The Internal Structure of Slavic Noun Phrases, with Special Reference to Bulgarian

Steven Franks
Indiana University, Bloomington

1. Introduction. This paper overviews some of the puzzles presented by Noun Phrases (NPs) in Slavic, and focuses in particular on the distribution of clitics inside NPs. One of the most striking facts about the use of clitics inside NPs is that these are not as free as one would expect, under the erroneous assumption that they are simply more laconic version of full NPs. Their distribution ranges from relatively unrestricted, as in Polish (Pol), to completely disallowed, as in Serbian/Croatian (SC). Most telling, however, is Bulgarian (Bg), and to a lesser extent Macedonian (Mac), where pronominal clitics are employed inside NPs, but with restrictions peculiar to this particular domain.

In what follows, I will show how the possibility of positioning clitics within the nominal domain raises some very interesting theoretical and descriptive issues. I begin by briefly surveying some of the other languages, then turn in detail to Bg. After reviewing the basic facts, I sketch some previous proposals about how the proper placement of clitics inside NPs in that language might be effected, and then develop my own, minimalist analysis. Ultimately, I will argue for an "interarboreal" merger account, which involves no actual movement of morphological material.

2. The extremes. I first turn to the systems in which NP-internal clitics either appear as regular arguments or are completely unacceptable. As argued in Franks (1998) and Franks and King (2000), Polish represents the former extreme. Although there are stylistic and prosodic constraints at work, clitics in this language behave essentially as full NPs. Some examples to illustrate the productivity of pronominal clitics as complements of deverbal nouns are provided in (1):[1]

(1)   a.   zrozumienie   **ci´**
              understanding   you.gen

        b.   przebaczenie   **mi**
              forgiving   me.dat

        c.   moje pomaganie   **mu**
              my   helping   him.dat

        d.   nasze oskar″enie **go**   o   zdrad´
              our   accusing   him.gen of treason

        e.   ich   zaproszenie **go**   na   obiad
              their inviting   him.gen to   lunch

---

[1]This discussion is drawn from Franks (1998); see also Franks and King (2000: 277-78).

There is moreover some freedom of word order that resembles the distribution of pronominal clitics in clauses. Katarzyna Dziwirek (pc) informs me that expressions such as *moje mu pomaganie* for (2c) are acceptable if poetic, although preposing clitic forms *jej* 'her.gen' or *ich* 'them.gen' is not possible because these are homophonous with the 3rd person possessive pronouns, which leads apparently to parsing difficulties. Adam Szczegelniak (pc) provides the judgments in (2), indicating that although locating the clitic lower than immediately following the noun is not possible, higher positions are acceptable with various degrees of felicity.

(2)   a. moje (**mu**)   pomaganie (**mu**)   we wtorki   (***mu**)
            my   him.dat helping   him.dat   on Tuesdays him.dat
            'my helping him on Tuesdays'

    b. nasze (**go**)   szybkie (**go**)   ratowanie (**go**)
            our   him.gen quickly him.gen saving   him.gen
            'our saving him quickly'

    c. twoje (?**go**)   stale   (**go**)   oglądanie (**go**)   w nocy (***go**)
            your   him.gen constantly him.gen looking-at him.gen in night him.gen

    d. twoje (??**go**)   w nocy (?**go**)   stale   (**go**)   oglądanie (**go**)
            your   him.gen in night him.gen constantly him.gen looking-at him.gen

While certain factors interfere, such as fronting the adverb in (2d), which apparently makes scrambling of *go* more difficult, argument clitics inside NPs clearly behave in a strikingly similar fashion both to pronominal clitics in clauses and to full phrases in NPs.

    To account for the Pol situation, I make the minimal assumption that pronominal clitics in this language *are* full phrases, whatever the maximal extended projection of N is. Thus, they can have their case checked in the same way. Elsewhere I have argued that Slavic noun phrases—with the exception of Bg and Mac, where they are clearly DPs—are in fact Kase Phrases (KPs), and I do not repeat those arguments here. Pol clitics are thus not simple K°s, but rather full KPs, i.e. K°s with empty NP complements, with the result that they have phrasal internal structure. For this reason, they move and have their case checked as phrases. Whatever licenses case on NP-internal (more properly "KP-internal") full KP elements also applies to pronominal clitics, from which they are technically indistinguishable. How is this case-checking implemented? For the purposes of this paper, I will assume the (admittedly passé) AgrP system, so that, in Polish, KPs containt AgrPs which can check genitive and dative case.

Since nothing more needs to be said about Pol, I put this language aside for the remainder of the paper. [Sorry!]

The opposite extreme is represented by languages such as SC, Slovenian, Czech and Slovak, which tolerate no NP-internal clitics.[2] In Franks and King (2000: 273) we provide the following illicit example:

(3) a. [Predstavljanje njega/***ga** Mariji]   je   iznenadilo   svakoga
        introduction   him.gen   Maria.dat aux.3sg surprised   everybody.acc
        'Introducing him to Maria suprised everybody'

Only the full KP pronominal *njega* (or nonpronominal, such as *Gorana*) can appear inside NP, never a clitic, such as *ga*. Presumably, this distinction has to do with the assumption that *ga* is a simple K°, whereas *njega* and its ilk have internal structure. So SC, like the other languages, has a mechanism to check case generally within NP, but this mechanism for some reason fails when the item being checked is a head. There are various brute force solutions to this puzzle, none of which I find persuasive. First, it could be that case-checking within NP is accomplished in some fundamentally different way than within the clause, such that within NP it does not involve Agr (or other) heads. Possibly, the difference lies in the categorial distinction between Ns and Vs, assuming that Agr checks accusative (and nominative), and the reason these are not found in NPs is because Agr itself is absent. If case-licensing via Agr (or whatever fuctional heads is implicated) requires V-features to activate it, and if clitics are pure functional heads, then inside NPs, unlike VPs, Agr could never in fact be activated.[3] This is however not a particularly desirable approach from a conceptual perspective, and also will raise questions once Bg and Mac are considered, since in these languages we would then need to say that NPs *are* like clauses, at least in that they seem to be able to contain an AgrIOP. This could only work either if the dative-like clitic in Bg and Mac DPs is not an instantiation of AgrIOP, or if Bg and Mac Ns differ from their counterparts in the other languages in being able to activate the case-checking property of AgrIO. Although certainly possible, this sort of approach involves brute force stipulation, at best. Since Bg and Mac instantiate verb-adjacent rather than 2P clitic systems, another possibility is that the correct account in some way exploits this distinction. In Franks (1998), Franks and King (2000), and Franks (2000), I argue that Bg and Mac pronominal clitics are actually Agr heads, with the form of K°s, hence generated outside VP,

[2]One exception is that clitics can appear inside NPs if they are arguments of appropriate adjectives or participles, as in SC *suvremeni* **mu** *pisci* 'contemporary to him authors'.

[3]Reflexive clitics seem to behave differently, which could follow from recent proposals that these instantiate a different kind of functional head than pure case clitics.

whereas in 2P systems they are generated in argument position, like full KP arguments. If so, then what is impossible in languages like SC is checking of case on a head, i.e. full KPs can move to [Spec, AgrP] for case-checking purposes, but K°s cannot (where they, or the relevant features, would have to move to Agr itself). Why this should be so is, however, unclear. The impossibility of having a clitic in 2P clitic languages in examples like (3) seems to be a striking striking and pervasive fact, one that if correct should find a straightfoward solution predicted by the architecture of the correct syntactic theory.[4] Unfortunately, at present no obvious way of capturing this comes to mind.

3. Basic Bulgarian and Macedonian facts I: the article. I now turn to the articulated forms in Mac and Bg. The former language has postpositive demonstrative forms, as in (4), which display a three-way spatial opposition,[5] whereas the latter simply has a postpositive definite article, as in (5).

(4) a. kniga-**va**   [Mac]
         book-this (here)

     b. kniga-**ta**
         book-the (that)

     c. kniga-**na**
         book-that (there)

(5) kniga-**ta**   [Bg]
     book-the

Since the chief substantive difference seems to be marking of proximity, in the following discussion I consider the two languages together (unless there is some special reason to distinguish them), and cite Bg examples exclusively, for ease of exposition.

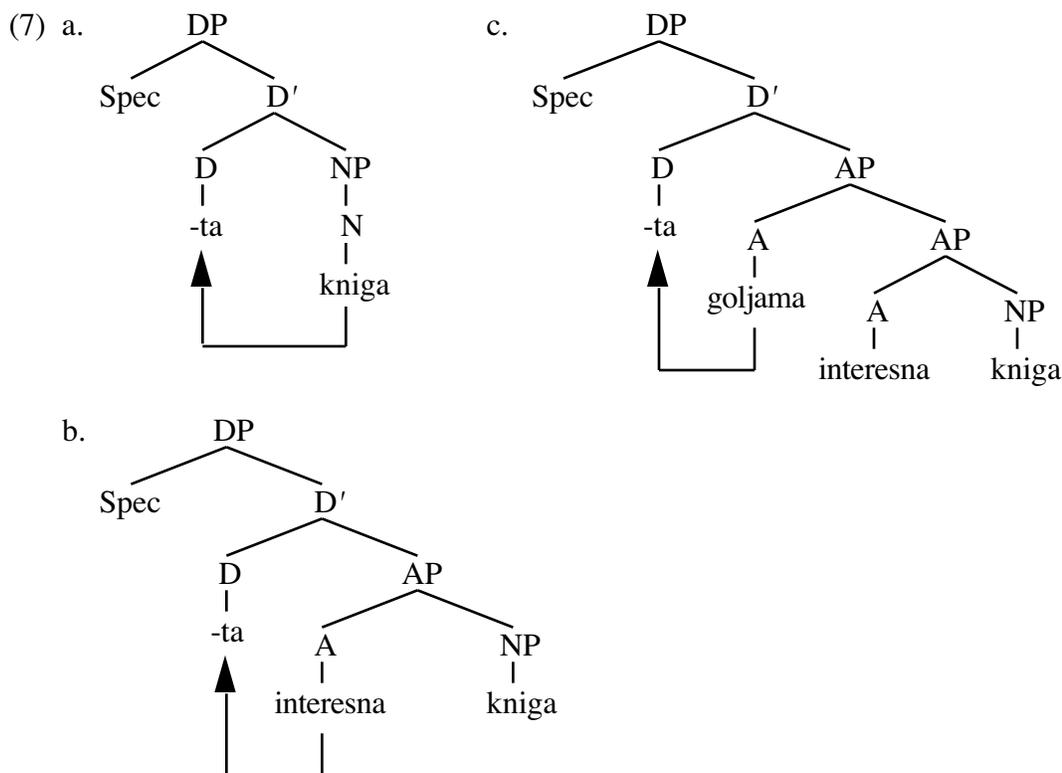   Some slightly more complex examples are provided in (6).

(6) a. kniga-**ta** 'the book'

     b. interesna-**ta** kniga 'the interesting book'

     c. goljama-**ta** interesna kniga 'the big, interesting book'

A sensible initial generalization to cover these examples is that the article attaches to the right of the first word in the nominal phrase. This might suggest an account in terms of head movement, as indeed has

---

[4]However, according to Roumyana Pancheva (pc), there were clitics inside the NP in Bg even before the 2P position period, i.e., when clitics appeared predominantly in a post-verbal position.
[5]See Koneski (1976: 225-237) for detailed discussion of the Macedonian facts.

been proposed, taking the article to be a D° and the highest substantive head beneath it raising to D°, roughly as in (7), from Fowler and Franks (1994).

(7) a.

```
            DP
          /    \
       Spec     D′
               /  \
              D    NP
              |    |
             -ta   N
              ▲    |
              |  kniga
              └────┘
```

c.

```
            DP
          /    \
       Spec     D′
               /   \
              D     AP
              |    /  \
             -ta  A     AP
              ▲   |    /  \
              | goljama A   NP
              |        |    |
              └──── interesna kniga
```

b.

```
            DP
          /    \
       Spec     D′
               /   \
              D     AP
              |    /  \
             -ta  A    NP
              ▲   |    |
              | interesna kniga
              └────┘
```

Note that in these structures I have adopted the account of the extended projection of NP in Abney (1987), according to which AP is actually a complement to D.

A little further investigation, however, reveals that this cannot be the correct analysis, for several reasons. For one thing, as we shall see in a moment, the article is an inflection, not a clitic. (Of course, inflected forms such as *knigata, goljamata* and *interesnata* could be inserted in D after head movement.) The real reason the head movement analysis must be rejected is that, as shown in the examples in (8), when the adjective is preceded by an adverbial modifier, the article must still follow the adjective. It can never simply follow the adverb.

