

On the Clausal and NP Structure of Serbo-Croatian

Željko Bošković
University of Connecticut

The paper has two goals: to discuss the basic clausal structure of Serbo-Croatian (SC) including basic operations that apply at this level, and the NP structure of SC, with the emphasis on the question of whether SC NPs have the DP layer. Much of the discussion also applies to other Slavic languages.

1 Clausal structure

1.1 V/Aux-movement

Consider the position of the verb. I will compare SC with French and English in this respect. As the ambiguity of (1) shows, V in French moves both above low, manner adverbs, and high, sentential adverbs, i.e. it moves to the highest X⁰ within split I. (2) is standardly taken to show English Vs don't raise outside of VP. (I ignore potential movement within VP/vP.)

- (1) Jean répond correctement à Marie.
Jean replies correctly to Marie
'Jean is giving Marie a correct answer.'
'Jean is doing the right thing in answering Marie.'
- (2) *John answered correctly Mary.

Stjepanović (1999b) notes SC Vs can cross manner but not sentential adverbs. This shows SC V is lower than French, but higher than English V.

- (3) Odgovara pravilno Mileni.
answers correctly Milena_{DAT}
'He is giving Milena a correct answer.'
'*He is doing the right thing in answering Milena.'

Assuming with Bošković (1997), who adopts Split I, that sentential adverbs are TP-adjoined, Stjepanović suggests SC Vs move to T. She also suggests the movement is optional due to *Pravilno odgovara Mileni*, where *pravilno* is ambiguous (it can have sentential reading).

As in English, there is a V/aux contrast in SC: in contrast to Vs, auxiliaries can precede sentential adverbs (clitics are given in italics).

- (4) Oni *su* pravilno odgovorili Mileni.
they are correctly answered Milena_{DAT}
'They did the right thing in answering Milena.'
'They gave Milena a correct answer.'

Bošković (2001) notes a difference between SC and English. In contrast to

English, sentential adverbs cannot precede subjects in SC. This can be accounted for by assuming (a) sentential adverbs can be either TP or AgrsP adjoined in English while in SC they can only be TP adjoined or (b) SC subjects are higher than English subjects (SpecIP could be filled by pro in SC; see Barbosa 1995 regarding Romance). I will proceed by adopting (a).

(5) Probably, they have beaten Peter.

(6) a. *Vjerovatno oni tuku Petra. b. Oni vjerovatno tuku Petra.
 probably they beat Petar

Putting all of this together, we get (7) for a basic SC clause.

(7) [_{AgrsP} Subject aux-clitic [_{TP} sent. adverb [_{TP} (finite main verb) [_{T'} [_{VP/AgroP} manner adverb [_{VP/AgroP} (finite main verb)

As for strong auxiliaries, Bošković (2001) notes that they pattern with such auxiliaries in English in that they cannot move across sentential adverbs. I suggest strong auxiliaries move to ΣP, which is located below sentential adverbs, possibly for semantic reasons (sentential adverbs may need to have scope over negative/emphatic aux).

(8) a. *Nisu/jesu vjerovatno poljubili Mariju.
 not+are/ARE probably kissed Marija
 ‘They probably did not/did kiss Marija.’

b. Vjerovatno nisu/jesu poljubili Mariju.

(9) a. They probably haven’t kissed Mary.

b. *They haven’t probably kissed Mary.

1.2 Clitics

SC clitics cluster in second position (2P). Until recently it has been standardly assumed that SC clitics cluster syntactically in the same head position. However, there is strong evidence against this position. E.g., Bošković (2001) shows that while aux clitics can (4), object clitics cannot occur above subject-oriented adverbs (10). This provides strong evidence that aux and object clitics don’t occur in the same head position (11).

(10) Oni *su joj* pravilno odgovorili.
 they are her_{DAT} correctly answered

 ‘They gave her a correct answer/*did the right thing in answering her.’

(11) [_{AgrsP} aux-clitics [_{TP} sent. adverb [_{TP} object clitics

Interestingly, *pravilno* still cannot intervene between *su* and *joj*.

(12) *Oni *su pravilno joj* odgovorili.
 they are correctly her_{DAT} answered

Bošković (2001) argues there is nothing wrong with (12) syntactically: it

is bad because it violates the 2P requirement, which is a PF, not a syntactic condition. (13)-(14) illustrate the 2P effect (placing *smo ga* in any other position would lead to unacceptability), which is traditionally stated in syntactic terms: clitics must be second within their clause.

- (13) Mi/zašto *smo ga* upoznali juče
 we why are him met yesterday
 ‘We met him yesterday./Why did we meet him yesterday?’
- (14) Ona tvrdi da *smo ga* upoznali juče.
 she claims that are him met yesterday

The traditional statement that SC clitics are second within their clause is clearly incorrect. As (15)-(17) show, certain elements, such as appositives, fronted heavy constituents, and parentheticals, can cause clitics to occur further than 2P of their clause.

- (15) Sa Petrom Petrovićem srela *se* samo Milena.
 with Petar Petrović met self only Milena
 ‘With Petar Petrović, only Milena met.’
- (16) Znači da, kao što rekoh, oni *će* sutra doći.
 means that as said they will tomorrow arrive
 ‘It means that, as I said, they will arrive tomorrow.’
- (17) Ja, tvoja mama, obećala *sam ti* sladoled.
 I your mother promised am you_{DAT} ice cream
 ‘I, your mother, promised you an ice cream.’

The distribution of SC clitics can be stated in very simple prosodic terms:

- (18) SC clitics occur in the second position of their intonational (I-) phrase.

Prosodic structure is determined by syntactic structure. It is standardly assumed that unless interrupted by an element that forms a separate intonation domain, each clause is mapped to a single I-phrase, with the CP edge corresponding to an I-phrase boundary. Some elements, such as appositives, parentheticals, and heavy fronted constituents, form separate I-phrases, evidence for which is provided by the fact that they are followed by pauses. Under the most natural pronunciation clitic second examples in (13) then contain only one I-phrase. In (15)-(17), on the other hand, the relevant clauses are parsed into more than one I-phrase, since the fronted heavy constituent, the parenthetical, and the appositive form separate I-phrases. This means a new I-phrase starts after these elements, which are obligatorily followed by a pause. Given this, the clitics are located in 2P of their I-phrase in (15)-(17). When we place a clitic in 3P of its I-phrase, violating (18), we get ungrammatical examples.

- (19) a. *Petra srela *je* samo Milena.
 Petar_{ACC} met is only Milena_{NOM}

- b. *Ja obećala *sam ti* sladoled.
 c. *Znači da oni *će* sutra doći.

The correct generalization regarding the distribution of SC clitics is then that they are second within their I-phrase, not their clause, which shows that the 2P effect is a PF effect.

A confirmation of (18) is provided by Bošković's (2001) (20)-(21).

- (20) *Ko koga *je* poljubio?
 who whom is kissed
 'Who kissed who?'
 (21) ?Koji čovjek, koju *je* knjigu kupio?
 which man which is book bought
 'Which man bought which book?'

Given Rudin's (1988) claim that fronted wh-phrases in SC don't form a constituent, (20) violates (18) (assuming straightforward mapping from syntactic to prosodic constituents). (20) improves with heavier wh-phrases (21). The first wh-phrase in (21) must be followed by a pause, an indication of an I-phrase boundary. As a result, *je* is located in 2P of its I-phrase. (18) easily captures (20)-(21). On the other hand, it is difficult to see how they can be accounted for under a purely syntactic account since the proposed analyses of MWF assign (20)-(21) the same syntactic structure.

Bošković (2001) gives an account of (18) on which SC clitics must encliticize to a constituent that is right-adjacent to an I-phrase boundary because of their PF lexical properties. As a result, they must be second within their I-phrase. The analysis forces phonological clustering of I-phrase-mate clitics, but not clause-mate clitics. It doesn't force their syntactic clustering in the sense that it does not force clitics to occur in the same head position. (22) is then ruled out in PF because the prosodic properties of *ga* are not satisfied. (*Ga* violates (18).)

- (22) ...*da *su* juče *ga* istukli.
 that are yesterday him beaten
 'that they beat him yesterday'

In Slovenian a clitic host also must be adjacent to an I-phrase boundary. However, Slovenian differs from SC in that its clitics can be either enclitics or proclitics. As a result, prosodically, nothing prevents breaking of a clitic cluster in Slovenian by an element that is adjacent to an I-phrase boundary. As noted in Bošković (2001), examples of this type are indeed acceptable in Slovenian (23). This confirms the relevance of prosodic requirements to clitic clustering in the languages in question.

