Reprinted From
Anthropological Linguistics
Fall 1983 (Vol. 25)

LANGUAGE USE IN THE BALKANS: THE CONTRIBUTIONS OF HISTORICAL LINGUISTICS

Brian D. Joseph
The Ohio State University

Historical linguistics is the branch of linguistics concerned primarily with the changes languages undergo and the reasons for these changes. The causes of language change generally fall into one of two categories,[1] those that are language-internal, i.e., involve responses to factors which are entirely internal to one linguistic system, and those that are language-external, i.e., involve contact between two or more independent linguistic systems. While there are numerous instances of purely system-internal developments within the languages of the Balkans,[2] it is the external causes of language change which are of most relevance to the question of what historical linguistics can contribute to the issue of language use in the Balkans.

The reason for this is fairly obvious: from the earliest times, the Balkans have been an important crossroads region, and numerous groups of peoples have moved into (and out of) the area, absorbing, or being absorbed by — linguistically as well as socially and culturally — already existing groups. This constant process of absorption has left its mark on the various languages currently and previously in the Balkans, so that it is possible for one to identify different layers of language contact; these layers, in turn, allow one to draw conclusions about shifts in Balkan language use from prehistoric times on.

A few comments about contact-induced language change are necessary, though, before some specific cases from the Balkans can be examined profitably. It is commonly thought that certain linguistic subsystems, for example the inflectional morphology of a language, are less likely than others to undergo contact-induced change,[3] and it is certainly true that the effects of language contact can be seen especially clearly in some areas of a language, most notably the lexicon in the form of loan-words and place-names, and only less so in others. However, it seems that no component of a linguistic system is impervious to language-contact influence. Thomason (1981) has shown numerous cases of contact change in all components of a language and has demonstrated that differences in the likelihood of contact change have more to do with the conditions under which the contact takes place than with any intrinsic properties of particular linguistic components.[4]

Accordingly, in looking into language contact in the Balkans, one must be prepared to examine all aspects of the various languages without preconceptions about which types of changes might be due to contact. The examples presented below

bear this out, for they reflect contact-induced change in several different linguistic components; all of them, moreover, allow for important inferences to be drawn concerning language use in the Balkans at different periods in history and prehistory.

As noted above, the lexical component of a language is the area in which influence from outside languages shows up most readily. Since virtually no language is free from contact with other languages in some way or another, the vocabulary of any particular language will include many nonnative words often from several different linguistic stocks. In the case of one ancient Balkan language, for example, namely Ancient Greek, it is possible to identify likely foreign vocabulary elements from Semitic (e.g., χιτών *tunic*, κυπάσσις *linen tunic*, σαγήνη *large drag net*), Anatolian (e.g., κυανός *dark-blue enamel, lapis lazuli*, κύμβαχος *helmet*) and Iranian (e.g., ἀγγοπηνία *honeycomb of bees*).[5] These loan words, however, tend to represent *cultural borrowings*, i.e., the borrowing of the foreign name for a particular foreign cultural item when the item itself is borrowed,[6] and so could well have entered Greek through chance contact with some peoples who happened to have had the item and the name for it. As such, they are not so valuable as a way of gaining insights into language use. However, other aspects of apparent foreign vocabulary in Greek are more useful.