(8) a. [mnogo xubavi]-**te** knigi 'the very nice books'

b. [dosta glupava]-**ta** zabeležka 'the quite stupid remark'

c. [priblizitelno sto]-**te** duši 'the approximately 100 people'

Neither is this phrase movement—the article does not simply appear after the AP, but rather, as shown by (9), it is precisely the *head* of the AP which is articulated:[6]

(9) a. [počti nerazrabotena-**ta**   u   nas]   problematika
         almost not-worked-out-**the**   by us   problematics
         'the problematics (which are) almost not worked out here [= in Bulgaria] '

     b. [polučena**ta**   sâs   mâka]   stipendija
         received-**the**   with pain   scholarship
         'the received with pain scholarship'

     c. [verni-**jat**   na demokratični idei]   prezident
         faithful-**the**   to democratic   ideas   president'
         'the president (who is) faithful to democratic ideas'

     d. [kupeni**te**   včera]   knigi
         bought-**the**   yesterday   books
         'the books (which were) bought yesterday'

     e. [zabranena**ta**   ot zakona]   kniga
         forbidden-**the** by law   book
         'the book (which is) forbidden by law'

Further examples that demonstrate that the head is targeted to bear the article will be provided in section 6, when the dative clitic is discussed; see also Franks (1998), Franks and King (2000), and Caink (2000) for statement of this descriptive generalization.

4.   Bulgarian and Macedonian articles are inflectional. In this section I first review arguments, summarized in Franks and King (2000: 278-284) which lead to the conclusion that the postpositive articles in Bg and Mac, despite their apparent mobility, are inflectional suffixes, not clitics. I then offer an appropriate structure to express this fact.

4.1. The form of the article. The article has the phonological properties of an affix, not a clitic. In particular, its form depends on its lexical host in ways that would be unexpected if it were a clitic. An initial descriptive generalization about the form of the article is that it reflects the final vowel of the word it attaches to if there is one, otherwise its form depends on the morphological gender of that word.[7] The

---

[6]Some of these examples are adapted from Mišeska-Tomić (1996), Dimitrova-Vulchanova and Giusti (1996), and Halpern (1995). Mišeska-Tomić (1996:521-522) notes that the Mac equivalent of (9a) is unacceptable, although this is independent of definiteness. The same would presumably hold for Bulgarian (9b), which she cites as (i) with the PP preceding the head A:

(i)   sâs maka polučena**ta** stipendija

Either order is acceptable, confirming my point that what is relevant in positioning the postpostitive demonstrative is the head of the phrase inflected for definiteness.

[7]See Mišeska-Tomić (1996) for discussion and instantiation of this principle and alternatives.

following set of Bug examples illustrates the range of complexity involved in determining the form of the article. The sets in (10) and (11) contain articulated nouns and adjectival modifiers of nouns which end in vowels, respectively.[8]

(10) a. glava**ta** 'the head'   [fem]

      b. dete**to** 'the child'   [neut]

      c. bašta**ta** 'the father'   [masc]

      d. čičo**to** 'the uncle'   [masc]

      e. ženi**te** 'the women'   [fem pl]

      f. mâže**te** 'the men'[9]   [masc pl]

      g. gradove**te** 'the cities'   [masc pl]

      h. sela**ta** 'the villages'   [neut pl]

      i. xora**ta** 'the people'   [pluralia tantum]

(11) a. nova**ta** kniga 'the new book'   [fem]

      b. interesno**to** selo 'the interesting village'   [neut]

      c. interesni**te** ženi/knigi 'the interesting women/books'   [fem pl]

      d. interesni**te** mâže/gradove 'the interesting men/cities'   [masc pl]

      e. dobri**te** deca/xora 'the good children/people'   [pluralia tantum]

      f. dve**te** knigi/sela 'the two books/villages'   [fem/neut pl]

      g. dva**ta**/tri**te** stola 'the two chairs'   [masc count form]

      h. dvama**ta** mâže 'the two men'[10]   [masc pl]

      i. mnogo**to** interesni knigi 'the many interesting books'   [fem pl]

---

[8]Friedman (1993: 264) describes the following situation for Macedonian: "In the singualr, masculines in a consonant take /-ot/, feminines in a consonant and all nouns in /-a/ take /-ta/... all remaining simgulars (and collectives, including l*ug'e* 'people') take /-to/. In the plural, nouns in /-a/ take /-ta/ , and all others take /-te/." See Koneski (1976: 228-229) for details and examples. The situation thus seems comparable to Bulgarian, other than the problem posed by collectives described below.

[9]Note that *mâže* is one of only five *-e* plural masculines, most monosyllabics taking the stem augment *-ove* and polysyllabics simply taking *-i*.

[10]This is a special numeral used for modifying masculine humans. It exists only for two through six, and can take regular plural or count form complements:

(i)   petima**ta**/šestima**ta** studenti/vojnici/studenta/vojnika
        'the five/six students/soldiers'

Crucially, the article is always **ta**, in keeping with the generalization provided below that when the host ends in *-a* the article is always phonologically determined (to be **ta**).

On the basis of these examples it seems as though the vowel of the article depends on the final vowel of the stem to which it attaches.

There are however two glitches, one minor, the other much more troublesome. First, stem-final *-i* is mapped into **e**, just as stem-final *-e* is, presumably **e** being the only front vowel option. This implies that selecting the appropriate vowel for the article requires inspection of the set of possible articles. Second, it looks as though when *-e* is a neuter singular ending, as in (10b), it is mapped into **o** rather than the expected **e**. Somehow these two types of *-e* must be differentiated.[11] It is necessary to assume reference to morphological information here to distinguish the two endings. Notice that a similar problem arises in (11f), where the fact that the *-e* of *dve* marks plural, rather than neuter, must be taken into consideration in selecting **e** as the vowel of the article. That this is not simply phonological is further demonstrated by *stote* 'the hundred' in (12).

(12) sto**te** sela/knigi/mâže 'the hundred villages/books/men'   [neut/fem/masc pl]

In (12) no appeal to phonology is possible. Instead, the relevant factor is that *sto* is plural rather than that it ends in an *o*. We are thus going to have to bite the bullet and admit morphological information in determining the form of the article. Compare this however with *mnogoto* 'the many' in (11i), where it indeed seems as though phonological factors prevail; *malkoto* 'the few' behaves similarly. The relevance of the final vowel of the host element, rather than its pluralness, is also made clear by the examples in (11g) and (11h). However, the fact that *mnogo* takes the neuter singular article **to** is probably not simply a phonological reflex, since in addition to these items *poveče* 'more' also requires **to** (with the meaning 'most'), as in (13):[12]

(13) poveče**to** sela/knigi/mâže 'most villages/books/men'   [neut/fem/masc pl]

Adopting a morphological solution might lead to the conclusion that Q(uantifiers) taking **to** are actually neuter nouns.[13] There is indeed a morphosyntactic distinction between Qs that take **te** and those that take

---

[11] One could conceivably maintain the the neuter singular *-e* is really a "basic **o**," along the lines of the approach to the Russian neuter singular ending in e.g. Townsend (1975). While credible for nouns like *četene* 'reading' or *momče* 'boy', which were after all originally *o*-stem neuters, this approach is slightly less motivated for others like *dete* 'child' and *vreme* 'time?', whose final *-e* is the historical reflex of the front nasal. Presumably all synchronic neuter nouns ending in-*e*, whether resulting from *o > e* after palatals or not, would have to be reanalyzed as underlying /o/ in order to force selection of the back vowel option for the article.

[12] Here as well, one could conceivably analyze the final *-e* as an underlying /o/ after a palatal, in an attempt to assimilate (13) to neuter nouns in *-e*.

[13] As such, they may be treated analogously to the *butilka vino* 'bottle (of) wine' type, where *butilko* seems to have the status of a Q, taking a bare NP complement.

to: the former admit complement NPs in the *brojna forma* 'count form', as in (14a), whereas the latter do not (14b), only accepting the regular plural (14c).[14]

(14) a. pet/sto vestnika/studenta 'five/a hundred newspapers/students'   [count form]

  b. *mnogo vestnika/studenta 'many newspapers/students'   [count form]

  c. mnogo vestnici/studenti 'many newspapers/students'   [masc pl]

   In the case of articulated nouns and modifiers of nouns ending in consonants the need to refer to morphological information is even more transparent. Masculine singular nouns generally end in consonants and correspondingly take **ât** or **a** as the article, as in (15).[15]

(15) grad**a**(**ât**) 'the city'   [masc]

The fact that masculines ending in vowels take the appropriate vowel suggests that (15) might have a phononological rather than morphological account; compare (15) with (10c, d) or (16):

(16) a. djado**to** 'the grandfather'   [masc]

  b. sâdija**ta** 'the judge'   [masc]

Thus, (15) too could conceivably be handled in phonological terms, assuming as in Scatton (1975) a stem final *jer*, and the masculine article then to be morphophonemically /t/ + *jer*. Even masculine singular adjectives might possibly be handled in a similar fashion, provided that they are analyzed as ending not only in a *jot* but also a *jer*:

(17) interesni**ja**(**t**) grad/bašta/čičo 'the interesting city/father/uncle' [masc]

   It turns out however that *feminines* ending in consonants take the ending appropriate to their gender (**tá**); with consonant-final feminines this ending is always stressed:[16]

(18) a. radost**tá** 'the joy' [fem]

  b. nošt**tá** 'the night' [fem]

We thus conclude that reference to morphological properties is inescapable. However, on the basis of all the examples examined so far, the need for such reference appears to depend on the host, as summarized in (19).

---

[14] *Vsički* 'all' only takes the plural article **te**: *vsičkite xubavi knigi* 'all the nice books'. In this respect it seems to behave look a regular plural adjective; cf. Dimitrova-Vulchanova and Giusti (1996) for examples and discussion.

[15] Literary Bg employs two distinct forms of the definite article for masculine singular consonant stem nouns: **ât** in oblique positions and **a** in nominative positions. Similarly, masculine singular adjectives will end in either **jat** or **ja**. This is a completely artificial situation, the result of drawing from two dialects in creating the literary norm, and is not reflected inordinary speech.

[16] In Mac, this combination is spelled with a single "t".

(19)   If the stem ends in an -*a*, then the article is always **ta**, otherwise it depends on morphological properties of the stem.

This interplay between phonological and morphological requirements is nicely demonstrated by the doublets in (20), where 'knees' and 'shoulders' have developed with two competing synchronic plurals (one the reflex of the old dual):

(20) a. kolene**te**   OR   kolena**ta**   'knees'

      b. ramene**te**   OR   ramena**ta**   'shoulders'

Similarly, when the articulated form is a numeric quantifier the article is usually **te**, as shown in (21).

(21) a. pe**tte** ženi/deca 'the five women/children   [fem/neut pl]

      b. dve**te** knigi/sela 'the two books/villages'   [fem/neut pl]

      c. tri**te** stola/konja 'the three tables/horses'   [masc count form]

      d. sto**te** mâže 'the hundred men'   [masc pl]

The sole exception, as expected given the generalization in (19), is when the numeral ends in -*a*, as demonstrated in (22).

(22) a. petima**ta** vojnici/vojnika 'the five soldiers'   [masc pl/count form]

      b. šestima**ta** studenti/studenta 'the six students'   [masc pl/count form]

      c. dva**ta** stola/konja 'the two tables/horses'   [masc count form]

    The combination of morphological and phonological factors is such that ultimately we are sometimes forced to allow specification of the articulated form in individual lexical entries. Another unpredictable fact is that, as can be seen from example (17), the masculine singular articulated adjective requires a special longer stem—the indefinite form would be simply *interesen*.[17] Further quirks exist, such as the "softening" article **ja**(**t**) used instead of **a/ât** with many stems that once ended in front *jers*, as in (23).

(23) a. kon ~ kon**ja**(**t**)   'the horse'

      b. pât ~ pât**ja**(**t**)   'the path'

      c. učitel ~ učitel**ja**(**t**)   'the teacher'

Once again, although this could be handled by retaining a final *jer* in these stems synchronically, that this is a lexical property is evidenced by the fact that any vocalic ending induces the same iotizing effect, as shown by the count form *konja* 'horses' in (22c). This is more or less expected if the article is essentially

---

[17] The same problem exists in Mac: *nov* '(a) new' but *noviot* 'the new'.

inflectional: there is one overrriding subregularity, that *-a* implies **ta**, then the form is selected in accordance with gender and number features, although occasional idiosyncrasies which, as always in morphology, block the regular rule from applying.

Finally, collectives in Mac might pose a special problem. As observed by Mišeska-Tomić (1996), they show plural agreement but take the neuter singular article **to**, as in Mac (24):

(24) a. lug'e**to** 'the people' BUT   dobri**te** ljug'e 'the good people'

　　 b. nivje**to** 'the meadow?'   BUT   ubavi**te** nivje 'the beautiful meadow?'

Morphologically these are neuter singular, but semantically they are plural, as evidenced by the form of the adjective. The only recourse is to indicate their mixed behavior in the lexicon, i.e. the have definite forms as indicated by are plural. If Bg had such forms, we would be faced with differentiating them on the one hand from the *mâže* 'men' type,[18] which takes plural adjectives and **te**, and on the other hand from the *dete* 'child' type, which is perfectly regular in taking singular adjectives and **to**. However, this may not be a problem, as speakers I consulted provided the following completely regularized forms; Bg (25a) assimilates *ljude* to the *mâže* type, (25b) treats *nivja* as a neuter plural, and in (25c) *nivi* behaves like a feminine plural:

(25) a. ljude**te**   AND   dobri**te** ljude

　　 b. nivja**ta**   AND   hubavi**te** nivja

　　 c. nivi**te**   AND   hubavi**te** nivi

Judgments were for some speakers somewhat strained, as these are archaic forms in Bg, *ljude* being regarded as high style and poetical, *nivja* as a dialectal form.[19] If there are transitional dialects with both patterns, then they would have to be regarded as lexical exceptions, which is to be expected if the selection of article is part of the morphology of the stem. I thus concur with Halpern (1995) that postpositive demonstratives in Bg and Mac must be analyzed as a special type of inflection.

