

[To appear in M. Rivero & A. Ralli, eds. *Comparative Balkan Syntax*  
(Oxford University Press, 2000)]

## **IS BALKAN COMPARATIVE SYNTAX POSSIBLE?**

Brian D. Joseph

The Ohio State University

Department of Linguistics  
222 Oxley Hall  
Columbus, Ohio USA 43210-1298  
Phone: 011-1-614-292-4981  
Fax: 011-1-614-292-4273  
e-mail: [bjoseph@ling.ohio-state.edu](mailto:bjoseph@ling.ohio-state.edu)

[Version of August 28, 1998]

## 1. Preliminaries

The question asked in the title to this paper might seem to have a self-evident answer, especially when viewed in the context of the other papers in this volume, all of which deal with some aspect of the syntax of two or more languages of the Balkans. However, it is not as irrelevant or trivial a question as it might at first seem, for despite the successes of these studies, and others to be cited below, with regard to taking a comparative perspective on the various languages in the Balkans, the logically prior question of whether it is in fact possible or even enlightening to do comparative Balkan syntax needs to be explored, as the discussion below indicates.

As a starting point, consider the following succinct statement of the goal of Universal Grammar: the determination of the ways in which all languages are alike and the ways in which they differ.<sup>1</sup> To that end, comparative syntax plays an important and even crucial role.

To illustrate the power of such an approach, one need only look at what has been accomplished regarding the *pro*-Drop (or “Null Subject”) parameter, to choose one area that has been heavily investigated.<sup>2</sup> Taking a comparative perspective on *pro*-Drop — the possibility some languages show of omitting overt expression of unemphatic subject pronouns in tensed clauses — has suggested various correlations to be found between the possibility of such omitted (“null”) subjects and the occurrence of other syntactic and morphosyntactic phenomena. For example, the occurrence of overt subjects in expletive constructions such as those involving weather-verbs or extraposition has been shown (Lightfoot 1991) to correlate with a language not allowing *pro*-Drop, as in English or French, and the absence of such subjects with a language allowing *pro*-Drop, as in Spanish. Similarly, Jaeggli & Safir (1989b:29-30) define “morphological uniformity” in verbal paradigms in terms of the structure of verb forms — a “uniform” paradigm being one in

which forms are either all stem + affix or all bare stem, but not mixed with bare-stem and some stem + affix — and suggest further that “null subjects are permitted in all and only languages with morphologically uniform inflectional paradigms.”

Clearly, if only one language were looked at, it would be harder to be sure that any property that was identified as correlating with the possibility or impossibility of *pro*-Drop was a significant one; thus comparison, and especially cross-linguistic comparison, is at the heart of the enterprise of universal grammar. Correlations such as these for *pro*-Drop may indeed be wrong; for instance, the notion of morphological uniformity that Jaeggli & Safir promote seems to be of dubious value,<sup>3</sup> and the very definition of a “*pro*-Drop language” is far from clear-cut, as argued elsewhere (e.g. Joseph 1994), especially in the face of sentences from English, a putative non-*pro*-Drop language, in which initial material, including subjects, can be deleted,<sup>4</sup> and of the construction-specific reversal of the usual *pro*-Drop setting in French and Greek.<sup>5</sup> Still, without comparative syntactic methodology, no one would be in the position to evaluate such claims and correlations.

Examining the nature of null subjects is an example of doing comparative syntax by focusing on a particular construction or syntactic feature. Yet, there are other ways of focusing this enterprise, including looking at genetically related languages — the basis for comparative syntax in the traditional sense, e.g. as practiced by Wackernagel 1892 in his classic study of an aspect of Indo-European word-order patterns — or a combination of the genetic and construction-specific approach, as with studies of clitics in Romance languages (e.g. by Kayne 1991, and others) or Verb-Second phenomena in Germanic (e.g. by Weerman 1989, the authors represented in Haider & Prinzhorn 1986, among others).<sup>6</sup>

Yet another basis for the comparison of languages has been areal, comparing languages that are geographically related, and much interesting work has been done in the comparative syntax of the languages of the Balkans. Under this rubric would be works examining, for example, from various perspectives within Generative Grammar, the structure of verb phrase

and the domain of Verb Movement (Rivero 1990, 1994), the properties of subjunctive clauses and modal inflection vis-à-vis control phenomena and clitic-climbing (Terzi 1992), the realization of Tough-Movement constructions in languages with finite subordinate clauses (Joseph 1980, 1983), and multiple WH-Questions (Rudin 1988), but also the uncovering, within a more traditional descriptive framework, of numerous shared morphosyntactic and syntactic features such as the structure of the future tense, a postposed definite article, the pleonastic use of weak pronominal forms as verbal markers, and the general absence of nonfinite complementation, all presented by Sandfeld in his classic work (Sandfeld 1930) on the Balkan languages and discussed, along with others, in the enormous literature on Balkan linguistics.<sup>7</sup>

While the results of areally based comparative syntactic investigations have often served to shed light on aspects of Universal Grammar, an areal perspective has been especially interesting when the languages in question show other common characteristics that unite them, that is, when they show traits linking them as a “Sprachbund,” to use the German designation as a technical term in English.<sup>8</sup> A Sprachbund is an area where long-term intense and intimate contact among speakers of several different languages has led to massive structural convergence in languages that were once quite different from one another.<sup>9</sup> The languages of the Balkans constitute perhaps the best-known and most deeply investigated case, but other examples include South Asia (Masica 1976), and Meso-America (Campbell, Kaufman, & Smith-Stark 1986). Among the syntactic characteristics of these other Sprachbünde are for South Asia, the use of conjunctive participles for serialization, SOV word order, and dative subject constructions, and for Meso-America, the occurrence of nominal possession of the type *his-dog the-man*), the absence of switch reference marking, and non-verb-final word order (vs. SOV languages in surrounding areas).<sup>10</sup>

Still, some of these results, especially those from the generatively-based investigations mentioned above, may well pose some problems when viewed from the Sprachbund

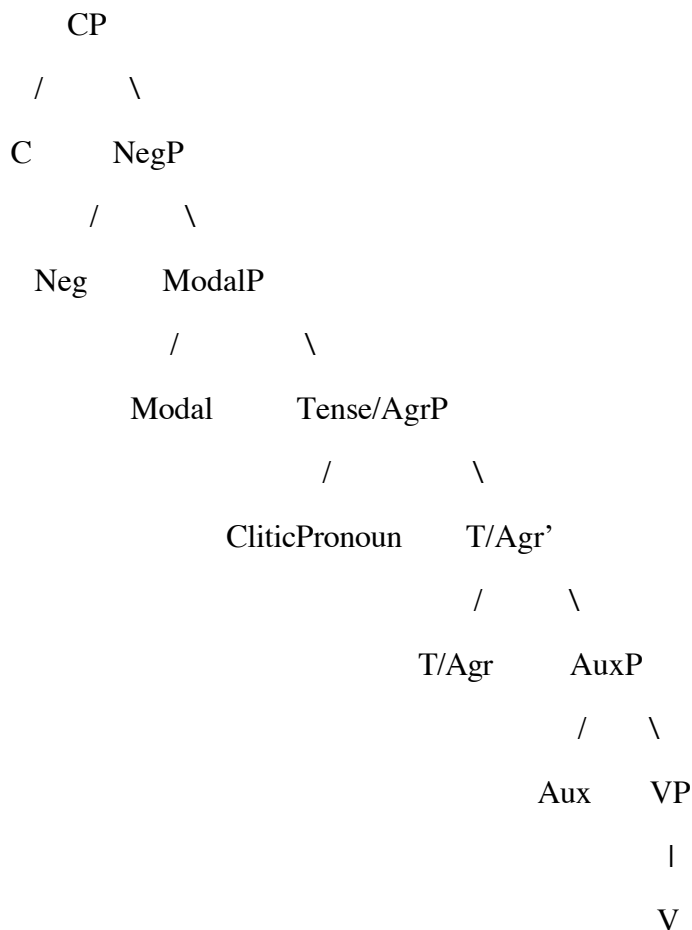
perspective, since Sprachbund phenomena generally are attributed to language contact in some form. The exact nature of the contact that leads to a Sprachbund is often a matter of some controversy, in that there is debate as to whether it is substrate influence from one population shifting to a target language, superstratal influence of one language over the others in the area, massive bi- or multi-lingualism often of an imperfect sort, sociolinguistic accommodation, or some combination of such situations, that has led to the convergence in the Sprachbund. Nonetheless, contact in some form is invariably responsible, and language contact would seem to be more of an accidental happenstance in the history of particular languages that could render comparative syntax less interesting than it might otherwise be.

It is valuable therefore to explore this issue more deeply; accordingly, in what follows, four ways in which the results of some types of comparative syntactic investigations are potentially problematic are discussed in some detail, with an eye towards determining whether the enterprise of comparative Balkan syntax is possible and if so, what the limits are that it is subject to. With those issues addressed, some specific case-studies involving negation in the Balkans are discussed, in order to illustrate what different approaches to comparative Balkan syntax might yield in the way of insights.

