To appear in *Linguistic Inquiry*

**Topicalization, Focalization, Lexical Insertion, and Scrambling**

Željko Bošković

University of Connecticut

Bailyn (2001) uses Russian data to argue against Bošković and Takahashi’s (BT) (1998) analysis of scrambling, which is based on Japanese. He also claims that BT’s analysis faces several theoretical problems. In section 1, I address the empirical problems that Bailyn claims Russian poses for BT’s analysis. In section 2, I discuss the main theoretical issues that Bailyn raises, which concern 0-relations, lexical insertion, and Last Resort. In this section, I also establish a crosslinguistic correlation between lack of articles and availability of scrambling. I give an account of the correlation under BT’s analysis and explore its consequences for the status of lexical insertion with respect to Last Resort.

1 Scrambling in Japanese and Russian

BT’s analysis of scrambling was intended to address certain problems that arise under the classical analysis of Japanese scrambling, which considers scrambling in Japanese to be an optional overt movement operation that applies for no reason at all (see, e.g., Fukui 1993, Saito 1992, 1994, Saito and Fukui 1998). In minimalist terms, the scrambling movement of *sono hon-o* ‘that book-ACC’ in (1) does not involve any feature checking, which raises an obvious problem for Chomsky’s (1986, 1995) conception of movement as a last resort operation, applying only when necessary.

(1) \([_{IP}Sono \, hon-o]\, \; [{_{IP}John-ga \, [{_{CP}Mary-ga \, [{_{VP}t, \, katta]}]}\, \text{to}] \, omotteiru}\].

‘That book, John thinks that Mary bought.’

BT propose an analysis of scrambling that replaces the optional overt movement of the classical analysis that violates Last Resort with an obligatory LF movement that fully conforms with Last Resort. They propose that the scrambled element in (1) is base-generated in its S-Structure position. If it were to remain in this position in LF, the derivation would crash because *sono hon-o* would not
be Case- and 0-licensed. *Sono hon-o therefore undergoes lowering in LF to a position where it can receive Case and a 0-role. The movement is obligatory in the sense that if it does not take place, the derivation will crash.

(2) a. SS: [\text{IP } Sono hon-o [\text{IP } John-ga [\text{CP } \text{Mary-ga [VP katta]]] to} \text{ omotteiru}]].

that book-ACC John-NOM Mary-NOM bought that thinks

b. LF: [\text{IP } John-ga [\text{CP [IP Mary-ga [VP sono hon-o katta]]] to} \text{ omotteiru}]

BT give a number of arguments for this analysis. For example, they show that it explains the otherwise puzzling undoing property of scrambling (radical reconstruction in Saito’s terms). Consider (3), where *daremo-ni ‘everyone-DAT’* must have narrow scope. (I indicate the position where the scrambled element is interpreted with *e*. Under BT’s analysis, this is the landing site of LF lowering, whereas under the classical analysis, it is the launching site of overt movement. For uncontroversial overt movements, I will use *t(race)*.)

(3) Daremo-ni dareka-ga [Mary-ga *e* atta to] omotteiru.  \( \exists \succ \forall; \forall \succ \exists \)

everyone-DAT someone-NOM Mary-NOM met that thinks
‘Everyone, someone thinks that Mary met.’

(Bošković and Takahashi 1998)

Why can’t the scrambled element take scope in its S-structure position? This puzzling fact is immediately explained under BT’s analysis: *daremo-ni* must lower in LF to the position where it is 0- (and Case-) marked. Since it necessarily lowers into the embedded clause, it cannot take scope over *dareka-ga ‘someone-NOM’*.

It is worth noting here that the undoing effect (for another illustration of it, see (25)) provides strong evidence against attempts to analyze scrambling as focus or topic movement. (Bailyn makes a suggestion along these lines. Miyagawa (1997) claims that Japanese long-distance scrambling involves focus movement.) What the undoing effect shows is that semantics does not “know” about scrambling (at least long-distance scrambling, which is what we are concerned with here), in other words, for semantics, scrambling does not exist. Now, if scrambling were focus movement, we would be dealing here with focus movement that semantics does not know about. This raises an obvious problem that the focus movement analysis has to address.\(^1\) (Notice also that focus generally
facilitates wide scope, which, as noted above, is unavailable for the scrambled NP in (3).) Until the problem is addressed in a satisfactory manner, it is difficult to see the focus movement analysis as a viable alternative.²

To return to BT’s analysis: among other things, it also accounts for the inability of adjuncts to undergo scrambling, illustrated by (4). (Following BT, I ignore quasi-argument adjuncts and short-distance scrambling of adjuncts, since in the latter case it is not clear whether we are dealing with scrambling or with base-generation even under the movement analysis of scrambling.)

   Mary-NOM John-NOM reason-even without that theory-ACC believes that thinks
   ‘Mary thinks that John believes in that theory without any reason.’


   (Saito 1985)

Under the assumption that scrambling is an optional movement operation applying without any driving force, the ungrammaticality of (4b) on the relevant reading is puzzling. Why is it that, in contrast to arguments, adjuncts cannot scramble? Under BT’s analysis, this fact is readily explained. Under this analysis, the adjunct is base-generated in its S-Structure position in (4b) and must be lowered to the embedded clause in LF to modify the embedded predicate. Note, however, that the adjunct is fully licensed in its S-Structure position. In contrast to sono hon-o ‘that book-ACC’ in (1), which has Case and θ-features that are not licensed in its base-generated, S-Structure position, the adjunct in (4b) possesses neither a Case feature nor a θ-role that could motivate its LF movement. Since there is no reason for the adjunct to lower into the embedded clause in LF, Last Resort prevents it from moving.

