Multiple Determination in Greek and the Balkans

Brian D. Joseph
The Ohio State University

1. Background on the Balkans and Multiple Determination

It is well-known that the Balkans form a special type of language contact zone, usually referred to as a *sprachbund*.[1] Several languages are to be found there, of different origins, representing five different branches of Indo-European — Albanian, Greek, Slavic (via Bulgarian, Macedonian and Torlak Serbian, within South Slavic),[2] Italic (via the Eastern Romance languages Aromanian, Megleno-Romanian and Romanian[3]), and Indic (via Romani) plus the non-Indo-European Turkish, especially in its Western Rumelian variety. And, significantly, although originally these languages had somewhat different structures and lexicons, they have come to show, through intense and sustained contact, a convergence in both structure and lexicon. The relevant time period for this convergence is the Ottoman Empire era of c. 1400-1900.

Among the convergent structural features found in the Balkan sprachbund is what may be called "multiple determination," the multiple marking within a noun phrase for some form of definiteness or deixis. This trait is found in some of the Balkan languages, as outlined below, but occurs most robustly in Greek, with two manifestations: the so-called "polydefinite" construction, in which an adjective is not adjacent to the definite noun it modifies and the article is repeated with the adjective, as exemplified in (1), and the "double determination" construction in which a demonstrative modifier and a definite article co-occur with a noun, as exemplified in (2):

(1) to   puli   to   mikro
     DEF.ART   bird   DEF.ART   small
     'the small bird'

(2) afto   to   puli / *afto puli
     DEMNS   DEF. ART   bird
     'this bird' (literally 'this the bird')

In both of these, there is a definite article that doubles ("multiplies") in some way the determination that is present in another part of the noun phrase, either the determination by a definite article itself, as in (1), or the determination by a demonstrative, as in (2).

It is important, when considering the extent of these types of multiple determination in the Balkans, to note that they both have been a part of Greek grammar since Classical times, so that their occurrence in Modern Greek represents an inheritance from earlier stages of the language. Such is not the case, as becomes apparent below, with the other languages in the Balkans that show multiple determination. Moreover, various pieces of evidence suggest that multiple determination can be affected by language contact. Thus, given the prevalence of contact-related effects on the grammars of various Balkan languages, it is essential to consider the role of contact in the appearance of multiple determination in the sprachbund. Accordingly, after some further background relating to contact and multiple determination, one aspect of possible contact-induced change in the Balkans, involving Greek and Aromanian with regard to demonstrative double determination, is explored in some depth and in all its ramifications.

2. Multiple Determination and Language Contact

As just noted, language contact can affect multiple determination, and seems to do so in various ways. There is reason to believe, for instance, that it can spread via contact, as several examples can be cited from the Balkans that point to the diffusion of multiple determination under conditions of language contact. For example, according to Koneski 1967, in southern Macedonian, presumably under Greek influence given the geography of the occurrence of this construction in the language, doubled articles of the polydefinite type (as in (1)) can be found, and Ugrinova-Skalovska 1960/1961 attests to this construction for colloquial Macedonian more generally. It is found as well in some varieties of Romani that have been strongly influenced by close contact with Greek, *e.g.,* the Romani of Agia Varvara (Igla 1996) and Serres (Sechidou 2011). Further, with demonstratives, the Greek pattern as in (2) seems to have spread into Balkan Slavic, for it is found in colloquial Macedonian (Friedman 2019, Rudin 2018, to appear) and in colloquial Bulgarian (Rudin 2018, to appear, Stojkov 1968:173), as well as Torlak Serbian (Belić 1905:447) and Agia Varvara Romani (Igla 1996). Examples of the double demonstrative determination from each language are given in (3):

