

Romanian and the Balkans: Some Comparative Perspectives

Brian D. Joseph

The Ohio State University

1. Introduction

Romanian occupies a position that makes it of particular interest to linguists with comparative interests, for it allows for several crucial types of comparisons to be made. First, comparisons of a genetic nature are possible, pertaining to Romanian in the context of those languages most closely related to it, and thus looking at Romanian vis-à-vis the other Romance languages. Second, comparisons can be made that are of a geographical nature, pertaining to Romanian in the context of its linguistic neighbors, and thus looking at it in relation to the other languages of the Balkans. These two types of comparison are in addition to general typological comparisons that can be made between any two (or more) languages, in the interests of pursuing linguistic universals and determining the nature of Universal Grammar.

In a sense, just as general typological comparisons can take any language(s) as a starting point, so too can the first two types of comparison, genetic and geographic, be made for any language. In particular, with the exception of the relatively few generally agreed-upon language isolates, e.g. Basque or Sumerian,¹ there will always be genetic relatives of a language with which to make comparisons.² Similarly, there are always linguistic neighbors somewhere within or alongside the territory in which a given language is spoken, if even only at its outer reaches.

What makes Romanian of special interest in this regard is the fact that its closest relatives — the members of the Romance group of languages — are so well-known and so well-studied that

¹Recent attempts to connect these languages to other known language families notwithstanding, I maintain here the *communis opinio* that these are genetically isolated languages, given our current state of knowledge. See Trask 1997 for some discussion of Basque as an isolate.

²And even for isolates there are comparisons to be made among different dialects of the language.

illuminating comparisons can be made on matters of linguistic detail that might not be possible within other genetic groupings that are less well-documented. As for its geographic setting, Romanian again is of special interest since along with most of its neighbors in the Balkans, it participates in numerous areally based similarities that define the Balkan *Sprachbund*,³ an area where long-term intense and intimate contact among speakers of several different languages has led to massive structural convergence.

The languages that participate to some significant extent in the Balkan *Sprachbund* convergences can be referred to as “Balkan languages”⁴ and besides Romanian include Albanian (both major dialects: Geg (North) and Tosk (South) but especially Tosk); Bulgarian; Greek; (Slavic) Macedonian;⁵ Romany (the language of the Indic Gypsies), Serbo-Croatian⁶ (with the

³I use *Sprachbund* as a technical term rather than any of its clumsy and infelicitous possible translations such as “linguistic union” or “language league”; “linguistic area” is sometimes used in English, but “convergence area” probably comes closest to being a suitably apt term.

⁴This is intended as a technical term, following the distinction drawn by Schaller 1975 between “Balkan languages”, those that are Sprachbund members, and “languages of the Balkans”, a purely geographic classification that includes all the Balkan languages but also (not counting very recent migrations, e.g. by Arab speakers into Greece, or adventitious occurrences, such as American English speakers residing in Greece): Armenian (spoken in Bulgaria); Circassian (Adygey variety; spoken in the Kosovo area of (former) Yugoslavia); German (spoken in Romania); Hungarian (spoken in Romania); Italian (spoken in the Istria area of (former) Yugoslavia); Judezmo (the dialect of Spanish spoken in parts of Greece by former Iberian Jews, also known as Judeo-Espagnol or occasionally, and perhaps erroneously, Ladino); Ruthenian (also known as Rusyn, spoken in Vojvodina area of (former) Yugoslavia and perhaps actually a dialect of Ukrainian); Slovak (in a small enclave in Vojvodina area of (former) Yugoslavia); and Slovenian.

⁵I say “Slavic Macedonian” to distinguish it from Ancient Macedonian, an entirely different language possibly closely related to Greek and spoken in the first millennium B.C. also in the

Torlak dialects of Southeast Serbia being most relevant), and Turkish (though mostly relevant for its lexical contributions). Actually, in counting Romanian among the Balkan languages, the greatest attention belongs to Aromanian, spoken in pockets in northern and central Greece, in Albania, and in Macedonia, and to Megleno-Romanian, spoken in a few areas in northern Greece; Daco-Romanian is a Balkan language to some extent, but not fully so, and Istro-Romanian is largely irrelevant as far as the Balkan *Sprachbund* is concerned. The extent of participation in the *Sprachbund* for some of these Romanian varieties⁷ is discussed below in section 3.

There are many features on which various of the Balkan languages agree and have converged,⁸ covering all levels of linguistic structure, but among the most salient and most widely discussed ones are those listed in (1) to (3); for a fuller discussion, the reader is referred to various handbooks such as Sandfeld 1930, Schaller 1975, and Banfi 1985, among others:⁹

Balkans; sometimes the autonym “Makedonski” is used in English for the modern Slavic language, with “Macedonian” being reserved for the ancient language.

⁶Now more usually referred to separately as Serbian and Croatian, as well as Bosnian; I adopt here the still traditional joint designation though the actual labels are not significant for the purposes of this paper.

⁷I take no position here on whether or not these varieties represent four distinct but closely related languages or instead are dialects of a single language. While I am inclined toward the former view, nothing crucial in this discussion hinges on that decision, and my use of the term “Romanian” herein should be taken to mean Daco-Romanian, in the unmarked case.