There are actually exceptions to the impossibility of adjunct scrambling. Thus, the adjuncts in (5)-(6) can undergo scrambling.

   why Mary-NOM John-NOM that theory-ACC believes Q knows
   ‘Mary knows why John believes in that theory.’

   (Bošković and Takahashi 1998)
Note that the adjunct in (5) has a *w*-feature that can be licensed only in the embedded Spec,CP and the adjunct in (6) is a negative polarity item (NPI), whose licensing negation is located in the embedded clause. The correct descriptive generalization concerning scrambling of adjuncts is that adjuncts can undergo scrambling if and only if they are subject to a formal requirement that can be satisfied only in a lower clause. The generalization immediately follows under BT’s analysis: the formal requirement is necessary to drive LF lowering. Thus, in contrast to the adjunct in (4b), the adjuncts in (5)-(6) do have a formal feature that cannot be checked in their base-generated, S-Structure position, namely, the *w*-feature and the NPI feature. The adjuncts in (5)-(6) have to lower into the embedded clause to check these features. BT’s analysis thus accounts both for the contrast between arguments and non-*w/-NPI adjuncts and for the contrast between *w/-NPI adjuncts and non-*w/-NPI adjuncts with respect to scrambling, both of which remain unaccounted for under the standard analysis. More generally, BT’s analysis straightforwardly captures the otherwise mysterious generalization that a phrase that undergoes scrambling (be it an argument or an adjunct) must have a formal requirement that can be satisfied only in a lower position.

As BT discuss, the LF lowering analysis captures several additional otherwise puzzling properties of Japanese scrambling. Among other things, it accounts for the fact that movement out of scrambled elements is possible although normally extraction out of heads of nontrivial chains is disallowed (see Takahashi 1994, Ormazabal, Uriagereka, and Uribe-Echevarria 1994), the fact that LF scrambling is disallowed, and the fact that a scrambled element that floats a numeral cannot serve as an anaphor binder. Oku (1998a,b) shows that BT’s analysis also explains why Japanese freely allows argument drop despite the absence of standard subject and object agreement.

Bailyn claims that BT’s analysis faces several problems when applied to Russian. Thus, he claims that scrambling has semantic import in Russian and that adjuncts can scramble in Russian. The former claim is illustrated by (7), where the universal can take wide scope (for discussion of (7), see also footnote 15), and the latter by (8). (7)-(8) contrast with Japanese (3) and (4b) in the relevant respect.
(7) Každого mal’čika kto-to xočet, čtoby Boris uvidel e.
   every boy someone wants that-SUBJ Boris saw
   ‘Every boy, someone wants Boris to see.’
   (Bailyn 2001)

(8) Ja bystro xoču, čtoby oni dopisali kursovyje e.
   I quickly want that-SUBJ they wrote papers
   ‘I want them to write their papers quickly.’
   (Bailyn 2001)

(7)-(8) raise an obvious problem for BT’s analysis. Bailyn observes that they can be accounted for under the overt movement analysis of scrambling, concluding that data concerning “scrambled” scope and scrambling of adjuncts favor this analysis. However, for the conclusion to go through, one would have to show that the overt movement analysis can handle not only Russian (7)-(8) but also Japanese (3)-(6),— that is, the contrast between Russian and Japanese. Bailyn, however, does not do that. He furthermore ignores other arguments for the superiority of the LF movement analysis given by BT, briefly mentioned above, as well as BT’s discussion of several issues that he raises as potential problems for their analysis.³

Returning to the data presented so far, it appears that the Japanese paradigm favors the LF movement analysis and the Russian paradigm favors the overt movement analysis. In other words, the full paradigm cannot tease the two approaches apart. In fact, it seems to me that the contribution of Bailyn’s article lies not in teasing apart different theories of scrambling, but in claiming that Russian scrambling is in some fundamental respects quite different from Japanese scrambling, a claim that casts doubt on the possibility of a uniform analysis of Japanese and Russian scrambling.

A word of caution is in order regarding the term scrambling, one of the most abused items in the linguistic vocabulary. In the current literature, the term is often used for expository convenience when authors are not sure what kind of movement they are dealing with, or when they want to avoid committing themselves to the issue, or merely to indicate that the movement in question is different from other, better-known instances of movement regarding languages/phenomena considered.⁴ As a result, almost every well-studied language, including
English, has been claimed to have scrambling. But this is not necessarily scrambling of the kind found in Japanese. The ease-of-exposition use of the term *scrambling* raises a serious problem in crosslinguistic studies of scrambling. Obviously, what one is not sure about in one language does not have to be the same thing one is not sure about in another language. So, we cannot simply rely on the term *scrambling* when comparing claims made regarding scrambling, especially not when comparing “scrambling” in different languages. It is necessary to conduct the relevant tests to make sure we are dealing with the same phenomenon. Bailyn’s article does not do this for Japanese and Russian. In other words, it does not show that the constellation of properties that is taken to characterize Japanese scrambling is actually found in Russian. Given the above differences between Japanese and Russian “scrambling,” we could simply conclude that they are actually different phenomena, brought together only by the unfortunate usage of the term *scrambling*. In fact, Bailyn’s article can be interpreted as showing exactly that, which would make it irrelevant to BT’s analysis of scrambling. This interpretation is particularly natural in light of the fact that the undoing property, illustrated in (3), is taken in a number of works, including BT 1998 (see also, e.g., Fukui 1993, Saito and Fukui 1998, Saito 1992, 2000), to be the defining and most interesting property of Japanese-style scrambling (JSS). Since Russian does not have it (according to Bailyn), it would then follow that Russian does not have JSS. However, there is reason to believe that Russian scrambling and JSS are not as different as Bailyn’s data would lead us to believe. In fact, Bailyn’s crucial data, which provide evidence that Russian scrambling and JSS are different phenomena, do not seem to involve scrambling at all.