(3)   a.   ovie idioti   ve   (Macedonian)
             these idiot.PL the. PROX
             'these idiots'

       b.   te   turcite   (Bulgarian)
             these Turks.DEF
             'these Turks'

       c.   taj   postarata   (Torlak Serbian)
             that   older.FEM.DEF
             'that older-one (FEM)'

       d.   kova o manuš / o manuš kova (Agia Varvara Romani)
             this the person/the person this
             'this person'

Since such a pattern is found nowhere else in Slavic, and not even elsewhere in South Slavic, it is reasonable to see its appearance in Balkan Slavic as being due to contact with Greek.[4]
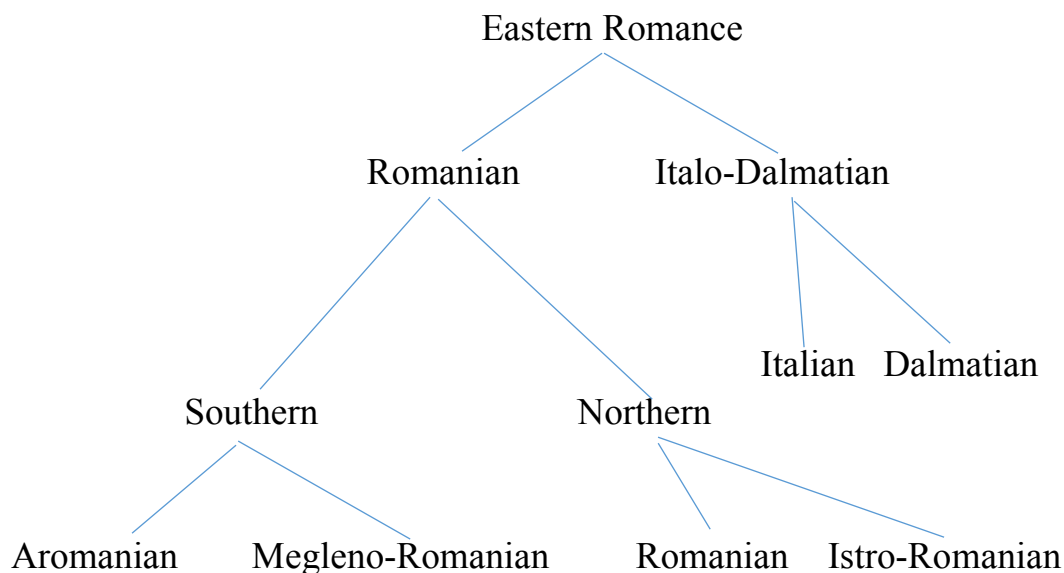
Besides the gain of double determination through contact with other languages, it is also the case that contact can lead to the loss of double determination. For instance, in Italo-Greek (Griko) dialects of the Salento area in Apulia, the inherited polydefinite construction has been lost, due to contact with local varieties of Italian, which lack this structure (Guardiano *et al.* 2016:113, Guardiano and Stavrou 2017).

Double determination can of course arise spontaneously in languages, as it must have in Classical Greek. One can point, moreover, to constructions like *this here* in English, or *celle-ci*, in French, where, for emphasis, there is added deixis through the use of *here* and *ci* respectively. Nonetheless, as the above examples show, multiple determination is not immune to the effects of language contact.

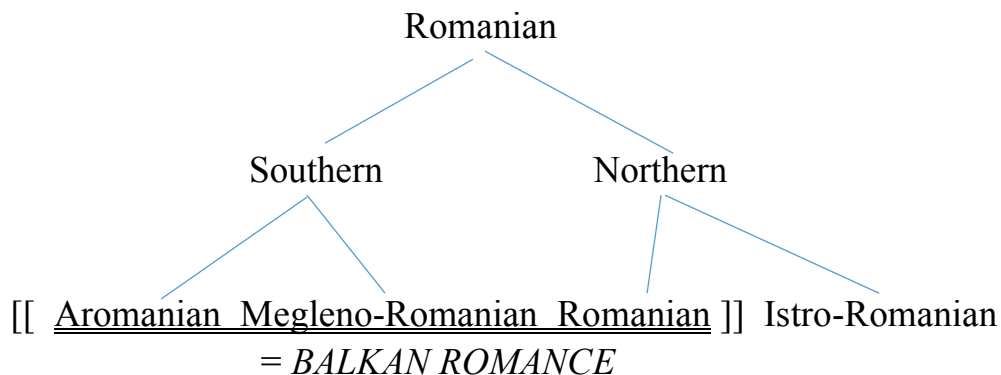*3. Romance Languages in the Balkans and Multiple Determination*