- (23) *So* včeraj *ga* pretepli?
 are yesterday him beaten
 'They beat him yesterday?'

I now turn to clitic placement. During the discussion below we will need to control for the 2P effect since an example violating (18) will be ruled out in PF independently of whether syntactic requirements of its clitics are met.

There is a lot of evidence for a height difference between aux and object clitics, which shows they don't cluster in the same head position. First, the adverb data in (4)/(10) quite clearly show aux and object clitics don't occur in the same head position. The same holds for Stjepanović's ellipsis data. Given that ellipsis affects constituents, it must be the case that the object clitics and *dali* in (24) form a constituent to the exclusion of the aux clitic, hence aux and object clitics cannot be in the same head position.

- (24) ?*Mi smo mu ga dali, a i vi ste ~~mu~~ ~~ga~~ dali.*
 we are him_{DAT} it_{ACC} given and also you are him_{DAT} it_{ACC} given
 'We gave it to him, and you did too.' (Stjepanović 1999b)

That aux clitics are higher than object clitics is confirmed by (25), where quite a bit of material occurs between the clause-mate clitics *su* and *se*. (Due to the parenthetical, which is followed by an I-phrase boundary, each clitic in (25a) is located in 2P of its I-phrase. Note that (25c) is unacceptable because *se* is not located in 2P of its I-phrase. The contrast in (25a)/(25c) shows I-phrase-mate, but not clause-mate clitics have to cluster together, indicating the clustering requirement is prosodic, not syntactic.)

- (25) a. *Oni su, kao što sam vam rekla, predstavili se Petru.*
 they are as am you_{DAT} said introduced self_{ACC} Petar_{DAT}
 'They, as I told you, introduced themselves to Petar.'
 b. **Oni se, kao što sam vam rekla, predstavili su Petru.*
 c. **Oni su predstavili se Petru.* (Bošković 2001)

Wilder and Čavar (1997) note speakers who allow VP fronting with aux clitics accept (26), which confirms aux clitics are higher than object clitics.

- (26) *Dali ga Mariji su Ivan i Stipe.*
 given it_{ACC} Marija_{DAT} are Ivan and Stipe
 'Give it to Marija, Ivan and Stipe did.'

There is also a height difference between pronominal clitics—they don't cluster in the same head position either. Thus, when applied to pronominal clitics, the ellipsis and the parenthetical split test show dative clitics are higher than accusative clitics. ((29)-(30) are from Bošković 2001.)

- (27) ?*Mi smo mu ga dali, a i vi ste mu ~~ga~~ dali.*
 we are him_{DAT} it_{ACC} given and also you are him_{DAT} it_{ACC} given
 (28) **Mi smo mu ga dali, a i vi ste ga ~~mu~~ dali.* (Stjepanović 1999b)
 (29) ?*Oni su mu, kao što sam vam rekla, predstavili ga juče.*
 they are him_{DAT} as am you_{DAT} said introduced him_{ACC} yesterday
 'They, as I told you, introduced him to him yesterday.'

(30) *Oni *su ga*, kao što *sam vam* rekla, predstavili *mu* juče.

Progovac (1993) shows clitic climbing is marginally possible out of some finite clauses. Stjepanović (1999b) notes that if only one pronominal clitic in a double object construction climbs it must be the dative, which follows if the dative clitic is higher than the accusative clitic.

- (31) a. ?Marija *mu* želi da *ga* predstavi.
Marija him_{DAT} wants that him_{ACC} introduces
'Marija wants to introduce him to him.'
b. *Marija *ga* želi da *mu* predstavi.

Putting all of this together, Bošković (2001) adopts the structure in (32).

(32) [_{AgrsP} aux-clitic [_{Agriop} dative clitic_i [_{AgrdroP} acc. clitic_j [_{t_i} main V t_j]]]

Turning to ethical dative, (33) shows that, in contrast to argumental dative (10), ethical dative can precede sentential adverbs, indicating the latter is higher than the former. (37), where the ethical dative must precede the argumental dative (ethical dative cannot be in the 3rd person), confirms this.

- (33) Oni *su ti* pravilno odgovorili Ani.
they are you_{DAT} correctly answered Ana_{DAT} (you=ethical dative)
'They did the right thing in answering Ana/gave Ana a correct answer.'
(34) a. Juče *sam ti joj* pomogla.
yesterday am you_{DAT} her_{DAT} helped (you=ethical dative)
'Yesterday, I helped her.'
b. *Juče *sam joj ti* pomogla.

Ethical dative clitics can then be incorporated into (32) as follows, where ΔP is a discourse-related projection.

(35) [_{AgrsP} aux-clitic [_{ΔP} ethical dative clitic [_{TP} sent. adverbs [_{TP} [_{Agriop} dative clitic [_{AgroP} accusative clitic [_{VP}

These data show the order of clitics within the cluster matches their height (if X precedes Y, X is higher than Y), which favors a structural account of the order over arbitrary morphological template accounts, where the order within the clitic cluster is stipulated in the morphology. In such an account, the correlation with syntactic height is completely accidental.

A standard argument for a morphological template analysis concerns *je*, which, in contrast to other aux clitics, follows object clitics.

- (36) a. Oni *su mu ga* predstavili.
they are him_{DAT} him_{ACC} introduced
'They introduced him to him.'
b. Ona *mu ga je* predstavila.
she him_{DAT} him_{ACC} is introduced

However, Bošković (2001) shows *je* is higher than object clitics in the syntax. The above tests conclusively show this (compare (40) with (10)).

- (37) Ona *mu ga je* predstavila, a i on ~~*je mu ga*~~ predstavio
 she him_{DAT} him_{ACC} is introduced and also he is him_{DAT} him_{ACC} introduced
 ‘She introduced him to him and he did too.’
- (38) ?On *je, kao što sam vam* rekla, predstavio *se* Petru.
 he is as am you_{DAT} said introduced self_{ACC} Petar_{DAT}
 ‘He, as I told you, introduced himself to Petar.’
- (39) Dao *ga* Mariji *je* Ivan.
 given it_{ACC} Marija_{DAT} is Ivan
 ‘Give it to Marija, Ivan did.’
- (40) Jovan *je* pravilno odgovorio Ani.
 Jovan is correctly answered Ana_{DAT}
 ‘Jovan gave Ana a correct answer/did the right thing in answering Ana.’

Bošković (2001) concludes *je* is not lower than object clitics in the syntax; there is in fact no difference in syntactic height between *je* and other aux clitics. As discussed in section 1.5, the word order difference is a PF effect.

I now turn to Prosodic Inversion (PI). Halpern (1995) argues that when SC clitics are sentence initial in the syntax, they move in PF looking for a host. This movement, PI, applies only when necessary and moves clitics the minimal distance necessary (after the first stressed word). Halpern proposes PI to account for cases like (41), where *su* seems to break a constituent. For him, *su* is sentence initial in the syntax, undergoing PI in PF.

- (41) Tog *su* čovjeka vidjeli.
 that are man seen
 ‘They saw that man.’
- (42) Syntax: *su* tog čovjeka vidjeli. PF: Tog *su* čovjeka vidjeli.

However, there is strong evidence against this analysis. It fails to capture the correlation between syntactic movability and the ability to host a clitic and overgenerates in that it rules in many cases where a clitic cannot occur following the first stressed word (see Wilder and Čavar 1994, Franks and Progovac 1994, Bošković 2001.) Notice first that we don’t need PI to derive (41). SC allows left-branch extraction, as shown by (43), which can’t be derived by PI and must involve left-branch extraction of *kojeg/tog*.

- (43) Kojeg/Tog_i tvrdiš da *su* t_i čovjeka vidjeli.
 which/that you-claim that are man seen
 ‘Which man do you claim they saw/That man, you claim they saw.’

Strong evidence against PI is provided by cases where a syntactically immobile element attempts to host a clitic. In (44) we have an element

that cannot move in the syntax. (45) shows *prema*, which is stressed, also cannot precede a 2P clitic. Given (46), it should be possible for the syntax to provide to PF the output in (47), with PI incorrectly deriving (45).

- (44) *Prema_i hodaju [_{PP} t_i Mileni].
 toward walk Milena_{DAT}
 ‘They are walking toward Milena.’
- (45) *Prema *su* Mileni hodali (juče).
 toward are Milena_{DAT} walked yesterday
 ‘Toward Milena they walked.’
- (46) cf. Juče *su* prema Mileni hodali.
- (47) SS: *su* prema Mileni hodali PF: Prema *su* Mileni hodali

Split names, discussed in Franks (1998) and Bošković (2001), confirm only elements that can be placed in front of clitics by syntactic movement can host them, which means syntax, not PF, provides a host for SC clitics. Consider (48)-(50). It is possible in some cases to inflect for structural case either one or both names in a first+last name complex. (Nom. is the default case in (48)-(50).) *Leo* can be separated from *Tolstoi* by movement only when they are both inflected for structural case. Significantly, cliticization patterns with movement. This is expected if only elements that can be base-generated or syntactically moved in front of a clitic can precede it. Under the PI analysis we would expect all the examples in (50) to be good, since nothing blocks the derivation in (51).