4.2. Some phonological arguments. I now turn to some phonological arguments that these postpositive articles are inflectional. As observed by Halpern (1995), and discussed also by Mišeska-

---

[18]Curiously, although accidently, Mac (24) does not present this problem, since the Bg *mâže* type plural is not found (this form is *maži* instead). However, one dilemma remains that cannot be resolved purely on the basis of phonology: Mac still has two plurals in *-e*, resulting from older duals, and these are articulated as expected: *racete* 'arms' and *nozete* 'legs'; cf. Koneski (1976: 228). The *ljug'eto* type in (24) therefore still needs to be differentiated from them.

[19]Other speakers however find the forms in (25) all perfectly acceptable.

Tomić (1996: 532-533) and Franks and King (2000: 283-284), there are several phonological factors which prove that Bg and Mac articles are not clitics. In addition to requiring the long form, as pointed out in connection with (17), two other phenomenona noted by Halpern as favoring the suffixal over clitic analysis are (i) blocking of final devoicing and (ii) triggering liquid vowel metathesis. Both final devoicing and metathesis are word-level processes, and consequently are expected to take the article into account. Final devoicing thus takes place in Bg (26a) and Mac (27a), but is blocked by the vowel-initial article in the (b) examples.

(26) a. bratovčed   [bratofčet]   'cousin'

     b. bratovčed**ât**   [bratofčedət] 'the cousin'

(27) a. maž   [maš]   'husband'

     b. maž**ot** [mažot]   'the husband'

On the other hand, with true clitics, such as the "dative" possessive to be discussed in the section 6, final devoicing still occurs. An example is given in Mac (28), cited by Mišeska-Tomić (1996).

(28) maž ì   [maši]   'her husband'

Unfortunately, because as we shall see the possessive clitic essentially occurs only in definite DPs, I have been unable to construct any comparable Bg examples, although (34) below comes close.[20]

    This same problem afflicts the liquid-metathesis diagnostic, with further complications, although I have been able to devise one relevant example to confirm the expectation that the article should not behave like a true clitic. Following Zec (1988), so-called "liquid metathesis," as in the Bg example in (29), actually results from schwa epenthesis.

(29) mlâk 'silence!'   ~   mâlkom 'silently'

In Franks (1990), I argue that this root is morphophonemically /mlk/, with the syllabification in (30) before epenthesis. A schwa is consequently inserted in the appropriate place to create a closed CVC syllable, where "•" demarcates syllable boundaries.

(30) •mlk•   ~   •ml•kom•

---

[20] Some Bg speakers accept the unarticulated variant *mâž ì* with the pronunciation [məši] , and contrasting as expected with definite *mâžât* [məžət].

With this in mind, note that addition of the article generally results in the same alternation. The examples in (32) thus have the non-syllabic representations in (31) before schwa epenthesis; "#" indicates an underlying jer.

(31) a. •grk•   ~   •gr•k#t•[21]

    b. •grb•   ~   •gr•b#t•

(32) a. grâk 'Greek'   ~   gârk**ât** 'the Greek'

    b. grâb 'back'   ~   gârb**ât** 'the back'

These alternations show unequivocally that resyllabification can take place after the addition of the article.[22]

What about dative possessive clitics? Although these require a definite context, many kinship terms, because they are inalienably possessed, take possessive clitics even when they lack an overt article. The Bg examples in (33) behave this way:

(33) a. majka mu 'his mother'

    b. brat ì 'her brother'

    c. žena mi 'my wife'

    d. čičo ni 'our uncle'[23]

    e. lelja ti 'your aunt'[24]

---

[21] Scatton (1975) would more precisely represent these words as /grk#-t#/ and /grb#-t#/, with the vocalized *jer* triggered by the *jer* of the article; cf. the discussion of example (15). Under this approach to the masculine article, Ernest Scatton (pc) observes that the segmental features of the *jer* must in some way "echo" those of the preceding stem desinence; perhaps some kind of autosegmental feaure spreading might turn out to be effective here. Although not generally accepted, this type of system seems to me to have much to recommend it, and of course requires that the article be treated as affixal. However, I abstract away from these details of analysis here.

[22] Franks (1990) opposes such words to those with underlying schwas, such as *krâg* 'circle' and *blâv* 'cliff', that show no metathesis. Other words, however, are able to participate in V~Ø alternations even though they exhibit no metathesis. The following examples are drawn from Aronson (1968: 121-127):

(i) a. teatâr 'theater'   ~   teatâr**ât** 'the theater'   ~   teatri 'theaters'
    b. otrasâl 'branch'   ~   otrasâl**ât** 'the branch'   ~   otrasli 'branches'
    c. lakât 'elbow'   ~   lakât**ât** 'the elbow'   ~   lakti 'elbows'
    d. kosâm 'strand (of hair)' ~   kosâm**ât** 'the strand'   ~   kosmi 'strands'
    e. potomâk 'descendant'   ~   potomâk**ât** 'the descendant'   ~   potomci 'descendants'

Following Scatton (1975), in Franks (1990) I posit an underlying # in these forms, which is deleted before endings beginning in vowels other than #. Thus, vocalization persists in the articulated forms (as well as in the count form and vocative for most such nouns), which can be treated as the failure of deletion before another #, whereas the plural ending -*i* causes deletion.

[23] Technically, *čičo* is the 'brother of one's father', although it is also used when the precise relationship is unknown as well as to address male adults in general; cf. Scatton (1993: 244).

[24] Although *lelja* means the 'sister of one's father or mother', it is also used when the precise relationship is unknown as well as to address female adults in general; cf. Scatton (1993: 244).

As demonstrated by Mišeska-Tomić (1996: 526-531), despite the lack of overt article these are always definite in interpretation, hence they cannot occur in indefinite contexts, such as after the indefinite article *eden* 'one'. Given this phenomenon, one marginally credible Bg example with liquid metathesis, due to Lily Grozeva (pc), is offered in (34):

(34) [Grâb ì] se viždaše otdalex. [grəpi] 'Her back was visible from afar.'

This example has poetical status at best; 'her back' in Bg must be expressed as *gârbât ì* if a clitic is used.The point is, however, that—to the extent there can be valid judgments about such hypothetical forms—the preference for *grâb ì* over the completely impossible alternative **gârb ì* reiterates both phonological observations about how the possessive clitic is different from the article: it fails either to trigger metathesis/resyllabification or to block devoicing.

Halpern (1995: 172-183) offers several other reasons for taking the article in Bg and Mac to be an inflectional suffix rather than a clitic. One is the fact in (33) that certain kinship terms have special unarticulated definite forms. These are intrinsically definite, and as shown by (35) it is impossible to inflect them overtly for definiteness.[25]

(35) a.   *majka**ta** mu 'the his mother'

b.   *brat**a/ât** ì 'the her brother'

c.   *žena**ta** mi 'the my wife'

Interestingly, as can be seen in (36), the article reappears once the noun is modified:

(36) a.   xubava**ta** mu majka 'his pretty mother'

b.   pomladi**jâ**(**t**) ì brat 'her younger brother'

c.   inteligentna**ta** mi žena 'my intelligent wife'

This shows that failure of the words in (35) to inflect overtly for definiteness is a lexical idiosyncracy of these words, as does the fact in (37) that the definite form of other kinship terms expresses the article:

(37) a.   djado**to** mu 'his grandfather'

b.   sin**â**(**t**) ì 'her son'

---

[25] As Halpern observes, this is comparable to the (inflectional!) plural of English *fish* and *sheep* being unmarked morphologically, as is the effect of attempting to add overt morphology:

(i)   a.   *fishes
       b.   *sheeps

That this is a lexical fact is further supported by Mišeska-Tomić′s (1996: 531) comment that there is

variation between Mac and Bg. She provides the contrast in (38), explaining that "unique reference" is

for some reason "more readily established" for Mac *drugarka* than for Bg *prijatelka*:

(38) a.   prijatelka**ta** ti   'your (best) girlfriend'   [Bg]

b.   drugarka ti   'your (best) girlfriend'   [Mac]

Note finally that omission of the article is dependent on the interpretation.[26] Whereas Bg (35c) *ženata mi*

is unacceptable in the meaning 'my wife', Halpern (1996: 173) points out that it is the correct way to say

'my maid/nanny'. Interestingly, Mišeska-Tomić notes that Mac (38b) with the article, i.e. *drugarkata ti*,

is acceptable in the meaning 'your wife'.[27]

5. Analysis of the article. I have presented various kinds of arguments in favor of the claim that Bg

and Mac articles are affixes on words rather than clitics situated in D. The first is that attempts to position

the necessary articulated form immediately to the left of D° are unworkable, given the descriptive

generalization that the article appears on the leftmost adjectival head (or noun, if unmodified) in the DP.

The second is that the article's form transparently depends on the stem to which it attaches, through what

generally seems to be a "vowel harmony" mechanism, slightly tempered by lexical and morphological

factors. The third is that available phonological tests show the article to form a single "word" with its

stem. Articulated items constitute lexical entries for the purposes of the phonology, feeding lexical and

cyclic rules.[28]

　　We want a structure that reflects the fact that the articulated form is invariably the head of its phrase,

i.e., that although definiteness is a feature associated with XP, it is realized as an affix on X. Since the

definiteness feature resides in D, and each of the XPs which we have seen mark definiteness can be

analyzed as the complements to D, what this means is that the highest phrase under DP is the one which

contains the articulated form as its head. That is, in structure (39), if D is [+definite} and if QP is

---

[26] It seems to me that the parallel with the English plural extends here to, in that *fishes* can be used to describe different kinds of fish.

[27] This is curious, since (in most societies) 'best friend' and 'wife' are equally unique in reference.

[28] True clitics do form a *prosodic* word with their hosts, and are subject to relevant post-lexical phonological rules, such as antepenultimate stress assignment in Mac. Thus, addition of the possessive clitic *mi* 'my' can shift the stress one syllable to right to preserve antepenultimate stress:

(i)   žéna**ta** 'the wife' BUT žená**ta** mi 'my wife'

present, then Q is articulated, if not, and if AP is present, then A is articulated, and if there is just an NP, then N is articulated.

(39)

```
        DP
       /  \
      D    QP
          /  \
         Q    AP
             /  \
            A    NP
                /_\
```
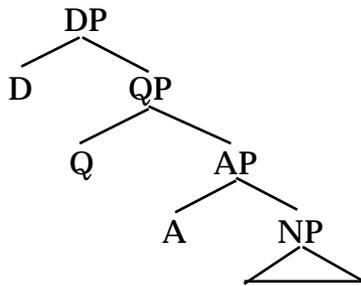
I assume, as in Abney (1987), that AP dominates NP. It takes the definiteness features of D to restrict everything in their scope, essentially the c-command domain of D, so that marking them on the highest head in this domain seems perfectly appropriate. I imagine the process of feature licensing to involve LF-movement of this highest head to D, with the raising taking place, in conformity with shortest move, in order to check off the definiteness feature of that head against the same feature in D.