In the preceding sections, little has been said about the multiple determination and the Romance languages in the Balkans. In this section, the relevant Romance data is considered.

First, though, it is useful to review the ways in which the Eastern Romance languages can be classified within the Balkans.[5] From a genetic/genealogical perspective, one can schematize the relations among the languages as in (4), based on Rosetti 1973 and Vrabie 2000:

```
                        Eastern Romance
                       /               \
               Romanian              Italo-Dalmatian
              /        \                    /      \
                                      Italian   Dalmatian
     Southern            Northern
     /      \            /      \
Aromanian  Megleno-Romanian  Romanian  Istro-Romanian
```

(4) Genetic/Genealogical Romance Subgroupings in the Balkans

However, it is also possible to view these languages from the perspective of contact-induced developments and thereby to identify a "Balkan Romance" that cuts across the genealogical southern/northern grouping within the Romanian branch of Eastern Romance; such a grouping is schematized in (5):

```
                    Romanian
                   /        \
           Southern          Northern
           /      \          /      \
[[ Aromanian  Megleno-Romanian  Romanian ]]  Istro-Romanian
              = BALKAN ROMANCE
```

(5) Contact-Related Romance Subgroupings in the Balkans

The reason for giving both groupings here is that multiple determination in Eastern Romance is found in Balkan Romance. The relevant multiple marking involves both the (enclitic) definite article occurring together with the so-called "adjectival article," an element which links an adjective to a noun.[6] This construction is found in Romanian, Megleno-Romanian and Aromanian, as shown in (6), (7) and (8), and its presence in all three languages indicates that it is an old structure in Balkan Romance:

(6)   băiat-ul   cel   mare (Romanian)
        boy-DEF.ART   ADJ.ART   big
        'the big boy'

(7)   fitšor-u   țela   bun   (Megleno-Romanian,
        son-DEF   ADJ.ART   good   Capidan:1925:150)
        'the good son'

(8)   om-lu   atsel   bun   (Aromanian, Capidan
        man-DEF   ADJ.ART   good   1932:400)
         'the good man'

Given that the etymology of the adjectival article, Romanian *cel* and Aromanian *atsel*, lies in demonstrative elements − "the endophoric distal demonstrative *acel*/*acela* (< ECCE/ECCUM + ILLUM)" (Nicolae 2013:310) − this construction can certainly be said to qualify as an instance of multiple determination.

     In addition, however, Aromanian in Greece shows a polydefinite, multiple-article, structure on top of the adjectival article demonstrative multiple determination. This polydefinite demonstrative construction, what might be characterized as "multiple multiple determination," takes the form shown in (9), a variant of (8) that Capidan records (1932:400):

(9)   om-lu   atsel   bun-lu
        man-DEF   ADJ.ART   good-DEF
        'the good man'

Moreover, in Arvantovlaxika, the label Campos (2005:313) uses to refer to "a dialect of Aromanian spoken in Thessaly, Greece," the same structure occurs:

(10)   pul'i-l'i   atsel'i   ñits-l'i
       birds-DEF.ART   ADJ.ART   small- DEF.ART
       'the small birds'