## **2. Some Problematic Aspects of Comparative Syntax in the Balkans**

First, it seems fair to ask whether the results that have been obtained from recent generative comparisons of the syntax of various Balkan languages are revealing beyond what might be found if one were to compare any arbitrary set of typologically related languages chosen on a basis other than geography. That is to say, especially when one realizes the role that language contact has played in shaping the Balkan languages, in what way does a claim about parallels in the structure of the clause among the Balkan languages advance our knowledge? For instance, the analysis given by Rivero 1994 for Verb Movement in the Balkans, which proposes the structure in (1) for the “Balkan clause”:

(1) Pan-Balkan Clause Structure (following Rivero 1994)



is interesting and well-argued in its own right, but what does it show? In particular, it seems not point to a uniquely Balkan clause-structure, but rather merely extends to these languages analytic principles — in this case, the “exploded INFL” analysis given by Pollock 1989 for clause structure — which are assumed to be part of Universal Grammar. In that case, however, such an analysis seems to call into question any special value that might be posited for Balkan syntactic parallels. That is to say, if the analysis assimilates Balkan verb structures to well-known universal principles, then this is really a matter of comparative syntax more generally, not comparative Balkan syntax in particular. Moreover, given that the Sprachbund effect is a matter of some type of language contact, one can legitimately wonder

if any of the results obtained by such an analysis are ones that provide insights into the Balkan Sprachbund as a contact-induced phenomenon.

Such criticisms are to be leveled not just at relatively recent work of this type. The same could be said of the results reported in Joseph (1980; 1983: 232ff.), for instance, where a possible parallel was identified between Greek and Romanian in the realization of Tough-Movement in a language without an infinitive. As indicated in (2), it was claimed that in each of these languages, as the earlier infinitive was replaced by finite subordinate clauses, constructions developed which avoided having a transitive complement verb which is both finite and objectless; this was achieved in Greek through a copying-type of Tough Movement, as in (2a), in which a pronominal object in the subordinate clause copying the “raised” nominal renders the transitive verb non-objectless, and in Romanian by a “passive”-type of Tough Movement, as in (2b), in which the subordinate clause is passivized through the reflexive passive construction and is thus detransitivized:

- (2) a. ta anglíká<sub>i</sub>                      íne              ðískola                      na  
           the-English/NTR.NOM.PL    are/3PL    difficult/NTR.NOM.PL    SUBJUNC  
       ta<sub>i</sub>                                      katalávi                      kanís  
           them/NTR.ACC.PL            understand/3SG    someone/NOM.SG

‘English is difficult to understand’ (literally = “the-English (things) are difficult that someone understand them”)

- b. asta nu-i              greu să              se              facà  
           this not is/3SG    hard    SUBJUNC    REFL    do/3SG

‘This is not hard to do’ (literally = “This is not hard that it be done”)

Still, even if of interest, the parallel seen here is a rather abstract and “deep” one. Moreover, it cannot be a matter of language contact since the form of the construction in each language is quite different, involving what can descriptively be called “copying” in Greek but a reflexive/passive formation in Romanian. Rather, the parallel seems to have to do with a

universal tendency in Tough-Movement constructions, whether the result of a syntactic rule or outcome of other rules, to “prefer” or select a nonfinite complement, as reflected in the prevalence of nonfinite subordinate clauses cross-linguistically in Tough-Movement sentences. Thus if due in large part to some aspect of Universal Grammar, however the cross-linguistic tendency is to be captured formally, the parallel is not very interesting from a Balkanological point of view, though it does provide additional input into the universal characterization of Tough Movement.

Such examples therefore do tell us something, but they do not reveal anything about Balkan-particular characteristics, or at least not in the same way as does finding parallel structures such as those noted above in section 1. That is, it has been taken to be quite significant with regard to their common development that many languages in the Balkans exhibit a future tense with an invariant prefix-like marker based — historically at least — on a verb meaning ‘want’, as in (3a), a definite article that is postposed within the noun phrase, as in (3b), the use of pleonastic weak personal pronouns coindexing object noun phrases and thus serving roughly as object agreement markers or transitivity markers on the verb, as in (3c), and finite (i.e. person-marked, tensed) subordinate clauses where English and many other European languages, at least, use nonfinite complementation, as in (3d), to take four widespread and commonly noted Balkan features:<sup>11</sup>

(3) a. Alb *do (të) punoj* = Blg *ste rabotja* = Grk *qa ðulévo* (< earlier *qe (na) ðulévo*) =

Rmn *o sà lucrez* ‘I will work (historically, “wants-(that)-I-work”)

b. Alb *ujk-u* = Blg *vùlk-ùt* = Rmn *lup-ul* ‘the wolf’ (literally, “wolf-the”)

c. Alb *e pashë Gjonin* = Blg *go gledax Ivan* = Grk *ton íða ton jáni* = Rmn *l’am vàzut pe Ion* ‘I saw John’ (literally, “him I-saw (the-)John”)

d. Alb *përpiqem të ndihmoj* = Blg *opitvam se da pomogna* = Grk *prospaqó na voiqíso* = Rmn *încerc sà ajut* ‘I try to help’ (literally, “I-try that I-help”)



Given the fact that these parallels are not the result of a common inheritance from Proto-Indo-European,<sup>12</sup> and that they represent a divergence from earlier stages of each of these languages, the convergence they show is striking and provides an important starting point for an investigation into the language contact situation that gave rise to them. Such, however, is not the case with parallels that can be attributed to the workings of Universal Grammar.

Second, as some of the features discussed here already show, the syntactic similarities found in Sprachbünde and other contact situations tend to be superficial in nature, and are really a matter of a convergence in surface structure, rather than in deep structure or a set of rules by which underlying forms are realized on the surface. The Balkan features illustrated in (3), for instance, can all be readily described in terms of gross surface patterns, what may be characterized as “target structures” that speakers aim at, and such is the case also with the convergent verb serialization structures which, as noted above, are found in many South Asian languages, with convergent word order patterns such as the South Asian OV structures or the Meso-American non-verb-final order, and with the *his-dog the-man* expression of possession in Meso-America, where the target structure is a desired output. Similarly, in the on-going contact situation in Kupwar village of the Maharashtra state in India involving Kannada, Marathi, and Urdu speakers, as described by Gumperz & Wilson 1971, among the convergences is the Kupwar Kannada use of ‘be’ after predicate adjectives, paralleling the Marathi/Urdu surface pattern and diverging from the Standard Kannada absence of ‘be’ in that context.

Languages in such situations may show “deeper” similarities or even differences in the ways these surface forms are generated synchronically or more generally are integrated into the grammars of individual languages, but the surface forms themselves, the output of generative rules of syntax, would seem to be the critical level at which to judge similarities that would reveal the existence of a Sprachbund. For instance, it would not matter what the processes are in individual languages that lead to basic verb-final order in South Asia, e.g.

whether it is an underlying SOV order, or an obligatory object-fronting process from an underlying SVO order, or whatever — as long as the surface similarity is there, the phenomenon will be salient for linguists, and presumably, and more importantly, will be so for the speakers, too. Alternatively, if the processes by which verb-final order is generated in a language are so constrained by universal grammar that there is really just one possibility and no other options, then all that matters is the presence of the target structure in the language output, not the processes that give rise to it. In such a case, Universal Grammar would be responsible for the deep similarity, and it would thus be uninteresting from the Sprachbund perspective of language contact.

Within the Balkans, this situation can be illustrated by the convergence involving perfect-tense formations with the verb ‘have’, where a superficial similarity has long been noted<sup>13</sup> between Albanian and Macedonian regarding the fact that they both have a past perfect consisting of the past tense of ‘have’ with a (generally passive)<sup>14</sup> participle, and Greek can be added to this as well,<sup>15</sup> e.g. Alb *kisha lidhurë* ‘I had tied’, Grk *íxa ðeméno* ‘I had tied’,<sup>16</sup> Mac *imav storeno* ‘I had made’. However, as Friedman 1983 remarks: “[although] the superficial resemblance between the Macedonian and Albanian forms has been noted at least since Sandfeld, ... these forms play very different roles in the structure of their respective languages,” especially in terms of the relationship to other verb tenses and formations in each language. Thus there are deeper differences in how these forms are embedded in their respective verbal systems, yet such differences are irrelevant to the similarity these forms show in terms of surface structure.

Such an importance on surface structures is really to be expected if the basis for the spread of such features — that is, the basis for the development of contact-induced areal convergences — is at least limited bilingualism, transfer, and reverse interference, for surface forms are the point of contact between speakers. It should also be noted, in this regard, that lexical borrowing, which is a quintessentially surface-oriented phenomenon that

is widespread in the Balkans, can easily shade off into construction borrowing and thus into the realm of syntax. This point can be illustrated by the Greek construction exemplified in (4a) consisting of repetitions of a perfective verb form sandwiched around the morpheme for ‘not’ but with the meaning ‘whether one VERBS or not’, for this construction, as noted by Banfi (1985:80), occurs with the verb ‘want’ in several Balkan languages, as in (4b), the form being third person plural in Albanian and the Turkish negation following the usual suffixal pattern for the language:<sup>17</sup>

(4) a. Grk *fīji ðe fīji* ‘whether one leaves or not’

b. Grk *qéli ðe qéli* = Blg *ste ne ste* = Rmn *vrea nu vrea* ‘whether one wants to or not’  
 ≈ Trk *ister istemez* ‘willingly or not’ (literally: “want-AORIST.3SG want-NEG.AORIST.3SG”) ≈ Alb *donin s’donin* ‘whether they want to or not’

One interpretation of these facts that suggests itself, in the face of the mild productivity it shows in Greek and the widespread occurrence of the ‘want’ formation, is that the ‘want’ construction is the starting point which has spread via loan translation throughout the Balkans, but that this borrowing has become the basis for extension to other verbs, with the result that it has become a syntactic pattern rather than just an isolated lexical form. In such a view, it is hard to distinguish something that is in essence lexical borrowing of a phrase from the borrowing of what is in essence a syntactic pattern. A surface borrowing can thus have repercussions into and throughout the syntax.