Željko Bošković (pc) points out a potential problem for this sort of account. What, he asks, technically prevents definiteness from being accidentally marked on (the head of) a lower X^{max} in structure (30), since a higher head that remains unmarked for definiteness should be irrelevant. My response to this objection is that *all* substantives, i.e. [+N], heads in Bg and Mac are relevant, in that they bear features for [±definite]. Every such word is listed in the lexicon with definite in definite forms. Thus, movement over any such head will indeed violate the Head Movement Constraint. In this way, the HMC highest head effect is reconciled with problem of locating the article on the correct morpheme.
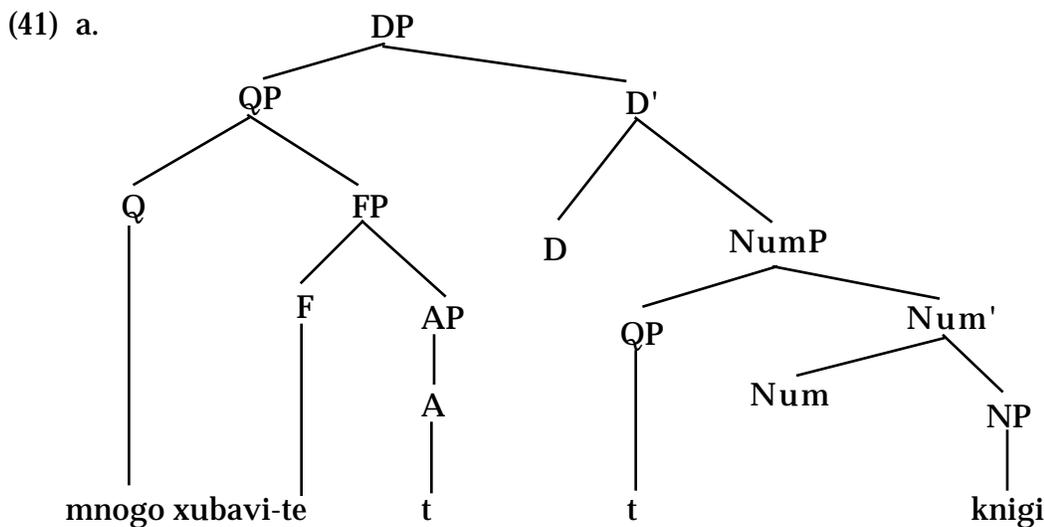
The analysis in Dimitrova-Vulchanova and Giusti (1996) is similar to mine in that they also claim that the articulated form "inflects for definiteness" on the basis of the fact, copiously demonstrated above, that the selection of article depends on the stem on which it appears. For them, however, this simply precludes an analysis in terms of A-to-D movement, which can be independently rejected on the basis of examples such as those in (8), e.g. *mnogo interesnata kniga* 'the very interesting book', where A clearly does not raise out of its AP. However, they take this to mean that the entire AP is in fact moving to [Spec, DP], assuming it to be generated in the next lower specifier position. One could, however, as an alternative to the Abney-type analysis in (39), claim that the XP in the specifier of the complement to D
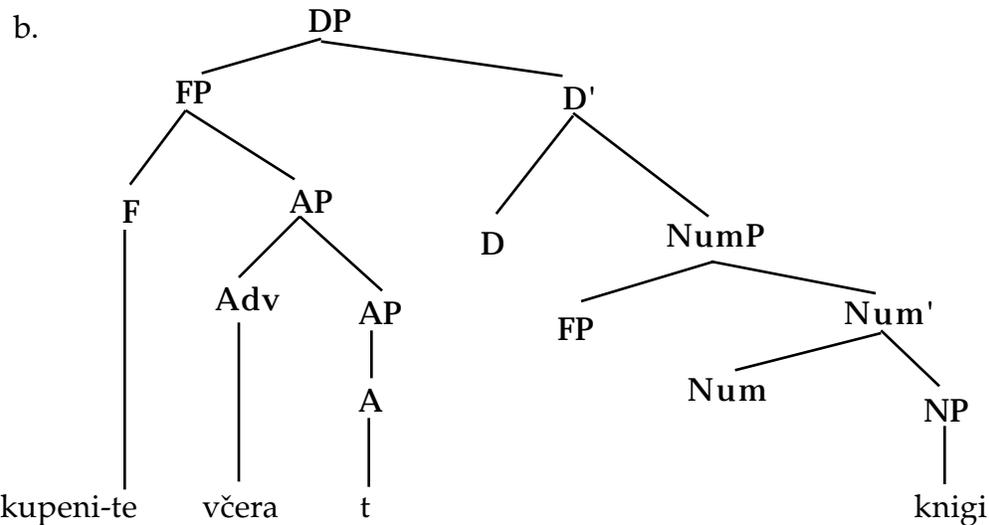
raises *overtly* to [Spec, DP] and then, through Spec-head agreement, XP is checked for definiteness against the features of D. Just as in the account proposed early without overt movement, the definiteness features of XP would be realized on its head X. Schematically, the structure in (40) would represent this scenario:

(40)
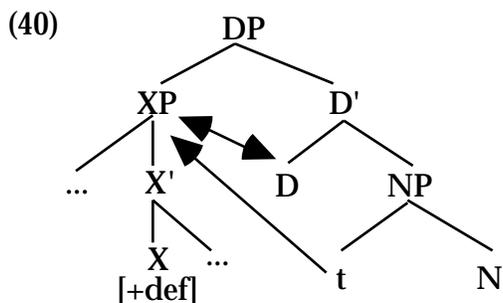
Unfortunately, (40) is *not* what Dimitrova-Vulchanova and Giusti propose. Instead—to the extent that I can figure out what their claims actually are—they try to replicate the "movement" analysis within XP. That is, they combine a Spec-head agreement system (to license definiteness) with a movement analysis (to position the elements properly). However, we have seen that attempts to place the right material in front of the article invariably fail.

Explicit structures provided by Dimitrova-Vulchanova and Giusti (1996: 48-49) for examples (8a) and (9d) are as in (41):

(41) a.

b.

```
                    DP
           ┌─────────┴──────────┐
          FP                    D'
       ┌───┴───┐          ┌──────┴──────┐
       F       AP         D            NumP
       │    ┌───┴───┐              ┌─────┴─────┐
       │   Adv     AP             FP          Num'
       │    │       │                      ┌───┴───┐
       │    │       A                     Num      NP
       │    │       │                               │
    kupeni-te  včera  t                            knigi
```

To summarize these representations, the article is generated as the head of FP, a generic functional category apparently part of the extended projection of A and immediately dominating AP. The adjective then undergoes head movement to left-adjoin to F. In (41a), FP is itself embedded in a QP in order to position the quantificational adverb *mnogo* 'very' in front of *xubavite* 'nice-the', whereas in (41b) *včera* 'yesterday' is placed in specifier position within AP. The extended projection of AP is then generated in [Spec, NumP], a phrase that intervenes between DP and NP, and raises from there to [Spec, DP].[29] The inconsistencies and complexities involved here should be obvious, FP serving no purpose other than providing an extra site around which to manipulate the structure. First, it is only the *highest* in a series of APs that will be contained in an FP. Second, although Dimitrova-Vulchanova and Giusti say nothing about articulated nouns, clearly in this case NP will need also to be contained in an FP which then raises to [Spec, DP], although again only in the absence of an AP.[30] Third, locating QP above FP to accommodate pre-head modifiers is I think suspect, given that they same problem arises regardless of

[29] It is unclear to me why Dimitrova-Vulchanova and Giusti require this NumP, as they could just as well have generated the AP in [Spec, NP]. Presumably there are theory-internal reasons for positing a NumP, although it seems superfluous in their analysis of Bg. One possible effect of NumP, although I doubt this is something Dimitrova-Vulchanova and Giusti had in mind, is that by providing a position for AP outside of NP it may facilitate an economy approach to ruling out *knigata interesna* 'book-the interesting', requiring the node to be contained in FP to be the higher one (AP, in [Spec, NumP]), not the lower one (NP, the complement of NumP) . However, this may not help, since they are probably equidistant from the target [Spec, DP] position anyway. (Note that if AP were generated in [Spec, NP] the highest/most economical place for FP would be above NP, in which case the entire NP would raise to [Spec, DP] and the strongly ungrammatical variant *interesna knigata* 'interesting book-the' would result.)

[30] It is possible that checking theory combined with shortest move requirements could force a single FP on the highest candidate, if extra FPs would not have their definiteness features checked and if generating FP in lower positions would result in less economical derivations.

what precedes that adjective—e.g. *včera* 'yesterday' could just as well precede *kupenite* 'bought-the' in (9d/41b) with no effect on article placement; cf. the examples in (9) and especially fn. 6.

A further interesting problem is raised by the possibility of having a demonstrative cooccur with the article. For one thing, Mišeska-Tomić (1996) cites the Mac examples in (42) and the Bg ones in (43).

(42) a. Me izede ovoj moj**ov** sin.   [Mac]
       me ate this my-this son
       'This son of mine is killing me.'

      b. Go sakam onoj sini**on** fustan.
       it want.1sg that blue-that dress
       'I want that there dress, the blue one.'

(43) a. Onĵâ moj**ât** sin mi izjade glavata.   [Bg]
       this my-the son me ate head-the
       'This son of mine is killing me.'

      b. Iskam onâzi sina**ta** roklja.
       I-want that blue-the dress
       'I want that dress, the blue one.'

Similar examples from colloquial Bg are cited by Caink (2000), who attributes them to Olga Arnaudova (pc and 1998), as in (44):

(44) a. tazi kniga**ta** 'this book-the'   c. tazi hubava**ta** kniga 'this nice-the book'

      b. tozi čovek**ja** 'this person-the'   d. tozi neprijatni**ja** čovek'this unpleasant-the person'

Sometimes the article is even obligatory, as in (45):

(45) a. tozi sâšti**ja** čovek 'this same-the man'   [Bg]

      b. tova cjalo**to** čakane 'this whole-the waiting'

With these adjectives the article cannot be omitted, despite the presents of the the demonstrative.

Such examples would be easily accommodated under my account, assuming the demonstrative to be generated in [Spec, DP] and definiteness, as always, to be realized as the article on the head of the complement to D. The definite features of the article would then raise to D at LF to be checked against the definiteness features of D. That is, in (42)-(45) definiteness would be doubly articulated, expressed as a demonstrative in [Spec, DP], whose features are checked under Spec-head agreement in the syntax, and as an article, whose features are checked under head-head adjunction at LF. This multiple checking, which following Chomsky (1995) is possible because definiteness is a semantically relevant feature hence not deleted when checked, has two interesting empirical consequences. Doubling is employed in

emphatic contexts and there is a matching requirement in proximity features between the demonstrative

and article. Note that the matching effect is only visible in Mac, since Bg only has the

single article, which is presumably non-distinct from both distant and proximate demonstratives, hence

e.g. **ta** 'the' can be checked against the same features that license both *tâzi* 'this' and *onâzi* 'that'. In

Mac, however, an article in -**n**- or -**v**- can only be used with a demonstrative in -**n**- or -**v**-; other non-

matching combinations for Mišeska-Tomić's *ovoj mojov* in (42a) would be impossible:

(46) a. ovoj moj**ov**   *toj moj**ov**   *onoj moj**ov**   [Mac]

     b. *ovoj moj**on**   *toj moj**on**   onoj moj**on**

This identity follows if both demonstrative and article are checked against the same set of definiteness

features.

Notice however that Mišeska-Tomić (1996) also cites Mac examples that do not conform to the

matching requirement, such as (47):

(47) celi**ot** ovoj ostrov 'this whole island'   [lit. 'whole-the this island']   [Mac]

She argues, I think convincingly, that here the article represented on the quantifier and the

demonstrative originate in separate DPs. This makes sense since the articulated form precedes rather than

follows the demonstrative. If so, I would analyze (47) with a layered structure, with the adjectival

quantifier *celiot* sandwiched between DPs.[31] There are then two realistic possibilities—the structure in

(48a), which generates *ovoj* is in [Spec, DP], or the structure in (48b), if there is never anything overt in

DP in Slavic (a typologically credible generalization, as is DP iteration), which respects the morphology

in treating *ovoj* just like any other adjective heading an AP:

(48) a. [DP [D +def] [QP celi**ot** [DP ovoj [D +def, +prox] [NP ostrov]]]]

     b. [DP [D +def] [QP celi**ot** [DP [D +def, +prox] [AP [A ovoj] [NP ostrov]]]]]]

The definiteness features of **ot** would be checked against the higher D at LF and the proximate definite

features of *ovoj* would be checked against the lower D.[32]

---

[31] I label it QP here, but morphological considerations lead me to think that technically non-numeric quantifiers are simply quantificational adjectives. The distinction is not vital to the discussion, so long as the articulated form is analyzed as the head of some phrase; cf. Franks (1995) for discussion of QP in Slavic.

[32] A similar approach could be employed for Mišeska-Tomić's (i), although it should be borne in mind that, unlike Bg *vsički*, Mac *si-* 'all' never occurs without the definite article, i.e. it is only *si**te***:

(i) si**te** ovie studenti 'all these students'   [lit. 'all-the these students']

Notice now that this need for layered DP structures when quantifiers (and conceivably other adjectival modifiers with wide scope) precede the demonstrative suggests DP interation as the correct analysis of *all* colloquial doubled forms. I leave the choice between these alternatives open for the time being. However, even if the correct analysis does turn out to involve this kind of layered multiple DP structure, there is still a crucial dependency in proximity features between the two DPs which must be captured. In all examples of demonstratives cooccuring with articulated forms, the only acceptable mismatch in proximity features that I am aware involves articles with the root -**t**-. This is explicitly pointed out by Koneski (1976: 231-232), who observes that although articulated forms in -**t**- can double demonstratives in -**v**- or -**n**-, articulated forms in -**v**- or -**n**- can only double demonstratives based on the same root. Thus, the row of *acceptable* mismatches in (49) can be added to the unacceptable ones in (46):

(49) c. ovoj moj**ot**   toj moj**ot**   onoj moj**ot**

This is however not especially surprising, if we take Friedman's (1993: 261) description of the various forms literally: Friedman characterizes articles in -**t**- as "unmarked", articles in -**v**- as "proximate," and articles in -**n**- as "distal." In terms of features, then, we can define these as [+def], [+def, +prox], and [+def, –prox], respectively. Notice now that articles in -**t**- are technically non-distinct from all three demonstratives. They are thus able to exhibit the same mismatch as is forced in Bulgarian by the absence of the [±prox] feature on articles; see Koneski for speculations as to the semantic factors involved.[33]

Consider now the analysis in Dimitrova-Vulchanova and Giusti (1996). Since this involves moving XP to [Spec, DP], on the face of it their analysis should encounter difficulties with this sort of construction, because [Spec, DP] appears to be already occupied by a demonstrative. They are however aware of the problem, citing examples such as (50), where the article is obligatory.

(50) tezi vsički*(**te**) knigi 'all these books' [lit. these all-the books']

The DP-iteration analysis would be reasonable here, with *vsičkite* sandwiched between the DPs, which would assure *vsičkite* rather than *vsički*; the only unresolved issue being why a single DP would not be an option. This could I believe be handled through selection, taking *vsički* to require a DP complement.

