old loans such as many verbs in -tr, e.g. shëtit 'stroll', which are common to all of Albanian.

40 A series on this subject by Qemal Murati appeared in the May 1990 issues of Flaka e vellacërimit.

41 This same type of insecurity occurs elsewhere in the Balkans. Thus, for example, Greek nationalists will become terribly upset on hearing the Slavic names Lenin and Kostur used for the towns known in Greek as Flovna and Kastoria, which are in Greek Macedonia but which were predominantly Slavic until their Hellentization during the course of this century (a process that is not yet complete). There was also an incident recently when the Turkish Consul in Komotini, the capital of Greek Thrace, used its Turkish name, Gumuchlence, in a document (Dialekti June, 1994, p. 4).
thought among some Macedonian intellectuals during the first years of codification (post-1944) that maintained that Turks should be encouraged and preserved because they were characteristic of folk speech and also emphasized Macedonian's differentiation from the other Slavic languages. Still, in many ways, the situation is the reverse of what it was forty years ago. Thus, for example, a year after the official recognition of Literary Macedonian, Konesi (1967) wrote an article in which he criticized a translation of Montel's Le Tartuffe for being full of Turkisms. It is true that these are in part a feature of the style chosen by his translators, but in part they reflect a recognition of the reality that the majority of the population in which the vocabulary is already so well-rooted. The use of a literary language means that the writer can use the language of the people in a way which is not immediately perceptible to the reader.

(25) „Toa znáči... da go sništ... istučeniet poetički jezik na Molitvi... da nivoeto na našoto balkansko, karščko, češtnico musubitite.“ (Konesi 1967: 16) It means lowering the refined poetic language of Molitve to the level of our Balkan, small-town, marketplace chitchat.

Three and a half decades later, Korubin (1981b) found it necessary to admonish translators not to become so carried away by purity that they translate Serbo-Croatian Turkmens such as Bujnim efendi. At your command, sir with the native but here inappropriate Povele gore dolupine.

The situation in Albanian is essentially the same as that in Macedonian. There has been a significant puristic movement, especially in recent years, that has sought to eliminate as many foreign elements as possible from Albanian, and Turks have been a particular focus of that campaign.

In Albanian, as in Macedonian, Turks are characteristic of colloquial style and also used for pejorative, historical flavor, and local color. The Albanian alphabet in Macedonia, however, does not seem to be favoring a return to Turks, although they are still quite common in the media when a colloquial or emotive effect is desired. Rather, they seem to be increasingly used in the use of G in serious contexts. Thus, for example, an article in Flaka e vëllezëritës (7; VII: 91—98) entitled „Dhëti e vërtet, pyetje nëse të përitë. „Day and night, sleeplessness and waiting" describing the anguish of parents waiting at the Macedonian Red Cross in Skopje for news of their children sent to the front in the Serbo-Croatian war, one of the parents was quoted in Gëg.

Albanians' own ethnonym Shqipar, but in South Slavic it has become a term of abuse (Cf. Skendi 1980: 52 cited in Byron 1985: 67, also Flaka e vëllezëritës 5: IX: 90—6, 3: II: 91—12) in a process comparable to the use of Polek for Pole or Yid for Jew in English. It has begun to appear again, however, in Slavic media (e. g. Nova Makedonija 19: VIII: 91—9). The Albanian derogatory term Çifet Jew (a Turkism, ultimately from the Hebrew Yhidim, 'tribesmen of Judah') has likewise reappeared (Flaka e vëllezëritës 14: IV: 91—2), the standard term being Ebrej. So far, Rom is still used for Romans, except when translating titles that originally contained the lexical equivalent of Gypsy (Macedonian Cigan, Albanian Evjig). The Albanian derogatory term for Slav, Shqak, has also not occurred in the media to the best of my knowledge.

Interestingly enough, Qemal Murati in one of his series of articles on calques in Albanian (Flaka e vëllezëritës 13: V: 90—5) recommends using the Turkism yli, 'fireman' rather than native hekur, 'iron' (both terms used in modern and other meanings, the former on the grounds that the use of fire to mean 'iron' is a calque on the Romance languages.