If similarities in contact situations are focused on the surface, it therefore becomes potentially problematic to view the syntactic similarities among Balkan languages in terms of deep syntactic features such as parameter settings, as Rudin 1988 did for instance for parallels in multiple *wh*-Question constructions in Bulgarian and Romanian. Sprachbund significance for such features would be inconsistent with their deep nature, since the “action” in language contact, so to speak, is at the surface, not at a deep level, yet contact is crucial for the development of a Sprachbund.

Third, and this is a problem that pervades much of Balkan linguistics, most of the relevant studies<sup>18</sup> have been based just on a comparison of the modern standard languages, when in fact the crucial period for the “Balkanization” of all the languages in the Balkan Sprachbund was some 400 to 700 years ago and involved contact at the level of the regional dialects not the standard languages. Instructive here is the observation made by Masica 1976 and seconded by Campbell et al. 1986 that some linguistic areas are “the relics of past contacts, no longer active and others ... are in the process of formation and extension because of on-going interaction and change” (Campbell et al., 533). Most of the features that make the Balkan Sprachbund interesting are ones that are the result of past contacts, not on-going continuing contact in the present day.

In a sense then, looking at an on-going contact situation such as that mentioned above involving Urdu, Marathi, and Kannada speakers in Kupwar is more crucial for understanding the Balkan Sprachbund than are constructs from modern syntactic theory. Similarly, current contact within Greece involving standard Modern Greek interacting with Arvanitika, the variety of Tosk Albanian spoken in Greece for some 600 years or more, or Aromanian, also known as Vlach, the variety of Romanian spoken in Greece for at least several centuries, provide important insights into the formation of the Balkan Sprachbund, for these typically village-based situations approximate the contact situation in the Balkans 600 years or so ago in ways that an examination or comparison of the various present-day, generally urban-based, standard languages cannot. What one sees in examining the urban standards is perhaps the aftereffects of the contact from several centuries ago, but it is not such a direct window on the conditions that gave rise to the Sprachbund effects.

As an example<sup>19</sup> of insights from such relatively recent contact in the Balkans, consider the following involving the assimilation of the interpretation of the understood subject with the gerund in Arvanitika to the current Greek pattern for the present active participle (or “gerund”) in *-ondas*. In what is the result of an innovation that most likely took place in

late-ish Medieval Greek, given the fixing of the current form of the gerund in that period,<sup>20</sup> the Modern Greek *-ondas* form permits an interpretation of its understood subject as coreferent only with a main-clause subject, as in (5):

- (5) o jánis<sub>i</sub>            íðe            ti maria<sub>j</sub>            perpatóndas<sub>i/\*j</sub> s to ðrómo  
the-John/NOM saw/3SG the-Mary/ACC walk/GERND on the-road  
‘John saw Mary while he/\*she was walking on the road’

In Standard Albanian, however, based on the Tosk dialect, either subject or object control of understood subject of a gerund, e.g. *duke ecur* ‘(while) walking’, is possible, as indicated in (6):

- (6) a. Njeri<sub>i</sub>                            afrohej                            [Ø<sub>i</sub> *duke ecur*]  
man/NOM.SG.INDEF approached/3SG.IMPF            GERND walk/PPL  
‘A man was approaching (while) (he was) walking’
- b. Vajza<sub>i</sub>                            pa            njeri<sub>j</sub>                            [Ø<sub>i/j</sub> *duke ecur*]  
girl/NOM.SG.DEF saw/3SG man/ACC.SG.INDEF            GERND walk/PPL  
‘The girl saw the man (while) (she/he was) walking.’

In Arvanitika, on the other hand, even though also part of the Tosk dialect group, one finds only subject control, as in Modern Greek. Therefore, influence from Greek on Arvanitika is likely, with the pattern of interpretation for the gerund in Arvanitika influenced by the Greek pattern, based on recognition by Arvanitika speakers that the Greek form and the Arvanitika form are parallel, to be identified cross-linguistically as being the same type of grammatical element. Significantly, such an identification is ultimately surface-based and is precisely the sort of development one expects to find in intense contact situations where there is at least some bilingualism. Although it is not clear exactly when the change in Arvanitika took place, it is most likely to be recent, once Arvanitika speakers became increasingly bilingual in Arvanitika and Greek; indeed, such bilingualism is the norm in virtually all Arvanitika communities nowadays, with the younger generation tending towards exclusive use of

Greek. Nonetheless, in the period of widespread bilingualism, these Arvanitika communities mirrored aspects of the multilingual villages prevalent in the Balkans in centuries past, and thus provide a window of sorts onto the conditions of the past.

Fourth, as the discussion of the ‘have’-perfect already demonstrates, much of what is attended to in Balkan linguistics is similarities among languages, without as much attention being paid to the differences these languages show. This is more true of traditional descriptive studies, perhaps, than more recent generative studies, for the latter generally attempt to develop a typology according to which the languages under investigation can be said to fall into one or the other class of languages. However interesting the differences might be, though, and however important it might be to investigate them — for only by knowing the extent of differences can we judge whether there really are significant similarities — the differences are not something that arise by language contact and thus in a sense fall outside of the purview of at least traditional Balkan linguistics.

These four issues loom large in any attempt at comparative Balkan syntax, but they are not insurmountable. In what follows, a path towards their resolution is charted.

### **3. Towards a Resolution**

Once one takes all of these problems into consideration, it becomes clear that a Balkan comparative syntax is indeed possible, but the success of the enterprise depends on what one’s goals are. For example, for the purposes of tying any results into the Sprachbund phenomenon so well documented for the languages of the Balkans, the most enlightening comparisons will be those involving surface phenomena which are likely to be transferred in language contact situations. On the other hand, comparisons involving parametric variation or parallels at deeper levels of structure are illuminating insofar as they shed light on Universal Grammar, for instance, or clarify the extent of a superficial similarity, although

they do not provide any input into an understanding of the contact that created the Balkan Sprachbund.

What is most useful here as a means towards a resolution is a distinction that draws on and is somewhat analogous to Schaller's 1975 distinction between "language of the Balkans," a purely geographic designation that takes in any language that occurs within the geographic bounds of the Balkans, and "Balkan language," a designation for those languages of the Balkans that participate in the Balkan Sprachbund and show parallels due to language contact. Using that dichotomy as a basis, one can distinguish between working on the "comparative syntax of the Balkan languages," i.e. examining the syntax of individual languages of the Balkans in comparison with other languages of the Balkans and elsewhere, and doing "comparative Balkan syntax," i.e. examining the syntax of Balkan languages, keeping the Sprachbund in mind. And, more generally even, one can further distinguish "linguistics of the Balkans" from "Balkan linguistics," the former being the analysis of the languages in and of themselves, the latter being the analysis with regard to Sprachbund.

The recognition of such a distinction means that the different aims of comparative syntax can be clarified. Just as the "language of the Balkans" versus "Balkan language" distinction is a useful one, so too are the ones proposed here for comparative syntax and for studying the languages in the Balkans more generally — the goals of each enterprise are different and thus success is measured in different ways.

With all this now in place, a discussion of some facts concerning negation in the Balkans can be examined, as case studies where both types of perspectives can fruitfully be taken.

#### **4. Two Case Studies in Balkan Negation**

**4.1.** The first area of interest starts with the formal parallels evident in one of the negation markers in Greek and in Albanian, as well as most non-Vlax dialects of Romany (the language of the Gypsies) spoken in the Balkans. In each of these languages, a negator beginning with [m] is found for nonindicative negation, as well as some other functions discussed later on; the forms in question are Modern Greek *mi(n)* (Ancient Greek [mév] < mhv >), Albanian *mos*, and Romany *ma*.

There are clear cognate forms elsewhere in Indo-European to these Balkan #*m*-negators, and the paths of development to the attested forms are well-understood. In particular, Sanskrit *má*, Avestan *má*, and Armenian *mi* all point to a Proto-Indo-European \**mé* as the source for the Greek, Albanian, and Romany forms, and Tocharian *má* is generally taken to do so also.<sup>21</sup> Ancient Greek *mé* continues PIE \**mé* directly, from which Modern Greek [mi] developed by regular sound change; the final *-n* found in some forms of *mi(n)*, especially those marking verbal negation, was added to inherited *mi* by analogy to the finite verbal indicative negator *Den*, which itself derives from Ancient Greek [oudén] (< oujdevn >) ‘nothing; not at all’. As for Albanian, *mos* derives from a composite \**mé-k<sup>w</sup>id* (‘not’ + ‘anything’) by regular sound change; *sorrë* ‘blackbird’ from \**k<sup>w</sup>érsná*, for example, provides examples of the vowel development of \**é* to *o*, the assibilation of \**k<sup>w</sup>* to *s* before a front vowel, and to a certain extent also the reduction of the final syllable needed to derive *mos*. An Indo-European “pedigree” for present-day Balkan *mi/mos/ma* guarantees that the Turkish general negation marker-*me/-ma-* is not in any way responsible for the occurrence of *m*-negators in these other Balkan languages, however unlikely such a scenario might be in any case.