- (48) a. Lava Tolstoja čitam.
 Leo_{ACC} Tolstoi_{ACC} read
 ‘Leo Tolstoi, I read.’
 b. ?Lava Tolstoj čitam.
 Leo_{ACC} Tolstoi_{NOM} read
 c. Lav Tolstoja čitam.
 Leo_{NOM} Tolstoi_{ACC} read
- (49) a. Lava čitam Tolstoja.
 b. *Lava čitam Tolstoj.
 c. *Lav čitam Tolstoja.
- (50) a. Lava *sam* Tolstoja čitala.
 Leo_{ACC} am Tolstoi_{ACC} read
 ‘Leo Tolstoi, I read.’
 b. *Lava *sam* Tolstoj čitala.
 c. *Lav *sam* Tolstoja čitala.
- (51) SS: Clitic Leo Tolstoi PF: Leo clitic Tolstoi

These data are the tip of the iceberg. It is easy to show with other examples that there is a correlation between syntactic mobility and the ability to host a clitic, which is totally unexpected under the PI analysis. In other words, adopting PI for SC is extremely problematic. (Bošković

2001 argues this in fact holds for Slavic in general, including the notorious *li*-construction.)

1.3 Multiple *wh*-fronting

I now turn to multiple *wh*-fronting (MWF). Rudin (1988) argues that despite superficial similarity, Bulgarian (52a) and SC (52b) have different structures. According to her, in Bulgarian all fronted *wh*-phrases are in SpecCP, while in SC only the first *wh*-phrase is in SpecCP.

- (52) a. Koj koga vižda? b. Ko koga vidi?
 who whom sees who whom sees
 ‘Who sees whom?’

Bošković (2002) argues no *wh*-phrase has to move overtly to SpecCP in SC (52b). One of my arguments concerns Superiority (ordering of *wh*-phrases). Rudin shows Bulgarian and SC behave differently regarding Superiority.

- (53) a. *Kogo koj vižda? b. Koga ko vidi?

Bošković (2002) shows this picture is more complicated. Bulgarian shows Superiority effects in all contexts. Russian doesn’t show them at all. SC, on the other hand, shows them in some contexts, namely exactly in those contexts where French must have *wh*-movement: embedded, long-distance (LD), and overt C (*li*) questions. I illustrate this here for LD questions.

- (54) a. ?Ko koga tvrđiš da je istukao?
 who whom claim that is beaten
 ‘Who do you claim beat whom?’
 b. *Koga ko tvrđiš da je istukao?
(55) a. *Jean et Marie croient que Pierre a embrassé qui?
 John and Mary believe that Peter has kissed who
 b. cf. Qui Jean et Marie croient-ils que Pierre a embrassé?
 c. cf. Pierre a embrassé qui?

There is then a correlation between Superiority in MWF languages and the contexts where non-MWF languages must have *wh*-movement: SC has superiority effects where French must have *wh*-movement, Bulgarian has them where English must have *wh*-movement, and Russian has them where Chinese must have *wh*-movement (i.e. never). This can be captured if SC/Bulgarian/Russian pattern with French/English/Chinese regarding when they have *wh*-movement; the former differ from the latter in that they have additional *wh*-fronting which I argue involves focalization. *Wh*-movement is then well-behaved with respect to Superiority. Anytime a MWF language must have *wh*-movement, it shows superiority effects. (See Bošković 1999 for explanation why, in contrast to *wh*-movement,

focalization doesn't show superiority effects. Richards 2001 proposes an alternative account, which however does not extend to all relevant contexts in SC and is based on certain incorrect assumptions about SC scrambling, see Bošković 1998).

Bošković (2003a) also shows there is variation regarding whether questions like (56) allow single-pair (SP) answers. While wh-movement languages like English and German don't allow them, wh-in-situ languages like Chinese, Hindi and Japanese allow them. Particularly interesting is French: wh-in-situ (57a) allows SP readings while (57b) does not.

(56) Who bought what?

(57) a. Il a donné quoi à qui?

he has given what to whom

b. Qu'a-t-il donné à qui?

Based on this, I conclude overt wh-movement has a damaging effect on SP answers (see Bošković 2003a for an account of this. Note we are dealing here with a one-way correlation which doesn't rule out the option of non-wh-movement languages disallowing SP answers.) Interestingly, SC allows a SP answer for (56), while Bulgarian doesn't, which confirms that, in contrast to Bulgarian, SC doesn't have to have wh-movement. As for other MWF languages, Polish, Czech, and Russian pattern with SC regarding both superiority and SP answers, while Romanian and Yiddish pattern with Bulgarian (see the references in Bošković *in press a*, which also includes discussion of speaker variation in SC and Russian that confirms the above correlation). The correlation between the availability of SP answers and the lack of Superiority effects is expected under Bošković's (2002) analysis, where they both indicate the lack of true wh-movement.

It is also worth noting that in Bošković (2003b) I argue the same mechanism is responsible for different behavior of English and French with respect to the obligatoriness of Inversion and wh-movement. Not surprisingly given the above discussion, Bulgarian and SC again pattern with English and French respectively (Inversion turns out to be irrelevant to the question of whether Russian has wh-movement; see Bošković 2002).

(58) a. Qui tu as vu?

b. *Who you have seen?

(59) a. *Kakvo toj dade na Petko/√Kakvo dade toj na Petko.

what he gave to Petko

'What did he give to Petko?'

(Bulgarian)

b. Šta on dade Ivanu?

what he gave Ivan_{DAT}

(SC)

Finally, recall Rudin argues all fronted wh-phrases are located in SpecCP in Bulgarian, forming an impenetrable cluster. Bošković (2003b) shows that when SC must have wh-movement, it switches to the Bulgarian paradigm,

with all fronted wh-phrases located in SpecCP. So, while in the contexts where SC doesn't have to have wh-movement a parenthetical can split fronted wh-phrases, in contrast to Bulgarian, in the contexts where SC must have wh-movement, SC patterns with Bulgarian. I illustrate this for LD questions (see Bošković 2003b for the full paradigm and an explanation).

- (60) Ko, po tebi, šta kupuje? (SC)
 who according-to you what buys
 'Who, according to you, is buying what?'
 (61) ?*Koj, spored tebe, kakvo kupuva? (Bulgarian)
 who according-to you what buys
 (62) *Ko, po tebi, koga vjeruju da tuče? (SC)
 who according-to you who believe_{3PL} that beats
 'Who, according to you, they believe beats who?'

1.4 Scrambling

Examples like (63) are often taken to show SC has scrambling. However, (63) doesn't necessarily show this since (63) is acceptable in English, and English doesn't have scrambling. Rather, (64) involves topicalization.

- (63) Ivana Marija voli.
 Ivan_{ACC} Marija_{NOM} loves
 (64) Ivan, Mary loves.

It is well-known that, in contrast to topicalization, scrambling Japanese is semantically vacuous. This is shown by (65), where the scrambled QNP cannot take wide scope, which the topicalized QNP in (66) can do. (All the Japanese data are from Bošković 2004 and Bošković and Takahashi 1998.)

- (65) Daremo-ni dareka-ga [Mary-ga *e* atta to] omotteiru.
 everyone_{DAT} someone_{NOM} Mary_{NOM} met that thinks
 'Everyone, someone thinks that Mary met.'
 (66) Everyone, someone thinks that Mary met.

Do Slavic languages then have Japanese scrambling? Bailyn (2001) notes that the fronted QNP can take wide scope in Russian (67).

- (67) Každogo mal'čika kto-to xočet, čtoby Boris uvidel *e*.
 every boy someone wants that_{SUBJ} Boris saw
 'Every boy, someone wants Boris to see'

Does this mean Russian doesn't have scrambling? Not necessarily. As Bošković (2004) notes, since Russian has topicalization (top) and focalization (foc) (67) may simply represent the top/foc option, making it irrelevant to the question at hand. The point extends to SC. To determine whether SC has scrambling in addition to top/foc, we need something that top/foc can't do, but scrambling can. One relevant test involves

relativized minimality (RM). It is well-known that, in contrast to, e.g. topicalization, scrambling is insensitive to RM. Thus, multiple scrambling and scrambling out of wh-islands is possible, while topicalization is disallowed in these contexts.