It has long been recognized that there are numerous words in Greek which are most likely of nonnative origin, but for which no apparent source language can be readily identified. As noted by Katičić (1976:53), in his summary of the literature on the ancient languages of the Balkans, one group of these words, those showing such suffixes as -νθ- or -σσ-, includes many toponyms and terms which "usually denote plants, animals, and material objects that belong to the Aegean flora, fauna, and material culture"; some examples are the toponyms Κόρινθος, Τίρυνς (genitive Τίρυνθος), and Κνωσσός, and words like μίνθη *mint*, ὄλονθος *wild fig*, κυπάρισσος *cypress*, and νάρκισσος *narcissus*, among others. The usual interpretation is that these words reflect a language spoken by the inhabitants of the area which is now Greece prior to the comings of the Greeks into the region.[7] Similarly, a second group of words, with certain characteristic phonological (e.g., aspirates where native Greek words have unaspirated stops) and morphological (e.g., the occurrence of certain apparent prefixal elements) features but of no definable semantic sphere, is also to be found in the Ancient Greek vocabulary, e.g., ἀλείφ-ω *anoint* (cf. native Greek λίπ-ος *fat*), φέλλ-ευς *stony ground* (cf. native Greek πέλλ-α *stone*), θεράπ-νη *dwelling* (cf. native Greek τέραμ-νον *dwelling*). These too seem to reflect a pre-Greek linguistic layer which may or may not be distinct from the -νθ-/-σσ- language noted above.[8]

What inferences might be drawn from these facts and interpretations about language use in the prehistoric Balkans, specifically in the southernmost region? First, from the evidence

of toponyms, it is clear that a change in language use occurred with the coming of the Greeks. Second, from the evidence of cultural words like μίνθη , etc., it seems that the incoming Greeks accepted autochthonous terms for new items they encountered; thus the coming of the Greeks to Greece, from a linguistic standpoint, was not unlike the coming of the Western Europeans to North America (see Footnote 6). Third, instead of it being the case that the autochthonous elements were simply pushed out of Greece, some must have been absorbed into the Greek tribes and thus have undergone a shift in their language use to Greek; still, they must have been present in sufficient numbers to effect the entry into Greek proper of numerous autochthonous words (άλείφω , etc.).[9] Thus the linguistic evidence discussed here gives a fairly clear picture of certain aspects of prehistoric language use and language shifts in the ancient Balkans.[10]

The previous example drew mainly on evidence concerning language contact at a lexical level. But, as already mentioned, contact-induced change can affect other areas as well. One that is on the borderline between lexical and phonological in that it involves layers of lexical items with a particular phonological shape is the status of the sounds ts and dz in Modern Greek; moreover, this example sheds some light on an interesting aspect of present-day language use in Greece.

It has been argued elsewhere (Joseph 1982a, 1982b) that the sounds ts and dz[11] have a special status synchronically in Modern Greek in that they mainly function "allolinguistically" in the overall language system. *Allolinguistic* is a term proposed by Wescott (1975) for all aspects of language which are "alienated from conventionally structured speech" (p. 497), and it refers specifically to elements like conventionalized child language forms, slang, onomatopoetic and generally sound-symbolic words, diminutives, expressive and affective vocabulary, and the like. In Greek, the sounds ts and dz occur quite frequently (though not exclusively) in just these types of lexical items. For example, the following forms, constituting just a small sampling of items with ts/dz, give an indication of the range of occurrence of these sounds: the diminutive suffixes -ίτσα , -ίτσι for nouns, -ούτσικος for adjectives,[12] independent hypocoristic diminutives like Μήτσος (from Δημήτριος), child language forms like τζίσια (with a variant τσίσια) *peepee*, τσιτσί *meat*, τσάτσα *aunty*, onomatopoetic words and words derived from noises such as τσάκ *sound of something breaking* and τσακίζω *I break* or γράτς (with variants κράτς/χράτς) *sound of scratching* and γρατσουνάω *I scratch*, expressive words like τζάμπα *for nothing, free* and τζουτζές *dwarf; ugly immoral person*, and a number of sound symbolic words such as τσιμπῶ *pinch*, τσιμ(π)ούρι *tick, pest*, τσιβίκι *tick*, τσουκνίδα *nettle*, where the initial τσV- appears to signal a stinging, biting sensation or figurative extension thereof. What is especially significant about the distribution of ts and dz is that the vast

majority of their occurrences in Greek lexical items is in such allolinguistic words; no other sounds in Greek have such a lexical distribution, i.e., one finds, for instance, a [k] in allolinguistic forms (e.g., in κράτς and the diminutive suffix -άκι) but it is not the case that [k] occurs primarily in allolinguistic forms, while this is so for ts and dz. Thus, one is justified in singling out ts/dz for allolinguistic status in Greek based on the range of lexical forms they occur in.