---

[33] There seem to be no examples of mismatch under non-distinctness in the other direction, with the article specified for [±prox] but demonstrative left unmarked: *toj mojo**n**/mojo**v**. Apparently the article cannot be more specified than the demonstrative. This state of affairs strikes me as strongly reminiscent of other Spec-head agreement phenomena, although I imagine it could equally well be handled by DP-iteration.

After all, it is never lower in the DP than immediately below DP. This assumption is supported by the examples in (51), also from Dimitrova-Vulchanova and Giusti (1996), which further illustrate the distribution of *vsički*.

(51) a. knigi**te** gi pročetox vsički**te**   'the books I read them all'

     cf. *knigi**te** gi pročetox xubavi**te**   '*the books I read them nice'

  b. vsički nas/nie   'all of us/we all'

  c. vsički(**te**) tezi knigi 'all these books'

  d. moi**te** vsički**te** knigi

The pair in (51a) suggests *vsičkite* constitutes a separate DP, something ordinary adjectives cannot do. The example in (51b), taking the pronoun to be a D (although with accusative and nominative variants), suggests that *vsički* is outside this DP, and in fact in apposition to it. Both (51c) and (51d) appear to require a higher DP, the first because articulation precedes the demonstrative, the second because of the double articulation.

Dimitrova-Vulchanova and Giusti (1996) however propose a completely different analysis, indicated in (52).

(52)
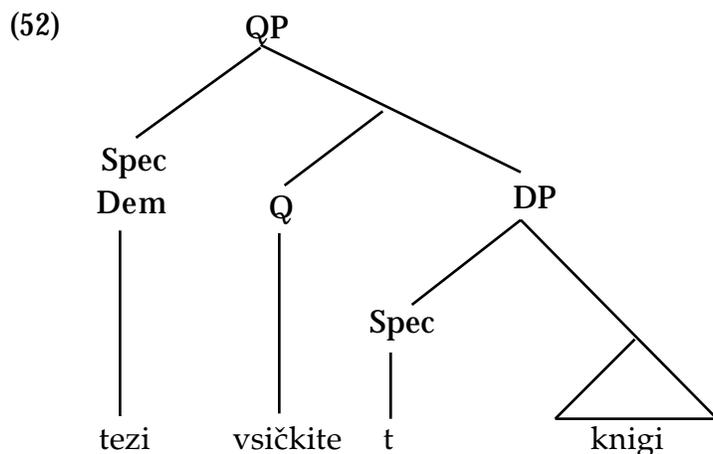
Dimitrova-Vulchanova and Giusti indicate that *tezi* 'these' raises from [Spec, DP] to [Spec, QP]. They then claim that the reason the article **te** appears on *vsički* 'all' is through Spec-head agreement with *tezi*. The connection between this structure and the FP method of deriving the article, adopted for all other examples, remains however completely mysterious. Moreover, to handle (53), where the article is

optional, they propose yet a third mechanism for achieving the article, whereby it is optionally inserted in D when [Spec, DP] contains a demonstrative and Num contains a cardinal numeral.

(53) tezi dva(**ta**) stola 'these two tables'

Their analysis is clearly missing a generalization about the distribution of the articulated form. This generalization, as I have already described, is simply that the article appears on the head of the complement to DP. While the possibility of [Spec, DP] also containing a demonstrative introduces certain complexities having to do with when doubling of the article is possible, impossible, or required, this does not cause serious problems for my analysis as it does for theirs. Indeed, even their system of multiple unrelated mechanisms to handle articulated forms will fail to handle examples with demonstratives *plus* articulated adjectives or even articulated nouns. Dimitrova-Vulchanova and Giusti claim these do not exist in Bg, citing (54), but the examples from Arnaudova and Caink, with both demonstratives and articulated adjectives, provide direct contradiction.

(54) tezi xubavi(****te**) knigi 'these nice books'

   Macedonian is even clearer in this regard. In addition to the examples of demonstratives plus articulated adjectives discussed above, even demonstratives plus articulated nouns are quite common in the colloquial language. Friedman (1993: 261) cites (55a) with the comment that this is not normative and Koneski (1976: 230-231) provides a number of comparable examples, as in (55b).

(55) a.   ovie deca**va** 'these children'

      b.   toj čovek**ot** 'this person'

Koneski states that this doubling is a way of emphasizing the specificity of the referent. It is even possible to iterate the article, as in (56).

(56) onoj moj**ot** najmali**ot** 'that there youngest (one) of mine'

Once again, this possibility might arise under the feature checking model, given the assumption that semantically relevant features are not deleted upon checking.[34] I conclude that my generalization about the distribution of the article is the correct one, although the limited and colloquial possibility of coupling articulation with a demonstrative presents some interesting problems.[35]

---

[34] Alternatively, it could be taken as indicative of further DP iteration, consistent with my claim that it is the highest head in the complement domain of D that realizes definiteness and proximity features, under an F checking relationship with D.

[35] This possibility is very much akin to clitic doubling, as in Mac (i):

We have just seen that one problem with treating the article as a clitic is that it can coocur with a demonstrative and even may be multiply expressed, as in (56). A corollary problem is that the definite article can sometimes remain unexpressed even in definite contexts, as in the examples in (33). It thus lacks many of the properties of true Determiners. There are finally important theory-internal reasons for not treating the article in Bg and Mac as an instantiation of D. In more familiar instances of N-to-D raising, such as the Italian cases discussed by Longobardi (1994), N moves over adjectives, as in (57a), but we have seen that this is impossible in Bg (57b):

(57) a.   il mio Gianni OR Gianni mio BUT *mio Gianni

       b.   *moja kniga**ta** AND *kniga**ta** moja BUT moja**ta** kniga

Another difference would be that in Bg proper names are the least likely to be articulated, and they certainly do not raise over modifiers, whereas in Longobardi's work on Romance these are shown to be the most likely to undergo N-to-D raising; cf. e.g. Dimitrova-Vulchanova and Giusti (1995) for comparison with Romanian. Moreover, in Romance, N raises when D is empty rather than when it is filled, as would have to be the case in Bg. Furthermore, if head movement were involved, we might expect the possessive clitic *ni* 'our' to induce a HMC violation in examples such as (58).[36]

(58)   stolica**ta** ni 'our capital'

The head N *stolica* 'capital' would have to skip the intervening clitic head *ni* in moving up to the putative D **ta**.

All these considerations support my conclusion that treating the article as a Determiner clitic is incorrect, and that it is best analyzed as an inflection for definiteness. This inflection invariably appears on the highest lexical head under DP, taking D itself to consist solely of features for specifity and proximity. Compelling confirmation of this distribution can be seen in the Bg possessive adjective paradigm. Consider the in itself curious fact that possessive adjectives, described by Corbett (1987) in remarkable detail for virtually all the Slavic languages, can be formed either from both members of

(i)   Daj mi ja knigana 'Give me it$_{ACC}$ that book'

The feminine singular accusative clitic *ja* doubles the definite object *knigana* 'that book'; such doubling is obligatory with definite objects in Macedonian, while it generally serves more stylistic purposes in Bulgarian, typically emphasis. Taking the clitic *ja* to be some kind of case/agreement head checked against features of AGR and *knigana* to be a DP checked in [Spec, AgrP], the parallelism between this relation and the article–demonstrative relation is quite striking.

[36] This is a patent defect with the analysis in Fowler and Franks (1994), under which N (and Q) undergoes head movement but AP raises to [Spec, DP]. I am grateful to Longobardi's student Monica Gambarotto (pc) for first drawing my attention to the problem posed by the failure of *ni* to block N-to-D movement in examples like (58).

compound names or just from the second member. Crucially, the placement of the definiteness morpheme depends on which strategy is adopted, as shown by the examples in (59):[37]

(59) a. čicova**ta** Stojanova kâšta   'Uncle Stojan's house'

      b. čico Stojanova**ta** kâšta

When the alternative whereby the first element is a possessive adjective is employed, it is this element which is inflected for definiteness, as in (59a). If, on the other hand, *čico Stojan* 'Uncle Stojan counts as a single word for the purposes of possessive adjective formation, as in (59b), then the article appears on the second element. Thus, even under these complex conditions, it is always the highest head that bears the definiteness suffix.

6. Basic Bulgarian and Macedonian facts II: the pronominal clitic. I now turn to the types of DP-internal clitic pronouns found in Mac and Bg. I shall refer to these historically dative forms as "oblique" pronominal clitics (and gloss them as such) in order to cover their two distinct functions. As we have seen, their basic function is that of a possessive pronoun. This will be the focus of sections 6.1 and 6.2. Somewhat parasitic on this function is the use of oblique pronouns to express what appear to be argument relations in deverbal nouns, as discussed in section 6.3. Once again, in the discussion I will focus on Bg, although my reason here is more practical: in Mac the use of DP-internal oblique clitics is highly restricted, essentially limited to the expression of possession of "family or family-like" relations; cf. Mišeska-Tomić (1996: 526-531) for comparison of the Bg and Mac systems.

6.1. Possessive clitics. The case systems of Mac and Bg bear an uncanny resemblance to that of English. There are three cases, nominative/subjective, accusative/objective, and dative/oblique. These are manifested most clearly on pronouns, the only distinctions on nouns being isolated historical remnants and two notable artifices of the literary languages: an optional nominative—non-nominative opposition realized on masculine names (and a few other words) in Literary Macedonian, and a similar case opposition realized in the selection of the masculine singular article in Literary Bulgarian, although this contrast is properly speaking not one of the noun *per se*, it just happens to appear on nouns as well as their modifiers; cf. also fn. 15. In the pronominal systems there are nominative pronouns for subjects, accusative pronouns for objects of verbs and prepositions, and dative pronouns for indirect objects. The

---

[37] Daniela Hristova (pc) suggests that (59a) is more appropriate if the speaker has only one uncle, whereas (59b) is preferable if she has more than one.

accusative and dative pronouns come in tonic or full (so-called "long") and atonic or clitic (so-called "short") varieties; cf. Friedman (1993: 264-265) and Scatton (1993: 203-204) for discussion of Mac and Bg, respectively.

The dative full forms are however optional in Mac and archaic in Bg, having been replaced by the preposition *na* plus a tonic accusative pronoun (e.g. *na nego* instead of *nemu*), where *na* has the general function of English *of*. On the other hand, the two sets of pronominal clitics are are extremely robust, with the accusative forms basically restricted to marking objects of verbs and the historically dative forms extended to "oblique" situations expressed with the generic preposition *na*. The table in (60) lists these oblique clitic forms (the same in both languages).

(60) mi 'me'   ni 'we'

    ti 'you' [sg]   vi 'you' [pl]

    mu 'him, it'   im 'them'

    ì 'her'   si 'self' [refl]

The basic DP-internal use of the clitics in (60) is to expression possession, comparable to the English genitive, as an alternative to a full-blown *na*-phrase (which is the way possession is expressed with nonpronominal possessors). As observed already, such use is limited to definite DPs, which are almost always articulated, as in the Bg examples in (61); note that in this section pronominal clitics rather than articles are indicated in boldface:

(61) a.   roditelite **mu** 'his parents'

      b.   knigata **vi** 'your book'

      c.   semejnija/ât **im** praznik 'their family holiday'

      d.   starata **si** kâšta 'self's old house'

      e.   mnogoto **ti** novi knigi 'your many new books'

      f.   večno mladata **ni** stolica 'our perpetually young capital'

      g.   tvârde bârzo preminalata **ì** mladost 'her too quickly passed youth'

The oblique clitic is clearly unlike the article in that it shows no phonological dependency on its host, other than combining with the element to its left to form a single prosodic domain, as is the general requirement for Bg pronominal clitics.

With this mind, observe that, although in form and phonological properties NP-internal pronominal clitics seem comparable to pronominal clitics in clauses, their distribution is quite different. An quick inspection of the examples in (61) reveals that the possessive clitic can follow the head noun (61a, b), the adjective (61c, d), the first in a series of modifiers (61e), or the first adjective phrase (61g). This pattern should by now be familiar, and leads to the descriptive generalization in (62):

(62) The oblique clitic appears immediately after the articulated element in the DP, *whatever that element happens to be*.

The generalization in (62) is striking and potentially extremely problematic. The reason is that, if I am correct in wanting to distinguish the articles from true clitics by treating them as suffixes and accomodating their distribution by inflecting the head of the XP complement to a [+definite] D, then I should paradoxically be led to treat pronominal clitics in the same manner. Such a solution is of course untenable, as I would lose my account of the differences between articles and atonic pronouns, and would be forced into the indefensible position of analyzing DP-internal atonic pronouns in a completely different way from clausal ones.

Nonetheless, a careful examination of these and even more complex examples invariably corroborates the generalization in (62). Consider (63), with a possessive clitic introduced into the constructions in (9):

(63) a. [počti nerazrabotenata   **mu**   u   nas] problematika
          almost not-worked-out-the him.obl   by   us   problematics
          'his problematics (which are) almost not worked out here [= in Bulgaria] '

      b. [polučenata **si**   sâs   mâka] stipendija
          received-the self.obl with pain   scholarship
          'self's received with pain scholarship'

      c. [vernijat   **ti**   na demokratični idei]   prezident
          faithful-the you.obl   to   democratic   ideas   president'
          'your president (who is) faithful to democratic ideas'

      d. [kupenite **ì**   včera]   knigi[38]
          bought-the her.obl   yesterday books
          'her books (which were) bought yesterday'

      e. [zabranenata **mi**   ot zakona]   kniga
          forbidden-the me.obl   by law   book

---

[38] In examples (63d, e) the dative clitics **ì** and **mi** can also be taken as arguments of the participle, with the meaning 'the books bought for her yesterday' and 'the book forbidden to me by law'. The argumental status of clitics inside DPs is treated in section 6.3.