⁸Indeed, one of the most fascinating aspects of the Balkan *Sprachbund* is that virtually all of the significant structural convergences represent also significant *divergence* from earlier states of each of the languages.

⁹There is considerable debate on the validity of several of these features as true *Sprachbund* features; for example, there may not be an adequate language contact scenario to account for the appearance of a mid-to-high central vowel in many of these languages and it does have different historical sources in the languages that show it. Also, some are found outside the Balkans (e.g.

- (1) a. Balkan Sprachbund phonological features
- absence of “overlay” features, e.g length and nasalization, in the articulation of vowels
 - presence of a mid-to-high central vowel
- b. Balkan Sprachbund morphological features
- enclitic definite article
 - invariant future tense marker derived from a verb meaning ‘want’
 - merger of genitive and dative case
 - analytic comparative
- c. Balkan Sprachbund syntactic features
- use of a special verb form to indicate confirmativity/evidentiality
 - cross-indexing (“doubling”) of direct and indirect objects by a weak (“clitic”) pronoun
 - use of finite complementation in place of infinitives

These features are not realized in all of the Balkan languages nor are they necessarily found uniformly in the languages that show them to any extent; still, even if no single one is criterial for membership in the *Sprachbund*, the overlapping clusters of these features in the various languages serve to define the *Sprachbund*.

Recognizing the Balkan *Sprachbund*, then, allows for comparisons of a particularly interesting sort to be made for Romanian, taking in not just genetic and typological comparisons but also those having to do with the special character of language contact in the Balkans and thus with what can be revealed about the historical development of Romanian on its own and in relation to the other Balkan languages.

Scandinavian languages and Northern Russian dialects have developed a postpositive definite article, and English has a future based on a volitional verb), so that presumably they can develop independently in a given language. Nonetheless, these features are listed here as representative of what is typically cited in the literature.

Accordingly, several instances of comparisons involving Romanian and its Balkan neighbors, especially Greek and Albanian, are discussed here. Inasmuch as the most interesting convergences in the Balkans involve morphology and syntax (see note 9 regarding some of the putative phonological Balkanisms), the examples presented here focus on (morpho-)syntax. These comparisons are undertaken with an eye toward illuminating which are of interest for the purely historical purpose of determining the diachrony of Romanian and its neighboring languages, which are of relevance to understanding the nature of language contact in the Balkans and in general, and which have significance for furthering our knowledge of Universal Grammar. These various goals for comparison cut across the three types of comparison discussed above: genetic, geographic, and typological, and thus can draw on all three.

2. A Possible Substratal Effect in Romanian Syntax

The first example to be discussed involves a syntactic anomaly within Romanian prepositional syntax that is paralleled in Albanian.¹⁰ Somewhat idiosyncratically for prepositions in Romanian, which normally allow only an indefinite noun phrase as object, as in (2):

- (2) a. *din cladire* ‘from a/the building’

from building/INDEF

- b. **din cladierea*

building/DEF

the preposition *cu* ‘with’ requires a definite article for an unmodified nominal object and allows a definite article for a modified object:

- (3) a. *cu prietenul* ‘with the friend’

with friend/DEF

- b. **cu prieten* ‘with a friend’

friend/INDEF

¹⁰Noted, for example, in Schaller (1975:169), with some discussion of relevant literature, and in Banfi (1985:§3.6.6); the Albanian facts are taken in part from Newmark, Prifti, and Hubbard 1982.

- c. cu un prieten bun ‘with a good friend’
 a friend/INDEF good
- d. cu prietenul bun ‘with the good friend’
 friend/INDEF good

While there is no apparent synchronic reason for this special behavior of *cu*, so that most likely it requires lexical stipulation as being exceptional to the general pattern of only indefinite objects with prepositions, the comparison with Albanian, especially when taken together with some other facts about Romanian and Albanian, suggests an historical reason for the anomaly, which moreover sheds some light on the historical antecedents of Romanian itself.

In particular, in Albanian also, the preposition for ‘with’, *me*, is unusual in requiring a definite object, even though the object of most prepositions can occur with the definite article or not, according to meaning. Thus the contrast between *pa* ‘without’ in (4) and *me* ‘with’ in (5) is instructive in this regard:

- (4) a. pa sheqer ‘without sugar’
 without sugar/INDEF
- b. pa mjetet ‘without the tools’
 without tools/DEF
- (5) a. me shqiptarët ‘with (the) Albanians’
 with Albanians/DEF
- b. *me shqiptarë
 Albanians/INDEF.

Moreover, as in Romanian, Albanian allows indefinite objects with *me* when the noun is modified:

- (6) një lëvizje me karakter politik
 a movement with character/INDEF political
 ‘a movement with a political character’

There may well be some differences in the range of nouns for which use of the definite article after *me* and *cu* is required or impossible, as suggested by Boretzky 1968. Still, this agreement that

Albanian and Romanian show with regard to a synchronic anomaly involving prepositions and definiteness is striking and therefore quite intriguing, and is something that demands an explanation at least in historical terms even if an illuminating synchronic account is not possible.