This Aromanian construction appears to be an extension of the structure seen in Romanian, with a repetition of the definite article distinguishing it from the more widespread Balkan Romance type in (6)-(8). At some point the second definite article must have been optional, to judge from the fact that Capidan gives the presence and absence of that article as variants. This suggests that adding it was not a necessary, possibly semantically driven, development, and it may even have been motivated by a kind of formal "harmony," inasmuch as the various elements in the noun phrase in this way come to end in the same way; still, it cannot be ruled out that the addition of some perhaps slight or nuanced degree of emphasis much like as in *this here* in English or *celle-ci* in French (as pointed out in *Section 2*) may have been at work as well. It can be noted, too, that the postposed nature of the Aromanian (and indeed Balkan Romance) definite article makes its addition at the end of the noun phrase a seemingly trivial and simple structural adjustment, perhaps even like an afterthought.

4. Origins of the Aromanian Structure: An Hypothesis and an Evaluation

Nonetheless, despite the availability of such a reasonably compelling language-internal account as that just sketched in the previous section for the innovation in the Aromanian multiple determination construction, its geographic limitation to Aromanian in Greece raises the possibility that it is due to language contact. In that case, its origin would be externally motivated.

        In fact, Campos (2005:318) opts for such an external account and hypothesizes that this structure in Aromanian is a borrowing from Greek. This is a plausible suggestion given that Aromanian has been in contact for centuries with Greek and shows numerous other effects of this contact. For instance, in lexis, some 27% of the 9,236 lexical items in Papahagi 1974 are of Greek origin (so Vrabie 2000:82) and among those loans are several morphological elements in the form of various derivational affixes, *e.g.,* the suffix *-aða*, for forming deadjectival nouns, and the prefixes *ksana-* 'again, back' and *para-* 'beyond, too much.' Furthermore, as *-aða* shows, there are phonological effects in Aromanian caused by the contact with Greek, in that Greek loans containing the fricatives [ð θ γ], which were alien to

the Balkan Romance sound system, were borrowed without adaptation, so that the Greek sounds entered the Aromanian phonemic inventory and in fact even spread to some native words (see Joseph 2009 for details and references).

Moreover, the demonstrative double determination construction of Greek gives a model that could be the basis for the Aromanian structure. A Greek noun phrase designed so as to match the Aromanian in (10) would be that in (11):

(11) to   puli afto to   mikro
      the   bird this the   small

with *afto*, the proximate demonstrative 'this' matching the adjectival article of Balkan Romance in this case.

Thus treating the pattern in (10) as the result of contact with Greek is certainly reasonable, especially since structural borrowing, which this would involve, has been argued by Thomason and Kaufman 1988 to be a reality of language contact effects, despite some earlier views to the contrary.

However, even though it is a reasonable hypothesis, there is good cause to reject it. In particular, most problematic for the Greek-influence hypothesis is the fact that the Greek structure does not have the same meaning as the Aromanian one. The Aromanian meaning is simply definite, meaning 'the small bird,' whereas the Greek construction is more overtly deictic, meaning 'this small bird.'

To verify that this is indeed the meaning of the Greek (11), some speakers of Standard Modern Greek (SMG) were surveyed, and the results proved to be somewhat surprising and ultimately very interesting. The survey, the results and what they mean are taken up in the next section.

5. More on the Greek — and Balkan — Dialectological Front

As noted, by way of trying to get a handle on the issues raised in *Sections 3* and *4* with regard to possible influence of Greek on the Aromanian structure in question, a limited survey was carried out, testing the meaning of (11) for speakers of SMG from different parts of the Greek-speaking world. In particular, one speaker was queried from each of the following regions: Cyprus, Athens, Thasos (a northern island, thus representing the northern dialect zone) and Thessaly (in central Greece). The specific issue was whether (11) has the definite meaning of the Balkan Romance

doubly determined NP, *i.e.,* 'the small bird' or instead a full demonstrative sense, *i.e.,* 'this small bird.'