- (68) *That book_j, John_i, Bill said that Mary handed $e_i e_j$.
 (69) Sono hon-o_i John-ni_j Bill-ga Mary-ga $e_j e_i$ watasita to itta
 that book_{ACC} John_{DAT} Bill_{NOM} Mary_{NOM} handed that said
 (70) ??That book, John wants to know whether Mary read.
 (71) Sono hon-o_i John-ga [Mary-ga e_i yonda ka dooka]siritagatteiru
 that book_{ACC} John_{NOM} Mary_{NOM} read whether wants-to-know

SC patterns with Japanese: Stjepanović (1999a) notes (72a) contrasts with wh-movement out of wh-islands (72b), which follows if it involves scrambling, like Japanese (71) and unlike English (70). That SC has scrambling is confirmed by (73), which patterns with (69) rather than (68).

- (72) a. Ovu knjigu_i Marko i Ivan znaju kada je Petar pročitao e_i .
 this book Marko and Ivan know when is Petar read
 b. ?*Kakvu knjigu_i Marko i Ivan znaju kada je Petar
 what book Marko and Ivan know when is Petar
 pročitao e_i ?
 read
 ‘What book do Marko and Ivan know when Peter read?’
 (73) Ivanu tu knjigu Marija daje.
 Ivan_{DAT} that book_{ACC} Marija gives

The conclusion is confirmed by radical reconstruction. Saito (1992) shows that, in contrast to topicalization (74), scrambling can take a wh-phrase outside of its scope (75). Stjepanović (1999a) shows SC allows examples similar to (75), where the wh-phrase is taken outside of its scope. (Due to MWF, the wh-phrase still has to be fronted. What is important is that (76) is interpreted like *Marko zna ko želi koliko novca potrošiti*.)

- (74) *[That Mary met who]_i I know who_j e_j believes e_i ?
 (75) ?[Mary-ga nani-o katta to]_i John-ga [Bill-ga e_i itta ka sitteiru]
 Mary_{NOM} what_{ACC} bought that John_{NOM} Bill_{NOM} said Q knows
 ‘John knows what Bill said that Mary bought.’
 (76) ?[Koliko novca potrošiti]_i Marko zna ko želi e_i .
 how-much money to-spend Marko knows who wants
 ‘Marko knows who wants to spend how much money.’

This shows that in addition to top/foc, SC has Japanese-style scrambling. (As for Russian, there is some controversy regarding the RM test data; see Bailyn 2001 and Bošković 2004. The wh-phrase-outside-of-its-scope test cannot be run in Russian due to an interfering factor; see Bošković 2004).

1.5 Pronunciation of lower copies

I now turn to pronunciation of lower copies (PLC), which plays an important role in SC syntax. Under the copy theory of movement a question arises which copy of a moved element should be pronounced. It is often assumed it is always the highest copy. However, Franks (1998) (see also Bošković 2001, 2002) makes an important modification of this assumption. He argues pronunciation of heads of chains is just a preference. A lower copy can be pronounced iff this is necessary to avoid a PF violation. Bošković (2002) provides evidence for this based on MWF. Consider Romanian (77)-(80).

- (77) a. *Cine a adus ce? b. Cine ce a adus?
 who has brought what
- (78) a. Ce precede ce?
 what precedes what
 b. *Ce ce precede?
 c. Ce ee_i precede ce_i?
- (79) Ce precede ce fără să influențeze?
 what precedes what without subj. particle influence_{3p.sg}
 ‘What precedes what without influencing.’
- (80) a. What did John file without reading?
 b. *Who filed what without reading?

(77) shows Romanian is a MWF language. However, there is an exception to the obligatoriness of MWF. When *wh*-phrases are homophonous, the second *wh*-phrase is pronounced in situ (78a). Many languages have bans on homophonous sequences of certain morphemes. Since the ban pays attention to pronunciation, it should be a PF constraint. This is what rules out (78b). What about (78a)? It seems a *wh*-phrase fails to do here the movement it normally must do in the syntax to avoid violating a PF condition. Since we normally don't find this kind of phonology/syntax interaction, I proposed an alternative account in Bošković (2002). Suppose that, as always, the second *wh*-phrase undergoes syntactic movement. We then get (78c). If we pronounce the head of the chain of the second *what*, we violate the PF constraint in question. But this is exactly the case when we can pronounce a lower copy. Under the PLC analysis, the “*wh*-in-situ” in (78a) undergoes overt *wh*-movement, just like *what* in *What did John buy*, it just happens to be pronounced in situ. There is strong evidence for this analysis. It is well-known that only moved *wh*-phrases can license parasitic gaps; a *wh*-in-situ cannot do that (80). Romanian *wh*-in-situ in question licenses parasitic gaps (79), just like overtly moved *wh*-phrases.

Returning to *je*, Bošković (2001) shows PLC enables us to explain the behavior of *je* noted above. Recall *je* precedes (it is higher than)

object clitics in the syntax, but follows them in PF. Following den Dikken (1994) I adopted (81), where *je* is generated below the SS position of object clitics, and then moves above them. I proposed a PF constraint requiring *je* to be pronounced last within the clitic cluster, which was shown to have independent motivation. Given this, we must pronounce lower *je* in (81). We then have an account of the dual behavior of *je*: it behaves as if it's higher than object clitics in the syntax because it is higher than they are. It follows them in PF because a PF constraint requires pronunciation of a lower copy of *je*.

(81) je_i [_{Agrio} dative clitic [_{Agro} accusative clitic [_{VP/AuxP} je_i ...]]]

In both the *je* and the *what..what* case, PLC provides us with an elegant way of capturing syntax-phonology mismatches, where X behaves as if it's higher than where it is pronounced. In Bošković (2001) I show PLC also enables us to turn a number of optional movements into obligatory movements. To account for (4) and (82), Bošković (1997) argued that after the participle moves in front of the aux clitic, establishing part-aux order, the aux optionally moves to Agrs, the option being taken in (4) but not (82).

(82) Odgovorili *su* pravilno Mileni.
 'They gave Milena a correct answer.'
 '*They did the right thing in answering Milena.'

Under PLC, aux movement can be considered obligatory. We then have (83), where aux always moves in front of part. If there is a pronounced element in front of the aux clitic we pronounce the higher aux (83a). If there isn't, pronunciation of the higher copy would induce a PF violation, which means we can pronounce the lower copy (83b). Part-aux order then arises via lower copy pronunciation, which occurs for PF reasons so that the aux clitic can be prosodically supported. The analysis makes a prediction. Since there is nothing wrong in PF if a non-clitic aux is sentence initial, we should always pronounce the higher copy of the strong aux, which means part-aux order should be impossible with a strong aux. The prediction is borne out, as (84)-(85) show (*su* is a clitic aux).

(83) a. X aux-clitic part. ~~aux-clitic~~ b. ~~aux-clitic~~ part. aux-clitic

(84) *Odgovorili nisu/jesu njoj.
 answered not+are/ARE her
 'They did not/DID answer her.'

(85) a. Nisu/jesu odgovorili ~~nisu/jesu~~ njoj. b. ~~su~~ odgovorili *su* njoj.

Consider now (86). (86) could be taken to indicate the subject optionally moves in front of the clitic. PLC again enables us to treat this as obligatory movement. Assume the subject always moves in front of the clitic. In (87a), we can, hence must, pronounce higher *oni*. But this is

impossible in (87b), since this would violate the 2P requirement on *su*. We then pronounce lower *oni* to satisfy the PF condition in question.

- (86) a. *Oni su zaspali.*
they are fallen-asleep
'They fell asleep.'
b. *Petar tvrdi da su oni zaspali.*
Petar claims that are they fallen-asleep

- (87) a. *Oni su ~~oni~~ zaspali.* b. *Petar tvrdi da ~~oni~~ su oni zaspali.*

PLC has extensive application in SC, often hiding overt movement effects (see Bošković 2001, Stjepanović 1999b). It is then important to bear it in mind when discussing examples where PF considerations may be relevant.

2 NP structure

I now turn to NP structure. I will start by establishing several generalizations involving articles, which will be shown to have important consequences for the structure of the traditional NP (TNP). (They could turn out to be strong tendencies, which would still call for an explanation.)

2.1 Generalizations

Languages differ regarding whether they allow left-branch extractions (LB) like the following.