This parallel seems to be too precise to be due to chance, and it certainly cannot be a matter of lexical borrowing since the form of the preposition is different in each language (*cu* vs. *me*). It is possible that it is a structural borrowing from one language into the other or even a calque (loan-translation), but most of the structural parallels that are found between Romanian and Albanian are ones that are shared with other Balkan languages (e.g. regarding the infinitive or the falling together of dative and genitive cases) and thus are more likely to be the result of the various processes that led to overall convergence in the Balkan languages, or are else have a Latin origin (thus inherited in Romanian and the result of heavy Latin influence on early Albanian). While it has been suggested (e.g. by Reichenkron 1962) that there is a common Romance source for the prepositional parallel under consideration here, one other possibility presents itself as a compelling source for this Romanian-Albanian convergence, namely that of spread from a substratum.

The spread of substratum elements into a second language presumably takes place through a period of language shift, as speakers of one language carry over traits of their primary language into the language they are shifting to. Although substratum effects are real enough, as is evident from observations of second-language acquisition, whether in tutored (e.g. academic) environments or in a natural untutored learning situation, substratum accounts in historical linguistics are often hard to maintain since they generally seem to be little more than a last resort “explanation” rather than anything substantive.¹¹ Still, Thomason & Kaufman 1988 have argued that in real cases of

¹¹Indeed, some substratum explanations have been given for various Balkan features, but without much credibility. For example, it has been claimed that the loss of the infinitive is due to the workings of a prehistoric Thracian or Illyrian substratum, but the chronology of the replacement of earlier infinitives by finite complementation argues against this, for that was an innovation that took place within the fairly recent (for the Balkans) historical period of the Post-Classical period and on

substratum interference, one typically finds effects in both syntax and phonology, and this observation then provides a basis for judging this particular Romanian-Albanian convergence, since overall the special parallels between these two languages fit this pattern fairly well.

In particular, in addition to this syntactic parallel involving certain prepositional objects, there is other evidence for the view expressed by Hamp (1989: 47) that “historically Romanian is Latin spoken with an Albanian stress system”, i.e. with the “Danubian Late Latin” of Dacia being filtered through the grammars of one group of Proto-Albanian (what Hamp sometimes calls “Albanoid”) speakers. This evidence consists of striking parallels in stress placement and stress shift, whereby both languages have stress on the final syllable of the stem of the lexeme, so that the stress shifts to the right when a derivational suffix is added, e.g. agentive *-tor-* in Albanian or deverbal adjectival *-tor-* in Romanian, but not when an inflectional ending, e.g. Albanian dative plural *-eve* or Romanian definite dative singular *-lui* or feminine singular *-e* is added:

(7) Albanian

katúnd ‘village/NOM.SG’ / katúndeve ‘to villages/DAT.PL’

púnë ‘work/NOM.SG’ / punëtór ‘worker/NOM.SG’ / punëtóreve (DAT.PL)

(8) Romanian

cîine ‘dog/NOM.SG’ / cîinelui ‘to the dog/DEF.DAT.SG’

a folosí ‘to use’ / folositór ‘useful/MASC.SG’ / folositoáre (FEM.SG)

Thus these striking parallels can be explained as the result of substratum influence in Romanian, as “Albanoid”/Proto-Albanian speakers shifted to the Latin spoken in Dacia and carried these aspects of their native syntax and prosodic system into emerging Romanian. The fact then that there is both an accentual parallel and a syntactic parallel is highly suggestive of a substratum, and the fact that only two languages are involved — Romanian as the shifted-to language and (modern) Albanian as the sibling language of that spoken by the shifting population — makes a substratum scenario quite plausible. Thus while a synchronic explanation for the definiteness effect with *cu*

into the Medieval period, inasmuch as infinitives persist into the Medieval period for all of the Balkan languages; see Joseph 1983, especially chapter 7, for discussion.

and *me* may still be lacking, some historical perspective is possible, thus enriching our understanding of the emergence of Romanian and its prehistoric connections with Albanian (or “Albanoid”) speakers. The comparisons made here with Albanian, both as to the primary focal point regarding ‘with’ phrases and to the secondary comparison made regarding accent rules, therefore together serve to illuminate an otherwise obscure aspect of Romanian prepositional phrase syntax.

3. Comparisons within a Sprachbund and a Romance Context

The second goal given above for comparison, namely extending our understanding of the Balkan Sprachbund and thus shedding light on the nature of language contact in the Balkans, emerges from the parallel seen in the formation of compound tenses in Romanian and other Balkan languages, especially in regard to perfects formed with ‘have’ as an auxiliary. A consideration of this feature provides some evidence bearing on the question of just how “Balkan” Daco-Romanian is, as opposed to other varieties of Romanian, and in particular Aromanian and Megleno-Romanian, and also clarifies the place of different varieties of Romanian in an overall Romance context. In this way, therefore, some perspective is gained on the characterization of Daco-Romanian as a Romance language as opposed to a Balkan language.¹²

At issue is the composition and value of a perfect tense formation made up of ‘have’ and a participle that independently functions as a passive participle to transitive verbs and an active participle to intransitives. Of particular interest is the fact that in Aromanian and Megleno-

¹²I speak here somewhat figuratively, for the terms of comparison are not equivalent nor are they mutually exclusive: “Romance” is a genetic classification while “Balkan” is rather a geographic and sociolinguistic classification (sociolinguistic in the sense of defining a zone of a special type of contact among speakers). A presupposition here, one which admittedly may not be warranted, is that there is a general “type” that one finds among Romance languages and which therefore is a norm against which all genetically Romance languages can be measured.

