

3/19/08

THE BALKAN LANGUAGES

By Brian D. Joseph, The Ohio State University

[FROM: International Encyclopedia of Linguistics, ed. by W. Bright (1992) Oxford University Press, Volume 1, pp. 153-155. (Revised version in 2nd Edn., 2003 (ed. by W. Frawley).)]

The Balkan peninsula, the rugged mountainous landmass in southeastern Europe bounded by the lower Danube and Sava rivers to the north (though by some accounts beginning even further north, in the vicinity of Vienna), the Adriatic Sea to the west, and the Black Sea up to the mouth of the Danube river to the east, and including all of Greece to the south, is home to a number of languages whose interrelationships present numerous elements of interest for questions of language contact, language change, areal linguistics, and sociolinguistics. The languages in this area come from several different language branches: from Indo-European come Albanian (two major dialects, Geg in the north and Tosk in the south), Greek, the South Slavic languages Serbo-Croatian (especially the southernmost dialects, often called Torlak), Bulgarian, and Macedonian, the Romance (modern Italic) languages Daco-Romanian, Aromanian (or Vlah), and Judeo-Spanish (also known as Ladino), and the Indic language Romani (the language of the Gypsies, particularly those dialects spoken in the Balkans), and from the Altaic family comes Turkish.

The most significant fact about these languages is that various sets of them share certain structural and lexical features that do not, in the case of the Indo-European languages, derive from their being genetically related. Such features do not represent common inheritances from Proto-Indo-European, but rather result from linguistic convergence over a period of intense intimate contact among the speech communities in this area. Moreover, genetic relatedness

3/19/08

would in any case not be a possible explanation for any similarities between the Indo-European languages in the Balkans and Turkish. The languages sharing significant numbers of these features are often designated "Balkan languages", as opposed to the merely geographic designation (see Schaller 1975) "languages of the Balkans" (which includes as well Slovenian, Armenian (spoken in Bulgaria), Circassian (spoken in the Kosovo area of Yugoslavia), Ruthenian (spoken in the Vojvodina area of Yugoslavia), and German and Hungarian (both spoken in Romania)).

Moreover, this fact of structural and lexical convergence has led to the characterization of the Balkan languages as a "Sprachbund" (also: "linguistic league" or "linguistic area"). The recognition of this fact, first noted as early as 1884 by Franz Miklosich, but systematically elaborated on in the classic 1926 work of Kristian Sandfeld (more widely known in its 1930 French translation), is based on the belief that the convergences observable in a comparison of these languages are not chance similarities but instead are the natural consequence of the close contact these speech communities have had with one another.

The most significant relevant features shared by various of these languages are as follows, covering phonology (a-b), morphology (c-f), syntax (g-i), and lexicon (j):

a. the presence of a (stressed) mid-to-high central vowel; this feature is found in Albanian, Romanian, Bulgarian, some dialects of Macedonian and Serbo-Croatian, some Romani dialects, and Turkish.

b. the presence of i-e-a-o-u in the vowel inventory without phonological contrasts in quantity, openness, or nasalization; this feature is found in Greek, Tosk Albanian, Romanian, Macedonian, Bulgarian, Serbo-Croatian, and Romani.

3/19/08

c. a reduction in the nominal case system, especially a falling together of genitive and dative cases; this feature is found in Greek, Albanian, Romanian, Bulgarian and Macedonian (though note that the latter two have eliminated other case distinctions as well).

d. the formation of a future tense based on a reduced, often invariant, form of the verb 'want'; this feature is found in Greek, Tosk Albanian, Romanian, Macedonian, Bulgarian, Serbo-Croatian, and Romani.

e. the use of an enclitic (postposed) definite article, typically occurring after the first word in the noun phrase; this feature is found in Albanian, Romanian, Macedonian, Bulgarian, and Torlak Serbo-Croatian.

f. the use of verbal forms to distinguish actions on the basis of real or presumed information-source, commonly referred to as marking a witnessed/reported distinction but also including nuances of surprise (admirative) and doubt (dubitative); this feature is found in Albanian, Bulgarian, Macedonian, and Turkish, and to a lesser extent in Romani, Serbo-Croatian, and Romanian (the presumptive).