- (88) *Expensive/That_i he saw [t_i car]
(89) *Skupa/Ta_i je vidio [t_i kola]* (SC)
expensive/that is seen car
(90) *Doroguju/Tu_i on videl [t_i mašinu]*
expensive/that he saw car (Russian)

Uriagereka (1988), Corver (1992) and Bošković (2005) establish (91): (Like most generalizations below, this is a one-way correlation; (91) doesn't say an articleless language must have LB.)

(91) Only languages without articles may allow LB examples like (89).

Bošković (2005) notes Bulgarian and Macedonian, the only Slavic languages with articles, differ from most other Slavic languages in that they disallow LB. Within Romance, Latin, which didn't have articles, differs from Modern Romance, which has articles, in that it had LB. Mohawk,

Southern Tiwa and Gunwinjguan also allow LB and lack articles (see Baker 1996).¹

- (92) a. *Novata_i prodade Petko [t_i kola].
 new-the sold Petko car
 ‘The new car, Petko sold.’
 b. Novata kola_i prodade Petko t_i.

Before proceeding, let me note that for the purpose of (91) and other generalizations below, I take articles to be unique, i.e. occur once per TNP. The *i* ending in (93) is then not considered to be an article.²

- (93) novi/nov crveni auto
 new_{DEF}/new_{INDEF} red_{DEF} car (SC)

This makes languages like Greek, where some speakers allow AP LB, irrelevant to (91). (The “article” in such examples would not be considered an article. See also Mathieu and Sitaridou 2002, who suggest that this type of “articles” in Greek are actually agreement markers (for definiteness).)

Consider now adjunct extraction from TNPs, which English disallows.

- (94) a. Peter met [_{NP} girls from this city]?
 b. *From which city_i did Peter meet [_{NP} girls t_i]?

Observing SC and Russian do and Bulgarian doesn’t allow extraction of adjuncts out of TNP, Stjepanović (1998) (see also Bošković 2005) establishes (100). Note Polish and Czech pattern with SC and Russian.³

- (95) Iz kojeg grada_i je Petar sreo [djevojke t_i] (SC)
 from which city is Peter met girls
 (96) Iz kakogo goroda ty vstrechal [devushek t_i]? (Russian)
 from which city you met girls

¹ Based on (i) Bašić (2005) argues Bulgarian allows LB. However, without extraction (i) is unacceptable, which suggests (i) involves an adjective that is base-generated in, not moved to, its SS position, i.e. it doesn’t involve LB.

(i) Nova ja prodade kolata (toj).
 new it sold car-the he
 (ii) *(Toj) (ja) prodade nova kolata.

² It should become clear from the discussion below that what is important is the existence of a definite article in a language, given that indefinite articles have often been argued to be located below DP even in languages that clearly have DP (see, e.g., Bowers 1987, Stowell 1989, Chomsky 1995, Bošković in press b).

³ Spanish allows (95). However, Ticio (2003) shows the phrase in question is an argument in Spanish. With clear adjuncts, such extraction is impossible.

- (97) *Ot koj grad, Petko [sreštna momičeta t,]? (Bulgarian)
 from which city Petko met girls
- (98) Z którego miasta spotkałeś dziewczyny? (Polish)
 from which city you-met girls
- (99) Z kterého města jsi řekl, že jsi potkal dívky?(Czech)
 from which city you-are said that you-are met girls
- (100) Only languages without articles may allow adjunct extraction out of TNPs.

In Bošković (2004) I also establish the generalization in (101).⁴

(101) Only languages without articles may allow scrambling.

As an illustration of (101), SC, Latin, Japanese, Korean, Turkish, Hindi, Chukchi, Chichewa, Mohawk, and Warlpiri all have scrambling and lack articles. Particularly interesting here are Slavic and Romance. Note, e.g., that Bulgarian has noticeably less freedom of word order than SC. As for Romance, all modern Romance languages have articles and lack scrambling, while Latin lacked articles and had scrambling.

Next, we have the rather interesting, new generalization in (102).

(102) Negative raising (NR) in examples like (103) is disallowed in languages without articles.

SC, Czech, Polish, Slovenian, Russian, Turkish, Korean, Japanese, and Chinese all disallow NR and lack articles. On the other hand, English, German, Spanish, French, Portuguese, Romanian, and Bulgarian have both articles and NR. In light of this, (102) may actually be a two-way correlation: languages without articles disallow NR, and those with articles allow it. There are two important points to note here. First, I consider here only NR out of finite clauses (with overt C if this is an option). Second, I have relied on the ability of NR to license strict clause-mate NPIs, such as those in (104)-(107) (note the contrast between *believe*, an NR verb, and *claim*, a non-NR verb), not the interpretation judgment regarding (103), where the negation is interpreted in the lower clause.

(103) John does not believe that Mary is smart.

(104) John didn't leave/*left [_{NPI} until yesterday]

(105) John hasn't/*has visited her [_{NPI} in at least two years]

⁴ By scrambling I mean here the kind of movement referred to as scrambling in Japanese, not German, whose “scrambling” is a very different operation with very different semantic effects from scrambling in Japanese. One of the defining properties of scrambling for the purpose of (101) is taken to be the existence of long-distance scrambling out of finite clauses, which German doesn't have. For relevant discussion of German, see Bošković (2004).

- (106) a. John didn't believe [that she would leave until tomorrow]
 b. John doesn't believe [that she has visited her in at least two years]
 (107) a. *John didn't claim [that she would leave until tomorrow]
 b. *John doesn't claim [that she has visited her in at least two years]

(108) gives a partial strict NPI paradigm for the languages in question.⁵

(108)

- a. Juan no cree/*dijo que María la ha visitado en al menos dos años.
 'Juan doesn't believe/*claim that Maria has visited her in at least two years.'
 (Spanish)
- b. O João não acreditou/??disse que a Maria vai sair até amanhã.
 'John didn't believe/say that Mary would leave until tomorrow.'
 (Brazilian Portuguese)
- c. Er hat *(nicht) sonderlich viel gegessen.
 he has not particularly much eaten.
 'He did not eat that much.'
- d. Ich glaube/*freue mich nicht dass er sonderlich viel gegessen hat
 I believe/*look.forward not that he particularly much eaten has
 (German)
- e. Ion nu a crezut/spus că Maria va pleca până mâine.
 'John did not believe/*say that Mary would leave until tomorrow.'
- f. Ion nu crede/*spus că Maria a vizitat-o de cel puțin doi ani.
 'John doesn't believe/*didn't say Maria has visited her in at least two years.'
 (Romanian)
- g. Az ne vjarvam/*kazah če Meri ja e poseštavala pone dve godini.
 'I don't believe/*didn't say that Mary has visited her in at least two years.'
 (Bulgarian)
- h. Jean ne croyait/*espérait pas que Marie parte avant demain.
 'Jean didn't believe/*hope Mary would leave until tomorrow.'
 (French)
- i. *Janez ne verjame, da jo je Marija obiskala že vsaj/najmanj dve leti.
 'John doesn't believe that Mary has visited her in at least two years.'
- j. *Janez ni verjel/ne verjame, da bo Marija odšla yse do jutri.
 'John didn't believe Mary would leave until tomorrow.'
 (Slovenian)
- k. *Ivan ne vjeruje da ju je Marija posjetila najmanje dvije godine.

⁵ I used 'believe' in all the examples. If there were no interfering factors I used the above NPIs, which are underlined and interpreted in the embedded clause, the relevant reading being 'John believed/claimed Mary would not leave until tomorrow' and 'John believes/claims Mary has not visited her in at least two years'. The judgments are given only for these readings. Several examples have other readings which I have ignored (e.g. 'return tomorrow' for 'leave until tomorrow'). For space reasons I omitted base-line data like (104-105). I gave both an NR and a non-NR verb for NR languages to show that we are dealing with clause-mate NPIs. (The distinction is not relevant in non-NR languages.)