Romanian, as in Albanian, Greek and Macedonian, there is a past perfect formation, utilizing the past of ‘have’ as the auxiliary verb preceding the participle, as exemplified in (9):

- (9) Albanian: kisha lidhurë ‘I had tied’
 Greek: ίχα ἔμμένο ‘I had tied’
 Macedonian: imav storeno ‘I had made’
 Aromanian: avea mîcatǎ ‘(s)he had eaten’
 Megleno-Romanian: vea durmit ‘(s)he had slept’

While other forms of such a ‘have’ perfect system can be found, e.g. a present perfect such as Greek *έχο ἔμμένο* ‘I have tied’,¹³ it is the past perfect in general that seems to be historically prior — and is demonstrably so in Greek at least (see Joseph 1983, 1997) — and which has been the focus of Balkanological interest.¹⁴

In Daco-Romanian, by contrast, the only relevant comparable verbal formation, consisting of an auxiliary form of ‘have’¹⁵ and a participle with a similar valency to the Greek, Albanian, etc. participle, is the *perfectul compus* (“compound perfect”), a full paradigm for which is given in (10) for the verb *a lucra* ‘to work’:

- (10) 1SG am lucrat 1PL am lucrat
 2 ai lucrat 2 ați lucrat
-

¹³Greek also shows a perfect formation, an innovation dating to the Medieval Greek period, 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 ἔési* ‘I had tied’, as discussed in Joseph 1983, 1997.

¹⁴The Balkan character of this parallel was discussed by Sandfeld (1930:105-6).

¹⁵The paradigm of ‘have’ in this formation overlaps with that of the present tense of the main verb ‘have’, differing in the 3SG (*a* vs. *are* ‘(s)he has’), 1PL (*am* vs. *avem* ‘we have’), and 2PL (*a,i* vs. *ave,i* ‘you (all) have’). The forms are historically related, to be sure, so that treating the auxiliary in this formation as belonging synchronically to the paradigm of ‘have’ is certainly plausible, though not necessarily required (see also note 16).

This tense has a similar form to the Balkan perfect in (9), though admittedly with an apparent present form of the auxiliary,¹⁶ but importantly, in general it instead has the meaning of a simple past, and not a perfect; thus, the forms in (10) mean ‘I worked; you worked, (s)he worked, etc.’ and not ‘I have worked, etc.’. In this way, the Daco-Romanian ‘have’ periphrastic tense is more like the French *passé composé*, e.g. *j’ai travaillé* ‘I worked’, which has ousted the simple (monolectal) preterite, the so-called *passé simple*, from “normal spoken French” (Harris 1987:221),¹⁷ and it also shows an affinity, moreover, with relatively recent developments in Castilian Spanish (so Green 1987:257) and in northern Italian dialects (Vincent 1987:279), in which a compound past tense that originated as a ‘have’-perfect has encroached upon the function of a simple preterite. In that way, therefore, Daco-Romanian seems more like a proto-typical member of the Romance language group than a member of the Balkan Sprachbund, in that it is taking part in this widespread Romance drift concerning the value of various tense formations¹⁸.

¹⁶I say this is an “apparent present form” since the paradigm is not identical with that of the present of ‘have’, though, as mentioned in note 15, it does overlap considerably with the present paradigm of the full verb. In any case, the auxiliary verb here is closer in form to the present of ‘have’ than to the past, which has the root *av-* running throughout the paradigm (e.g. 1SG *aveam* ‘I had’, 2PL *aveați*, etc.).

¹⁷To be sure, as Harris points out, the French *passé composé* can also have the value of a true perfect, but the fact that it has replaced the *passé simple* is what is significant here by way of comparison with Daco-Romanian.

¹⁸I say “drift” since the chronology of these changes in French, Spanish, and Italian — the Castilian changes are quite recent, for instance — means that they must be taken as independent but parallel innovations in these languages. Nonetheless, one can speculate that there is something in the set of structural oppositions in the verbal system inherited from Proto-Romance that might make such developments particularly natural, hence the product of a Romance “drift”.