'my book (which is) forbidden by law'

These are the only possibilities for placement of the article and clitic, i.e. together and immediately after the adjective or participle that heads the AP, regardless of whatever freedom of location other elements in that AP may enjoy.

6.2. Fine-tuning the distribution. Before discussing how the oblique clitic is extended to express relations beyond mere possession in Bulgarian DPs, let us consider a few more facts about their placement. The correct generalization, following observation also made by Caink (2000), is that (62) needs to be refined as in (64).

(64) The oblique clitic only appears in definite DPs. It immediately follows the articulated element in the DP, if there is one; if none is present, then it follows the demonstrative.

Consider first the example in (65).

(65) čičovata **mi** kâšta 'my uncle's house'

What is curious about (65) is the crossing of dependencies. Whereas the article *-ta* applies to the head noun *kâšta*, the oblique clitic *mi* applies to the possessive adjective *čičo*. Now, if one takes the paradigm in (59) and introduces a possessive clitic, we find, as expected, that the clitic necessarily follows the articulated form. This is illustrated in (66).

(66) a. čicovata **ti** Stojanova kâšta 'your Uncle Stojan's house'

 b. čico Stojanovata **ti** kâšta

Both variants are possible, and placing the clitic *ti* elsewhere in these expressions is not permissible. This corroborates the observation in (62) that the clitic follows whatever element happens to be articulated, regardless of other considerations.

Next consider the fact, as noted in Franks (1998), Arnaudova (1998) and Caink (2000), that in colloquial Bg the oblique clitic can appear in definite DPs that lack an articulated form, in which case they must follow the demonstrative. This is shown in (67).[39]

(67) tezi   **ti**   (?xubavi) knigi
      these   you.obl nice   books
      'these (nice) books of yours'

---

[39] Franks (1998) comments that, for some speakers, the adjective causes these examples to degrade.

Interestingly, if there is both a demonstrative *and* an articulated form, as in the colloquial examples in (44), then it is the articulated form rather than the demonstrative that determines the position of the clitic. This is shown in (68).

(68) tazi   neprijatnata **ti**   žena
      this   unpleasant   you.obl   wife
      'this unpleasant wife of yours'

The fundamental analytical question, to be addressed in section 7, is how can we relate the position of the article and the clitic, such that they both target the same phrase, although in different ways. Descriptively, the solution is simply to take the article as an inflection of a head, as I have done, and to take the clitic as adjoined to this head. Neither of these moves are especially controversial: the first I have already argued for at length, and the second assumption I believe to be independently necessary for Slavic clitics in general, as discussed in much of my other work on Slavic clitics. The major remaining difficulty is to explain how and why definiteness feature checking and clitic placement target the same head.

6.3. Clitics that mark argument functions. Before turning to a discussion of possible analyses, I briefly mention the fascinating and relatively unexplored topic of oblique clitics in DPs which appear to bear argument roles. As mentioned in Franks and King (2000: 56-57), oblique clitics can be used to correspond to arguments when in Bg DPs headed by deverbal nouns, as in the example in (69).

(69) interesnoto   **ì**   pojavjavane
      interesting-the her.obl   appearance
      'her interesting appearance (on the scene)'

This same relation would be expressed using a *na*-phrase if there were a full DP argument, as in (70).

(70) interesnoto   pojavjavane na Katia
      interesting-the appearance   of Katia
      'Katia's interesting appearance (on the scene)'

If the verb takes multiple arguments, the clitic is potentially ambiguous, in the same way that the corresponding *na*-phrases would be, as in (71).

(71) a. ranoto   **mu**   prepluvane (na kanala)
          early-the   him.obl   swimming   (of channel-the)
          'his swimming (of) the channel early' OR 'the channel's early swimming'

      b. pisaneto   **mu**
          writing-the   him.obl
          'his writing' OR 'the writing of it'

    c. peeneto   **ì**
       singing-the her.obl
       'her singing' OR 'the singing of it'

These arguments correspond to agents/subjects or themes/direct objects. Interestingly, true dative arguments (goals/indirect objects) can never be indicated with a clitic, even though these are also expressed using *na*-phrases (both in the VP and NP).[40] Thus, in (71b) *mu* can never correspond to the goal argument; this can only be done with a *na*-phrase, as is true for the verb *pisa* 'write'.

    Some further examples of deverbal nouns based on verbs that take dative arguments are given in (72).

(72) a. pomaganeto   **mu**
       helping-the   him.obl
       'his helping' BUT NOT *'the helping of/to him'

    b. predstavjaneto   **mu**
       introducing-the him.obl
       'his introducing' OR 'the introducing of him' BUT NOT *'the introducing to him'

In (72a) the clitic *mu* can only correspond to the nominative argument, not the dative one (which would have to be expressed as *na nego*). Similarly in (72b), which is a nominalization from of a verb that takes a nominative, accusative, and dative argument, the clitic can correspond to the nominative or accusative, argument, but once again, not the true dative one.

    Note that in example (71a) *mu* can double the argument DP *na kanala*. The possibility of clitic doubling inside DPs generally exists, both for possessors, as in (73a), and for arguments, as in (71a) or in (73b), which is an emphatic way of expressing (70).

(73) a. interesnata **ì**   misâl   na Katia
       interesting her.obl thought of Katia
       'Katia's interesting thought'

---

[40] However, experiencer datives can be expressed by clitics, with or without doubling:

(i)   interesnoto   **ì**   hrumvane na Katia
     interesting-the her.obl occurring to Katia
     'Katia's interesting idea'

The clausal counterpart to this would be, as in (ii), with a dative "subject"; as is always the case in Bg, oblique subjects require clitic doubling:

(ii) Na Katia **ì**   hrumna.
     to   Katia her.obl occurred
     'Katia had the idea.'

Thanks to Daniela Hristova (pc) for pointing this correspondence out to me.

b. interesnoto **ì**   pojavjavane na Katia
    interesting-the her.obl appearance of Katia
    'Katia's interesting appearance (on the scene)'

This fact strengthens the parallelism between the use of oblique clitics in DPs and in clauses, where

doubling appears to play a similar role; cf. Franks and King (2000) or Rudin (1997) for discussion.

Consider, finally, participial phrases that contain clitics, as in (74).

(74) a. [izgarjaštata go]   strast
         burning-the him.acc passion
         'the passion which burns him'
    b. [upravljavaštoto ja]   množinstvo
        ruling-the   it.acc majority
        'the majority which rules it'

The question is what happens when one forces an oblique clitic into this DP, as the possessor of *strast*.

The two possibilities are as in (75).

(75) a. %[izgarjaštata go]   **ti**   strast
         burning-the him.acc your.obl   passion
         'your passion which burns him'

    b. izgarjavaštata **ti**   go]   strast
        burning-the   your.obl him.acc passion

    d. %[upravljavaštoto ja   **im**   množinstvo
        ruling-the   it.acc   their.obl majority
        'their majority which rules it'

    e. upravljavaštoto **im**   ja   **im**   množinstvo
        ruling-the   their.obl it.acc   majority

Speakers seem to vary considerably on this. Some speakers accept both, others balk at either variant,

some prefer (75a, c), and yet others disallow this clitic order entirely. While the construction clearly

requires further investigation, it seems to me that these two possibilities reflect two distinct analyses of

the "participle + clitic" structure. In (75a, c), the partiple plus the clitic has been analyzed as a single,

complex head. This head is formed before the oblique clitic is position, so that it must necessarily

follow, with the bizarre result that (i) the possessive clitic is not immediately adjacent to the articulated

word and (ii) the surface order of clitics is accusative before dative, an order which is otherwise

completely disallowed everywhere else in Slavic.[41] In (75b, d), on the other hand, the oblique clitic

---

[41] The one exception is Polish, in which, according to Franks and King (2000) among others, pronominal clitics are not actually heads and so need not respect the standard ordering template imposed by the universal hierarchical arrangement of functional categories.

follows the highest head, as is the general rule, with the object clitic merged subsequently. The order is thus comparable to other complex modifer structures, such as those in (63).

7. Getting the oblique clitic in the right place. We now turn to the central analytical question of how, using the formal mechanisms of minimalist phrase structure, these collocations can be constructed. Recall that the fundamental problem is getting the clitic and the article to target the same head, given that the former is syntactic and the latter morphological. In this section I will consider three kinds of approaches, two that involve movement of the clitic and one that does not. I will ultimately argue for the latter approach.

7.1. A raising analysis.

Clitics in Slavic, and probably universally, are always functional heads.[42] In particular, I would analyze the pronominal clitics as heads of Agreement Phrases. Much recent work, such as Rudin (1997) in her treatment of Bg clitics, argues for their Agr status when they appear in clauses. The simples proposal is therefore that pronominal clitics should receive the same analysis in the nominal domain. One might therefore posit an AgrIOP immediately below DP, revising the tree in (39) as in (76):[43]

(76)
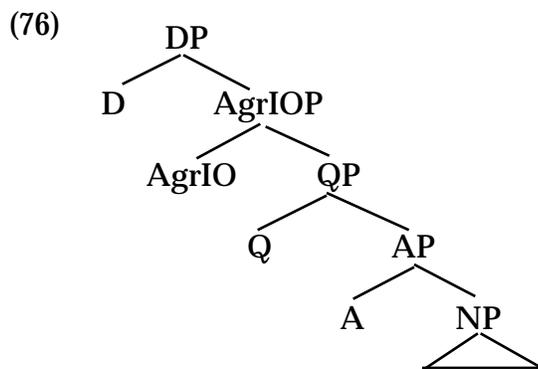
The structure in (76) represents the possible projections in a Bg DP; AgrIOP is limited in Mac to when N is a kinship term. Notice, however, that placing AgrIOP higher than the XP whose lexical head is

---

[42] I have in mind, in addition to the pronominal , auxiliary verb, and definiteness markers discussed here, potentially also markers of mood, tense, voice, negation, focus, interrogation, aspect, etc. As noted by Sadock (1991: 112), "clitics always seem to represent closed lexical classes. They are frequently encountered among determiners, auxiliary verbs, prepositions, complementizers, conjunctions, and pronouns, but I have no good examples of clitic main verbs, clitic nouns, or clitic adjectives. I will refrain from speculating here why this should be so, ..." It seems to me a natural characterization of this state of affairs is that only functional heads can be clitics.

[43] As an alternative to the nested structure in (76), one might introduce the higher phrases as specifiers of functional categories that serve no purpose other than the formal one of "gluing" the phrase structure together, e.g. [XP AP [X' X NP]] instead of [AP [A' A NP]]. So long as AP could be determined to be higher than NP in this structure, I can think of no considerations for choosing between them that are not theory-internal.

articulated is going to require some adjustment to the analysis. Recall (40), in which the articulated phrase was argued to raise to [Spec, DP], where it could be checked for definiteness features under Spec-head agreement with D. It turned out that the only real problem with this structure was targetting the correct phrase to raise to [Spec, DP]. If, however, we adopt Abney's nested structure, as we had to anyway, then the raising account is equivalent to the *in situ* one, except for the issue of whether [Spec, DP] is ever needed for other material (which in turn depended on the viability of DP iteration). It seems to me that only the raising analysis is going to be compatible with (76), since the complement to D is now AgrIOP. In other words, the proposal is that the highest phrase XP to the right of AgrIO raises to [Spec, DP], where it is checked for definiteness (and proximity, in Mac) features, morphological realized on its head X.

This scenario of course introduces some new problems. Let us explore some of the consequences of (76). First of all, it is clear that, if nothing else, moving the highest XP to the right of AgrIO directly to [Spec, DP] might be a violation of some general locality constraint, since there is an intervening slot that is skipped over, e.g. Rizzi's Relativized Minimality or the Minimal Link Condition of Chomsky (1995). The solution to this is to let XP stop in [Spec, AgrIOP] as an intermediate landing site. Now notice that this provides an opportunity for AgrIO to cliticize to X, as shown in (77).