All of the speakers queried had the full demonstrative sense and most rejected the simple definite sense, as can be seen in *Table 1*, where * indicates the impossibility of a reading and √ indicates that the reading is available:

	Definite	Demonstrative
SMG speaker (Cyprus)	*	√
SMG speaker (Thasos)	*	√
SMG speaker (Athens [Corfu])	*	√
SMG speaker (Thessaly)	√	√

Table 1:
Greek Survey Results

The outcome shown in *Table 1* reveals two different aspects about the polydefinite construction with a demonstrative in Greek. First, the demonstrative meaning is most certainly a feature of the Greek construction. Second, only for the speaker from Thessaly is a simple definite reading possible. What is interesting about these results for Campos's hypothesis of Greek influence is first of all that the Greek construction, while a reasonable model for the Aromanian on the formal side, is most assuredly not a good model on the semantic side. Second, only in Thessaly might the semantics of the Greek be appropriate, that is, the definite reading of (11) appears to be a feature of Thessalian Greek, and this is a particularly interesting result, since Thessaly is where the highest concentration of Aromanian speakers is to be found. Thus, Campos could well be right, despite the unavailability of the definite reading in Greek more broadly.

However, it is fair to ask where this feature of definite-only semantics for the construction with both a demonstrative and multiple articles come from in Thessalian Greek. Given the distribution of the "demonstrative=definite" interpretation in this article, being restricted in Greek only to the area where there is a significant number of Aromanian speakers, one might well hypothesize that it is Thessalian Greek that has been affected by contact, here with Aromanian, rather than vice-versa. This could then be the result of transfer/substratum interference if Aromanian speakers shifted to Greek and imported this structure into their Greek.

Alternatively, Greek speakers could simply have calqued this structure, replicating it in Greek terms, with Greek material; it would then be rather like the borrowing of double determination discussed in *Section 2*.

    This interpretation has interesting consequences for Greek dialectology, and for Greek *vis-à-vis* Balkan dialelectology. In terms of Greek dialectology, it is exactly like the situation with the 'feels like VERB-ing' construction, consisting of an impersonal nonactive verb with a dative (or dative surrogate) weak personal pronoun. This is found in the Greek of the region around the city of Kastoria in the north of Greece (Papanastassiou and Papadamou 2013), but nowhere else in the Greek-speaking world; it is exemplified in (12)[7]:

(12) a.   mi   trojiti
            me. ACC   eat.3SG.PRES.NONACT
            'I feel like eating' (literally: '(to-)me (it-)is-eaten')

      b.   mi   piniti
            me.ACC   drink. 3SG.PRES.NONACT
            'I feel like drinking' (literally: '(to-)me (it-)is-drunk')

This construction is not possible in SMG (*mu trojete/pinete*) but is present in a regional dialect just exactly in an area where there are other languages — specifically Albanian, Aromanian and Macedonian — with that same construction, as shown in (13)[8]:

(13) a.   më   hahet   (byrek)   (Albanian)

      b.   nji   -si mãcã   (burec)   (Aromanian)

      c.   mi   se jade   (burek)   (Macedonian)
            me.DAT   eat3SG.PRES.NONACT (burek)
            'I feel like eating (burek).'

Since this 'feels-like' impersonal construction is found in Greek only in a region where influence from other languages with that particular structural manifestation together with that particular meaning is possible, it is reasonable to attribute its

occurrence in the Kastoria area to the effects of language contact with one or more of those languages.

A similar conclusion can then be reached for the geographically highly restricted polydefinite demonstrative construction of Thessalian Greek, which matches the Aromanian structure. Moreover, the fact that the Aromanian innovation requires just a small adjustment leading to what amounts to an end-rhyme harmony within the noun phrase, while the Greek construction has no such rhyming possibility, makes it more likely that the initial innovation emerged in Aromanian and then spread to Greek.[9]

6. Conclusion

Based on the foregoing, it emerges that Greek dialectology is in a position to shed some light onto this thorny issue in Balkan linguistics but, more likely, the Balkan situation sheds light on Greek dialect divisions. Moreover, if the account argued for here, of an internal innovation in Aromanian providing a model for Greek on a highly localized basis, is right, it shows the importance of looking to regional dialects when charting the distribution of shared Balkan features; we cannot rely on the evidence of the standard languages alone (as emphasized in Friedman and Joseph 2017).