As noted above, BT, Fukui (1993), Saito and Fukui (1998), and Saito (1992, 2000) all take the undoing property to be the defining characteristic of JSS. In delimiting the nature of the phenomenon, these authors pay particular attention to differentiating JSS and English-style topicalization, the main distinction being the undoing property: since topicalization has semantic import (i.e., it establishes an operator-variable relation), it is not undone, unlike JSS. Thus, in contrast to the scrambled NP in (3), the topicalized NP in (9) can have wide scope.

(9) Everyone, someone thinks that Mary met.

A factor that interferes with Bailyn’s conclusions regarding Russian scrambling based on (7)-(8) is
that the language uncontroversially has topicalization as well as focalization (see, e.g., King 1993),
a fact that Bailyn disregards. Now, Japanese also has topicalization. However, topicalized elements
in Japanese have a special topic marker, *wa*. Since *sono hon-o* ‘that book - ACC’ in (1) is not *wa-
marked, it unambiguously undergoes scrambling; it could not have undergone topicalization.⁵ Unlike
topicalization in Japanese, topicalization in Russian is not accompanied by special morphology. The
same holds for focalization.⁶ There is then no way to rule out the topicalization/focalization option
for *každago mał’čka* ‘every boy’ in (7). Consequently, the fact that the quantifier can take wide
scope is not surprising: it patterns in the relevant respect with the topicalized quantifier in English
(9). Given the availability of the topicalization/focalization derivation, (7) thus does not tell us
anything about whether Russian scrambling has the undoing property, that is, whether Russian has
JSS. (The same holds for all the examples Bailyn gives to support his claim that Russian argumental
scrambling has semantic import; see his page 643.) The adverb-fronting example in (8) is also
irrelevant: all the example tells us is that adverbs can be topicalized/focalized, which is well known.

The question now arises whether Russian has JSS at all. Could it be that all the freedom of
word order in Russian is the result of topicalizing/focalizing movements, possibly coupled with some
optionality regarding subject and object A-raising? The data considered above cannot answer the
question. If dislocated elements in constructions like (7) could undergo JSS as well as
topicalization/focalization, they should be able to do everything that scrambled phrases can do and
everything that topicalized/focalized elements can do. Above, we tapped the latter. What about the
former? We can test the former with respect to islandhood, more precisely, relativized minimality
(RM). (See below for other islands. Note that when not committing myself to whether the Russian
operation under consideration involves topicalization, focalization, or JSS, I will simply refer to it
as dislocation.)

The RM data BT discuss indicate that Russian has JSS. Consider (10)-(13).

(10) a. *Kto, ty videl kogda t, pod’ezžal?
   who you saw when came

   b. *Čto, vy videli kak zapakovali t,?
   what you-PL saw how (they-)did-up
   (Müller and Sterenefeld 1993)
(11) a. Ty doktor, videl kogda e, pod’ezžal?
   you doctor saw when came
   ‘Did you see when the doctor came?’

   b. Vy pocalyku, videli kak zapakovali e,.
   you-PL parcel-ACC saw how (they-)did-up
   ‘You saw how they did up the parcel.’

   (Zemskaja 1973)

(12) a. ??Kakvu knjigu Marko i Ivan znaju kada je Petar pročitao t,?
   what book Marko and Ivan know when is Petar read
   ‘What book do Marko and Ivan know when Peter read?’

   b. Ovu knjigu Marko i Ivan znaju kada je Petar pročitao e,.
   this book Marko and Ivan know when is Petar read
   ‘Marko and Ivan know when Petar read this book.’

   (Stjepanović 1999a)

(13) *That doctor, you wonder when Peter fired t,.

(10a-b), which involve Ā-movement across an Ā-element, show that Russian wh-movement is subject to RM islands. Still, (11a-b) are acceptable. A parallel contrast is found in Serbo-Croatian (SC), another Slavic language, whose freedom of word order is similar to that of Russian, as shown in (12). Given that, as indicated by English (13), topicalization is sensitive to relativized minimality (more precisely, w/h-islands), (11) then should not involve topicalization on the derivation that yields a fully acceptable outcome. It is well known that focalization is also subject to the Wh-Island Constraint crosslinguistically. (In fact, if Russian w/h-fronting actually involves focus movement, as argued in Bošković 2002b, (10) illustrates the sensitivity of focus movement to w/h-islands.) The obvious conclusion, then, is that (11) involves scrambling.

