DIALECTAL BASE, VARIATION AND CODIFICATION OF THE BALKAN LITERARY LANGUAGES: ALBANIAN, BULGARIAN, MACEDONIAN AND ROMANY

by Victor Friedman

We may say that the language league created by the intensive contacts between the languages of the Balkans during the period of the Ottoman rule—and both before and after that—practically represents an association of dialects. General characteristics can be found in the languages themselves, i.e. as collective entities, and these characteristics have certainly attracted a great deal of attention since the times of Kopitar (1829), Mikloshich (1884), Papahaji (1908), Seliščev (1925) and Sandfeld (1930) up to the present day. Nevertheless, if we analyze these features in their specific dialectal details, we shall find that the picture is much more complex. We would certainly not like to fail to see the wood because of the trees, as the saying goes, but, in order to go beyond the well-known generalizations, we should, first of all, make a thorough analysis of dialectal data before proposing new concepts. Thus, for instance, in Thracian Albanian dialects (as well as in their descendents around the Black Sea) we can find a special word order of the genitive-core type, for example Jani Janakít pláka mëma, i.e. “Jani Janaki’s old mother,” instead of the strictly defined core-genitive word order in the Albanian language (as well as the more usual noun-adjective word order), i.e. mëna e plaka e Jani Janakít (Derdzavini 1934, Sokolova 1983). Comparing Macedonian structures with Turkish equivalents, e.g. Ha yapot majka my with Sultanin annesi, Golab (1960) concludes that this word order in the Macedonian language is an imitation of Turkish. Popov (1979; 248) has come to the same conclusion about the Bulgarian
language, quoting examples like Na mъжът хубост не се гледа. On the other hand, in south-western Turkish dialects there is the opposite phenomenon, i.e. the word order with the genitive coming after, e.g. *familiaris Adānam*, literally “family of man’s,” obviously influenced by Macedonian and Albanian (Katona 1969). On the other side of the Balkano-Turkish lingual territory, i.e. in the Gagauz dialects in Dobruja and Moldavia, we also find such Slavicized syntax, e.g. *gečmiš jašamasy insamaryn*, literally “the past life of men’s” (Покровская 1979; In literary Turkish *insanların gečmiş yaşaması*). Such dialectal variation shows, among other things, that the processes of lingual contacts have moved in different directions and that they have been stimulated by various sociolinguistic, demographic and structural causes, and may give opposite results in various dialects of the same language.

One of the tasks of standardization is precisely the comparison of choices and compromises among such dialectal variations. Here, too, there are more important and less important distinctions. Thus, as far as our example with the word order in the genitive constructions is concerned, the processes of standardization have reflected the distribution of dialects. So the genitive-core word order is entirely excluded from standard Albanian in the same way that the opposite core-genitive word order is not accepted in standard Turkish, i.e. in both languages variations in such syntagms are not allowed, perhaps because of the marginality of the respective dialects. On the other hand, in Macedonian and Bulgarian, although both possibilities are accepted in the standard language, the imitated word order is more characteristic of the colloquial style.

I have so far mentioned several literary languages in the Balkans which have achieved a certain degree of standardization. I shall later examine the differences among them in this respect. But, before doing that, as regards the question of the word order in the genitive constructions, I would like to present the situation in the Romany language. Unlike the rest of the aforementioned languages, the status of the so-called genitive in the Romany language is rather controversial, and, besides, shows a great degree of dialectal variation. Let us first briefly examine the opposing arguments concerning the question of whether the Romany genitive is a case suffix, postposition or derivative suffix for adjectives (cf. Friedman, in print).

Taking into account our examples from the Skopje Arli dialect, we may say that the basic arguments for each point of view are the following. The possibility that it is an adjective is supported by the fact that the genitive agrees with the core: -koro if it is masculine, nominative and singular, -kiri if it is feminine, nominative and singular, but -kere in all other instances, e.g. in e romeskoro dad ‘the Romany’s father,’ e romeskiri daj ‘the Romany’s mother,’ e romeskere chave, dadestar, čhavendar... ‘the Romany’s children, of Romany’s father, of Romany’s children...’ etc.

Those who maintain that the genitive, just like any other case suffix in the Romany language, is a postposition and not a case, are under the influence of the diachronic situation. In all Indian languages, including Romany, almost all the old Sanskrit cases have been lost, and postpositions have taken their place. Starting from the new Indian languages on the Indian subcontinent, some linguists think that, just like in those languages, a kind of postpositional structure has also been retained in Romany.

Against these two points of view, i.e. in favour of an analysis of the Romany genitive as a case suffix, we can quote the following arguments:

1. In front of the Romany genitive there is always the kosven (indirect) form of the definite article and other modifiers (namely e, in contrast to o nominative (for masculine singular) or i (for feminine singular), for instance o rom gelo ‘the Romany came,’ but e romeskoro dad gelo ‘Romany’s father came.’ This is not the case with simple adjectives, for instance o lačho rom gelo ‘the good Romany came.’

2. The genitive suffix, like all other case suffixes, has grown into the kosven base of the noun, for instance: rom ‘Romany’ nominative romes- kosven, while adjectives, like all other derivatives, derive from the nominative, for instance:
   - manus-es-koro ‘to the man,’ but
   - manus-ano, -ikan ‘man’s’ or ‘manly’
   - manus-ipe ‘manhood’
   - manus-oro ‘little man’
3. The genitive can be used with non-attributive functions; for instance, as an object of preposition: bi akale phurane gadengoro ‘without these old clothes’ (Cortiade, in print).