The question of language contact in connection with this special status of ts/dz comes into play in two ways. First, many of the words with these sounds are borrowed from neighboring languages; these include both allolinguistic words such as τζάμπα , from the Turkish caba ([ǰaba]) *thrown into the bargain*, and τζουτζές from Turkish cüce ([ǰüǰe]) *dwarf*, and also nonallolinguistic words such as παπούτσι *shoe*, from Turkish papuç *shoe* ([papuč] , ultimately of Arabic origin). Furthermore, even these nonallolinguistic loan words with ts/dz function in such a way as to reinforce the special status of those sounds. That is, the Greek lexicon is stratified into words of different stylistic levels; in particular, a "higher style" and a "lower style" can be recognized,[13] and there are often pairs of roughly synonymous words which differ only in style, e.g., λίθος *stone* and πέτρα *stone*, λευκός *white* and ἄσπρος *white*. In the case of words with ts/dz, it turns out that where these words have synonyms in Greek, the ts/dz word is always low-style while the synonym is higher style; examples which bear this out include τζάμπα ~ δωρεάν *gratis, for nothing*, παπούτσι ~ ὑπόδημα *shoe*, and φάτσα ~ πρόσωπο *face*. The relative stylistic value of the individual members of these pairs enhances the special status claimed here for ts/dz, for again it seems that no other phonological units in Greek pattern in this same way. Thus loan words of all sorts containing the sounds ts and dz contribute to the allolinguistic status of these sounds.

The second way language contact is relevant here is more historical in nature. Marchand (1953:59) has demonstrated that tš and dž (spelled <ç>, <c>) occur in numerous words in Turkish which he terms generally "lautsymbolisch," including words for murmured and vibrating noises in the case of tš and words of "affective" origin in the case of dž. From Marchand's discussion, it is clear that these words are allolinguistic, in Wescott's sense. Although it is not the case that Turkish c̱ and c have allolinguistic status within their overall phonological system in the way that Greek ts and dz do, inasmuch as they occur in many ordinary nonallolinguistic words and thus have a different lexical distribution from the Greek sounds, still they do have some sound-symbolic value. Thus it is possible that the special status of ts and dz within Greek is at least in part due to outside influences, since corresponding sounds in Turkish, a language which had a considerable lexical influence, especially of what linguists call an *intimate* (i.e., noncultural) nature, over Greek within the past 500 years, have

a similar value.[14] This hypothesis, while highly speculative, receives some suggestive support from the fact that many of the sound symbolic words with tsV- designating stinging and the like seem to have developed from Ancient Greek words with initial k(C)i- sequences, by various sound changes or morphological reshapings,[15] e.g., τσιμ(π)οῦρι < κίμμυρος (Hesychius μικρολόγος *counting trifles, caring about trifles*), τσουκνίδα < κνίδη *nettle*, τσιβίκι < κίμβιξ *skinflint*. This suggests that the sequence k(C)i- may have had some sound symbolic value in Ancient Greek, so that the change to tsV- as the bearer of sound-symbolism for this lexical group would have occurred only in the Post-Classical era, i.e., possibly only after contact with a language in which similar sounds had a somewhat special status or occurred in words with a special status.

In any case, the status of ts/dz in Greek phonology raises numerous questions relevant to the issue of language use in one of the Balkan states, namely Greece. Moreover, historical linguistics, with its concern over borrowed versus native elements and over the development of individual native words, contributes much to the understanding of these questions.

The final example to be examined here involves a linguistic feature at the morphosyntactic level, and thus shows the same type of methodology applied to something other than the phonology or the lexicon of a language; in addition, it focuses on a different time period in the history of the Balkans. The linguistic feature in question is the complete or partial absence of the verbal category *infinitive*.