Notice that, as BT’s (14) shows, JSS is indeed not sensitive to w/h-islands. On the other hand, like w/h-movement in Russian and SC, w/h-movement in Japanese is sensitive to w/h-islands, as (15), involving null operator movement, shows. (Kikuchi (1987) convincingly demonstrates that comparative deletion in Japanese involves null operator movement. See also Watanabe 1992
concerning *wh*-island effects in Japanese questions.) Japanese thus patterns with Russian and SC in the relevant respect.

(14) Sono hon-o, John-ga [Mary-ga e_i yonda ka dooka] siritagatteiru
    that book-ACC John-NOM Mary-NOM read whether wants-to-know
    ‘That book, John wants to know whether Mary read.’

    Bill-NOM Mary-NOM read whether wants-to-know than John-TOP more-GEN
    hon-o yonda
    book-ACC read
    ‘John read more books than Bill wants to know whether Mary read.’

It is worth noting here that BT use these data as an argument against the overt movement analysis of scrambling. On this analysis, long-distance scrambling is treated as A-movement. We should then expect it to pattern with other overt A-movement operations, like topicalization and *wh*-fronting, in that it should not be able to take place across an A-specifier.

The data concerning RM in Russian are, however, conflicting. Another difference between topicalization and scrambling discussed by BT is that, as noted by Fukui (1993), Saito (1989, 2000), and Saito and Fukui (1998), multiple scrambling is possible, whereas multiple topicalization is not.8

(16) *To John, that book, (Bill said that) Mary handed t_i t_j.

(17) Sono hon-o, John-ni, Bill-ga Mary-ga e_j e_i watasita to itta
    that book-ACC John-DAT Bill-NOM Mary-NOM handed that said
    ‘That book, to John, Bill said that Mary handed.’

(Bošković and Takahashi 1998)

According to Bailyn, Russian disallows multiple dislocation, the most natural interpretation of which would be that Russian dislocation is always topicalization/focalization—in other words, that Russian does not have JSS.9 My informants, however, find multiple dislocation examples like slightly modified Bailyn’s (18) acceptable. Müller and Sternefeld (1993) and Müller (1995) also claim that such examples are acceptable, citing (19)a-b. (Stjepanović (1999a) observes that examples like (19)a-b are also acceptable in SC.) This is consistent with the conclusion that Russian has both
topicalization/focalization and JSS, as a result of which dislocated elements in Russian can do everything that both topicalized/focalized and scrambled elements can do.\textsuperscript{10}

(18) (*) On Saše$_i$ kassetu$_j$ xočet [čtoby Boris peredal $e_j e_i$].

\[\text{he Sasha-DAT cassette-ACC wants that-SUBJ Boris gave}\]

\[\text{‘He wants Boris to give the cassette to Sasha.’}\]

(19) a. čto ty$_i$ menja$_i$ vižu [čto $e_i$ ljubiš’ $e_j$]

\[\text{that-IND you-NOM me-ACC I-see that-IND love}\]

\[\text{‘that I see that you love me.’}\]

b. čto knigu$_i$ mne$_j$ Maksim dal $e_j e_i$.

\[\text{that-IND book-ACC me-DAT Maxim-NOM gave}\]

\[\text{‘that Maxim gave me the book.’}\]

It is worth noting here that in their discussion of islands, BT focus on RM islands, which can be considered well understood in the current framework, and stay away from islands that because of their ill-understood nature cannot be used to tease apart the overt movement analysis and BT’s analysis of scrambling. Consider, for example, the Adjunct Condition (AC). In the current framework, it is not at all clear what is responsible for the descriptive generalization that crossing an adjunct boundary results in degradation. Note that under both the overt movement analysis and BT’s analysis, scrambling “out of” adjuncts involves movement crossing an adjunct boundary: under the former analysis, the crossing takes place during raising, and under the latter analysis, during lowering. To determine whether or not this should make a difference, we need to understand the nature of the AC better. If the very act of crossing an adjunct boundary is what leads to degradation, then we might expect to get AC effects with scrambling under both analyses. Bailyn argues that we would expect to get AC effects only under the overt movement analysis because BT’s scrambling movement does not leave a trace, given his assumption that a trace of an element that crosses an adjunct boundary is crucially involved in the violation. However, in the current theory, there is no reason to believe that this must be the case.

The above problem regarding the AC arises with respect to most of the putative syntactic locality phenomena Bailyn appeals to. In fact, some of them—for instance, the Coordinate Structure
Constraint (CSC)—most likely do not involve syntactic locality at all. Thus, Munn (1993) argues quite convincingly that the CSC is a constraint on semantic interpretation rather than a constraint on movement. (The CSC never actually felt comfortable as a syntactic movement constraint. Note, for example, that even the barriers system (Chomsky 1986), the most comprehensive account of locality conditions on movement and licensing of traces, failed to subsume the CSC. The well-known exceptions to it (see Postal 1998) and the existence of across-the-board movement also represent serious impediments to any syntactic treatment of the CSC.)

Returning to phenomena that are ill understood but still less controversially involve movement violations such as the AC, it is worth noting that in an early draft, BT observed that most of their consultants find constructions like (20) fully acceptable, the rest finding them only slightly degraded. Even for these speakers, such constructions are clearly better than constructions involving extraction out of adjuncts in English. (Note that a bilingual speaker of Japanese and English we consulted found Japanese (20) to be clearly better than the corresponding English construction involving topicalization out of the adjunct.) Consider also the Japanese comparative construction, which, as discussed above, involves null operator movement to Spec,CP. Japanese speakers judge the comparative (21), which involves movement to Spec,CP out of an adjunct, to be significantly worse than the scrambling example in (20).