- ‘Ivan doesn’t believe that Mary has visited her in at least two years.’
- l. *Ivan nije vjerovao da će Marija otići sve do sutra.
 ‘Ivan didn’t believe that Mary would leave until tomorrow.’ (SC)
- m. *Jan nevěří, že Marie ji navštívila nejméně dva roky.
 ‘John doesn’t believe Mary has visited her in at least two years.’ (Czech)
- n. *Jan nie wierzył, że Maria wyjedzie aż do jutra.
 ‘John didn’t believe that Mary would leave until tomorrow.’ (Polish)
- o. *Yuehan bu/cai, xiangxin Mali zhidao mingtian hui likai.
 ‘John didn’t believe that Mary would leave until tomorrow.’ (Chinese)
- p. *John [Mary o-nu en az iki yi] ziyaret et-ti] san-mı-yor.
 ‘John doesn’t believe that Mary has visited her in at least two years.’
- q. *John [Mary varin-a kadar ev-den ayrıl-acak] san-ma-di.
 ‘John didn’t believe that Mary would leave until tomorrow.’ (Turkish)
- r. *Jon-wa [Mary-ga ashita made syuppatsu suru darou to] sinzi-nakatta.
 ‘John didn’t believe Mary would leave until tomorrow.’ (Japanese)
- s. ??John-un [Mary-ka eccey-kkaci-to ttena-l kes-irako] mitci ahn-ass-ta.
 ‘John didn’t believe that Mary would leave until tomorrow.’ (Korean)
- t. *Ivan ne veril, što Marija uedet až do završnega dnja.
 ‘Ivan did not believe that Mary would leave until tomorrow.’
- u. Ivan palec o palec ne udaril, čtoby mne pomoč’
 relevant reading: ‘Ivan did not do anything to help me.’
- v. *Džon ne verit, što Ivan palec o palec udaril, čtoby mne pomoč.
 ‘John does not believe that Ivan did anything to help me.’ (Russian)

Interestingly, even in languages where the NPI licensing under NR test fails, negation seems to be interpretable in the lower clause. Thus, (109) allows the “atheist” (i.e. non-agnostic) interpretation “Ivan believes God does not exist”. (The same holds for Korean, Japanese, Turkish, Chinese, Polish, Russian, and Slovenian). Still, (108k-l) are ungrammatical.

- (109) Ivan ne vjeruje da bog postoji.
 Ivan neg believes that God exists (SC)

This suggest that there is actually a three way split among verbs with respect to NR: (a) negation interpreted in the lower clause and strict NPIs licensed under NR (possible only for some verbs in languages with articles) (b) negation interpreted in the lower clause, strict NPIs not licensed under NR c. no NR at all. In work in preparation with J. Gajewski we argue the lower clause negation interpretation is actually a pragmatic effect along the lines of Horn (1989), whereas strict NPI licensing is a semantic effect (assuming a semantic approach to NPI licensing). The reader should bear in mind the above restriction regarding what I consider NR in (102).

Next, there is the generalization in (110).

(110) MWF languages without articles do not display superiority effects in examples like (52)/(53).

Recall MWF languages differ regarding whether they show Superiority effects in examples like (52)/(53). Interestingly, MWF languages without articles (SC, Polish, Czech, Russian, Slovenian, Mohawk) don't show them. MWF languages that do show them all have articles (Romanian, Bulgarian, Macedonian, Basque, Yiddish). Hungarian is an exception (it has articles and no superiority), which, however, doesn't violate (110).⁶

Another new generalization concerns clitic doubling. It is allowed in only two Slavic languages, Bulgarian and Macedonian (cf. *Ivo go napisa pismoto* 'Ivo it wrote the letter'), which also have articles. Slavic languages that do not have articles disallow it. More generally, all clitic doubling languages I am aware of (Albanian, Macedonian, Bulgarian, Greek, Somali, Spanish, French (some dialects), Catalan, Romanian, Hebrew, Arabic, Dutch (some dialects)) have articles. We then have (111).

(111) Only languages with articles may allow clitic doubling.

Turning to adnominal genitive, Willim (2000) notes English, Arabic, Dutch, German, and Catalan, all article languages, allow two lexical genitive arguments of the noun, where the genitive is realized either through a clitic/suffix or a dummy P. On the other hand, articleless languages Polish, Czech, Russian, and Latin disallow two lexical genitives. The same holds for SC, Chinese, Quechua, and Turkish. (Compare German *Hannibals(gen) Eroberung Roms(gen)* 'Hannibal's conquest of Rome' with Polish **podbicie Rzymu(gen) Hannibala(gen)*, which is unacceptable regardless of the word order.) Willim's observation leads to the generalization in (112).⁷

(112) Languages without articles do not allow transitive nominals with two lexical genitives.

Next, Živanović (2006) notes (114) has the majority reading where more than half the people drink beer. The reading is missing in Slovenian (113), which has the reading where more people drink beer than any other drink though it could be less than half the people (the plurality reading. Beer is focused.) Živanović notes German, Dutch, Hungarian, Farsi, Macedonian, and Bulgarian, which have articles, allow the

⁶ There is some idealization of the judgments here, since I ignore some speaker variation within particular MWF languages. Note also that there is an issue with respect to Hungarian since Watanabe (2003) suggests the traditional definite article in Hungarian is not a D-element (the status of Hungarian is thus unclear).

⁷ (112) concerns only nominal arguments, not possessives. I ignore for obvious reasons languages (e.g. Japanese) allowing multiple identical case constructions.

majority reading. The reading is disallowed in Czech, Polish, SC, Chinese, Turkish, and Punjabi, which lack articles and allow only the plurality reading. This then leads to (115).

(113) Največ ljudi pije pivo.

(114) Most people drink beer.

(115) Only languages with articles allow the majority superlative reading.

Finally, two correlations that don't concern Slavic. There is a locality distinction among languages with head-internal relatives (HIR): HIR in Japanese, Quechua, Navajo, and Mohawk display island sensitivity, which is not the case with Lakhota and Mojave (see Bošković in preparation and references therein). Interestingly, the former group lacks articles, while Lakhota and Mojave have them. We then have (116). Finally, Baker (1996) notes (117).

(116) Head-internal relatives display island-sensitivity in languages without articles, but not in languages with articles.

(117) Polysynthetic languages do not have articles.

The above generalizations lead to the following conclusion: There is a fundamental difference between TNP in English and articleless languages like SC which cannot be reduced to phonology (overt vs phonologically null articles). If we posit DP for both, we need to make a radical principled distinction between D in English and SC. Appealing to phonological overtness will not work since English, e.g., disallows LB (88), adjunct extraction from TNP, and scrambling even when D is null. Moreover, we are dealing with syntactic/semantic, not phonological phenomena here. It is often assumed TNP should be treated in the same way in articleless languages and English for the sake of uniformity. However, the argument fails on empirical grounds: it is simply a fact that there are radical differences between the two—there's no uniformity here. Bošković (2005, in preparation) shows there is an easy way of capturing the differences: they can be captured if there is DP in the TNP of English, but not articleless languages like SC.⁸ As shown in Bošković (2004) for scrambling, Bošković (2005) for LB, and Bošković (in preparation) for other relevant generalizations, all the generalizations in question can be deduced under the DP/NP analysis. In the next section I briefly summarize

⁸ I don't rule out the possibility that the differences could be captured in a uniform DP analysis. Such an analysis would have to posit a radical difference in the syntax/semantics of DP in English and languages like SC. However, I am not aware of such uniform DP accounts. In fact, uniform DP accounts generally ignore the above generalizations, which are the most serious problems for them.

my (2005) account of LB, developing further an argument from this work. For deductions of other generalizations, see the works cited above.

2.2 *Back to left-branch extraction*

Bošković (2005) gives two accounts of (91). The first one is based on the Phase-Impenetrability Condition (PIC), which says only the head and the Spec of a phase are accessible for movement outside of the phase. (This means phrasal movement out of XP must proceed via SpecXP if XP is a phase.) On a par with Chomsky's (2000) proposal that CP but not IP is a phase, I suggest DP is a phase, but NP isn't. Given the PIC, XP can then move out of DP only if it first moves to SpecDP. There are two more ingredients of the analysis: the traditional assumption that AP is adjoined to NP and the Anti-Locality hypothesis (the ban on movement that is too short), which is derivable from independent assumptions and argued for by a number of authors (e.g., Bošković 1994, 1997, Abels 2003, Grohmann 2003, Ticio 2003, Boeckx 2005, Jeong 2006). Like most other approaches to anti-locality, the version of anti-locality adopted in Bošković (2005) requires movement to cross at least one full phrasal boundary (not merely a segment of a phrase). AP then cannot move to SpecDP in (118) due to anti-locality. Given the PIC, it cannot move directly out of DP either (119). Anti-locality/PIC thus prevent AP extraction from DP, banning AP LB in English. They don't ban all movement out of DP: (120) is still allowed.

(118) * $[_{DP} AP_i [_{D'} D [_{NP} t_i [_{NP} \dots$

(119) * $AP_i [_{DP} [_{D'} D [_{NP} t_i [_{NP} \dots$

(120) Who_i do you like $[_{DP} t_i [_{D'} D [_{NP} \text{friends of } t_i]]$?

The ban on adjunct extraction from TNP in English can be accounted for in the same way as the ban on AP LB, given that NP adjuncts are also adjoined to NP. Moreover, the PIC/anti-locality problem doesn't arise in SC, since DP is lacking in the relevant examples.