It could be argued that this quotation was used for emotive purposes, but if so, it clearly shows that in such stressful situations Gëg is felt to have more emotional power.
One difference between the Albanian of Albania and that of the Yugoslav lands, especially Macedonia, but also parts of Kosovo (e.g., Prizren) is the significant presence of Turkish-speaking population in the former and its absence from the latter (cf. Németh 1961), although the recent opening of Albania and movement of populations across the border with Macedonia has brought reports of a Turkish minority. Thus, for example, Sotir Sterjoški, a Macedonian refugee from Albania in Ohrid reported to the Macedonian Turkish-language paper Bitik (16: III: 91–6) that there are Turks in the Elbasan, Saranda, and Fier regions who still speak Turkish but who have no schools. Another Macedonian refugee, Esmeralda Skrbik reported that she had Turkish neighbors. The Geg dialect columns of Flaka evëllazërimi in Tërk is that are not found in the Standard Albanian dictatories, e.g. mënunpleased‘(Turkish memnun, Standard Albanian i kënaqur).

In the southern Balkan dialects of Romani, Turkisms still comprise a significant portion of the vocabulary since Romani has remained for a longer time strictly in the realm of oral communication. Thus, although the actual number of Turkisms in Romani is no greater than that in any other colloquial Balkan language (cf. Friedman 1989), the colloquial base of Literary Romani leaves open the possibility that a greater number of Turkisms will find their way into the codified version. Turkisms are extremely rare in the Romani dialects spoken outside the boundaries of the former Ottoman Empire, however, which means that the fate of Turkisms in Literary Romani will depend in some degree on the extent to which the Romani dialects of the Balkans serve as the basis for the currently emerging Romani literary language (Friedman 1989). Romani, however, is in a different position from the rest of the Balkan languages in this respect. Upon leaving the Indo-European group the Romans came into direct contact with Persian and borrowed a number of important words. Some of these Persian words are found throughout the Romani dialects and also in Turkish, which undoubtedly reinforced their retention in the Balkan Romani dialects, e.g., baxt ’luck, happiness’, others occur in Balkan Romani and Turkish but not in all Romani dialects, e.g. lâf (Turkish lâf) ’word’ (English Romani lâv but elsewhere vorba, svato, dama, thavati, etc.) Although the conscious attempt to eliminate Turkisms (or Arabo-Persianisms) in the Balkan literary languages has resulted in their stylistic lowering or marginalization, the same motivations need not apply to Literary Romani. Romani language planners have fewer nonlinguistic reasons for eliminating words of Turkish origin. As was mentioned, many of these words are ultimately of Persian origin and therefore more closely related to Romani than to the other Indo-European languages of the Balkans and some were borrowed directly from Persian during the early history of the Romani people. In addition to this, the elimination of Turkisms from the other Balkan languages was in part motivated by political independence from the Ottoman Empire. In the case of Romani, the Turkish language does not occupy a similar political position as distinct from other non-Romani languages.

5. CONCLUSION

As was indicated in the introduction, this survey of salient features of variation in Macedonian, Albanian, and Romani by no means exhausts the inventory of features in the individual languages, or even those that are comparable among them. Rather, the goal here has been to present a selective survey of comparable features that suggests both similarities and differences in the relation of dialect to standard in the literary languages of Macedonia and also indicates directions for future research.

As Byron (1976b: 118–20) points out for Albanian, standardizers have tended to favor Tosk over Geg and innovative phenomena over conservative in making dialect choices, to which can be added Gjinari’s (1989: 267–86) observation that in some cases the relative extent of a phenomenon has influenced its selection. This same observation appears to hold true for Macedonian.

Thus, for example, in both languages standard phonology is firmly rooted in the selected dialectal base (West Central Macedonian and Northwestern Tosk Albanian), and in both cases this base represents considerable innovations. Nonetheless there have been compromises, e.g., in the treatment of intervocalic /l/ in Macedonian and of final voiced consonants and the selection of stress alternates in Turkish loan words in Albanian. There are also areas where prescription is meeting resistance such as antepenultimate stress rules in Macedonian phraseological units or the Kosovian insistence on the use of length in their literary variant of Albanian. In Romani in Macedonia the tendency is to favor the Arilja dialect, which in certain respects is phonologically innovating (e.g., loss of intervocalic /l/ in inflections). Romani language codifiers outside of Macedonia, however, have shown a tendency to favor phonological conservatism (cf. Cortiade Forthcoming).