(20) (?)[IP Sono hon-o, [IP Bill-ga [Mary-ga e, yonda kara] odoroita]]
that book-ACC Bill-NOM Mary-NOM read because was-surprised
‘That book, Bill was surprised because Mary read’

Bill-NOM Mary-NOM read because was-surprised than John-TOP more-GEN
hon-o yonda
book-ACC read
‘John read more books than Bill was surprised because Mary read.’

In fact, a number of authors have claimed that even in Russian, scrambling is rather unconstrained and differs from wh-movement with respect to a number of putative locality phenomena, contrary to what Bailyn argues. Thus, Zemskaja (1973), Müller and Sternefeld (1993), Müller (1995), and
Yadroff (1991) claim that Russian scrambling is not sensitive to a number of phenomena that Bailyn considers to involve syntactic locality and that \(wh\)-movement in Russian is sensitive to (see also Stjepanović 1999a regarding SC). In (22)-(23), I give examples from Müller and Sternefeld 1993 involving extraction out of indicative clauses (\(wh\)-movement can take place out of subjunctive but not indicative clauses in Russian) and clausal subjects. (See also (19a) for the former and (11) for \(wh\)-islands. (11b) may also involve extraction out of an adjunct.)

(22) a. ?*[Kakuju knigu], ty dumaeš’ [čto Petr pročital \(t_i\)]?
   which book you believe that-IND Peter read
   ‘Which book do you believe that Peter read?’

   b. cf. [Kakuju knigu], ty dumaeš’ [čtoby Petr pročital \(t_i\)]?
   which book you believe that-SUBJ Peter read
   ‘Which book do you believe that Peter read?’

   c. On skazal čto noski, on rad [čto kupil \(e_i\)].
   he said that-IND the-socks he is-glad that-IND he-bought
   ‘He said that he is glad that he bought the socks.’

(23) a.?*Kogo tebe kažetsja [čto \(IP\) otpustit’ \(t_i\) odno tak pozdno]]bylo by bezumiem?
   who-ACC you-DAT seems that-IND to-let-go alone so late be would insanity-INSTR
   ‘Who does it seem to you that to allow to go alone so late would be insane?’

   b. Mne Katju, kažetsja [čto \(IP\) otpustit’ \(e_i\) odnu tak pozdno]] bylo by bezumiem.
   me-DAT Katja-ACC seems that-IND to-let-go alone so late be would insanity-INSTR
   ‘It seems to me that to allow Katja to go alone so late would be insane.’

Because of the ill-understood nature of the phenomena under consideration, I hesitate to interpret the above data as an argument for the superiority of BT’s analysis over the overt movement analysis, which treats scrambling in (20), (22c), and (23b) as involving overt A-movement just like the movement of the \(wh\)-phrase/null operator in (21), (22a), and (23a). However, in light of the above discussion, it seems safe to conclude, contra Bailyn, that locality constraints do not provide an argument for the overt movement analysis. If anything, they favor BT’s analysis of scrambling.
Returning now to whether Russian has JSS, another test that could help us answer the question concerns the undoing effect, the defining characteristic of JSS. Saito (1989, 1992) shows that, in contrast to topicalization and \textit{wh}-movement, scrambling can take a \textit{wh}-phrase outside its scope in overt syntax. Notice first that a \textit{wh}-phrase in Japanese can be interpreted only if it is within a CP headed by a +\textit{wh} C. Saito and Fukui (1998) refer to the constraint in question as the \textit{Wh-Q} Constraint (following Harada 1972) and assume that it applies in LF. (Given that Japanese interrogative clauses are marked with the question markers \textit{ka} and \textit{no}, the only +\textit{wh} C in (24) is the embedded clause C.)

(24) *Dare-ga [John-ga sono hon-o katta ka] siritagatteiru?
   who-NOM John-NOM that book-ACC bought Q wants-to-know
   ‘Who wants to know [Q John bought that book]?’

Significantly, in (25), where the most embedded CP containing a \textit{wh}-phrase is scrambled to the matrix clause, the \textit{wh}-phrase can still take scope in the intermediate CP. As Saito (1989, 1992) observes constructions like (25) are not perfect. However, such constructions, in which scrambling removes a \textit{wh}-phrase from its +\textit{wh} CP, are clearly better than (24), where a \textit{wh}-phrase is base-generated in its \textit{0}-position outside its +\textit{wh} CP.

(25) ?[Mary-ga nani-o katta to], John-ga [Bill-ga e, itta ka] sitteiru.
   Mary-NOM what-ACC bought that John-NOM Bill-NOM said Q knows
   ‘John knows what Bill said that Mary bought.’
   (Bošković and Takahashi 1998)

(26) shows that \textit{wh}-movement and topicalization differ from scrambling in this respect. (26a) is marginal because of a \textit{wh}-island violation. What is important for our purposes is that it cannot at all have the interpretation on which first \textit{who} takes embedded scope. The same holds for (26b), where topicalization of a phrase containing \textit{who} places \textit{who} outside the only +\textit{wh} CP in the sentence.\textsuperscript{11}

(26) a. ?[Which picture of who]$_j$ do you wonder who, $t_i$ bought $t_j$?
   b. *[That Mary met who]$_j$, I know who, $t_i$ believes $t_j$?