Due to a lack of sufficient information about the range of uses of *ma* in Romany, attention hereafter is focused on the Greek and Albanian *m*-negators. Similarly, even though a form [mi] occurs in the expression of a negative imperative (i.e. a prohibition) in the Macedonian spoken in the area around Thessaloniki, at least into the first half of the 20th



century, this occurrence seems clearly to be a matter of the borrowing of the Greek formation into Macedonian, as discussed most recently by Topolinjska (1995:310), most likely through the medium of bilingual speakers, and thus is not of immediate concern here.

Besides the formal parallels, there are a number of functional parallels between Modern Greek *mi(n)* and Albanian *mos*. The various functions these elements fulfill are given in (7), and examples of these uses are given in (8), with the two displays following the same order of presentation for these uses; *mi(n)* and *mos* are glossed as *mi* and *mos* respectively, and some relevant explanatory details about various of the uses are included in parentheses:<sup>22</sup>

(7) Functions of Balkan *m*-negators

- a. modal negator (in Grk, of subjunctive clauses; in Alb, of subjunctive and optative verbs)
- b. nonfinite negator (in Grk, of active participles; in Alb, of active participles (gerundives) and the infinitival formation)
- c. introducer of prohibitives and negator of hortatives (in Grk, with finite verb forms, not with imperatival forms *per se*; in Alb, with imperatives and hortatives)
- d. introducer of negatively evaluated clausal complements to verbs and nouns of fearing (in Grk, on its own as complementizer or with another morpheme in *mípos*; in Alb, with complementizer *se* (as *se mos*), though cf. (h) regarding another interpretation of *se mos*)
- e. introducer of tentative main-clause questions (in Grk, with variant *mípos*)
- f. independent utterance expressing negative actions (i.e. prohibitions)
- g. negative combining-element in word-formation (in Grk, in isolated formations; in Alb, more productively)
- h. pleonastic negator in clausal complements to heads with negative force (in Grk, e.g. *embodízo* ‘prevent’; in Alb, e.g., *frikë* ‘fear’, thus overlapping somewhat with (d))
- i. negator of ellipped (i.e. “understood”) elements

j. negator of nonverbal lexical items and constituents (not in Alb, unless (g) belongs here, or vice-versa)

(8) Examples of uses in (7) (i = Grk; ii = Alb)

a. i. borí na min éxun kimiθí  
can/3SG SUBJUNC *mi* have/3PL slept

‘It is possible that they haven’t gone to bed yet’ (lit., “It can that they have not slept”)

ii. sikur të mos jetë bujku usta  
if SUBJUNC *mos* be/3SG.SUBJ farmer/NOM.DEF craftsman

‘if the farmer were not a craftsman’

b. i. min éxondas iDéa ja óla aftá, o jánis tin pandréftike  
*mi* have/ACT.PPL idea/ACC about all-these the-John/NOM her/ACC married/3SG

‘Not having any idea about all these things, John married her’ (Veloudis 1982:22)

ii. për të mos e marrë / duke mos marrë asgjë  
INFINITIVAL *mos* him take/PPL GRDV *mos* take/PPL anything

‘in order not to take him’ / ‘(while) not taking anything’

c. i. min to petáksis!  
*mi* it/ACC throw/2SG

‘Don’t throw it out!’

ii. mos u bëni merak  
*mos* NONACT make/2PL care

‘Don’t worry!’

d. i. to éskase apó fóvo min ton xtipísun  
it/ACC burst/3SG from fear/ACC *mi* him/ACC beat/3PL

‘He ran off for fear that they might beat him’ (Mackridge 1985:300)

ii. kam frikë se mos na shajë

have/1SG fear that *mos* us/ACC scold/3SG

‘I fear lest he scold us’

e. i. *min íDes* to *peDí?*

*mi* saw/2SG the-child/ACC

‘Did you perhaps (happen to) see the child?’

ii. *mos e njihni atë?*

*mos* him know/2PL him/ACC

‘Do you (perhaps) know him?’

f. i. *mi!* (NB: \**min!* (with final *-n*))

‘Don’t!’

ii. *mos!*

‘Don’t!’

g. i. *míte* ‘not even; neither’ (cf. *úte* ‘not even; neither’ for segmentability); *miDén*

‘nought; zero’ (cf. the finite indicative negator *Den*); *miDé* ‘not even; neither’

(infrequent; cf. *uDé* ‘not even; neither’); *mípos* (variant of *mi(n)* in main-clause

tentative questions and with complements to verbs and nouns of fearing, and cf.

complementizer *pos* ‘that’); *mígár(is)* ‘perhaps’ (in tentative questions; rather

infrequent — note that *mígaris* also occurs, even more rarely)

ii. *mosbarazi* ‘inequality’ (cf. *barazi* ‘equality’); *mosbesim* ‘mistrust’ (cf. *besoj* ‘I

trust’); *mosnjohje* ‘ignorance’ (cf. *njoh* ‘I know’); *mosqeni* ‘nonexistence’ (cf.

*qeni* ‘being’), etc.

h. i’. *fováme na min érqí* (Veloudis 1982:11)

fear/1SG SUBJUNC *mi* come/3SG

‘I am afraid that he may come’ (NB: ≠ ‘I am afraid he may not come’)

i’’. *De se emboDízo na min milás* (Thumb 1964:200)

NEG you/ACC prevent/1SG SUBJUNC *mi* speak/2SG

‘I do not prevent you from speaking’ (NB: ≠ ‘I do not prevent you from not speaking’)

- ii. kam frikë se mos na shajë  
have/1SG fear that *mos* us/ACC scold/3SG  
‘I fear lest he scold us’

- i. i.’ parkarizména ke mi aftokínita ítan pandú  
parked/NTR.PL and *mi* automobiles/NTR were everywhere  
‘Parked and unparked cars (i.e. ‘cars that are parked and (ones that are) not (parked)’ were everywhere’ (based on Mackridge 1985:244)

- i.” mi ta xérja su ékso  
*mi* the-hands/ACC your outside  
‘Don’t (put) your hands out!’ (Mackridge 1985:244)

- i.” mi xirótera  
*mi* worse/NTR.PL.COMPVE  
‘What next? God forbid!’ (literally: “(May) not worse (happen)!”)

- ii. si mos më keq  
how *mos* COMPVE bad  
‘in a lamentable state’ (literally: “how (might) not worse (happen)?”)

- j. i.’ se períptosi mi pliromís tis epitajís  
in case/ACC *mi* payment/GEN the-check/GEN  
‘... in (the) case of nonpayment of the check’

- i.” i mi kapnistés káqonde eðó  
the *mi* smokers/NOM sit/3PL here  
‘Non-smokers sit here’

- ii. NO EXAMPLES (UNLESS SOME OF (g) BELONGS HERE)

The lists and examples in (7) and (8) show that there are some rather striking parallels between Greek and Albanian with regard to the use of their respective *m*-negators. In fact, only the last, constituent negation, is found just in Greek, and otherwise the overlap is considerable. Still, there are some differences as well to note in their use, beyond any signaled in the parenthetical notes in (7).

For one thing, as a word-formative element, the *m*-negators show differences in productivity. In particular, the *míte* type formation is rather limited in Greek, but *mos-* is a fairly productive derivational element in Albanian, especially with deverbal nouns in *-im*. If, however, *mos-* in this function is paralleled actually by *mi* as a *constituent* negator, as in (7/8j), e.g. *mi pliromí* ‘nonpayment’, *mi kapnistís* ‘nonsmoker’, then both are fairly productive, and Albanian would then have the full range of uses found in Greek.

Second, Albanian *mos* is used for negation in conditionals, e.g. *në mos gaboj* ‘if I am not mistaken’, while Greek now uses the finite indicative negator *Den* in such constructions, e.g. *an Den se pistépsō* ‘if I don’t believe you’. For Greek, this use of *Den* is found for at least (20th century) demotic Greek — the situation in Ancient Greek and in at least early 20th century katharevousa Greek was different,<sup>23</sup> with conditional clauses negated with the *m*-negator. The causes for the change to use of the indicative negator in conditionals in Greek may be tied up with the development of moods, and thus is tangential to the matters at hand here, but in terms of what to compare between Greek and Albanian, this change means that care must be taken. Similarly, with regard to verbal moods used with the *m*-negators, in prohibitives (cf. (7/8c)), Albanian *mos* is used with imperative mood forms, while Greek *mi(n)* is used with nonimperative forms; given this distribution, the Albanian prohibitive usage could be taken simply to be a case of nonindicative negation, as in (7/8a), while Greek shows a special usage that does not reduce to nonindicative negation, inasmuch as *mi(n)* cannot be used with the imperative. Moreover, following up on prohibitive uses, it should be noted that independent *mos*, besides the prohibitive value it has (cf. (7/8f)) which is

paralleled in Greek, can also have nonprohibitive exclamatory value, as in (9), while in Greek independent *mi* has only prohibitive value:

- (9) Është vrarë      Kajoja!      —      Mos!  
is/3SG slain/PPL Kajo/NOM.DEF      *mos*  
'Kajo has been slain! OH NO!')