Bošković (2005) observes LB in traditional A-as-the-head examples is allowed in SC, which also follows given that AP is not a phase.

(121) Novim_i je on $[_{AP} [_{A'} \text{zadovoljan} [_{NP} t_i [_{NP} \text{poslom}]]]$.

new is he content job

'He is content with his new job.'

Interestingly, AP LB is banned in the presence of another adjective. (Bošković 2005 notes that the ban doesn't hold for all classes of As and that strong contrastive focus on one A improves unacceptable examples; see Bošković 2005 for an account of these facts.)

(122) *Visoke je on vidio lijepe djevojke.

tall is he seen beautiful girls
 (123) cf. Visoke je on vidio djevojke.
 'He saw tall girls.'

Bošković (2005) gives an account of (122) based on McGinnis's (1998) Principle of Lethal Ambiguity, which says two elements equidistant from K are lethally ambiguous for attraction by K if they are featurally non-distinct. Since double AP LB involves a lethal ambiguity configuration ($[_{NP} AP [_{NP} AP [_{NP} N]]]$), LB of either AP is banned.

In Bošković (2005) I also propose an alternative account of AP LB based on the proposal that both the traditional structure where NP covers AP, and Abney's (1987) A-as-the-head analysis are correct, but for different languages. In particular, in English A takes NP as its complement (the AP option), while in SC N takes AP as its Spec (the NP option; NP adjunction would also work). The parametric difference is tied to DP. I assume the AP option is the default, but AP cannot be an argument. This means that when DP is lacking, as in SC (but not English), NP must dominate AP. This gives us a very simple account of English: AP LB is impossible in English because it would involve extraction of a non-constituent (AP is not a constituent to the exclusion of the NP in $[_{DP} D[_{AP} A [_{NP} N]]]$). The problem doesn't arise in SC, where the structure is $[_{NP} AP N]$. (The analysis, however, doesn't extend to the ban on adjunct extraction from TNP.)⁹

I also gave several arguments for an A/N difference in the headedness of TNP in English and SC. (124) shows prenominal adjectives disrupt case assignment in English (*him* bears default acc instead of nom). This is easily accounted for in Abney's system, where A shields the pronoun from outside case assignment as an intervening head.¹⁰ SC (125) differs from (124), suggesting Abney's analysis shouldn't be applied to SC. Note that the case of the pronoun changes in an acc. context, which shows we aren't dealing with a default case (nom. is impossible in (125b)). Note also that Russian behaves like SC.

(124) The real him/*he will never surface.

(125) a. Pravi on se nikad neće pojaviti.
 real_{NOM} he_{NOM} refl never neg+will show-up

⁹ Note that some DP languages, e.g. German (see Čavar and Fanselow 2000), allow an NP modified by an adjective to move alone (this is not fully acceptable in SC, see Bošković 2005). This is not surprising: since NP is the complement of A, AP cannot be extracted without NP, but NP is in principle extractable out of AP (provided there are no other interfering factors) in DP languages.

¹⁰ An A of a DP language doesn't seem to disrupt Case assignment to the N it modifies. I speculate the N gets its case via agreement with the D of the DP dominating the A, i.e. the V directly Case-marks the D, not the N.

- ‘The real him will never show-up.’
 b. Vidjeli smo pravog njega.
 seen are real_{ACC} him_{ACC}
 ‘We saw the real him.’ (SC)
- (126) a. Sil’naja ja smogu ego preodolet’.
 strong_{FEM.NOM} I_{NOM} will-manage him overcome
 ‘The strong me will be able to overcome him.’
 b. On ne smožet preodolet’sil’nuju menja.
 he neg will-manage overcome strong_{FEM.ACC} me_{ACC}
 ‘He will not be able to overcome the strong me.’ (Russian)

As expected, in Macedonian, which has articles hence should be an AP language, an intervening A does disrupt case assignment—the pronoun must bear the default case, which is nom. (The case doesn’t change in (127b).) Interestingly, if the pronoun is fronted (127c), it can bear structural acc. This is not surprising, since as a result of the fronting, the A no longer intervenes between the V and the pronoun. The contrast in (127b-c) confirms the intervention analysis (see Bošković 2005 for more evidence for the A/N difference in the headedness of TNP in English and SC).

- (127) a. Vistinskiot toj nikogas ne ke se pojavi.
 the-real he never neg will refl. show-up
 ‘The real him will never show up.’
 b. Go vidov vistinskiot toj/*nego.
 cl. saw the-real he/him
 ‘We saw the real him.’
 c. Go vidov nego vistinskiot. (Macedonian)

There are two alternative analysis of LB. Franks and Progovac (1994), who adopt Abney’s analysis for SC, propose a remnant movement (RP) account (see also Abels 2003, Bašić 2005), where LB involves NP movement followed by remnant AP movement (128). Ćavar & Fanselow (2000) propose a copy and delete (CD) analysis, where split constituents are derived via scattered copy deletion rather than subextraction (129).

- (128) [_{AP} Lijepe t_i]_j on gleda t_j [_{NP} kuće]_i.
 beautiful he is-watching houses
 (129) [Lijepe kuće]_i on gleda [_{NP} ~~lijepe~~ kuće]_i.

The analyses fail to capture the relevance of presence/absence of DP for LB and fail to extend to adjunct extraction. As shown in Bošković (2005), they face numerous additional problems. To mention just one, the RP analysis fails to account for the contrast in (121)/(122), while the CD analysis seriously overgenerates in that it rules in a number of unacceptable split-constituent examples (it is simply way too unconstrained). Consider also Bošković’s (2005) extraordinary LB, where a P+A complex is fronted.

- (130) *[Pravo u veliku sobu] je on ušao [pravo u veliku sobu].
 straight in big room is he entered
 (131) U veliku je on ušao sobu.

Clitic placement in (131) shows the P+A complex is a constituent. Bošković (2005) argues the constituent is created via movement (which doesn't depend on the clitic status of P) internal to the extended projection of the PP (ExPP); basically, the adjective moves to SpecExPP, the P then adjoins to it, so that further movement of it carries the P along. (130) is ruled out because *pravo* is located in SpecExPP, where the P+A complex is formed. On the other hand, (130) is unaccounted for under the CD analysis, given the indicated deletion. In fact, (130) provides evidence that extraordinary LB doesn't involve PP movement, as in the CD analysis.

2.3 Looking for D in the traditional NP in SC

Let us now consider arguments against DP in TNPs of articleless languages that are independent from the generalizations in section 2.1. I will discuss the issue with respect to SC. First, SC lacks articles, the prototypical D⁰. Though SC doesn't have articles, it does have items like *that*, *some*, as well as possessives. However, there is a lot of evidence that these items are adjectives in SC. First, they are morphologically adjectives (132). (Occasional departures from this pattern, such as those found in Russian, don't necessarily show the elements in question are not adjectives in Russian, just like the *go-went* pair doesn't show *go* is not a verb.)

- | | | | | |
|----------|------------------------------|-----------------------------|-----------------------------|------------------------------|
| (132) a. | <i>tim</i> | <i>nekim</i> | <i>visokim</i> | <i>djevojkama</i> |
| | those _{FEM.PL.INST} | some _{FEM.PL.INST} | tall _{FEM.PL.INST} | girls _{FEM.PL.INST} |
| b. | <i>tih</i> | <i>nekih</i> | <i>visokih</i> | <i>djevojaka</i> |
| | those _{FEM.GEN.PL} | some _{FEM.GEN.PL} | tall _{FEM.GEN.PL} | girls _{FEM.GEN.PL} |

Second, in contrast to English, the SC elements in question can occur in typical adjectival positions. Thus, in (133) a possessive occurs in the predicate position of a copula. (For English examples, see the glosses.)

- (133) Ova knjiga je moja.
 *this book is my

Third, unlike in English, these elements can stack up in SC, just like Adjs.

- (134) ta moja slika
 *this my picture

They also have some freedom of word order. While in English DP elements must precede adjectives, SC allows adjectives to precede some DP elements from English. (As is well-known, adjectives also have some freedom of word order (cf. *tall angry men* vs. *angry tall men*)).

- (135) a. Jovanova bivša kuća b. Bivša Jovanova kuća
 Jovan's former house *former John's house

Order permutations can have a semantic effect. So, (135b) can only refer to the house John formerly owned. To refer to an object John now possesses and that was once formerly a house (135a) must be used. (Russian *Byvšij Mišin dom/Mišin byvšij dom* pattern with (135) in this respect.) Note also that I am not saying here that the order of the SC elements in question, or adjectives in general, is completely free (contrary to what is reported in Pereltsvaig 2005). What is important is the contrast between SC and English regarding the permutability of true adjectives and some traditional “D” elements. The order of true adjectives with respect to each other, which follows from semantic and prosodic (not syntactic) factors (see the data in Pereltsvaig 2005), is not expected to be any freer in SC than in English.