As is well known, the languages of the Balkans share, to differing degrees, a considerable number of linguistic features. These include the absence of a distinction between genitive (i.e., possessive) and dative (i.e., indirect object) nominal cases, the presence of a future tense formation based on a form of a verb meaning *want*, the presence of a mid-to-high central vowel such as [ə] or [ɨ], an enclitic rather than proclitic definite article, special formations for the numerals between 11 and 19, and the absence of the infinitive coupled with the use instead of finite verbal forms. Not all of the Balkan languages have all of these features — Greek for example, has a proclitic definite article and no mid-to-high central vowel — but there is enough common linguistic ground among these languages to be noteworthy,[16] and to raise speculation as to what the causes of this linguistic convergence might be. While it is not clear what has led to the convergence for all of these features, a scenario has been developed (Joseph 1981, 1983) for the absence of the infinitive which has interesting consequences for the question of language use in the Balkans.

First, though, the basic facts concerning the infinitive in the Balkan languages need to be sketched.[17] In Modern Greek, except for lexicalized, unproductive remnants of earlier infinitives, one finds the infinitive preserved, in a transformed

shape (i.e., with ending [-(s)i] as in γράψει *write* instead of Post-Classical -(s)ein as in γράψειν and Ancient Greek -sai as in γράψαι), only in the formation of the perfect tense system with ἔχω *have* (e.g., ἔχω γράψει *I have written*); this one stronghold for the earlier infinitive, however, is not sufficient to warrant setting up a grammatical category of infinitive, so that Modern Greek effectively is a language with no infinitive; finite subordinate clauses, with the verb often marked with the participle νά,[18] are used instead of infinitives. In both main dialects of Albanian, an analytic infinitive consisting of a particle — the preposition për in the southern dialect Tosk and the particle me in the northern dialect Geg — plus a verbal nominalization is to be found; however, the Tosk infinitive is somewhat limited in use (finite subordinate clauses are more common) and the analytic form in both dialects suggests a relatively recent innovation, so that at least Tosk and possibly Geg as well may have undergone a prehistoric[19] reduction of the domain of the infinitive in favor of finite complementation. The Balkan Slavic languages show the absence of an infinitive and the corresponding wider use of finite clauses to varying degrees — it is completely absent from modern Macedonian, and present only in traces in standard Bulgarian, being restricted to use as complement to only four predicates and only optionally at that; in Serbo-Croatian, the Torlak dialects of southeast Serbia show the complete absence of an infinitive, and in general the Štokavian dialects located for the most part in the eastern part of modern Yugoslavia, shows less use of the infinitive than the more westerly Čakavian and Kajkavian dialects. Finally, in Romanian, the infinitive survives in numerous nominalizations and in a few grammatical constructions (e.g., singular prohibitions, the conditional, two future tense formations) and as complement to a few verbs, but has given way to finite subordinate clauses even in these contexts and in others as well. Thus one finds a continuum in the Balkans regarding the infinitive, with Macedonian at one extreme and Romanian (and non-Štokavian Serbo-Croatian) at the other, with Greek, Bulgarian, Tosk Albanian and Geg Albanian in between the extremes, in that order. This distribution suggests that the loss of the infinitive is best taken as a central Balkan feature, since the peripheral languages, especially Romanian and non-Štokavian Serbo-Croatian,[20] retain the infinitive to the greatest degree. Moreover, to judge from the historical documentation available for these languages, infinitives are used to a greater extent in relatively early documents and less so in later ones, so that the loss of the infinitive can be dated to approximately the 10th to 15th centuries.[21]