\textsuperscript{11}
The fact that scrambling can take a *wh*-phrase outside its scope, in contrast to *wh*-movement and topicalization, provides further evidence for the undoing property of scrambling (in fact, this was Saito’s original argument for it). Given that, in contrast to *wh*-movement and topicalization, scrambling can be—in fact, must be (see (3))—undone, the *wh*-phrase is within its scope in (25) in LF after the undoing of scrambling (i.e., after LF lowering in BT’s analysis) so that the construction does not violate the Wh-Q Constraint, in contrast to (24) and (26a-b) (on the relevant reading of (26a)). (Note that (26) shows that the Wh-Q Constraint is operative in English.)

This argument for the undoing property of scrambling is different from the one discussed with respect to (3) in that the scrambling derivation yields an acceptable sentence that cannot be derived under the topicalization/focalization derivation. Unfortunately, we cannot use the test in question to determine whether Russian has JSS, because of an interfering factor. Russian is a multiple *wh*-fronting language, which means that, aside from a few exceptions noted in Bošković 2002b, one of which is discussed below, all *wh*-phrases in Russian must front and establish an operator-variable relation in overt syntax, the movement in question involving either focus movement or *wh*-movement (see Bošković 2002b, Stepanov 1998). There is even a stronger requirement on Russian *wh*-phrases. Russian *wh*-phrases, including those that do not move to Spec,CP overtly, must be clausemates in overt syntax with the +*wh* C heading the CP where they are interpreted. Thus, as Stepanov (1998) observes, (27a-b) are unacceptable. (Note that, as discussed in Bošković 2002b and Stepanov 1998, although Russian *wh*-phrases must undergo Á-movement in overt syntax, as the *wh*-phrases in (27) do, they do not have to move to an interrogative Spec,CP overtly. Note also that the English counterpart of (27a), given in the translation, is grammatical and that the subjunctive counterpart of (27a), *Kto xočet čoby kogo videl Petr?* ‘who wants Peter to see who?’, is not.)

(27) a. *Kto dumaet čto kogo videl Petr?*
    who thinks that-IND whom saw Peter
    ‘Who thinks that Peter saw whom?’

b. ?*Ivan i Marija dumajut čto kogo videl Petr?*
    Ivan and Marija think that whom saw Peter
    ‘Who do Ivan and Marija think that Peter saw?’

14
Note also that (28) is unacceptable on the matrix reading of either of the embedded *wh*-phrases; that is, it has to be interpreted as a multiple indirect question, in contrast to its English counterpart. (Given that Russian questions do not have to involve overt *wh*-movement, *kogda* ‘when’ can be lower than Spec,CP. As discussed in Bošković 2002b and Pesetsky 1987, 1989, D-linked *wh*-phrases are exceptional in that they do not have to move overtly. Note, however, that D-linked and non-D-linked *wh*-phrases behave in the same way with respect to (27)-(28), apart from the irrelevant fact that D-linked *wh*-phrases do not have to front.)

(28) Kto znaet kogda ty videl kakogo doktora.

who knows when you saw which doctor

The clausal mate requirement interferes with conducting Saito’s test regarding the undoing property of scrambling in Russian. However, the test can be conducted in SC, a Slavic language similar to Russian in many relevant respects. Although SC is a multiple *wh*-fronting language like Russian (which means that non-D-linked *wh*-phrases in SC undergo either *wh*-movement or focus movement overtly; see Bošković 2002b, Stjepanović 1999b), its *wh*-phrases are not subject to the clausal mate requirement. Citing the results of this test in SC, Stjepanović (1999a) in fact argues that SC has JSS. Consider (29)-(30).

(29) Ko kaže da je koga pitao šta je ona uradila?

who says that is whom asked what is she done

‘Who says that he asked whom what she did?’

(Stjepanović 1999a)

(30) ?[Koliko novca potrošiti] Marko zna ko želi €?

how-much money to-spend Marko knows who wants

‘Marko knows who wants to spend how much money.’

(Stjepanović 1999a)

(29) contains two interrogative CPs, the matrix one and the most embedded one. Nonetheless, *koga* ‘whom’ must take matrix scope, the embedded clause reading being completely unavailable (i.e., (29) can be only a multiple direct question, not a multiple indirect question). The reason for this is
that the interrogative clause within which koga is contained in (29) is the matrix one, not the embedded one. The fact that koga cannot be interpreted in the most embedded CP indicates that the Wh-Q Constraint is operative in SC. Turning to (30), notice that clausal fronting in (30) takes the wh-phrase outside the scope of the embedded Q. Stjepanović observes that the wh-phrase can still be interpreted in the embedded clause Spec,CP; in other words, (30) can be interpreted as a multiple indirect question. In fact, it can be interpreted in the same way as (31) in this respect. (Note that although (30) is not perfect, it is much better than (29) on the multiple indirect question reading. The contrast in question thus parallels the contrast between Japanese (24) and (25).)\(^{13}\)

(31) Marko zna ko želi koliko novča potrošiti.