Thus Daco-Romanian really does not match up well with the other Balkan languages with regard to showing evidence of a ‘have’ perfect system, since it has a formation that is superficially somewhat parallel in terms of its form but quite different in terms of its meaning. Moreover, when approached from the perspective of meaning, it is noteworthy that the Daco-Romanian form corresponding in meaning to a past perfect is not a periphrastic tense at all, but rather is a synthetic, i.e. monolectal, formation, as in (11):

- (11) 1SG *lucrasem* 1PL *lucraseăm* ‘I had worked, you had worked, etc.’
 2 *lucraseai* 2 *lucraseați*
 3 *lucrase* 3 *lucraseau*

Again, this situation seems more characteristic of Romance languages in general; Portuguese, for instance, has a synthetic past perfect (e.g. *falara* ‘I had spoken’) and a periphrastic perfect (with the verb *ter* ‘have’ as the auxiliary, e.g. *tenho tomado* ‘I have taken’).¹⁹ Furthermore, in the one part of what might be viewed as a perfect “system” in Daco-Romanian in which one finds a periphrastic formation, namely the future perfect, the auxiliary is the verb ‘be’, *a fi*, and not ‘have’, as in (12):²⁰

- (12) *voi fi lucrat* / *vom fi lucrat* (etc.)
 will/1SG be/INF work/PPL will/1PL be/INF work/PPL
 ‘I will have worked’ ‘we will have worked’

¹⁹Interestingly, as Parkinson (1987:269) notes, there is a periphrastic past perfect in Portuguese, but it “is rarely used in colloquial registers”, its synthetic counterpart being far more common. See also note 21 below.

²⁰Dialectally within present-day Daco-Romanian, and indeed also in earlier stages of Romanian, perfect forms with ‘be’ as the auxiliary occur in parts of the system other than the future perfect, e.g. Transylvanian *am fost cântat* ‘I have sung’ (literally: “I-have/AUX been sung”); see Rosetti 1968 and the very useful summary of verbal forms and dialectal perfects in Stefanescu (1997:277ff.).

Such a use of ‘be’ in active forms within a system that can be characterized as marking “perfect” (as opposed to perfective aspect) is reminiscent of the Romance (e.g. French and Italian) distinction between ‘have’ and ‘be’ active compound tenses (e.g. French *j’ai travaillé* ‘I worked, with ‘have’, versus *je suis tombé* ‘I fell’ with ‘be’) but more likely represents the result of influence from surrounding Slavic languages, especially Bulgarian; compare (13) with (12):

(13) *Țte sȃm pisal*
 FUT am/1SG write/PPL
 ‘I will have written’.

Still, of the varieties of Romanian other than Istro-Romanian, Daco-Romanian is the least like the core Balkan languages — and the most like other Romance languages — while Aromanian and Megleno-Romanian are far more Balkan in their character, at least insofar as the perfect and the preterite tenses are concerned. Affinities that Daco-Romanian may show with Balkan languages in this regard (cf. (12) and (13)) have not been such as to give all parts of its perfect and preterite system an overall Balkan character.

Sandfeld (1930:105) points out another related way, pertaining to the perfect and the preterite, in which Aromanian, as opposed to Daco-Romanian, shows a more Balkan-oriented, and less Romance-oriented, character. He notes that southern varieties of Aromanian show a preference for the simple preterite over the composite preterite, saying “c’est surtout vrai pour les cas où le préterit simple a le sens d’un ‘perfectum praesens’” and assessing this preference as “sans doute sous l’influence de l’aoriste grec”.²¹ He cites as an example Aromanian *nu s mârtaî sor mea* ‘my sister is not married’ (i.e. ‘has not married’, literally “not self married/PRET sister my”), which he contrasts with Daco-Romanian *nu s’a mârîtat sora mea*, with the *perfectul compus* (and thus literally “not self has/AUX married sister my”). In this way, Aromanian aligns itself with Greek,

²¹So also Mallinson (1987:314); see however note 19 for a parallel in Portuguese to this Aromanian preference for the synthetic form, which could suggest an independent origin for this Aromanian development..

since in Greek the simple past is far more common than the present perfect,²² and has been for centuries; indeed, 17th century Greek grammarians (Germano 1622, Portius 1638) give past perfect forms such as *eîkha grápei* ‘I had written’ as the equivalent of a Latin *plus-quam-perfectum* (i.e., pluperfect or past perfect), but list the simple past (corresponding to the Ancient Greek aorist) *égrapsa* ‘I wrote; I have written’ as the equivalent of a present perfect, and Modern Greek usage confirms this distribution (see Thumb 1912:§229).

Overall, therefore, this exercise in areal and genetic comparisons involving equivalents of perfects and simple preterites reveals that Daco-Romanian is less fully a Balkan language, where “Balkan” marks a language as having taken part in the contact situations that led to the massive structural convergences characteristic of the *Sprachbund*, than Aromanian or Megleno-Romanian, and further that it shows more affinities with other Romance languages in this domain than do either Aromanian or Megleno-Romanian. The comparisons made here thus clarify the place of Daco-Romanian and the other varieties of Romanian both within the Balkans and within the Romance group, and further point to Aromanian and Megleno-Romanian as having taken part more fully in the intense and intimate contact that gave rise to the Balkan *Sprachbund*.²³

4. Universally Oriented Comparisons

As noted in the introduction, comparisons can also be made on matters of linguistic typology for the purpose of illuminating some aspect of Universal Grammar. Such comparisons are of course possible with any set of two or more languages, whether or not they are genetically or

²²Also, the past perfect is more common in Greek than the present perfect, a fact which constitutes some of the evidence for taking the past perfect to be historically prior to the present perfect, as discussed by Joseph 1983, 1997.