The question that naturally arises is why the infinitive was lost in these several languages which are geographically clustered. Most explanations have sought to draw either on purely language-internal or purely language-external developments. As an example of the former reasoning, Togeby (1962)

believed that the phonological merger of the infinitive with a third person singular finite form was the underlying cause for the loss of the infinitive in Greek, Romanian, Bulgarian and Macedonian, while others have drawn on the existence of constructions[22] in which infinitives could alternate with finite clauses in Vulgar Latin (Barić 1961) and Koine Greek (Hesseling 1892) as the source of the infinitive-loss in Romanian and Greek, respectively. Among those who argued from purely language-external developments, Sandfeld (1930) saw the overall Balkan absence of the infinitive as the result of the influence of Greek on its neighboring languages, while others, such as Gabinsky (1967), have seen a prehistoric Balkan substratum with no infinitive[23] as the source of the infinitive-loss in each of the present-day Balkan languages, and still others, such as Rozencvejg (1969, 1976), have attributed this loss to the development of a "mediator" language, which effects a compromise between synthesis (efficient language production) and analysis (efficient language processing), in a bilingual contact situation such as would have arisen in the Balkans; the use of finite complementation, it is claimed, would tend to aid analysis and thus enhance communication in such contexts.

It seems clear, though, that neither type of explanation is wholly satisfactory in and of itself, for the language-internal developments generally cited either occurred too late or are in themselves not compelling enough to have caused the widespread demise of this category,[24] and the language-external accounts fail to take into consideration actual individual developments in each language which must have been contributory to the loss of the infinitive. Therefore the best explanation for this historical development within these Balkan languages seems to lie in a combination of language-internal factors with some degree of language contact.

In order to draw on both types of factors, though, one has to envision a central Balkan society, for example in Northern Greece, in the 8th to 12th centuries, which is emerging as a bi- or even multi-lingual society. The advent of contact among Balkan peoples would no doubt have led to some simplification along the lines posited by Rozencvejg (1969, 1976) increasing the frequency of finite complementation in the contact languages. In some of these languages, such simplification might have enhanced already-existing tendencies for replacing infinitives by functionally equivalent finite clauses, such as that noted by Barić (1961) and Hesseling (1892). At the same time, though, in this emerging bi- or multi-lingual society, second-language learning, partly of an imperfect nature, may be assumed to have been going on. Schumann (1979:56-57) has noted that second-language learners often go through a stage which is characterized in part by the use of the "unmarked form of the verb"; among such speakers, therefore, there would have been a tendency to use a single unmarked, and thus presumably non-finite, verbal form in main and subordinate clauses.

These conflicting tendencies, the one leading to wider use of finite forms and the other to wider use of nonfinite (invariant) forms, would have given a high degree of flexibility to the linguistic systems at this time. Various functionally equivalent but syntactically and no doubt stylistically distinct types of complementation would have been available under differing sociolinguistic conditions.[25] Thus this scenario for the loss of the infinitive requires one to make certain assumptions about the way various languages were used in the Balkans during the period in question. The society must have been a multi-lingual one in which speakers with widely different abilities in the target languages had to communicate with one another; moreover, speakers must have employed a number of competing strategies for such communication, resulting in complex code-switching and perhaps even the assignment of different values to these different strategies and to the different languages involved.

Without these assumptions, it is difficult to construct a plausible course of events which could have led to the present-day situation with the infinitive in the Balkans and in the Balkan languages.[26] Once again, then, methods of historical linguistics, in this case aimed at developing a coherent account of an historical development in several languages, have led to inferences about language use in the Balkans.

Somewhat deliberately, the examples discussed here — Pre-Greek lexical traces in Ancient Greek, the loss of the infinitive in the Balkan languages, and the status of the sounds ts and dz in Modern Greek — have been chosen from three different time periods within the Balkans — prehistoric times, Medieval times, i.e., the beginnings of the modern historical era, and modern times respectively. Moreover, they each involve linguistic features from different levels of linguistic organization — the lexicon, the syntactic component, and the phonological component respectively. They do share certain things, however. Each example has a focus on Greek, although of necessity other Balkan languages are brought into the discussion; furthermore, and perhaps most important for the question of language use, they all share the property of showing how the methods of historical linguistics, especially with regard to different aspects of language contact and contact-induced changes, can be applied to particular developments to allow inferences about how the languages in question were used at the time the changes were occurring. Historical linguistics, therefore, can form an integral part of any research into this complex question of Balkan use, regardless of the time period of the particular language one is interested in.