g. the reduction in use of a nonfinite verbal complement (generally called an "infinitive" in traditional grammar) and its replacement by fully finite complement clauses (see Joseph 1983); this feature is found most intensely in Greek, Macedonian, Bulgarian, Serbo-Croatian (especially the Torlak dialects), and Romani, but also in Albanian (especially Tosk) and Romanian.

h. analytic adjectival comparative structures; this feature is found in Greek, Albanian, Romanian, Bulgarian, Macedonian, and Romani, as well as in Turkish.

i. the pleonastic use of weak object pronominal forms together with full noun phrase direct or indirect objects; this feature is found in Greek, Albanian, Romanian, Bulgarian, and Macedonian, dialectally in Serbo-Croatian, and to a limited extent in Romani.

3/19/08

j. lexical parallels, including shared phraseology (e.g. a phrase that is literally "without (an)other" meaning 'without doubt'), and numerous shared loan words many of which are from Turkish.

These features--known as "Balkanisms"--define the Balkan Sprachbund, even though, as the above catalog indicates, not all are found in all the languages (see also Hamp 1979). Most of these features represent synchronically valid statements about the superficial structure of these languages, but several of them are diachronically oriented, thus reflecting the view that one of the significant facts about the Balkan languages is their movement away from stages in which they were typologically more different from one another towards the present state of structural convergence after centuries of contact.

The causes of the linguistic convergence in the Balkans must undoubtedly be sought in language contact, although some linguists have viewed various developments in each language as independent of the situation found in the other languages. The nature of the contact leading to the convergence is a matter of some controversy, with some scholars opting for a substratum explanation drawing on facts about languages, e.g. Thracian and Illyrian, spoken in the Balkans in ancient times (see Katicic 1976), others for an adstratum explanation with the influence of Greek being especially important, and still others for different types of contact models, e.g. based on bilingualism or even a form of pidginization. Many details concerning the social situation in the medieval period, when the most significant contact occurred among these languages, are still needed before the matter can be fully understood.

3/19/08

References and Additional Bibliography

- Assenova, Petja (ed.). 1983. *Linguistique balkanique: bibliographie 1966-1975*. Sofia: Bulgarian Academy of Sciences.
- Friedman, Victor. 1981. Admirativity and Confirmativity. *Zeitschrift für Balkanologie* 17.1.12-28.
- Feuillet, Jack. 1986. *La linguistique balkanique*. (Cahiers balkaniques No. 10). Paris: INALCO.
- Hamp, Eric P. 1979. Linguistic areas or clusters? *Quatrième congrès international des études du sud-est européen: Abrégés des communications*. Ankara: Türk Tarih Kurumu, pp. 282-283.
- Joseph, Brian D. 1983. *The Synchrony and Diachrony of the Balkan Infinitive. A Study in Areal, General, and Historical Linguistics*. (Cambridge Studies in Linguistics, Supplementary Series No. 1). Cambridge: Cambridge Univ. Press.
- Katicic, Radoslav. 1976. *Ancient Languages of the Balkans. Part One*. The Hague: Mouton.
- Miklosich, Franz. 1884. *Die türkischen Elemente in den südost- und osteuropäischen Sprachen (Griechisch, Albanisch, Rumunisch, Bulgarisch, Serbisch, Kleinrussisch, Grossrussisch, Polnisch)*. Vienna: Denkschriften der Kaiserlichen Akademie der Wissenschaften, Philosophisch-historisch Klasse, 34, pp. 239-338.
- Sandfeld, Kristian. 1930. *Linguistique balkanique. Problèmes et résultats*. (Collection linguistique publiée de la Société de Linguistique de Paris 31.) Paris: Librairie ancienne Honoré Champion. [First published in Copenhagen in 1926 as *Balkanfilologien: en oversigt over dens resultater og problemer*.]
- Schaller, Helmut. 1975. *Die Balkansprachen. Eine Einführung in die Balkanphilologie*. Heidelberg: Carl Winter Universitätsverlag.
- Schaller, Helmut. 1977. *Bibliographie zur Balkanphilologie*. Heidelberg: Carl Winter Universitätsverlag.
- Solta, Georg R. 1980. *Einführung in die Balkanlinguistik mit besonderer Berücksichtigung des Substrats und des Balkanlateinischen*. Darmstadt: Wissenschaftliche Buchgesellschaft.