²³A further conclusion would be that the sort of contact Daco-Romanian speakers had with speakers of other Balkan languages, e.g. Bulgarian, was of a different character, perhaps in a different social setting (e.g. more focused on church-related matters than on day-to-day, side-by-side living).

geographically related, and in fact unrelated languages are often preferred, by way of diversifying the sample. Still, Universal Grammar must be broad enough to take in facts from any natural language, and to be sure, many claims about the universal behavior of clitic pronouns, for instance, have been based, at first at least, on their behavior in Romance languages. Thus, comparisons made with languages within the Balkans can also further this goal and indeed, researchers have not shied away from typologically and universally oriented comparisons, including those involving Romanian; one such case is presented here in some detail.²⁴

For example, Joseph (1980a, 1983:232ff.) noted a parallel between Greek and Daco-Romanian in the realization of *tough*-Movement type sentences, pointing out that in such sentences in each of the languages, the finite, tensed complement clause — the only possibility in Greek and one of two possible forms of the complement in Romanian, the other being the nonfinite supine, shown in (14c) — cannot be both transitive and objectless. In Greek, shown in (14a), this state is achieved by the occurrence of a resumptive pronominal object in the complement clause corresponding to the moved nominal, and in Romanian, shown in (14b), by the occurrence of a passive complement verb:²⁵

²⁴Other studies can be cited. Rudin 1988, for example, investigates the properties of different strategies for the multiple fronting of *wh* elements cross-linguistically, and attempts to “integrate them into a general typology of *Wh*-movement and multiple-*Wh*-constructions” (p.488); Romanian figures prominently here for Daco-Romanian and its Balkan neighbor Bulgarian provide the only examples in her survey of multiple *Wh*-Fronting languages with only syntactic *Wh*-Movement.

²⁵The transitivity part of this bipartite condition — i.e. both not transitive and not objectless — would cover the reflexive passive, under the assumption that it is a detransitivizing construction. However, if the Romanian reflexive passive in the complement of (14b) is analyzed so that the ostensible reflexive pronoun *se* is taken as a realization of the logical object of ‘do’ (as is the case in some treatments of reflexive passives, as opposed to regular passives, in Relational Grammar), then the bipartite condition can be reduced to a single condition, namely “not objectless”.

- (14) a. ta angliká; íne žískola na
 the-English/NTR.NOM.PL are/3PL difficult/NTR.NOM.PL SUBJUNC
 ta; 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 facă
 this not-is/3SG hard SUBJUNC REFL do/3SG
 ‘This is not hard to do’ (literally = “This is not hard that it be done”)
- c. asta nu-i greu de făcut
 this not-is/3SG hard do/SUPINE
 ‘This is not hard to do’ (literally = “This is not hard that it be done”)

This parallel was taken to be a reflection of a universal tendency, evident from cross-linguistic comparisons (Joseph 1978/1990, 1980b), for the complement clause in *tough*-Movement constructions to be nonfinite when objectless, as with the Romanian supine type in (14c) or the English translations in (14). The Romanian evidence, especially when taken together with the Greek,²⁶ was viewed as supporting a putative linguistic universal that *tough*-Movement cannot deprive a finite verb of its object, and provided an important dimension to the investigation of this hypothesized universal constraint by furnishing examples of the construction with a finite complement clause, thus showing that the universal cannot simply be stated as a requirement for a nonfinite complement in this construction. While other research casts doubt on the validity of such

²⁶As Joseph (1980b; 1983:227) notes, Albanian also shows a finite complement clause in *tough*-Movement sentences with a resumptive pronoun corresponding to the moved nominal, thus providing another Balkan point of comparison here.

a constraint,²⁷ or on the need to single out *tough*-Movement constructions in the statement of the constraint,²⁸ the value of the cross-linguistic comparisons is undeniable, as is the contribution of the Romanian evidence to the investigation of this constraint. Moreover, even if the constraint is not valid, the Balkan-internal parallel that Romanian and Greek (and Albanian — see note 26) offer evidence of is noteworthy.

However, it must be recognized that if the goal of these comparisons is to elucidate some aspect of Universal Grammar, then the actual terms of the comparison, that is the particular languages being compared, are perhaps less important to the investigation than the mere fact of comparison. In a sense, the role in any such comparisons of Romanian itself and of Romanian in relation to other Balkan languages is really accidental, since the purpose of developing a linguistic typology can be served by comparisons between Romanian and any other, non-Balkan, language, or for that matter, as noted above, by comparisons between any arbitrarily selected languages.

Moreover, if there is a Balkanological goal to the comparison involving Romanian, then it is not clear that anything significant can emerge by comparisons inspired by typological and universalist considerations; what led to a special “Balkan” character to Greek, Albanian, Aromanian, and to a lesser extent Daco-Romanian, is the special language contact situation in the Balkans and there is no reason to believe that Universal Grammar plays a role in language contact except in the most trivial sense that natural languages are involved.²⁹ Thus the account given above of *tough*-

²⁷Bach & Horn (1976:271-2), for instance, have cited as acceptable the following English sentence with “*tough*”-Movement out of a finite complement clause:

(i) Walter_j is hard for me to imagine that anyone would look at Ø_j.