WORKS CITED

Andriotis, N. 1967. Etymologiko Lexiko tēs Koinēs Neoellēnikēs. Thessaloniki: Aristotelian University, Institute of Modern Greek Studies.

Barić, H. 1961. La perte de l'infinitif dans les langues balkaniques. Godišnjak Balkanološki Institut II:1-11.

Gabinsky, M. 1967. Vozniknovenie infinitiva kak vtoričnyj balkanskij jazykovoj process. Na materiale albanskogo jazyka. Leningrad.

Georgacas, D. 1982. A Graeco-Slavic Controversial Problem Reexamined: The -ITΣ- Suffixes. (Pragmateiai tēs Akadēmias Athēnōn 47), Athens.

Georgiev, V. 1958. Das Pelasgische. In Proceedings of the Eighth International Congress of Linguists, Eva Sivertsen, ed. Oslo: Oslo University Press, pp. 406-13.

_____. 1966. Was stellt die Pelasgertheorie dar? Lingua 16:263-73.

Grannes, A. 1978. Le redoublement turk à *m*-initial en Bulgare, Balkansko Ezikoznanie 21:37-50.

Herzfeld, M. 1981. Meaning and morality: A semiotic approach to evil eye accusations in a Greek village. American Ethnologist (Special Issue: Symbolism and Cognition) 8:560-74.

Hesseling, D.-C. 1892. Essai historique sur l'infinitif Grec. In Études de Philologie NéoGrecque, J. Psicharis, ed. pp. 1-43 (Paris).

Householder, F. 1964. Three dreams of Modern Greek phonology. In Papers in Memory of George C. Pappageotes (Publications of the Linguistic Circle of New York 5, Supplement to Word 20.3), pp. 17-27.

Joseph, B. 1981. Multiple causation in language contact change. Paper read at Tenth Annual Linguistics Symposium (Topic: Language Contact), University of Wisconsin-Milwaukee (available in microfiche from ERIC (Educational Resources Information Center) Clearinghouse on Languages and Linguistics, Document #ED205021).

_____. 1982a. Phonesthemes in Modern Greek and the Balkans. Paper read at Third Balkan and South Slavic Linguistics Conference, Indiana University (April 1982).

_____. 1982b. Ya tin idieteri θesi tu [ts]/[dz] stin eliniki fonologia (with English summary). In Proceedings of the 3rd Annual Meeting of the Department of Linguistics, University of Thessaloniki (to appear).

_____. 1983. The Synchrony and Diachrony of the Balkan Infinitive: A Study in Areal, General, and Historical Linguistics. Cambridge: Cambridge University Press (Cambridge Studies in Linguistics, Supplementary Series).

Katičić, R. 1976. Ancient Languages of the Balkans. (Trends in Linguistics, State-of-the-Art Reports 4), Pt. 1. The Hague: Mouton.

Levy, H. 1980. An Anatolian language-trait in Byzantios' *Babylonia* and Parallel Traits on Three Continents. MGSA (Modern Greek Studies Association) Bulletin 12:47-55.

Marchand, H. 1953. Über die Bildung der lautsymbolischen Wörter im Türkischen. Oriens 6:50-62.

Meillet, A. 1921. Linguistique historique et linguistique générale. Paris.

Rozencvejg, V. 1969. Infinitivnye konstrukci i balkanskije jazykovye kontakty. Slavia 38:189-209.