Finally, the question-particle use in (7/8e) is broader for Albanian *mos* than for Greek *mi(n)*. In particular, *mos* can have overt negative dubitative value while *mi(n)* is only dubitative (and thus at best only weakly negative).

- (10) Mos është e fortë  
*mos* is/3SG strong/FEM  
'She isn't strong, is she?'

The approach taken here in the presentation of these similarities and differences has primarily been of a pretheoretical, somewhat descriptive and informal sort. Still, these facts are of some interest in regard to formal and comparative issues. For instance, they raise interesting questions concerning the extent to which these functions are all really separate or instead can be collapsed, e.g. does the question usage in (7/8e) involve some negative force, especially for Greek, in the same way that the modal negation does? Also, are the *m*-elements that are employed here all the same formal element in some relevant sense? It is noteworthy in this regard that in Greek some instantiations, particularly those attached to verbs, allow a final *-n* before vowels and some consonants,<sup>24</sup> as the examples in (8a-e) show, while some, e.g. the independent prohibitive utterance (8f), prohibit it, and so forth. Ultimately, a theory of morphology and indeed even a semantic theory should have something to say about such questions,<sup>25</sup> but they are relevant too for issues in the formal syntax of these languages.

The independent word status of the prohibitive utterance in (8f), for instance, has been taken by Rivero & Terzi 1994 as part of the evidence for treating negation in Greek as

formally distinct from clitic pronouns in terms of their blocking properties in Verb Movement to Comp. There are indeed such differences that permit such an interpretation, and further there is good evidence that the clitic pronouns of Greek are best analyzed as affixes, as morphological entities and not syntactic ones (as argued in Joseph 1988, 1990). Still, if the independent prohibitive utterance is a distinct element from the verbal negator, as their differential behavior regarding final *-n* could suggest, then part of the evidence for treating negation in the verb phrase as having special properties distinct from the so-called clitic pronouns evaporates. Albanian would be more amenable to such an argument, since there are no formal differences between the independent prohibitive utterance *mos* and the verbal negator.

Returning to the matter of the potential Balkanological import of these comparisons made between Greek and Albanian with regard to their *m*-negators, crucial to any insights here is consideration of the historical development of these various functions in these languages. The following observations are critical in the evaluation of these facts.

First, all of the functions in (7) for *mi(n)* are found in Ancient Greek for *mé* except for (7f): in the entirety of the Ancient Greek corpus, there are no instances of the independent usage of *mé* expressing negative actions, i.e. as prohibitions, except in ellipsis where it occurs with other words, as in (7i); an example of the elliptical prohibitive is given in (11):

(11) *mé moi sú* (Euripides *Medea* 964)

*mé* me/DAT you/NOM

‘None of that to /for me!’ (literally: “Not to-me you” (with an understood 2SG verb such as ‘give’ or ‘do’))

Moreover, there are parallels elsewhere in Indo-European to most of the other functions, except again the independent prohibitive utterance usage, though the non-prohibitive exclamatory value of the independent negative utterance has a parallel in Sanskrit,<sup>26</sup> the tentative question usage, and the use with complements of fearing. Given that Ancient Greek

had both the question usage and the “fear”-complement usage and that there are numerous uncertainties about the prehistory of Albanian, it is tempting to think of these Greek-Albanian parallels as innovations that spread from Greek to Albanian, but such a spread would have occurred, if at all, in an early, pre-Balkanizing, period of contact between the languages. Alternatively, the occurrence of both the question usage and the “fear”-complement usage in Ancient Greek and Albanian could be taken to warrant positing these as inheritances from Proto-Indo-European, even if they are not found elsewhere in the Indo-European family.

What all this means is that of the various uses without solid comparative justification as inheritances from Proto-Indo-European, the independent prohibitive utterance use (7f) has the best chance of being a real Balkan innovation, since it clearly must have arisen in Greek after Ancient Greek. It may well have come about as an extension of the elliptical use understood as a prohibition, as in (7/8i), though it is not clear whether it spread from Greek to Albanian or vice-versa or was an independent creation in each language..

It is important to note, however, that this is exactly the sort of word that one might suppose would be very frequent in everyday contact situations, so that it is a good candidate to have spread in the intense contact and (often imperfect) bilingualism that gave rise to the Balkan Sprachbund. A speaker of one of these languages, when confronted with a parallelism between their *mi* and another’s *mos* (or vice-versa) could easily have noted a difference in the extent of usage of the form in the other language and could have used that as the model for extending their use of their own native element. As a calque, then, a sort of loan-translation, it would have been transmitted superficially, but could be integrated into the receiving language in a way different from the way it fits into the syntactic structure of the model language. Thus, what would be most salient from the point of view of the speakers who are in contact with one another would be the function that particular surface forms have, for that is where the model for the calque and extension would be found.



In such a case, therefore, working out the formal and sometimes abstract details of where each element fits in its respective system is certainly important, but more for the syntax of the individual languages, for comparative syntax of the Balkan languages; the comparative Balkan aspect, in terms of what is revealed for the Balkan Sprachbund, that is, for comparative Balkan syntax, focuses more on the surface and on the function. Nevertheless, tracing the history of the forms and their respective functions in each language highlights a possible Sprachbund feature,<sup>27</sup> the independent prohibitive use of an *m*-negator, thus contributing to the goal of comparative Balkan syntax in the sense developed above. Even if it should turn out that this usage arose independently in each language, so that it is not a contact-induced feature in one of the languages, the identification of a possible shared feature through comparative syntax is crucial to determining the extent to which the feature is a syntactic Balkanism, an aim of a true comparative Balkan syntax.

**4.2.** Another set of negation facts to be considered here is more of the other sort, i.e. interesting from the perspective of comparative linguistics of the Balkans (as opposed to comparative Balkan linguistics).

At issue here is a phenomenon which can be referred to as “negative fusion,” i.e. the joining of a negative marker with a verb to form a single word unit. An example from earlier stages of English is Old English *nille* ‘not wants’ which represents the negative marker *ne* fused with *wille*, ‘he/she wants’. To a certain extent, negative fusion is found in Modern English too, under the analysis of Zwicky & Pullum 1983 whereby *-n’t* is not merely a syntactically generated clitic form of *not* but rather is a morphologically generated affix; thus, *won’t*, in their view, does not synchronically represent a reduction of *will* + *not*, but rather is an independent formation with *n’t* being an affixal realization of the feature of negation.<sup>28</sup> Moreover, the degree of fusion can vary within a language, for alongside the fused *-n’t* in English there is the more independent free form *not*, seen in (12a/b); however, even *not* shows some degree of dependency, as it is generally separable from most

auxiliaries, as in (12a/b), but nonetheless can not be separated from *do*, as shown by (12c) and the contrast between it and (12a/b/d):

- (12) a. John will definitely not win  
b. John must definitely not win  
c. \*John did definitely not win  
d. John definitely did not win.

Within the Balkans, a wide range of evidence for negative fusion is available. In South Slavic, for instance, there are isolated grammaticalized and fully unverbated forms in which the negative marker has fused with the verb, e.g. the prohibitive marker *nemoj* found in Macedonian, Bulgarian, and Serbo-Croatian, and the negative future marker based on ‘have’ found in Bulgarian in the form *njama* and Macedonian as *nema*. In addition, though, there are more productive ways in which negative fusion is evident in South Slavic. Alexander 1994 has shown, for example, that the prosodic behavior of the negative morpheme *ne* in verbal groups and clitic sequences exhibits some degree of fusion in that it forms a single prosodic domain with other elements but is nonetheless quite readily analyzable as a separate element; the Serbo-Croatian negative future based on ‘want’, e.g. *neću* ‘I won’t’, is a case in point, since *ću* in general shows some synchronic clitic-like properties (though not necessarily in this combination) making the synchronic analysis of the form quite transparent. Some combinations are less parsable, such as the Serbo-Croatian negative of ‘be’ (e.g. *nísam* ‘I am not’, etc.), which has a full paradigm and a clear connection to nonnegated clitic forms of ‘be’, but shows a contraction of the negative with ‘be’ that gives a synchronically unpredictable result.<sup>29</sup>

In Greek, negative fusion is found with the indicative negator *ðen*, which, following the Zwicky & Pullum criteria, is best analyzed as verbal affix because, as discussed in Joseph 1990, it is fixed in its position, occurring on left margin of verbal complex, it is restricted to being only a verbal negator, and it shows some semantic idiosyncrasies, e.g. in the

expression *ðen mu les* which means ‘by the way’ but is literally “you don’t say,” and thus is negative in form without any negative semantics.

Moreover, a further argument can be developed for the affixhood of *ðen*: it cannot be doubled in and of itself even when the semantics of doubled negation are appropriate; thus, the Greek equivalent of ‘I don’t not smoke for health reasons but because I hate the taste’ cannot have *\*ðe(n) ðe(n)*, but must resort to a circumlocutory paraphrase. Also, the doubling of *ðen*: is not possible even though Greek allows the occurrence of two “slots” for negation elements when each has a different form; thus, *ðen* can co-occur with the nonindicative negation element *mi* in the combination *na mi ðen* that can occur for some speakers in a negative complement to *fováme* ‘I fear’, as in (13):<sup>30</sup>

(13) *fováme na mi ðen érqí* (Veloudis 1982:11)

fear/1SG SUBJUNC *mi* not come/3SG

‘I am afraid that he may not come’

This situation is parallel to the argument Zwicky 1987 gave for the English possessive ‘s as a (phrasal) affix, based on what he terms a “shape condition” that blocks a phonological form, something that in his conception of grammar, with a “phonology-free syntax” in which syntactic rules cannot make reference to phonological elements, ought to be a matter of morphology and not syntax.