Notes

1. I adopt here the practice of Friedman and Joseph 2020 in treating the originally German word *sprachbund* as an assimilated loanword in English, and so write it without a capital letter.
2. I am following the lead of Friedman and Joseph (2020:Chapter 1) in excluding most of the "successor languages" to Serbo-Croatian — Bosnian, Croatian, Montengrin and Serbian — from the Balkan sprachbund. Even though they show a few of the characteristics that define the sprachbund, they do not exhibit anywhere near the extent of convergence that the other languages show.
3. I use the name "Romanian" for the national language of Romania, but, as seen below in (4) and (5), I also use it as the name for the node in the genealogical tree that includes the non-Italo-Romance languages of the Balkans.
4. Double demonstrative determination is found in Albanian, though it is not usual. Given the late attestation of Albanian (15th century), it must be admitted that the question of whether this construction was native to Albanian or due to contact cannot easily be decided.

5. I am leaving Judezmo (Judeo-Spanish) out of the picture here because it is a Western Romance language, and also because it shows no multiple determination structures.

6. This "adjectival article" has other functions as well; for instance, it can nominalize an adjective, so that *cel mare* can occur without a noun in the meaning 'the big one.'

7. These examples are given in their dialect form and thus show the effects of the raising of mid-vowels to high and the use of the accusative case for a dative experiencer; it is thus as if they were standing for an SMG *me trójete/pínete*.

8. This construction is found in Slavic more broadly, so that (South) Slavic may be the source ultimately for all of this in Balkans. See Friedman and Joseph (2020:Ch. 7.8.2.2.5.1) for discussion.

9. The Aromanian endings for masculine and feminine article largely coincide with the endings for the demonstrative (see Vrabie 2000: 47-48, 52), whereas such is not the case for Greek. Although neuter forms would have *to* (singular) and *ta* (plural) for the article and *afto* (singular) and *afta* (plural) for the demonstrative, in the masculine and feminine gender, there is less similarity across the endings: in the nominative, for instance, masculines have *o* (singular) and *i* (plural) for the article and *aftos* (singular) and *afti* (plural) for the demonstrative, and feminines have *i* (singular) and *i* (plural) for the article and *afti* (singular) and *aftes* (plural) for the demonstrative.

References

Campos, Hector. 2005. "Noun Modification, Pseudo-Articles, and Last Resort Operations in Arvantovlaxika and in Romanian," *Lingua* 115, pp. 311-47.

Capidan, Theodor. 1925. *Elementul slav în dialectul aromân*. Bucureşti: Cultura naţională.

Capidan, Theodor. 1932. *Aromânii, dialectul Aromân: Studiu lingvistic*. Bucharest: Monitorul Oficial şi Imprimeriile Statului, Imprimeria Naţională.

Friedman, Victor A. 2019. "Turkisms in the Macedonian '*Bombs*' of 2015: On Language as Flag in the Balkans in the 21st Century," *Celebrating Twenty Years of the Kenneth E. Naylor Lectures in South Slavic and Balkan Linguistics* (= *Balkanistica* 32:1), Dyer, Donald L., Brian D. Joseph and Mary-Allen Johnson (eds). Oxford, MS: University Printing Services.

Friedman, Victor A., and Brian D. Joseph. 2014. "Lessons from Judezmo about the Balkan Sprachbund and Contact Linguistics," *International Journal of the Sociology of Language* 226, pp. 3-23.

_____. 2017. "Reassessing Sprachbunds: A View from the Balkans," *Handbook of Areal Linguistics*, Hickey, Raymond (ed.). Cambridge: Cambridge University Press, pp. 55-87.