In morphology, both Macedonian and Albanian have made conscious efforts to integrate forms from outside the region serving as the dialectal base. Here, too, innovative forms seem to be favored, as is the case with the Macedonian third person singular present marker -a as opposed to the Western -i/ or the Albanian first person singular present marker i/-j as opposed to North Tosk i/-nj/. On the other hand, Macedonian has integrated the Eastern (and older) shape of the masculine definite article with the Western (and newer) tripartite distinction on the basis of the relatively broad range in which these two phenomena occur, whereas Albanian has not integrated dialectal variation in the shape of the definite article because the variants were too marginal. Variation in the Albanian indefinite article was also excluded, but out of consistency with the dialectal base. These morphological phenomena demonstrate a correlation between the time of codification and the incorporation of variants. In Bulgarian, where the fixing of standard norms was achieved earlier, an artificial grammatical solution was codified in to the language (see note 31), whereas in Albanian, which achieved a unified standard at a later date, a choice was made and only one form entered the standard, although the Geg indefinite article is still in common use by Geg speakers even in formal situations. The Macedonian
standard achieved integration without artificial distinctions, while in Romani the process of selection is still underway.

On the lexical level, Macedonian and Albanian have had in common the attempt to restrict or marginalize Turkisms, a solution that tends to break up the linguistic unity of the Balkans. A recent tendency toward colloquialization, however, may ultimately alter this direction. To this can be added the fact that Macedonian and the Albanian dialects of the Yugoslav lands also work on resisting pressure from Serbo-Croatian. Literary Romani is showing similar puristic and neologizing tendencies, but for the time being is still close enough to its colloquial base to have many Turkish loans and is still young enough that it must rely on Slavic for glosses. In syntax, it can be seen that both Romani and Macedonian use variations in genitive word order to distinguish colloquial and formal levels, while Albanian does not due to the marginality of the variants. Reduplication, however, shows variation in Albanian, Macedonian, and Romani, and while Macedonian and Albanian both have codified norms relating to it, actual usage shows more variation than is permitted by the norms. The question has yet to be definitively addressed in the context of the codification of Romani.

At present all three languages also display distinct differentiation in the use of the standard and of dialects. Both Macedonians and Albanians use the standard in formal and public contexts and their own dialects at home and informally. For both languages, those with less education have less control of the standard. For Albanians, however, the difference between the standard and the dialects is greater than for Macedonian, and the boundary between informal and formal in Albanian, i.e. between use of Gëg and use of the standard, seems to be broader in favor of Gëg, which until 1968 was the standard in Yugoslavia, while the same type of boundary for Macedonian dialects is more restrictive. In the case of Romani, the development of the standard is still in such an early stage that this type of differentiation is not clearly developed. For all three of these languages, their location within Yugoslavia has had crucial impact on the relation of dialectal differentiation to codification. In the context of current socio-political changes taking place in the Balkans, the relationship of dialect to standard in the languages of the Balkans appears to be changing as well and merits considerable future study.

University of North Carolina, Chapel Hill

REFERENCES


Diaspora and Questions of Standardization in Macedonian...


Popov, Konstantin. 1979. Štěrbenec bulgarskij sintaksis. Sofija: Nauka i izkustvo.


САДРЖАЈ

ЧЛАНЦИ И РАСПРАВЕ
 Victor A. Friedman: Dialect Variation and Questions of Standardization in Macedonia, Albanian and Romani .......................... 7
 Марта Биетич: Српскохрватско челибарнићко .................. 37
 George Thomas: Convergent Trends in the intellectual Lexicon of the Slovene, Croatian and Serbian Standard Languages in the nineteenth Century .................... 47
 Нада Арсенијевић: О поетским и егзистенцијалним removeca sa glagolom iman i njihovim madarskim еквивалентима .................. 61
 Снежана Петровић: Историја и стање процјењивања турскиха у српскохрватском језику .............................. 71
 Александар Младеновић: Неке српскохрватске особине у рукописима дијака Симона. (Друга половина XVI века) .................. 129
 Александар Младеновић: О једном дечанском запису из 1690. године .................. 137
 Александар Младеновић: О писму патријарха Арсенija III загребачком епископу Игнацију 1691. године .......................... 145
 Вера Јерковић: Морфолошки однакоје динака рукописа Косовски бој .................. 151
 П. А. Димитриев: Из набројеног над синтаксисом српскохрватског језика .................. 169
 Wayles Browne: Variation in Serbo-Croatian noun phrase constituents: declinable and indeclinable items ................. 173
 Милош Ковачевић: Распоред клеуза у независној режици .................. 181
 Асим Идо: Из деклинационе проблематике босанско-херцеговачких говора .................. 207

КРИТИКЕ И ПРИКАЗИ
 Milorad Radić: O svima knjigama Milorada Papovca .................. 233
 Данко Шинка: O глаголима — прелимина .................. 238
 Danko Šipka: Novi habštinski časopis .................. 240
 Vera Vasić: Svenska Savić, Diskurs analiza .................. 242