Nonetheless, despite being generally affixal and thus a fused element, *ðen* does seem to have some independence in the sense that it can be picked out in a “mention” function; that is, one can say *íse ólo “ðen” símera* ‘You are quite negative today’ (literally: “You are all “ðen” today”), a possibility which seems somewhat anomalous for affixes, though admittedly perhaps not impossible. The English equivalent *You are all “n’t” today* is distinctly odd, suggesting that fused (i.e. morphological) forms in general are not available to be mentioned.<sup>31</sup>

Thus in Balkan Slavic and Greek, at least, negative fusion can be found to varying degrees. Such a situation typically makes one wonder if there is something “Balkanological” going on here, that is something of interest for comparative Balkan syntax in the sense defined above in section 3. Here, the answer is probably not, for the usual reasons: the observed negative fusion could be an inherited tendency and moreover it is typologically quite a “natural” phenomenon.

For one thing, a consideration of a broader range of Slavic data shows that fusion between the negative marker and a verb, especially with the verbs ‘have’ and ‘be’, is widespread in non-Balkan Slavic, including both East and West Slavic. Thus one finds *njama* ‘there is no’, from the verb ‘have’, in Belorussian, *ne* treated as a prefix on the verb in both Czech and Slovak, special negated forms in Sorbian of ‘want’, ‘have’, ‘can’, and ‘be’, and so forth.<sup>32</sup> Moreover, this situation is not surprising, since the oldest available Slavic, Old Church Slavonic, shows fused forms of ‘be’, e.g. *nfsmb* ‘I am not’ (= \**ne* + *esmi*).<sup>33</sup>

Looking more widely yet, one can note cases of negative fusion throughout Indo-European, such as Latin *nolo* ‘I do not want’ (= \**ne* + *wolo* ‘want’), or Old Irish *ní* ‘is not’ (from \**nést* < \**ne* + *est*). Sometimes, the details of the fusion, even the apparently very common sort with ‘be’, are such that it must have happened independently. Thus, Lithuanian *nerà* ‘is not’, with a synchronically irregular contraction of *ne* ‘not’ plus *yrà* ‘is’, occurred within Baltic after the innovative replacement of inherited \**H<sub>1</sub>esti* (cf. Old Lithuanian *èsti* ‘is’) by *yrà* as the third person singular present form of ‘be’. Finally, the fusion was not complete in all the languages: in Gothic, the negative marker *ni* is dependent, occurring as an enclitic between a lexical preverb and the verb in forms such as *miπ-ni-qam* ‘did not come with’.

Therefore, negative fusion can occur independently, and so could have arisen on its own in each of the Balkan languages that show it. Furthermore, there is direct evidence that it has

occurred independently in the Balkans, for in early 20th century Tsakonian, often called a dialect of Greek but divergent enough to perhaps warrant at that time at least being called a different language, a “negative” auxiliary verb ‘be’ developed. This negative auxiliary is a crucial piece of verbal system, since ‘be’ is used with a participle to form ordinary present tense. The relevant forms are given in (14), where some irregularities can be noted that argue against a ready synchronic analysis of at least some negated forms into the synchronic negative marker *o* with a positive form of ‘be’.<sup>34</sup>

(14)	Positive				Negative (= NEG + Positive)			
	1SG	éni	1PL	éme	1SG	óni	1PL	óme
	2	ési	2	ét <sup>h</sup> e	2	ósi	2	ót <sup>h</sup> e
	3	éni	3	ín’i	3	ón’i	3	ún’i

The synchronic negator *o* derives from earlier *u* (Ancient Greek < ouj >) and most likely was extracted out of a contraction of / u / plus vowel-initial forms of verbs (e.g. auxiliaries in the present and imperfect tenses, the prefixed “augment” past tense marker in the simple past). Thus the contraction in the negative forms is regular diachronically, deriving from \*/u + e/ and 3PL \*/u + i/, and for all but the third person plural form is synchronically regular also as /o + e/; but for the third person plural form, the contraction is not regular synchronically, for /o + i/ would not be expected to yield [u]. Similarly, reduced forms of ‘be’ show synchronic irregularities: the third person singular positive form is *én* or *n*, but the corresponding negative is *ó*, with no [n].

It must be concluded, therefore, that these facts concerning negative fusion in the Balkans are interesting from the perspective of the “linguistics of the Balkans” or the “comparative syntax of the Balkans,” but not from the perspective of “Balkan Linguistics” or “comparative Balkan syntax.” Each language reveals an interesting phenomenon but its occurrence in each language need not be attributed to language contact

in any form, and thus is not immediately relevant to the concerns of the investigation of the Sprachbund as a contact-induced phenomenon.

## **5. Conclusion**

The extended examples in the previous section show the virtues of keeping both types of pursuits in mind, and clearly, researchers must be cognizant of both. Both enrich our understanding of language in general and of the languages in the Balkans specifically. The more we know about language in general, the better able we are to judge the particulars of the languages of the Balkans, both in terms of how these languages fit into the general domain of natural human languages and in terms of how they fit into the more specific domain of “Balkan languages.” Moreover, the two approaches work well together; pursuing the comparative syntax of Balkan languages identifies possible candidates for comparative Balkan syntax, and when working in the Balkans one must always keep the Sprachbund in mind, even if we end up learning more about individual languages when we feel free to reject language contact and focus just on the language-internal syntax of one language in comparison with the language-internal syntax of another.

## **Acknowledgments**

This paper was developed for the “After-GLOW” Workshop on Balkan Syntax, held in Athens in April of 1996. I would like to thank the workshop organizers, Angela Ralli and Irene Philippaki-Warbuton, for the invitation that gave me the opportunity to try these thoughts out on a knowledgeable audience, whose insightful comments were useful to me as I revised the paper for this volume. I also owe a debt of gratitude to Victor Friedman of the University of Chicago and Nick Nicholas of the University of Melbourne, who both also provided numerous important suggestions for improving the paper. Naturally, all are absolved of any complicity in remaining errors.

## References

- Alexander, Ronelle. 1994. "The Balkanization of Wackernagel's Law," *Indiana Slavic Studies* 7.1-8.
- Arvaniti, Amalia & Brian D. Joseph. To appear. "Variation in voiced stop prenasalization in Greek." To appear in *Glossologia* 12 (1998)
- Banfi, Emmanuel. 1985. *Linguistica Balcanica*. Bologna: N. Zanichelli.
- Campbell, Lyle., Terence Kaufman, & Thomas Smith-Stark. 1986. "Meso-America as a linguistic area," *Language* 62.3.530-570.
- Comrie, Bernard & Greville Corbett (eds.). 1993. *The Slavonic Languages*. London/New York: Routledge.
- Duro, Ilo & Ramazan Hysa. 1981. *Fjalor Shqip-Englisht*. Tiranë: Shtepia Botuese "8 Nëntori."
- Friedman, Victor. 1983. "Grammatical categories and a comparative Balkan grammar." In Norbert Reiter, ed., *Ziele und Wege der Balkanlinguistik*. Berlin: Südosteuropa Gesellschaft.
- Gumperz, John & Robert Wilson. 1971. "Convergence and creolization. A case from the Indo-Aryan/Dravidian border in India." In Dell Hymes, ed., *Pidginization and Creolization of Languages*. Cambridge: Cambridge University Press, pp. 151-67.
- Haider, Hubert & Martin Prinzhorn (eds.) 1986. *Verb Second Phenomena in Germanic Languages*. Dordrecht: Foris.
- Horrocks, Geoffrey. 1997. *Greek. A history of the language and its speakers*. London & New York: Longman.
- Jaeggli, Oswaldo & Kenneth Safir (eds.). 1989a. *The Null Subject Parameter*. Dordrecht: Kluwer.
- Jaeggli, Oswaldo & Kenneth Safir. 1989b. "The null subject parameter and parametric theory." In Jaeggli & Safir (1989a:1-44)
- Janda, Richard D. & Brian D. Joseph. 1986 "One Rule or Many? Sanskrit Reduplication as Fragmented Affixation." In *Studies on Language Change. Ohio State University Working Papers in Linguistics* 34.