²⁸For example, Grosu & Horvath 1987 see the constraint as subsumed under a more general principle that “complement clauses with an empty operator in their COMP binding a syntactic variable are non-finite”, though see Joseph (1978/1990:197n.B) for some counter-discussion.

²⁹That is, whatever state results in a language from contact with speakers of another language is by definition a natural human language, one that must therefore be accommodated into Universal Grammar; the absence of a role for Universal Grammar in guiding the direction of borrowing is

Movement structures in the Balkans does not really relate in any way to the Balkan Sprachbund, however (potentially) interesting it might be for understanding the nature of *tough*-Movement and related constructions universally.³⁰

5. Comparison for the Purpose of Finding Differences

Most of the comparisons discussed here have focused on similarities between Romanian and other languages, though in some instances, the comparisons revealed points of difference as well. Since Universal Grammar is as much defined by accounting for the differences among languages as it is by accounting for their similarities, by way of conclusion, an example is provided here of a comparison that leads to the identification of a point of grammar on which Romanian seems to differ from all other languages within the Balkans and within Romance, and quite possibly as well from all other languages period.

The particular feature in question concerns the placement of weak, so-called “clitic”, object pronouns. For the most part, weak object pronouns are positioned in Romanian in reference to the inflected verb, generally occurring as proclitics, even in compound tenses, but enclitic to imperatives and gerunds:

(15) a. îmi dau cadoul

especially evident if, as Thomason & Kaufman 1988 have argued, there are no linguistic constraints on what can be borrowed, only social constraints.

³⁰Furthermore, the parallel noted here between Romanian and Greek is a rather abstract and “deep” one. It is thus of a sort that seems not to emerge from language contact, for contact-induced structural parallels generally depend on shared surface structures resulting from intense and sustained contact, often with some degree of bilingualism. Numerous examples of such surface-oriented parallels are evident in the Balkan *Sprachbund* (see (1c) above and the discussion in section 2, and the structural convergences found in Kupwar village in India, as described by Gumperz & Wilson 1971, are of the same type; see also Joseph 1996 for some discussion of the role of surface structures in contact situations.

me/DAT give/3PL present/DEF

‘They give the present to me’

b. mi- l dau³¹

me/DAT it/ACC give/3PL

‘They give it to me’

c. mi- l au dat

me/DAT it/ACC have/3PL given/PPL

‘They gave it to me’

d. spune- i

tell/IMPV.SG him/DAT

‘Tell him!’

e. cerîndu- le

ask/GER them/ACC

‘asking them ...’

In this way, Romanian seems quite parallel to other patterns of clitic placement found in Romance languages, e.g. French, and in the Balkans, e.g. Greek.³²

One idiosyncratic exception to this placement, however, occurs with the feminine singular accusative weak pronoun *o*, and only with that pronoun. What is found in formations such as the *perfectul compus* is that instead of *o* being positioned before the tensed verb, as with *mi* or *l* in

³¹The reduction of *îmi* to *mi* is regular in the presence of another weak pronoun, here *l*.

³²In Greek, as argued in Joseph 1978/1990, 1980b, 1983 and elsewhere, the decisive factor in determining the placement of weak object pronouns is finiteness; the patterns of placement are similar in Romanian but finiteness has a different role in the overall verbal system, inasmuch as in Romanian the category of infinitive (i.e., a verbal noun used in complementation) is more robust than in Greek (where it is absent altogether). Thus the precise characterization of the relevant triggering features for weak object placement may differ between Greek and Romanian but the surface patterns are quite similar.

(15c), it occurs after the participial part of the compound tense; this contrasts with the placement of *o* before the verb in a simple tense, as in (16c):

- (16) a. *am văzut- o*
have/1SG seen/PPL her/ACC.WK
'I saw her'
- b. **o-am văzut*
- c. *o găsim*
her/ACC.WK find/1PL
'We find her'.

Although there may be some phonological issues that play a role in this placement (see Wanner 1993), it is still the case that only the feminine singular accusative weak pronoun shows this behavior. Such an idiosyncratic exception, and particularly one centered on the feminine singular pronoun, is unique within the Balkans and seems also to be unique within the Romance group of languages.³³ Moreover, given the highly specific nature of this idiosyncrasy, coming just within the placement of weak pronouns, just for the feminine singular form, and just in formations such as the compound past, is so specialized that it would be hard to imagine that there could be another language outside of Romance and outside of the Balkans that has precisely the same irregularity.

What this example shows, therefore, is that even with the quest for generalizations, especially those that cut across languages and are thus of interest from a universalist perspective or the perspective of language contact, there are still going to be highly particularized aspects of languages that are real, even if messy. They therefore cannot be ignored. Romanian, through the comparisons

³³Piedmontese Italian does show postverbal positioning of weak pronouns in compound tenses, so that the Romanian situation in general is not unparalleled in Romance; however, this postpositioning in Piedmontese is not restricted to just the feminine singular forms, so that Daco-Romanian is unique in the particular restriction it shows. See Parry 1995 for presentation and discussion of the Piedmontese facts. I am grateful to my colleague Dieter Wanner for clarification of these facts and for directing me to appropriate sources.

explored here, reveals itself as a member to some extent of the Balkan Sprachbund, as a member of the Romance group, and as a member of the set of human languages more generally, but at the same time these comparisons show that it has some unique characteristics that are not paralleled elsewhere.