4. There are also many arguments in favour of the assertion that the former postpositions in the Romany language have been transformed into agglutinative case suffixes. Here we shall quote only two of them. Firstly, from the point of view of the language structure, modern Indian languages are almost purely postpositional. i.e. they have no prepositions at all, except for two or three loanwords from Arabic or Persian, while in the Romany language, as opposed to case suffixes, there are only prepositions, and not postpositions. Even loan postpositions have been transformed into prepositions, for instance the Turkish iki aydan sonra, literally ‘Two months later,’ has become in Romany sono duaj čonendar (Messing 1988). So, case suffixes represent a closed, paradigmatic whole. Secondly, from a morphophonological point of view, the case suffixes act as flexive affixes, not as clitics. Thus, for instance, at the borders with clitics, we note regressive assimilation in voicelessness, and only in those cases when it is a distinctive voiceless, while in the flexion we have progressive assimilation in voicelessness, no matter whether it is distinctive or not, for instance les-da ka have ‘I shall eat it, too’ is pronounced [lez-da ka have], while we have an-fa mange ‘Come on, bring me’ without changes. But in the past participle we have, for instance: su-to ‘asleep,’ šun-do ‘heard,’ simultaneously with romni-tar ‘from the Romany woman,’ romni-nitar ‘from the Romany women.’

Another two important complications in connection with the Romany genitive are its dialectal diversity and its place in relation to the core. The main variant of the Romany genitive is the compressed form, without the syllable closing by -r, i.e. romesko, -eski, -eske. The place of the genitive is usually in front of the core, as we have already seen in the given examples, but we can also find it after the core, for instance: o kher me phralskoro ‘the house of my brother.’ In the southern Balkan Romany dialects, which are very conservative in general, we see very great variations in the order of the genitive constructions.

---

*There are other long variants, e.g. -ep-, -p-, but they are not relevant to our aim.*
BALKAN FORUM

Bulgarian lingual territory, as well as on the Macedonian, we may find, among other things, opposition between one-article and three-article systems, and also between the systems with the basic masculine definite article ending in a consonant (namely, -r) or a vowel (i.e. -o, -a, etc). In the Bulgarian literary language, the codification of which is the oldest, the one-article system was chosen as a result of the choice of the north-eastern dialectal base, but an entirely artificial difference was created with the introduction of two geographically distant forms of the masculine definite article, namely, the one with a consonant was proclaimed as nominative, and the one with a vowel as oblique, although there is no such difference in any dialect (Friedman 1976). A similar artificial difference was proposed by Faik Konica in the initial stages of the formation of the Albanian standard language, i.e. the Gheg një to be used with nouns of feminine gender, and the Tosk një with nouns of masculine gender (Byron 1976). But his proposal was not accepted, and they are now used in accordance with the respective dialectal base. In the codification of the Macedonian literary language, the eastern form of the masculine definite article (-r) was integrated with the western three-article system, so that, in the modern Macedonian literary language, we have a newly-formed compromise system, which nevertheless makes no artificial differences. In Romany dialects there are also various article systems where the main questions concentrate around the presence or absence of the initial l- in the oblique forms and the shape of the article or articles in the nominative plural (whether it should be the same as the nominative singular or as the oblique article and whether it should differentiate gender). These questions have still not been resolved, for instance in Jusuf and Kepeksi’s Grammar (1980), the Arli and Jambaz systems are mixed.

There are certainly many other processes and problems connected with the relations between the dialects of the Balkan language league and the formation of the Balkan literary languages which might lead to various typological analyses. For example, all the codified Balkan languages show a tendency towards the elimination or at least limitation of the so-called Turkicisms (in the Turkish language the same words have often been rejected because of their Arabic or Persian origins), a tendency common to all Balkan codifications, which simultane-

Victor Friedman: Dialectal Base, Variation and Codification 

ously brings about a disintegration of the Balkan language league. In the southern Balkan dialects of the Romany language, since they have existed for a long time as a means of oral communication only, there are still many so-called Turkicisms, but the same ones are not present in the northern Romany dialects. Thus, their fate in the standard Romany language will perhaps depend on which dialectal base is accepted as basic, a process which is still going on (Friedman 1989). In the relations between orthography and dialects, particularly in the Albanian, Macedonian, Romanian and Romany languages there are some general problems, for instance, the place and the presentation of the schwa vowel [a], as well as the palatal sounds, all of which have led to different solutions in each codification.

Thus, a parallel analysis of the different solutions and variations in the standardization would lead to a typological scheme with crossed criteria and measures. Taking into account four of the languages of the Balkans, namely Albanian, Bulgarian, Macedonian and Romany (i.e. the languages spoken in the central part of the Balkan Peninsula, which are already codified or are in the process of standard codification), we can distinguish the following elements of that typological scheme. From a typological and methodological point of view, if we accept genitive constructions and article diversity as arguments, we can see correlations between the time of codification and the way of integration of dialectal forms. As regards the Albanian genitive variation, as well as the Bulgarian article variation regarding the number of articles, we see the complete exclusion of the phenomena beyond the regional borders of dialectal bases, while, in both languages, there has been an inclination towards artificial integration of various article forms located on the territories of the dialectal bases. Such a solution was accepted in the earlier fixed Bulgarian language, while in the Albanian standard only one form was allowed at a given period of time. In the Macedonian standard there was integration without artificial differences, while in Romany the development is still under way.

On the other hand, the general orthographic and lexicographic problems have produced – perhaps even intentionally – solutions leading to the disintegration of the Balkan language league.