- 84-107. [Revised version of paper first published in *ESCOL '85. Proceedings of the Second Eastern States Conference on Linguistics*, 103-119. Columbus: OSU Department of Linguistics].
- Janda, Richard D. & Brian D. Joseph. 1989. "In Further Defense of a Non-Phonological Account of Sanskrit Root-Initial Aspiration Alternations." In *ESCOL '88. Proceedings of the Fifth Eastern States Conference on Linguistics*, 246-260. Columbus: OSU Department of Linguistics.
- Janda, Richard D. & Brian D. Joseph. 1992. "Meta-Templates and the Underlying (Dis)Unity of Reduplication in Sanskrit." In *ESCOL '91. Proceedings of the Eighth Eastern States Conference on Linguistics*, 160-173. Columbus: The Ohio State University Department of Linguistics.
- Janda, Richard D. & Brian D. Joseph. 1996. "On Constellations in Morphology: Crosslinguistic Evidence and a Case Study from Modern Greek," unpublished manuscript.
- Janda, Richard D. & Brian D. Joseph. 1997. "The Modern Greek negator mi(n)(-) as a morphological constellation." Paper presented at the 3rd International Conference on Greek Linguistics, Athens (September 1997). To appear in the Conference Proceedings.
- Joseph, Brian D. 1980[1982]. "A new convergence concerning the Balkan loss of the infinitive." *Indogermanische Forschungen* 85.176-187.
- Joseph, Brian D. 1983. *The Synchrony and Diachrony of the Balkan Infinitive. A study in Areal, General, and Historical Linguistics*. Cambridge: Cambridge University Press.
- Joseph, Brian D. 1988. "Pronominal Affixes in Modern Greek: The Case Against Clisis," *Papers from the 24th Regional Meeting of the Chicago Linguistic Society*, 203-15. Chicago: Chicago Linguistic Society.
- Joseph, Brian D. 1990. "The Benefits of Morphological Classification: On Some Apparently Problematic Clitics in Modern Greek." In Wolfgang U. Dressler, Hans C. Luschützky, Oskar E. Pfeiffer, & John R. Rennison, eds., *Contemporary Morphology (Trends in Linguistics. Studies and Monographs 49)*, 171-181. Berlin: Mouton de Gruyter.



- Joseph, Brian D. 1991. "A diachronic phonological solution to the syntax of Vedic negative particles." In Hans Henrich Hock, ed., *Studies in Sanskrit Syntax. A Volume in Honor of the Centennial of Speijer's Sanskrit Syntax*, 113-122. Delhi: Motilal Banarsidass Publishers.
- Joseph, Brian D. 1992. "Diachronic Perspectives on Control." In Richard Larson, Sabine Iatridou, Utpal Lahiri, & James Higginbotham, eds., *Control and Grammatical Theory*, 195-234. Dordrecht: Kluwer Academic Publishers.
- Joseph, Brian D. 1994. "On Weak Subjects and Pro-Drop in Greek." In Irene Philippaki-Warbuton, Katerina Nicolaidis, & Maria Sifianou, eds., *Themes in Greek Linguistics*. Amsterdam: John Benjamins Publishing Co., pp. 21-32.
- Joseph, Brian D. 1999. "Textual Authenticity: Evidence From Medieval Greek." To appear in Susan Herring, Piet van Reenen, & Lena Schoesler (eds.), *Textual Parameters in Ancient Languages*. Amsterdam: John Benjamins Publishing Co.
- Joseph, Brian D. To appear. "Processes of Spread for Syntactic Constructions in the Balkans." Paper read at Conference on Balkan Linguistics: Synchrony and Diachrony, University of Thessaloniki, 1997; to appear in C. Tzitzilis, ed. *Balkan Linguistics: Synchrony and Diachrony* (University of Thessaloniki, Greece, 1999)
- Joseph, Brian D. & Richard D. Janda. 1988. "The How and Why of Diachronic Morphologization and Demorphologization." In Michael Hammond & Michael Noonan, eds., *Theoretical Morphology: Approaches in Modern Linguistics*, 193-210. San Diego: Academic Press.
- Kayne, Richard. 1991. "Romance Clitics, Verb Movement and PRO." *Linguistic Inquiry* 22.4:647-86.
- Lightfoot, David. 1991. *How to Set Parameters. Arguments from Language Change*. MIT Press.
- Mackridge, Peter. 1985. *The Modern Greek Language*. Oxford: Oxford University Press.
- Mackridge, Peter. 1990. "Katharevousa (c. 1800-1974): An obituary for an official language." In Marion Sarafis & Martin Eve, eds., *Background to Contemporary Greece*, 25-51. London: Merlin Press.
- Masica, Colin. 1976. *Defining a Linguistic Area*. Chicago: The University of Chicago Press.
- Morin, Yves-Charles. 1985. "On the two French subjectless verbs *voici* and *voilà*." *Language* 61.777-820.

- Morin, Yves-Charles. 1988. "French *voici* and *voilà*: A reply to Bouchard." *Language* 64.101-103.
- Nevis, Joel A., Brian D. Joseph, Dieter Wanner, & Arnold M. Zwicky. 1994. *Clitics. A Comprehensive Bibliography 1892-1991* Amsterdam & Philadelphia: John Benjamins Publishing Co.
- Newmark, Leonard., Philip Hubbard, & Peter Prifti. 1982. *Standard Albanian. A Reference Grammar for Students*. Stanford: Stanford University Press.
- Pernot, Hubert. 1934. *Introduction à l'étude du dialecte tsakonien*. Paris.
- Pollock, Jean-Yves. 1989. "Verb movement, Universal Grammar, and the structure of IP." *Linguistic Inquiry* 20.365-424.
- Rivero, Maria-Luisa. 1990. "The Location of Non-Active Voice in Albanian and Modern Greek," *Linguistic Inquiry* 21.135-46.
- Rivero, Maria-Luisa. 1994. "Clause Structure and V-Movement in the Languages of the Balkans," *Natural Language & Linguistic Theory* 12.1.63-120.
- Rivero, Maria-Luisa. & Arhonto Terzi. 1994. "Imperatives, Illocutionary Force, and V-Movement," unpublished manuscript.
- Roberts, Ian. 1997. *Comparative Syntax*. London: Arnold.
- Rudin, Catherine. 1988. "On Multiple Questions and Multiple WH Fronting," *Natural Language & Linguistic Theory* 6.445-502.
- Sandfeld, Kristian. 1930. *Linguistique balkanique. Problèmes et résultats*. Paris: Honoré Champion.
- Schaller, Helmut. 1975. *Die Balkansprachen. Eine Einführung in die Balkanphilologie*. Heidelberg: Carl Winter.
- Schaller, Helmut. 1977. *Bibliographie zur Balkanphilologie*. Heidelberg: Carl Winter.
- Terzi, Arhonto. 1992. *PRO in Finite Clauses. A Study of the Inflectional Heads of the Balkan Languages*. City University of New York Ph.D. Dissertation.
- Thrasher Jr., Randolph H. 1974. *Shouldn't Ignore These Strings: A Study of Conversational Deletion*. University of Michigan Ph. D. Dissertation.

- Thumb, Albert. 1964. *A Handbook of the Modern Greek Language. Grammar, Texts, Glossary*. Chicago: Argonaut, Inc. Publishers (reprinting of 1912 edition).
- Topolinjska, Zuzanna. 1995. *Macedonian Dialects in the Aegean Part of Macedonia, Book 1: Syntax, Volume I*. Skopje: Macedonian Academy of Sciences and Arts.
- Tsitsipis, Lukas. 1981. *Language Change and Language Death in Albanian Speech Communities in Greece. A Sociolinguistic Study*. University of Wisconsin-Madison Ph.D. Dissertation.
- van Windekens, Albert J. 1976. *Le tokharien confronté avec les autres langues indo-européennes. Volume I: La phonétique et le vocabulaire*. Louvain: Centre International de Dialectologie Générale.
- Veloudis, Ioannis. 1982. *Negation in Modern Greek*. University of Reading Ph.D. Dissertation.
- Wackernagel, Jacob. 1892. "Über ein Gesetz der indogermanischen Wortstellung," *Indogermanische Forschungen* 1.333-436.
- Weerman, Fred. 1989. *The V2 Conspiracy. A Synchronic and a Diachronic Analysis of Verbal Positions in Germanic Languages*. Dordrecht: Foris.
- Zwicky, Arnold M. & Geoffrey K. Pullum. 1983. "Cliticization vs. inflection: English *n't*," *Language* 59.502-13.
- Zwicky, Arnold M. 1987. "Suppressing the Z's," *Journal of Linguistics* 23.1.133-148.

---

<sup>1</sup>. My source for this characterization of Universal Grammar is David Perlmutter, based on class lectures he presented at MIT over 20 years ago which I was privileged to have been able to sit in on. I do not know if this view is original with him or if he was passing on what he had learned from someone else, but I am pleased to be able to acknowledge the role he played in sharpening my understanding of linguistic theory through such statements.

<sup>2</sup>. See, for instance, the papers in Jaeggli & Safir (1989a).

<sup>3</sup>. For example, it is not at all obvious why there should be any link between the structure of verbal paradigms and the possibility of null subjects. Moreover, Jaeggli and Safir themselves, in the face of a counter-example from Dutch (discussed in their footnote 19),

---

retreat somewhat and suggest that perhaps “up to one stem identical form [in a paradigm with stem + affix forms otherwise/BDJ], excluding imperatives, is permitted” (p. 40).

<sup>4</sup>. As discussed by Thrasher 1974, sentences like:

- (i) Seems like no one cares! (= It seems like no one cares!)
- (ii) Can't get there from here! (= You can't get there from here!)

suggest that English has some pro-Drop-like structures, but as he points out, more than just subjects can be deleted:

- (iii) Gotta run! (= I 've gotta run)
- (iv) Cold? (= Are you cold?)
- (v) Guy over there is crazy! (= The guy over there is crazy!).

Moreover, such strings are not possible in subordinate clauses (e.g. \*John warned Mary that  $\emptyset$  can't get there from here) suggesting that pro-Drop is not at work in these English utterances. Still, when faced with just sentences like (i) and (ii), an English-speaking child conceivably could develop an analysis akin to a Null-Subject analysis, making it unclear what it means to speak of a “pro-Drop” or “Null-Subject” language.