Next, a SC prenominal possessive (*susjedov* in (136)) cannot be modified by a possessive, or more generally, an adjective. ((136) is acceptable on the implausible reading where *moj/bogati* modifies *konj*.)

- (136) *moj/bogati susjedov konj
 my/rich neighbor's horse

Assuming an adjective cannot be modified by an adjective, (136) follows if SC possessives are indeed adjectives. Note also that although Russian behaves like SC in this respect (**moj/bogatyj sosedov kon'*), Pereltsvaig (2005) argues such examples are irrelevant in Russian since they are ruled out independently because a possessor cannot be modified in Russian (even by an adverb). Note, however, that the simple possessor requirement clearly doesn't hold in SC. In fact, it doesn't seem to hold in Russian either.

- (137) Etot mjač nemnožko tvoj, nemnožko mamin. Net, etot mjač
 this ball a-little yours, a-little mom's no, this ball
 tol'ko mamin/Net, eto tol'ko mamin mjač.
 only mom's.

Elements that function as Ds in English are thus either missing or clearly not Ds in SC, which should be taken as an argument in favor of the no-DP analysis of SC. Notice also that Chierchia (1998) convincingly shows the DP layer is not needed for argumenthood, as is often assumed, which removes a potential semantic argument for DP in SC. Most importantly, while I am unaware of any explanations of the generalizations from section 2.1 under the universal DP analysis, they can all be explained under the DP/NP analysis, as shown in Bošković (2005, in preparation) and section 2.2.

References

- Abels, Klaus. 2003. Successive cyclicity, anti-locality, and adposition stranding. Doctoral dissertation, University of Connecticut.

- Abney, Steve. 1987. The English noun phrase in its sentential aspect. Doctoral dissertation, MIT.
- Bailyn, John. 2001. On Scrambling: A reply to Bošković and Takahashi. *Linguistic Inquiry* 32: 635-658.
- Baker, Mark. 1996. *The polysynthesis parameter*. New York: Oxford U. Press.
- Barbosa, Pilar. 1995. Null Subjects. Doctoral dissertation, MIT.
- Bašić, Monika. 2005. Split phrases in Slavic. Presented at FDSL 6, Potsdam University.
- Boeckx, Cedric. 2005. Some notes on bounding. Ms., Harvard University.
- Bošković, Željko. 1994. D-structure, θ -criterion, and movement into θ -positions. *Linguistic Analysis* 24: 247-286.
- Bošković, Željko. 1997. *The syntax of nonfinite complementation: An economy approach*. Cambridge: MIT Press.
- Bošković, Željko. 1998. Wh-movement and wh-phrases in Slavic. Presented at the Comparative Slavic Morphosyntax Workshop, Bloomington.
- Bošković, Željko. 1999. On multiple feature-checking: Multiple wh-fronting and multiple head-movement. In *Working Minimalism*, ed. S. Epstein and N. Hornstein, 159-187. Cambridge: MIT Press.
- Bošković, Željko. 2001. *On the nature of the syntax-phonology interface: Cliticization and related phenomena*. Amsterdam: Elsevier Science.
- Bošković, Željko. 2002. On multiple wh-fronting. *Linguistic Inquiry* 33: 351-383.
- Bošković, Željko. 2003a. On the interpretation of multiple questions. *Linguistic Variation Yearbook* 1: 1-15.
- Bošković, Željko. 2003b. On wh-islands and obligatory wh-movement contexts in South Slavic. In *Multiple wh-fronting*, ed. C. Boeckx and K. Grohmann, 27-50. Amsterdam: John Benjamins.
- Bošković, Željko. 2004. Topicalization, focalization, lexical insertion, and scrambling. *Linguistic Inquiry* 35:613-638.
- Bošković, Željko. 2005. On the locality of left branch extraction and the structure of NP. *Studia Linguistica* 59: 1-45.
- Bošković, Željko. in press a. A note on wh-typology. *Proceedings of FDSL* 6.
- Bošković, Željko. in press b. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry*.
- Bošković, Željko. in preparation. What will you have, DP or NP? *NELS* 37.
- Bošković, Željko, and Daiko Takahashi. 1998. Scrambling and last resort. *Linguistic Inquiry* 29: 347-366.
- Bowers, John. 1987. Extended X-bar theory, the ECP, and the left branch condition. *Proceedings of WCCFL* 6, 47-62.
- Ćavar, Damir, and Gisbert Fanselow. 2000. Discontinuous constituents in Slavic and Germanic languages. Ms., Universität Hamburg and Universität Potsdam.
- Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6: 339-405.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge: MIT Press.
- Chomsky, Noam. 2000. Minimalist inquiries. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. R. Martin, D. Michaels, and J. Uriagereka, 89-155. Cambridge: MIT Press.
- Corver, Norbert. 1992. Left branch extraction. *Proceedings of NELS* 22, 67-84.
- Dikken, Marcel den. 1994. Auxiliaries and participles. *NELS* 24, 65-79.

- Franks, Steven, and Ljiljana Progovac. 1994. On the placement of Serbo-Croatian clitics. *Indiana Linguistic Studies* 7: 69-78.
- Franks, Steven. 1998. Clitics in Slavic. Presented at the Comparative Slavic Morphosyntax Workshop, Bloomington.
- Grohmann, Kleanthes. 2003. *Prolific domains: on the anti-locality of movement dependencies*. Amsterdam: John Benjamins.
- Halpern, Aaron. 1995. *On the placement and morphology of clitics*. Stanford: CSLI.
- Horn, Laurence. 1989. *A natural history of negation*. Chicago: U. of Chicago Press.
- Jeong, Youngmi. 2006. The landscape of applicatives. Doctoral dissertation, University of Maryland.
- Mathieu, Eric and Ioanna Sitaridou. 2002. Split wh-constructions in Classical and Modern Greek. *Linguistics in Potsdam* 19: 143-182.
- McGinnis, Martha. 1998. Locality in A-movement. Doctoral dissertation, MIT.
- Pereltsvaig, Asya, 2005. On the universality of DP. Ms., Cornell University.
- Progovac, Ljiljana. 1993. Locality and subjunctive-like complements in Serbo-Croatian. *Journal of Slavic Linguistics* 1: 116-144.
- Richards, Norvin. 2001. *Movement in language*. Oxford: Oxford University Press.
- Rudin, Catherine. 1988. On multiple questions and multiple wh-fronting. *Natural Language and Linguistic Theory* 6: 445-501.
- Saito, Mamoru. 1992. Long distance scrambling in Japanese. *Journal of East Asian Linguistics* 1: 69-118.
- Stjepanović, Sandra. 1998. Extraction of adjuncts out of NPs. Presented at the Comparative Slavic Morphosyntax Workshop, Bloomington.
- Stjepanović, Sandra. 1999a. Scrambling: Overt movement or base generation and LF movement. *Journal of Slavic Linguistics* 7: 305-324.
- Stjepanović, Sandra. 1999b. What do second position cliticization, scrambling, and multiple wh-fronting have in common? Doctoral dissertation, University of Connecticut.
- Stowell, Timothy. 1989. Subjects, specifiers, and X-bar theory. In *Alternative conceptions of phrase structure*, ed. M. Baltin and A. Kroch, 232-262. Chicago: University of Chicago Press.
- Ticio, Emma. 2003. On the structure of DPs. Doctoral dissertation, University of Connecticut.
- Uriagereka, Juan. 1988. On government. Doctoral dissertation, University of Connecticut.
- Watanabe, Akira. 2003. Inderminates and determiners. Presented at the Workshop on Altaic in Formal Linguistics, MIT.
- Wilder, Chris, and Damir Čavar. 1994. Long head movement? Verb movement and cliticization in Croatian. *Lingua* 93: 1-58.
- Wilder, Chris and Damir Čavar. 1997. Verb movement, cliticization, and coordination. Ms., ZAS, Berlin and Universität Hamburg.
- Willim, Ewa 2000. On the grammar of Polish nominals. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. R. Martin, D. Michaels, and J. Uriagereka, 319-346. Cambridge: MIT Press.
- Živanović, Sašo. 2006. The stowaway of the *the*: On different readings of superlative determiners. Presented at FDSL 6.5, University of Nova Gorica.

zeljko.boskovic@uconn.edu