Department of Linguistics

222 Oxley Hall

The Ohio State University

Columbus, Ohio USA 43210-1298

bjoseph@ling.ohio-state.edu

References

- Bach, Emmon & George Horn. 1976. "Remarks on 'Conditions on Transformations'". *Linguistic Inquiry* 7:265-299.
- Banfi, Emanuele. 1985. *Linguistica balcanica*. Bologna: Zanichelli.
- Boretzky, Norbert. 1968. "Der bestimmte Artikel nach Präpositionen im Albanischen". *Zeitschrift für vergleichende Sprachforschung* 82:129-152.
- Comrie, Bernard, ed. 1987. *The World's Major Languages*. London: Croom Helm.
- Germano, Girolamo. 1622. *Grammaire et vocabulaire du grec vulgaire publiés d'après l'édition de 1622*. (ed. by Hubert Pernot, 1907). Paris.
- Green, John N. 1987. "Spanish". In Comrie, ed., pp. 236-259.
- Grosu, Alexander & Julia Horvath. 1987. "On Non-Finiteness in Extraction Constructions". *Natural Language and Linguistic Theory* 5:181-196.
- 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 Univ. Press, pp. 151-67.

- Hamp, Eric P. 1989. "Yugoslavia — a crossroads of Sprachbünde", *Zeitschrift für Balkanologie* 25:44-7.
- Harris, Martin. 1987. "French". In Comrie, ed., pp. 210-235.
- Joseph, Brian D. 1978/1990. *Morphology and Universals in Syntactic Change: Evidence from Medieval and Modern Greek*. Ph. D. Dissertation, Harvard University; printed and distributed by Indiana University Linguistics Club [Expanded and updated version published (1990) in Outstanding Dissertations in Linguistics series, New York: Garland Publishing, Inc.].
- Joseph, Brian D. 1980a. "A New Convergence involving the Balkan Loss of the Infinitive". *Indogermanische Forschungen* 85:176-187.
- Joseph, Brian D. 1980b. "Linguistic universals and syntactic change", *Language* 56:345-370.
- 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. 1996. "Is Balkan Comparative Syntax Possible?", paper presented at "After-GLOW" Workshop on Balkan Syntax, Athens, April 1996 (written version to appear in conference volume).
- Joseph, Brian D. 1997. "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.
- Mallinson, Graham. 1987. "Rumanian". In Comrie, ed., pp. 303-321.
- Newmark, Leonard, Philip Hubbard, & Peter Prifti. 1982. *Standard Albanian. A reference grammar for students*. Stanford: Stanford University Press.
- Parry, M. Mair. 1995. "Some observations on the syntax of clitic pronouns in Piedmontese". In John C. Smith & Martin Maiden, eds., *Linguistic Theory and the Romance Languages (Current Issues in Linguistic Theory, 122)*, pp. 133-160. Amsterdam: John Benjamins Publishing Co.
- Parkinson, Stephen. 1987. "Portuguese". In Comrie, ed., pp. 260-278.

- Portius, Simon. 1638. *Grammatica Linguae Graecae Vulgaris* (ed. by Wilhelm Meyer, 1889). Paris: F. Vieweg.
- Reichenkron, Gunther. 1962. "Der Typus der Balkansprachen." *Zeitschrift für Balkanologie* 1:91-122.
- Rosetti, Al. 1968. *Istoria Limbii Române*. Bucharest: Editura pentru Literatura.
- Rudin, Catherine. 1988. "On Multiple Questions and Multiple WH Fronting". *Natural Language and Linguistic Theory* 6:445-501.
- Sandfeld, Kristian. 1930. *Linguistique balkanique. Problèmes et résultats*. Paris: Honoré Champion.
- Schaller, Helmut. 1975. *Die Balkansprachen. Eine Einführung in die Balkanphilologie*. Heidelberg: Carl Winter Universitätsverlag.
- Stefanescu, Ioana. 1997. *The Syntax of Agreement in Romanian*. Ph.D. Dissertation, City University of New York Graduate Center.
- Thomason, Sarah G. & Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Thumb, Albert. 1912. *A Handbook of the Modern Greek Language. Grammar, Texts, Glossary*. Chicago: Argonaut, Inc. Publishers [reprinted 1964].
- Trask, Robert Lawrence. 1997. *The History of Basque*. London: Routledge.
- Vincent, Nigel. 1987. "Italian". In Comrie, ed., pp. 279-302.
- Wanner, Dieter. 1993. "Multiple Clitic Linearization Principles". In William Ashby, Marianne Mithun, Giorgio Perissinotto, & Eduardo Raposo, eds., *Linguistic Perspectives on the Romance Languages (Current Issues in Linguistic Theory, 103)*, pp. 281-302. Amsterdam: John Benjamins Publishing Co.