Clausal dislocation in (30) thus patterns with JSS rather than topicalization in that it can take a wh-phrase outside its scope. Stjepanović therefore concludes that clausal dislocation in (30) involves JSS: like JSS, it does not create an operator-variable relation, and it is undone in LF. After the clause is moved to its 0-position in LF, the wh-phrase in (30) is within its scope, just like the wh-phrase in (31). The Wh-Q Constraint is therefore not violated in (30).\(^{14}\)

I conclude therefore that Slavic has JSS. The Russian examples that Bailyn uses to argue against BT’s analysis of scrambling cited in (7)-(8) are irrelevant to that analysis, in fact any analysis of scrambling, because they do not involve scrambling on the relevant derivations.\(^{15}\)

### 2 Theoretical Issues: Movement into 0-Positions and Lexical Insertion

I now turn to theoretical issues Bailyn raises as problems for BT’s analysis. Bailyn argues against this analysis because it is inconsistent with Chomsky’s (1995) position that movement into 0-positions is disallowed. One of the main goals of BT 1998 is precisely to argue against this position on both conceptual and empirical grounds. The position is a relic of the preminimalist system that assumed D-structure. As BT discuss, in a system without D-structure, such as minimalism, the ban is a blatant stipulation, hence should be dispensed with on conceptual grounds.\(^{16}\) Furthermore, since Bošković 1994, probably the first minimalist attempt at legitimizing movement into 0-positions,
many researchers have argued for movement into 0-positions and/or that 0-roles are features (an assumption that naturally leads to endorsing movement into 0-positions). See arguments by Boeckx (in press), Bošković (1997b), Hornstein (1998, 1999, 2001), Hoshi (in press), Lasnik (1999), López (2001), Kang (2002), Kayne (2003), Kim (1997), Manzini and Roussou (2000), Roehrs (2002), Saito and Hoshi (2000), Saito (2001), Stateva (2002), and Watanabe (1999), among others. Empirically, this has been a very fruitful line of research, whose accomplishments are yet to be comprehensively addressed by those who would like to maintain the stipulatory ban on movement into 0-positions. To repeat one argument for movement into 0-positions from Bošković 1994, consider the Chilean Spanish example in (32) from Gonzalez (1988).

(32) A Juan le quiere gustar Marta.

to Juan cl. wants to-please Marta
‘Juan wants to like Marta.’

A Juan bears the experiencer 0-role of the embedded verb. It is also interpreted as the wanter. A ‘to’ is the overt instantiation of the inherent Case gustar ‘to please’ assigns to the NP bearing its experiencer 0-role (see Belletti and Rizzi’s (1988) discussion of psych verbs). The presence of a provides strong evidence that at some point in the derivation, Juan was located in the embedded VP and assigned the experiencer 0-role and inherent Case by gustar. Note that a controlled PRO analysis would not work for (32) since under this analysis we have no way of accounting for the presence of a on Juan. In contrast to gustar (33a), querer ‘to want’ cannot assign the inherent Case in question to the element bearing its subject 0-role (33b). In other words, a is a label telling us that Juan in (32) has moved into the matrix clause from the embedded clause.

(33) a. A Juan le gusta la musica.
‘Juan likes music.’

b. *A Juan le quiere la fama/comer la torta.
‘Juan wants fame/to eat the cake.’

Given these facts, I argued in Bošković (1994) that Juan in (32) is inserted into the position within the embedded VP that is assigned the experiencer 0-role of gustar, as a result of which it acquires
a; it then moves to the matrix clause, where it is assigned the subject 0-role of *querer*. In other words, (32) involves movement into a 0-position. See Bošković (1994) and the works cited above for additional examples of movement into 0-positions, which show that in principle not banning movement into 0-positions is not only conceptually preferable but also empirically necessary.

Another theoretical issue Bailyn raises as a problem for the BT’s analysis concerns optionality of lexical insertion. Bailyn observes that under BT’s analysis, *sono hon-o* ‘that book-ACC’ in (1) can be inserted either in its 0-position or in the “scrambled” position, and considers this kind of optionality to be problematic. By assuming that allowing options for lexical insertion is a problem, Bailyn seems to depart from the standard syntactic reasoning that everything that is not blocked by a principle is allowed. The question is not how to allow options for lexical insertion (nothing has to be done to do that), but how to block them. To block options for lexical insertion would require a condition stating that there can be only one possibility for lexical insertion, a stipulative condition that as far as I know has never been proposed and that would clearly be empirically inadequate. Consider (34), for example, where there is more than one option for lexical insertion of *John* and *Bill*.\(^{19}\)

(34) a. John hit Bill.

b. Bill hit John.

Note also that Chomsky (1995, esp. pp. 226-227) argues that lexical insertion (including all aspects of lexical insertion: numeration formation, Select, and Merge) is not subject to usual economy considerations Move is subject to. In other words, it is costless (see Chomsky 1995 for reasoning behind the assumption.) From this perspective, transferring an instance of optionality from Move to lexical insertion, which is what BT do with scrambling, is a significant step when it comes to the recent research effort to eliminate optionality. In other words, BT’s analysis indeed resolves the problem that the apparent optionality of scrambling raises for Last Resort, contrary to what Bailyn argues.\(^{20}\)

Note, however, that a descriptive generalization concerning scrambling argued for in Bošković 2002c (see also Boeckx 2003) sheds a new light on the issue of how lexical insertion (more precisely, pure Merge) should be treated with respect to Last Resort under BT’s analysis of
scrambling. In Bošković 2002c I establish a two-class distinction among the Romance and Slavic languages. Almost all Slavic languages have extremely free word order, and almost all lack overt articles. Bulgarian is exceptional in that it has articles and its word order is more rigid than the word order of other Slavic languages such as the closely related SC, behavior I interpret as indicating that Bulgarian does not have scrambling. When it comes to Romance, modern Romance languages do not have scrambling and do have articles. Latin, on the other hand, had scrambling and no articles. Given these facts, it appears that there is a correlation between the availability of scrambling and the absence of articles.21

(35) Scrambling languages lack articles.