_____. 2020. *The Balkan Languages*. Cambridge: Cambridge University Press.

Guardiano, Cristina, Dimitris Michelioudakis, Andrea Ceolin, Monica Alexandrina Irimia, Giuseppe Langobardi, Nina Radkevich, Giuseppina Silvestri, and Ioanna Sitaridou. 2016. "South by Southeast: A Syntactic Approach to Romance and Greek Microvariation," *L'Italia Dialettale* LXXVII, pp. 96-166.

Guardiano, Cristina, and Melita Stavrou. 2017. "Syntactic change under the pressure of contact: the loss of polydefiniteness in Italiot Greek." Paper presented at DiGS 19 Workshop on Language Variation and Change in Contact Situations, Stellenbosch University.

Igla, Birgit. 1996. *Das Romani von Ajia Varvara: deskriptive und historisch-vergleichende Darstellung eines Zigeunerdialekts*. Wiesbaden: Harrassowitz.

Joseph, Brian D. 2009. "Broad vs. Localistic Dialectology, Standard vs. Dialect: The Case of the Balkans and the Drawing of Linguistic Boundaries," *Language Variation – European Perspectives II: Selected Papers from the 4th International Conference on Language Variation in Europe (ICLaVE 4), Nicosia, June 2007*, Tsiplakou, Stavroula, Marilena Karyolemou and Pavlos Pavlou (eds). Amsterdam: John Benjamins, pp. 119-34.

Koneski, Blaže. 1967. *Gramatika na makedonskiot literaturen jazik*. Skopje: Nicolae Kultura.

Nicolae, Alexandru. 2013. "Means of Encoding Nominal Phrase Internal Possession," *The Grammar of Romanian*, Pană Dindelegan, Gabriela (ed.). Oxford: Oxford University Press, pp. 335-49.

Papahagi, Tache. 1974. *Dicționarul dialectului aromân*. București: Editura Academiei Republicii Socialiste România.

Papanastassiou, George, and Eleni Papadamou. 2013. "The Position of the Northern Dialects of the Prefecture of Kastoria." *MGDLT 5. Proceedings of the 5th International Conference on Modern Greek Dialects and Linguistic Theory (University of Ghent, 21 September 2012)*, Janse, Mark, Brian D. Joseph, Angela Ralli and Metin Bagriacik (eds), pp. 389-400; also available on line at [http://lmgd.philology.upatras.gr/files/MGDLT5_proceedings.pdf].

Rosetti, Alexandru. 1973. *Brève histoire de la langue roumaine des origines à nos jours*. The Hague: Mouton

Rudin, Catherine. 2018. "Multiple Determination in Bulgarian and Macedonian: An Exploration of Structure, Usage, and Meaning," *V zeleni drželi zeleni breg: Studies in Honor of Marc L. Greenberg*, Dickey, Stephen M., and Mark Richard Lauersdorf (eds), Bloomington, IN: Slavica Publishers, pp. 263-86.

_____. To appear. "On DP Structure in Balkan Slavic: Evidence from Multiple Determination," *Formal Approaches to Slavic Linguistics, Stanford 5/1-3/2018*. Ann Arbor, MI: Michigan Slavic Publications.

Sechidou, Irene. 2011. *Balkan Romani: The Dialect of Ajios Athanasios/Greece*. Munich: Lincom Europa.

Stojkov, St. 1968. *Bălgarska dialektologija*. Sofia: Nauka i izkustvo.

Thomason, Sarah G., and Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley, CA: University of California.

Ugrinova-Skalovska, Rada. 1960/61. "Dve stilski osobenosti na našiot govoren jazik," *Makedonski jazik* 11/12, pp. 105-10.

Vrabie, Emil. 2000. *An English-Aromanian (Macedo-Romanian) Dictionary: With Two Introductory Sketches on Aromanian*. University, MS: Romance Monographs.