<sup>5</sup>. See Joseph 1994 regarding the retention of weak subject pronouns in one Greek construction, a locative interrogative construction with *pún* ‘where is/are?’ (e.g. *pún dos* ‘Where is he?’), even though Greek generally is a well-behaved typical pro-Drop language; similarly, Morin 1985, 1988 analyzes *voici* and *voilà* ‘here is/are’ in French as subject-less predicates in an otherwise non-pro-Drop language.

<sup>6</sup>. See Nevis et al. 1994 for access to the literature on Romance clitic studies and Germanic V2 up through 1992. Roberts 1997 provides some updating on more recent work, with discussion.

<sup>7</sup>. See for instance Schaller 1975, 1977 and Banfi 1985 for references.

---

<sup>8</sup>. There is no really suitable widely-agreed upon English term, though “linguistic area” is sometimes used (cf. Campbell, Kaufman, & Smith-Stark 1986, for example).

“Convergence area” probably conveys the meaning best, but I adopt the German term nonetheless, following the vast majority of scholars writing in English in this practice.

<sup>9</sup>. It is often the case that languages in a Sprachbund are not related to one another or at least not closely related; in the case of the Balkans, although most of the relevant languages are Indo-European, they represent different subgroups (branches) of the Indo-European family.

<sup>10</sup>. As this last trait shows, part of what makes languages in a Sprachbund of considerable interest is not just the fact that they converge on one another but that this convergence represents a divergence from their previous stages and from other genetically and geographically related languages. Furthermore, for Sprachbund members that are genetically related, the convergent features are not a matter of a shared inheritance from their common ancestor.

<sup>11</sup>. Language abbreviations here are “Alb” for Albanian, “Blg” for Bulgarian, “Grk” for Greek, “Rmn” for Romanian; elsewhere, “Mac” is used for Macedonian and “Trk” for Turkish.

<sup>12</sup>. As Eric Hamp has remarked on occasion, Proto-Indo-European did not even have a definite article, so the postposed article cannot possibly be an inheritance from the common ancestor to these languages. Similar observations hold for the other features cited here.

<sup>13</sup>. For example, by Sandfeld (1930: 106), who discussed this convergence in the context of seeking an explanation for the apparently innovative occurrence of a ‘have’-plus-participle pluperfect in Aromanian and Megleno-Romanian, the most centrally Balkan varieties of Romanian, as opposed to its absence in Daco-Romanian). The existence of parallel forms in Greek, Albanian, and Macedonian led Sandfeld to suspect language contact as the source of

---

the Aromanian and Megleno-Romanian formation, and he eventually concluded that “il semble préférable de penser à une influence grecque” (p.106).

<sup>14</sup>. That is, passive when formed from a transitive verb, and active when formed from an intransitive verb.

<sup>15</sup>. Admittedly, in the context of a comparison with English, the occurrence of a past perfect with ‘have’ and a participle may not seem remarkable, but it is noteworthy that Romance languages typically use ‘have’ with a participle for a regular past tense (e.g. the French *passé composé*, the (Daco-)Romanian *perfectul compus*) and not for a perfect tense, and that the ‘have’ perfect is not found elsewhere within Slavic.

<sup>16</sup>. The Greek pluperfect cited here consists of the past of ‘have’ with a (generally passive) participle; there is also a pluperfect, innovated in the Medieval Greek period, formed with ‘have’ followed by an invariant verb form which historically continues the older infinitive, but which synchronically may be nothing more than a variant participial form, e.g. *ίχα δέσι* ‘I had tied’, as discussed in Joseph 1983. There are present perfect forms corresponding to these pluperfects in Albanian, Macedonian, and Greek (and even a future perfect in Greek), but in Greek at least it seems that pluperfect was the starting point for this type of perfect (see Joseph 1999 for some discussion).

<sup>17</sup>. In Turkish, the usual meaning of the juxtaposition of the aorist with a negative aorist form is ‘as soon as ...’, e.g. *gelir gelmez* ‘as soon as (s)he comes’. While this anomaly regarding *ister istemez* may point to that pattern with that particular root being a borrowing into Turkish, it must be noted that the verb *ol-* ‘be’ also shows an anomalous meaning for the aorist-plus-negative-aorist construction (*olur olmaz* means ‘any old ...; just any ...’), suggesting that irregularities here may well be a function of the high frequency of these verbs. Clearly, the Balkan side of this construction needs further investigation (see Joseph (to appear) for some suggestive discussion).

---

<sup>18</sup>. Including much of my own earlier work, I regret to relate.

<sup>19</sup>. The Arvanitika facts and the analysis given here are taken from Tsitsipis (1981:347); see also Joseph 1992 for further discussion.

<sup>20</sup>. The invariant ending *-ondas* is the old accusative singular of the present active participle apparently with the masculine nominative singular ending *-s* added on (though Horrocks 1997 also suggests that the *-s* may be an adverbial marker, as in dialectal Greek *tótes* ‘then’ versus Standard Greek *tóte*, which would be motivated by the circumstantial use of the participle). The accusative origin of the ending indicates that nonsubject control was indeed once possible (as a participle the form agreed in case with the nominal it was associated with), and it can be speculated that the innovative further characterization of the ending with a nominative desinence coincided with the form being restricted to subject-only control of its understood subject.

<sup>21</sup>. Admittedly, the development of the Tocharian form is not completely straightforward, according to some accounts of Tocharian historical phonology (especially that of van Windekens 1976; see Joseph 1991 for some discussion). Whatever the prehistory of Tocharian *má*, the reconstruction of PIE *\*mé* is secure.

<sup>22</sup>. The Albanian forms and sentences in (8) are taken from Newmark, Hubbard, & Prifti 1982 and Duro & Hysa 1981; the sources for the Greek are given where appropriate, with all other Greek data coming from consultation with native speakers.

<sup>23</sup>. The terms “demotic” and “katharevousa” refer respectively to the low and high style varieties of Greek that functioned in a diglossic relationship for much of the 19th and 20th centuries in the Greek-speaking world. The details of this relationship, which permeates Greek linguistics even today when katharevousa no longer has the official status it held into the 1970’s, are not of concern here; see Mackridge 1990 for some recent discussion of the history and resolution of Greek diglossia.

---

<sup>24</sup>. Basically the voiceless stops, which typically become voiced after the nasal and induce place assimilation on the nasal; for some speakers, the nasal can then be deleted under complex partly sociolinguistically governed conditions — see Arvaniti & Joseph (To appear) for some discussion. Before fricatives, the nasal can appear, mainly in careful speech.

<sup>25</sup>. Janda & Joseph 1996, 1997 discuss this issue, proposing the use of the construct they refer to as the “(morphological) rule constellation” to capture the simultaneous similarities and differences in the various realizations of *mi(n)*; see, e.g., Janda & Joseph 1986, 1989, 1992, and Joseph & Janda 1988 for more discussion of “constellations.”

<sup>26</sup>. Thus, *má* can occur independently (though often repeated, as *má má*) but only in the meaning ‘Not so!’, a somewhat emphatic negation, not a prohibitive. There is also an elliptical use of *má* that is prohibitive in value, e.g. *má svabdam* ‘Not a word!’ (where the accusative form of *svabda-* suggests a missing governing verb).

<sup>27</sup>. Admittedly, this feature would be found just in two languages, though with possibly more information Romany could be added to the list; still, there are Sprachbund features that are not found in all the languages, e.g. the postposed definite article is absent from Greek, so that it is not essential that a “Sprachbund feature” be found all over the region.

<sup>28</sup>. I find Zwicky and Pullum’s argumentation quite compelling, and so opt for the affixal analysis of *-n’t* here; facts such as the totally idiosyncratic morphophonemics for the shape of *will* in combination with *-n’t* are part of their evidence, for such idiosyncrasies are more typical of affixal combinations than clitic ones. The historical origin of *won’t* as a reduction of *will not* is irrelevant to the synchronic analysis.

<sup>29</sup>. In that way, *nísam* is somewhat like *won’t* in English, discussed in the previous footnote.

<sup>30</sup>. As discussed in Section 4.1, the *mi* that occurs in these complements to verbs of fearing is pleonastic.



---

<sup>31</sup>. Bound forms can, however, occasionally be “liberated” and take on independent status, as with the use of *ism* in English as a free noun meaning a ‘a distinctive doctrine, system, or theory’ (*American Heritage Dictionary* s.v.), extracted from nouns such as *socialism*, *communism*, etc. Thus mention might be possible in principle.

<sup>32</sup>. See the sketches in Comrie & Corbett 1993 for relevant details.

<sup>33</sup>. The pan-Slavic nature of negative fusion, especially with ‘be’ and ‘have’, need not mean that any given fused form in a Slavic language must be an old feature, and it is not clear just how much of this phenomenon can be or even should be reconstructed for Common Slavic. Still, its occurrence in earliest attested Slavic and its widespread nature would suggest that at least some of the fused forms are old, though they may well have provided a pattern for the innovative spread of fusion to other verbs. I am grateful to Daniel Collins and Charles Gribble for helpful discussion on this issue.

<sup>34</sup>. All Tsakonian data, and most of the interpretations, are taken from Pernot 1934.