_____. 1976. Linguistic Interference and Convergent Change. The Hague: Mouton (Janua Linguarum, Series Maior, 99).

Rosetti, A. 1977. Thrace, Daco-Mésien, Illyrien, Roumain et Albanais. Quelques Précisions. Balkansko Ezikoznanie 20:79-82.

Sandfeld, Kr. 1930. Linguistique Balkanique. Problèmes et Résultats. Paris: Librairie Ancienne H. Champion.

Schaller, H. 1975. Die Balkansprachen: Eine Einführung in die Balkanphilologie. Heidelberg: Carl Winter Universitätsverlag.

Schumann, J. 1979. The Genesis of a Second Language. In The Genesis of Language (The First Michigan Colloquium), K. Hill, ed. Ann Arbor: Karoma Publishers, Inc.

Szemerényi, O. 1974. The Origins of the Greek Lexicon: Ex Oriente Lux. Journal of Hellenic Studies 94:144-57.

Thomason, S. 1981. Are there linguistic prerequisites for contact-induced language change? Paper read at Tenth Annual Linguistics Symposium (Topic: Language Contact), University of Wisconsin-Milwaukee.

Togeby, K. 1962. L'infinitif dans les langues balkaniques. Romance Philology 15:221-33.

Wescott, R. 1975. Allolinguistics: Exploring the peripheries of Speech. The Second LACUS Forum, Peter Reich, ed. Columbia, S.C.: The Hornbeam Press, pp. 497-513.

NOTES

1. As is argued at the end of this paper, in some cases it is more realistic and more satisfactory to recognize multiple causation, often a combination of external and internal factors. See also Joseph (1981, 1983:Chapter 7) for more discussion.

2. For instance, the levelings that have taken place in most of the Modern Greek nominal paradigms, giving forms such as nominative singular πατέρας *father*, based on the accusative singular πατέρα , replacing the Ancient Greek nominative πατήρ , are purely system internal, as are sound shifts involving natural classes of segments, as when the Ancient Greek voiceless aspirated stops φθχ ([p^h, t^h, k^h], respectively) all became voiceless fricatives ([f θ x] respectively) in Modern Greek.

3. For example, as Thomason (1981:2) points out, Meillet (1921:87) "believed that *grammatical* loans were possible only between very similar systems, especially dialects of a single language."

4. For example, the social context can play a role. As Thomason (1981:5-6) points out, early loans from Russian into Asiatic Eskimo were altered so as to fit the native phonological system (e.g., Russian čaj *tea* → Eskimo saja since Eskimo had no [č]), while more recent ones, which have come within the social context of Soviet control of the area and also instruction in Russian in schools, have instead led to the introduction of new sounds (e.g., [č]) into the language (cf. modern čaj *tea*).

5. All of this is summed up, with meticulous documentation, in Szemerényi (1974), though it is important to note that not all scholars agree with Szemerényi's proposals. The ones given here are representative and differences of opinion do exist even with these.

6. Examples from English would be the borrowing of American Indian (specifically Algonquian) terms like *moccasin, toboggan,* or *moose* to designate these newly encountered objects, or the borrowing of French terms like *crêpe* or *aubergine* to designate those particular foods, introduced through the French.

7. See Katičić (1976:39-55) for details; these terms may well be Anatolian in origin, as Szemerényi (1974:153) and others believe, but could still represent a linguistic layer on the Greek mainland before the coming of the Greeks.

8. The Bulgarian linguist V. Georgiev is the source for much of this work on the language represented by the second group here; see Georgiev (1958, 1966), for example, and the summary of his work in Katičić (1976: 71ff.). He believes the two groups of words are from the same language, which he labels Pelasgian. As in the case in virtually all aspects of Greek etymology (see Footnote 5, for instance), it must be pointed out that Georgiev's conclusions are not universally accepted, and many scholars take these forms as native Greek. They are presented here for their value in showing how historical linguistics methods can shed light on prehistoric language use.