(77)
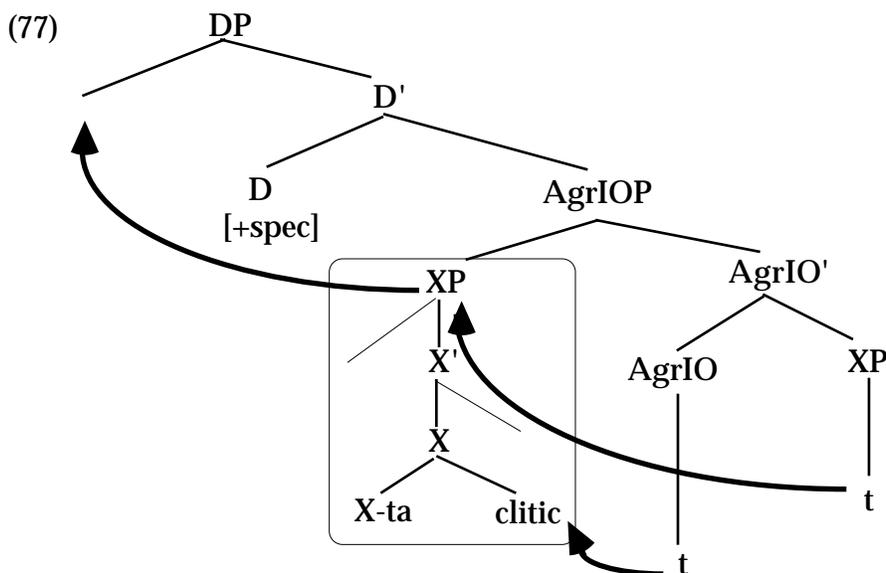
As before, it is the highest XP whose head is inflected for definiteness (and, in Mac, proximity). This head is pulled out of the lexicon with the appropriate features. The dative clitic adjoins to the head,

crucially outside the definiteness inflection (since inflection is technically part of the lexical item X), and then the entire XP continues on to [Spec, DP] to check its definiteness features. In this manner, the clitic ends up immediately following the inflected head, even though it is positioned syntactically rather than by the morphology. This would be the gist of the analysis.

The main difficulty with (77) is the movement of the clitic to the head of its specifier, as the phrase in this position travels through [Spec, AgrIOP] on its way to [Spec, DP]. It is therefore important that movement to this site be motivated. Principles of economy might be invoked as XP raised to [Spec, DP], or the movement to [Spec, AgrIO] itself might be feature driven by the need to check off features against the possessive clitic. It is also to my mind an open question, whether the second step, that of raising XP from [Spec, AgrIO] to [Spec, DP], takes place overtly or covertly. It depends on the same considerations as before, i.e. whether or not [Spec, DP] is independently needed to host the demonstrative in doubled constructions and whether or not material with phonological content is ever generated in the DP system in Slavic. Let us grant, then, that stopping in [Spec, AgrIO] is expected. Consider now (68), repeated as (78).

(78) tazi   neprijatnata **ti**   žena
      this   unpleasant   you.obl   wife
      'this unpleasant wife of yours'

To derive this using the structure in (77), *taxi* would need to be generated in [Spec, DP] amd *neprijatnata* would need to stop in [Spec, AgrIOP], with its definiteness features being checking in LF against those of D(P). It is not clear whether either conclusion can be substantiated.

Another controversial step in (77), it seems to me, might be performing cliticization precisely at this point in the derivation. Here I see two kinds of approach to the forcing cliticization. One would be simply to stipulate that cliticization takes place as soon as it can. In the particular case, since the Slavic pronominal clitics are Agr elements, and Agr is a head, as soon as there is an X° available on which to cliticize, the proposal would be that cliticization must take place. This idea has two obvious potential defects however. The first is empirical: unless we can define what an "available" X° is so as to exclude everything else, it seems likely that there going to be other heads around which could serve as possible targets for cliticization before the right one is in place. The second is more conceptual: it is mysterious that cliticization, of all operations, should occur as early as possible in the derivation. The problem

concerns Chomsky's (1995) economy principle of "Procrastinate," which requires movement to take place as late as possible. Clearly, if (77) is to respect Procrastinate, there must be some reason why immediate cliticization is forced. The answer may lie in Shortest Move: it is a syntactic requirement of a head clitic that it adjoin (or be adjoined to) another head. This requirement must be met in the most economical way possible. If, therefore, cliticization were to wait until later in the derivation, it would presumably involve a less economical movement.[44] Since the standard cliticization scenario in the clause is for the verb to move up through the various Agr heads, this movement automatically satisfies the clitic's requirement of being adjoined to a head. Failing this, as is the case inside DPs, one might claim that the movement in (77) constitutes the Shortest Move.[45]

7.2. A lowering analysis. I next turn to an alternative that avoids the problem of movement to the head of a specifier, and instead makes use of overt syntactic lowering. This approach was put forward in Franks (1998) and Franks and King (2000). In brief, if AgrIO were to move down to next lowest head, then it would always be placed correctly. Unlike in the clausal domain, where the verb raises to AgrIO, no head moves up to adjoin to the clitic. The clitic is thus stranded and has to move itself to be supported (taking movement to be the less economical option, meaning that it only occurs when it has to). In this analysis, the clitic moves, but down, rather than up, lowering onto the next head down, as in (79a) or, if the clitic is really somehow under D, (79b).[46]
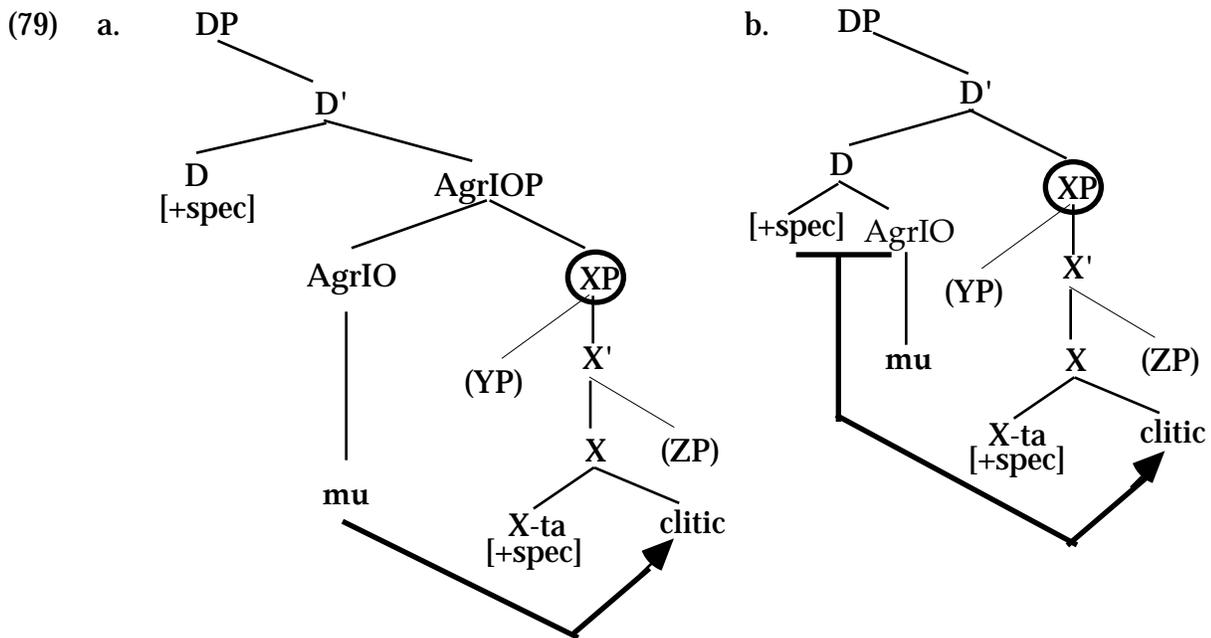
---

[44] I am grateful to Željko Bošković (pc) for suggesting this approach to forcing cliticization in (77).

[45] The possibility of Agr adjoining to next higher head needs also to be considered. It seems to me that in this particular case the fact that D° is always phonologically null might be a problem, although why this phonological fact should be relevant to the syntax is unclear.

[46] It might also move up, if Stateva (2000 and work in progress) is correct that DP-internal clitics, as in canonical (i), can escape the DP and move to a higher, clausal position, as in (ii).

(i) Vidjah   knigata   **mu**.   (ii)   Vidjah **mu** knigata.
     saw.1sg   book-the   him.dat
     'I saw his book'

(79) a.

```
              DP
               \
                D'
              /   \
           D        AgrIOP
        [+spec]    /      \
              AgrIO        (XP)
                |         /    \
               mu      (YP)    X'
                                |
                              X    (ZP)
                            /   \
                        X-ta    clitic
                      [+spec]
```

b.

```
              DP
               \
                D'
              /   \
           D       (XP)
        [+spec] AgrIO  /  \
                     (YP)  X'
                      |    /  \
                     mu   X    (ZP)
                        /   \
                    X-ta    clitic
                   [+spec]
```

Lowering to the first available head, namely to the X° to its right, is presumably the only thing the clitic can do, since there is no higher head to raise to. Note that, as in the analysis of the preceding section, we must discount D in (79a) as a possible target for raising. More problematic for (79a), as observed by Caink (2000: 176), are examples such as (67), repeated as (80).

(80) tezi   **ti**   (?xubavi) knigi
     these   you.obl nice   books
     'these (nice) books of yours'

It is unclear why lowering should not take place here. Clearly, the reason lowering does not occur is because there is no articulated element to the right of *ti*. Observe that, as in (78), when there is such an element, lowering once again becomes obligatory.

The variant of the lowering analysis in (79b), where the clitic is generated alongside definiteness features in D, somewhat avoids these problems. D is of course no longer a possible raising target. The failure of lowering in (80) could follow from *texi* being being generated in [Spec, DP], where its definiteness is checked against [+spec] on the head, obviating lowering. However, (78) then becomes problematic; we would need to assume that in (78) the definiteness of *tazi* is first checked against D°, containing *ti*, and then this lowers to check the definiteness of *neprijatnata*. The structure in (79b) does, however, have some interesting properties. Most significantly, it explicitly expresses the connection between definiteness and the clitic. However, since it also does not posit an AgrP, any parallelism with

dative clitics as clausal AgrIO would be lost.[47] Also, [Spec, DP] would then need to be used to check any phrase that the possessive clitic may double, as in the examples in (73).

7.3. An interarboreal analysis. In this final section I propose an alternative to the raising and lowering accounts that does not involve overt movement at all. This account takes bottom-up merge seriously and exploits the "interarboreal" method of constructing phrase structure trees argued for by Bobaljik and Brown (1997), among others.

Bobaljik and Brown (1997) point out that merge needs to be able to work on separate subtrees before combining these such that they ultimately meet the requirement of having a single root. The note that merger of two independent subtrees is obvious for such structures as branching subjects, as in (81):

(81) [[the monster] [ate [the pizza]]]

Clearly, the DP*the monster* must be created before merging with the VP *ate the pizza*. Note, however, that this latter merger still extends the tree, in conformity with the Extension Condition of Chomsky (1995). What Bobaljik and Brown argue is that an element from one merger can be copied and re-merged elsewhere. That is, whereas in normal Move an element from the tree that is being extended is copied and "re-merged" at the root nood of the tree in question, they propose that an element from a distinct subtree, which is simultaneously being constructed, can also be copied and re-merged. They do this specifically to accommodate head movement, which otherwise does not extend the tree (since the target head would have already been merged). Thus, for a structure as in (82), under their analysis the "I + V" complex is created *before* merger with VP. In this way, every operation obeys the Extension Condition.

(82)
```
              IP
            /    \
          I        VP
        /   \     /   \
       I     V   V     DP
```
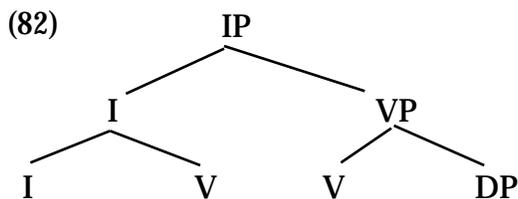
In order to to this, however, the "interarboreal" operation of copying an already merged V is going to be necessary.

---

[47] This may not actually be a disadvantage, given the facts in (72). The DP internal "dative" clitic thus only realizes what would be genitive in a language like Polish, and in this respect differs enough from its clausal counterpart to shed some doubt on calling it a true AgrIO°. Possibly, however, this difference derives from the idea that the specific content of various Agrs is not intrinsic, but is rather relationally determined. Thus, having only one Agr in DP means that it will be structural in nature (although for some reason dative in form). These are exciting questions, but unfortunately too involved to explore properly here.

How can this be applied to the problem at hand? First of all, we have established that all substantive heads in Bg (and Mac) come with features for specificity. That is, definiteness in these languages is inflectional, and most [+N] categories have forms with the features [–specific] or [+specific] (as well as [±proximate] and [±distal] for Mac) in the lexicon. This feature is then checked against D and, since D attracts specificity features, the closest [+N] head will need to match D, which gives rise to the highest head effect as an instance of Superiority. Given this, crucial observation is that the oblique clitic is always adjacent to a [+specific] head. I take this as a simple selectional requirement on the clitic: descriptively clitics must merge with something specific.[48] The reason, I suppose, is that pronominal clitics themselves are specific, hence cannot be in a checking relationship with anything nonspecific.[49] Thus, when they double DPs, either in the nominal or clausal domain, they double specific ones, and when they merge with a head, as in the DP-internal structures, they merge only with specific heads.