9. Katičić (1976:37) does note that the spread of Greek over the entire Aegean area suggests that the Greek invaders were quite numerous; still, without a somewhat substantial autochthonous population, lexical elements of the noncultural sort would probably not have survived.

10. Similar remarks can be made about other languages of the Ancient Balkans. Of particular interest and controversy is the question of Albanian and Romanian ties and the relation of these languages to the now-extinct and poorly documented Thracian and Illyrian languages. The evidence bearing on this point is extensive and beyond the scope of the present paper; moreover, it is not clear that any special insights regarding language use per se come out of this issue beyond points similar to those brought out already in the discussion of Greek and Pre-Greek languages. For a survey of the literature on these questions, see Katičić (1976:128-88), and for some brief remarks, see Rosetti (1977).

11. It is controversial whether these represent clusters of two segments or unitary segments; Householder (1964) discusses some of the phonological evidence bearing on this issue and in current research I am examining the phonetic evidence. I use the designation "sounds" here so as to be neutral on this question.

12. See Georgacas (1982) for a thorough discussion of these and other suffixes with -τσ- , especially with regard to their origin.

13. For a somewhat different but nonetheless analogous type of lexical stylistic differentiation, see the discussion in Herzfeld (1981:564-6) of the connotations of various terms for the evil eye in a Rhodian village. The distinction under consideration here often coincides with the dichotomy within Greek of *katharevousa* (Puristic) versus *dhimotiki* (Colloquial) stylistic levels but does not necessarily do so.

14. That a phenomenon which is allolinguistic in nature should be able to spread from one language to another and in particular from Turkish to Greek is not necessarily unusual. Affective reduplication of the type kitap-mitap *books and things like that* which is found in Turkish (and seems ultimately to stem from more Eastern sources) made its way into Greek and into other Balkan languages (and eventually even into other Slavic languages, German, Yiddish and English). See Levy (1980) and Grannes (1978) for some discussion concerning this trait in Greek and Bulgarian respectively.

15. These etymologies, which are not without some controversy, are from Andriotis (1967).

16. See Sandfeld (1930) for the classical statement regarding these correspondences and Schaller (1975) for a more recent treatment.

17. This survey, like those in the handbooks (e.g., Sandfeld (1930) or Schaller (1975)), is vastly oversimplified. Fur fuller discussion regarding this phenomenon in each language separately and in the Balkan languages as a whole, see Joseph (1983).

18. The presence of νά (as opposed to other subordinating particles like ὅτι or πώς) is conditioned in part by the governing verb and in part by the type of construction.

19. In the case of Albanian, "prehistoric" means before the 15th century since the language is attested only from 1462 (in a one-sentence baptismal formula) and the first substantial Albanian text, the Missal of John Buzuku, dates only from 1555.

20. Greek and Albanian are ambiguous in this regard for Albanian now, especially in Tosk, is undergoing a revival of the infinitive and Greek during the Medieval period underwent a similar but ill-fated renewal of the category infinitive.

21. Any dating must be considered very tentative, for one has to filter out the general conservatism of written records, purely literary revivals of the infinitive, and occurrences of infinitives in increasingly restricted contexts after this period.

22. Compare alternations in English of sentences such as *I expect (that) I will win* with a finite complement clause with one like *I expect (myself) to win* with an infinitival complement.

23. Recall the discussion earlier concerning the Pre-Greek languages; the lack of an infinitive for the ancient autochthonous Balkan languages, however, is purely conjectural, resting on no empirical evidence.

24. For example, in both English and German, there has been a formal merger of infinitives with some finite forms, but both languages retain the grammatical category of infinitive.

25. The conflicts between these opposing tendencies would have been resolved on a language-particular basis, affected by such additional factors as continued contact, homophony of infinitives with finite forms leading to reanalysis of the nonfinite forms as finite, and the like.

26. For further amplification of this point, see Joseph (1983: Chapter 7).