Once we see this as a selectional requirement, the next step that needs to be taken is to merge the oblique clitic *directly* with its host, creating a complex head, which as a matter of selection it requires to be [+specific]. Its features can then raise to AgrIO, if a structure as in (79a) is adopted, or to D, if a structure as in (79b) is adopted (but with covert upwards rather than overt downwards movement). The account thus resembles the approach in Bobaljik and Brown (1997) to head raising, since a complex head is created before that head is merged with its complement. To see how a simple derivation proceeds, consider the example in (83) with a phrase structure as in (84).[50]

(83)   xubavoto **mu** momiče
        pretty-the his girl
        'his pretty girl'

---

[48] Caink (2000: 180-181) take a similar stance, stating this restriction as a subcategorization frame on the clitic: "+ [+DEF] ___".
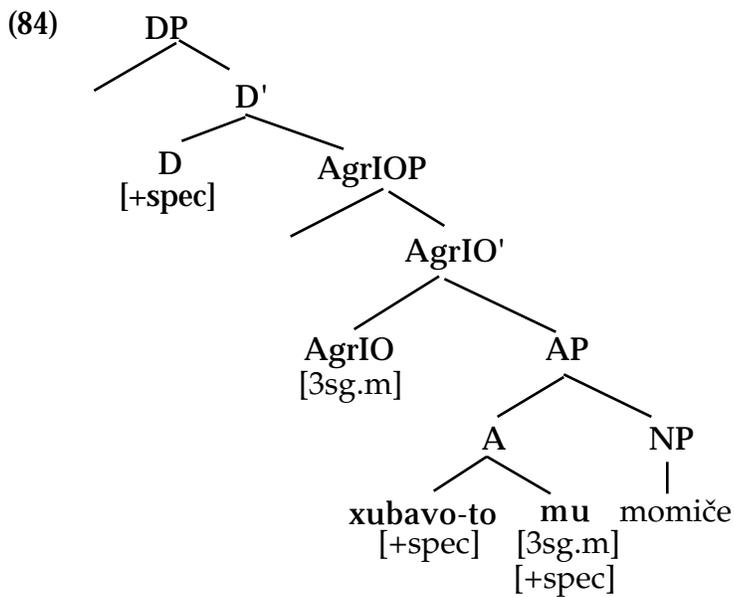
[49] For the definiteness of the specifier to determine the definiteness of the entire phrase is not unusual. Consider the interpretations of English (i) and (ii).

(i)   [[a man] 's hat]]   (ii)   [[the man] 's hat]]

In (i) the entire DP assumes the indefiniteness of *a man*, whereas in (ii) it assumes the definiteness of *the man*. This is presumably the result of Spec-head agreement between *a man/the man*, in [Spec, DP], and D°.

[50] So far as I can tell, these examples cannot be accommodated using the kind of Agr-less structure in (79b).

(84)

```
                    DP
                   /  \
                       D'
                      /  \
                     D    AgrIOP
                  [+spec]    \
                              AgrIO'
                             /      \
                          AgrIO      AP
                         [3sg.m]    /  \
                                   A    NP
                                  / \    |
                          xubavo-to  mu  momiče
                           [+spec]  [3sg.m]
                                    [+spec]
```

Then adjective *xubavoto* is drawn from the lexicon in its articulated [+spec] form. The clitic *mu* merges it with it, matching in definiteness features, as required. Next, this complex A merges with the noun *momiče*. This is the lexical core of the projection. The resultant phrase must be dominated by a [3sg.m] AgrP and a [+spec] DP, given that it contains these features. The formal features associated with A, namely [+spec] and [3sg.m], then must move up. First, Agr attracts the formal features of A, checking off its [3sg.m], and then D attracts those features up from Agr, checking off its feature, [+spec]. Since this is clearly not overt category movement, the relevant features must be weak.[51] In this way, the necessary ordering relations are obtained without overt movement of syntactic material. Moreover, the inflectional article and the syntactic clitic both target the same head is accommodated without lowering, and in accordance with the minimalist requirement that each merger extends the tree.

The colloquial and problematic possibilities in (68) and (80) still deserve some comment. They are repeated as (85a, b), respectively.

(85) a. tazi neprijatnata **ti**   žena   BUT *tazi **ti** neprijatnata žena
         this unpleasant   you.obl wife
         'this unpleasant wife of yours'

     b. tezi   **ti**   (?xubavi) knigi
         these you.obl nice   books
         'these (nice) books of yours'

---

[51] Under the traditional minimalist model this implies LF movement. Under the system in Franks (2000), based on alternative approaches to weak features, the formal feature movement may be over, but the category is still pronounced in its base position.

The dilemma here is the following: if *ti* can have its selection requirements satisfied by adjoining to *tezi*, as in (85b), why can it not do this in (85a), at least as an option? For (85a), I assume that the demonstrative*tazi* is generated directly in DP, either in [Spec, DP] or D° itself, in a structure such as (84). If so, it is merged *higher* than AgrIOP, hence there would be no way to check the features of *ti*, if the 2sg oblique clitic were merged instead directly with *tazi*. On the other hand, when *ti* merges with the [+spec] adjective *neprijatnata*, since this is below AgrIOP, the relevant formal features can be attracted upwards, first to AgrIO, and then to D. Given this account, what about (85b)? First note that speakers find this kind of example less acceptable, and some reject it outright, or perhaps there is dialectal variation here. To derive it, however, I suggest that the oblique clitic *ti* is inserted directly in AgrIO, where it checks off features, exactly as an indirect object does in the clause; cf. e.g. Rudin (1997), Franks (2000), or Franks and King (2000).It, or rather its formal features, then move up to D° to satisfy the specificity requirement.

Putting this complication aside, let us finally consider a few more conceptual aspects of the direct merger analysis of this Bg construction. It actually turns out not to be necessarily exactly "interarboreal" in the sense of Bobaljik and Brown (1997). Note that Freidin (1999: 118) objects that interarboreal operations run counter to the formulation of Merge in Chomsky (1995), since they require Merge to look inside of a single syntactic object. Once created, the pieces of a syntactic object should be inaccessible to further operations. While I am not sure that this is technically correct (otherwise, how could the construction in (65)/(66) arise?), and I am also not sure that interarboreal operations need to access pieces of syntactic objects (they could be accessing the Numeration instead), the analysis I have proposed for Bg avoids Freidin's objection, since it does not involve movement of a piece of a syntactic object. Instead, it is conceptually comparable to (81), although it is a complex head that is created prior to merger, rather than a phrase.Of course, subsequent feature checking must be able to access features such as 3sg.m for *mu* in (84), which are technically inside the complex A. But this is not the same as merging a piece of a structure, as Bobaljik and Brown require for (82). The difference of course is that I use the mechanism to avoid head lowering, whereas they use it to handle head raising without violating the Extension Condition. Since no exporation is involved, the counterpart to Freidin's objection to an interarboreal apporach to incorporation simply does not arise.

8. Summary. In this paper I have surveyed a variety of issues relating to the internal structure of Slavic Noun Phrases. A number of problems where identified, some of which remain largely unexplored. To conclude, I summarize three key puzzles enountered in this survey:

- •Why do some languages allow clitics in NPs and others do not?

- •Why do articles and clitics target the same head in Bg and Mac?

- •Why do "dative" clitics in Bg never correspond to clausal indirect objects?

I have considered a number of technical solutions, in particular, to the second issue, of how NP-internal oblique clitics might be position, and I have argued that substantives are drawn from the lexicon with features for specificity and that oblique clitics are merged directly with these substantives. However, explanation of the other questions identified unfortunately eludes me. It is my hope at least that the data touched upon here will stimulate someone more inventive than myself to elucidate these and other puzzles about the internal structure of Slavic NPs.

References

Abney, Steven (1987) *The English Noun Phrase in its Sentential Aspect.* Unpublished MIT Ph.D. diss.

Arnaudova, Olga (1996) "The Bulgarian Noun (Adjective) Movement to D", in Mila Dimitrova-Vulchanova and Lars Hellan, eds., *Papers from the 1st Conference on Formal Approaches to South Slavic and Balkan Languages, University of Trondheim Working Papers in Linguistics* **28**, pp. 1-28.

Arnaudova, Olga (1998) "Demonstratives and the Structure of the Bulgarian DP", in Mila Dimitrova-Vulchanova, Iliyana Krapova, Ivan Kasabov and Lars Hellan, eds., *Papers from the 2rd Conference on Formal Approaches to South Slavic and Balkan Languages, University of Trondheim Working Papers in Linguistics* **31**, pp. 1-12.

Aronson, Howard (1968) *Bulgarian Inflectional Morphology*, Mouton: The Hague.

Bobaljik, Jonathan and Samuel Brown (1997) "Interarboreal Operations: Head Movement and the Extension Requirement," *Linguistic Inquiry* **28**, 345-356.

Caink, Andrew (2000) "In Favour of a 'Clitic Cluster' in the Bulgarian and Macedonian DP," in Mila Dimitrova-Vulchanova, Iliyana Krapova and Lars Hellan, eds., *Papers from the 3rd Conference on Formal Approaches to South Slavic and Balkan Languages, University of Trondheim Working Papers in Linguistics* **34**, pp. 170-182.

Chomsky, Noam (1995) *The Minimalist Program.* MIT Press: Cambridge.

Corbett, Greville (1987) "The Morphology/Syntax Interface: Evidence from Possessive Adjectives in Slavonic," *Language* **63.2**, 299-345.

Dimitrova-Vulchanova, Mila and Giuliana Giusti (1996) "Quantified Noun Phrase Structure in Bulgarian," *University of Venice Working Papers in Linguistics* **5.1**, 43-65. [identical to Giusti and Dimitrova-Vulchanova (1994)

Fowler, George and Steven Franks (1994) "DP in Bulgarian," *Indiana Slavic Studies 7, Proceedings of the 9th Biennial Conference on Balkan and South Slavic Linguistics, Literature and Folklore*, pp. 58-68.

Franks, Steven (1990) "Vowel-Zero Alternations and Syllable-Counting Morphology" in Henry Cooper et al., eds, *Indiana Slavic Studies* **5**.

Franks, Steven (1995) *Parameters of Slavic Morphosyntax*. Oxford University Press: New York.

Franks, Steven (1998) "Clitics in Slavic". Position paper for *Workshop in Comparative Slavic Morphosyntax*, held at McCormick's Creek State Park, 5-7 June. To appear in *Proceedings*, Slavica: Bloomington, Ind.

Franks, Steven (2000) "Clitics at the Interface" In Frits Beukema and Marcel den Dikken, eds., *Clitic Phenomena in European Languages*, John Benjamins: Amsterdam, pp. 1-46.

Franks, Steven and Tracy H. King (2000) *A Handbook of Slavic Clitics*, Oxford University Press: New York.

Freidin, Robert (1999) "Cyclicity and Minimalism," in Samuel Epstein and Norbert Hornstein, eds., *Working Minimalism*, MIT Press: Cambridge, pp. 95-126.

Friedman, Victor (1993) "Macedonian" in Bernard Comrie and Greville Corbett, eds.,*The Slavonic Languages*. Routledge: London, pp. 249-305.

Giusti, Giuliana and Mila Dimitrova-Vulchanova (1994) "Quantified Noun Phrase Structure in Bulgarian," in Jindřich Toman, ed., *Proceedings of FASL III: The College Park Meeting*, University of Michigan Slavic Materials: Ann Arbor, pp. 123-144. [identical to Dimitrova-Vulchanova and Giusti (1996)]

Longobardi, Giuseppe (1994) "Reference and Proper Names" *Linguistic Inquiry* **25**, 609-665.

Halpern, Aaron (1995) *On the Morphology and Placement of Clitics*, CSLI Publications: Stanford. [slightly revised version of 1992 Stanford diss.]

Koneski, Blaže (1976) *Gramatika na makedonskiot literaturn jazik*, Kultura: Skopje.

Mišeska-Tomić, Olga (1996) "The Balkan Slavic Nominal Clitics" in A. Halpern and A. Zwicky, eds., *Approaching Second: Second Position Clitics and Related Phenomena*, CSLI Publications: Stanford, pp. 511-536.

Sadock, Jerrold (1991) *Autolexical Syntax* University of Chicago Press: Chicago.

Schick, Ivanka (2000) "The Phenomenon of Possessive Clitics in the Balkan-Slavic Languages" in Mila Dimitrova-Vulchanova, Iliyana Krapova and Lars Hellan, eds., *Papers from the 3rd Conference on Formal Approaches to South Slavic and Balkan Languages, University of Trondheim Working Papers in Linguistics* **34**, pp. 183-195.

Rudin, Catherine (1997) "AgrO and Bulgarian Pronominal Clitics" in Martina Lindseth and Steven Franks, eds., *Proceedings of FASL V: The Indiana Meeting*, University of Michigan Slavic Materials: Ann Arbor, pp. 224-252.

Scatton, Ernest (1975) *Bulgarian Phonology*, Slavica: Cambridge.

Scatton, Ernest (1993) "Bulgarian" in Bernard Comrie and Greville Corbett, eds. *The Slavonic Languages,* Routledge: London, pp. 188-248.

Stateva, Penka (2000) "How Optional 'Optional' Movement Is" in Mila Dimitrova-Vulchanova, Iliyana Krapova and Lars Hellan, eds., *Papers from the 3rd Conference on Formal Approaches to South Slavic and Balkan Languages, University of Trondheim Working Papers in Linguistics* **34**, pp. 109-120.

Townsend, Charles (1975). *Russian Word Formation*, McGraw Hill: New York.

Zec, Draga (1988) "Bulgarian ə Epenthesis: A Case for Moraic Structure" in J. Blevins andJ. Carter, eds. *Proceedings of NELS* **18**, Vol. II, pp. 553-566.