Several authors have argued that Slavic languages that do not have overt articles do not have DP at all—in other words, that the traditional NP is really an NP in these languages. (See Corver 1992, Zlatić 1997, 1998, Stjepanović 1998, Willim 2000, Bošković 2003a; for opposing views, see, e.g., Progovac 1998, Rappaport 1998, and Leko 1999. See also Chierchia 1998 for convincing arguments that DP is not necessary for argumenthood.) It is true that although the languages in question do not have articles, which are the prototypical instantiation of D₀, they do have lexical items corresponding to that, some, and so forth, as well as possessives. However, as the above-mentioned authors discuss, there is a great deal of evidence that in the relevant languages, these items are adjectives. First, they are morphologically adjectives, as illustrated by SC (36) with respect to a partial case paradigm. (I use SC here as the representative of the languages in question.)

(36) a. nekim mladim djevojkama

some.FEM.PL.INSTR young.FEM.PL.INSTR girls.FEM.PL.INSTR.

b. nekih mladih djevojaka

FEM.GEN.PL.

Furthermore, in contrast to their English counterparts, the SC elements in question can occur in typical adjectival positions, as shown in (37), where a possessive occurs in the predicate position of a copula construction. (For the English examples corresponding to the SC examples in (37)-(41), see the glosses.)
(37) Ova knjiga je moja.
   *this book is my

Another contrast between English and SC “D”s indicating that SC “D”s are adjectives concerns the fact that, unlike in English, the elements in question can stack up in SC, just like adjectives.

(38) ta moja slika
   *this my picture

Moreover, their order is relatively free in SC, unlike in English, where it is fixed. This is not surprising under the D-as-A analysis, since the relative order of adjectives is also relatively free.

(39) a. Jovanova skupa slika
    John’s expensive picture

b. skupa Jovanova slika
   *expensive John’s picture

(40) tall angry men vs. angry tall men

Another argument for the D-as-A analysis concerns the impossibility of modifying an SC prenominal possessive (susedov ‘neighbors’s’ in (41)) by a possessive, or more generally, an adjective. ((41) is acceptable only on the pragmatically implausible reading on which moj/bogati ‘my/rich’ modifies konj ‘horse’ instead of susedov.)

(41) *moj/bogati susedov konj
    my/rich neighbor’s horse

Assuming that an adjective cannot be modified by an adjective, (41) immediately follows if SC possessives are indeed adjectives. I therefore conclude that elements that function as Ds in languages that uncontroversially have DP are either missing or are clearly not Ds in the languages under consideration, which should be interpreted as an argument in favor of the no-DP analysis of these languages.

Another argument for the no-DP analysis of the Slavic languages that do not have articles comes from left-branch extraction and adjunct extraction out of NPs. The impossibility of left-branch
extraction of adjectives in English (see (42a)) and the impossibility of extracting adjuncts out of the traditional NP in English (see (42b)) have been attributed to the presence of D (on the former, see Corver 1992, Bošković 2003a; on the latter, Culicover and Rochemont 1992, Stjepanović 1998). The fact that they are both allowed in the Slavic languages that do not have articles, as shown by SC (43) (but crucially not by Bulgarian, a DP Slavic language, as shown in (44)), provides additional evidence that these languages have no D.  

(42) a. *New, he sold [cars].
   b. *From which city did Peter meet [girls]?  

(43) a. Nova, je prodao [cars].
   new is sold cars
   ‘New cars, he sold.’
   b. Iz kojeg grada, je Petar sreo [djevojke]? from which city is Petar met girls
   ‘Girls from which city did Petar meet?’

(44) a. *Novata prodade Petko [cars].
   new-the sold Petko car
   b. *Ot koj grad, Petko [sreština momići]? from which city Petko met girls

In light of the above discussion, suppose that the Slavic languages that do not have articles indeed have no DP and that the same holds for Latin. We can then reformulate (35) as (45).

(45) Only NP languages may allow scrambling. (+scrambling → −D) 

Given (45), the presence of DP implies the impossibility of scrambling. In other words, scrambling languages do not have DP. (Note that we are not dealing here with a two-way correlation.) Can the generalization be deduced from independent assumptions?

Under BT’s analysis, (45) entails that DPs, but not necessarily NPs, must establish a θ-relation as soon as possible, namely, in overt syntax. This can be ensured given certain assumptions concerning lexical insertion and Last Resort. As noted above, Chomsky (1995) assumes that no aspect of lexical insertion, including pure Merge, is subject to Last Resort. On the other hand,
Chomsky (2000) suggests that pure Merge is subject to Last Resort, an assumption that leads to a considerable enrichment of the theory of selection. In Bošković 1997b:37-39, I take a position that falls between Chomsky’s (1995) and (2000) positions: I suggest that only pure Merge of functional elements is subject to Last Resort. The literature contains a number of appeals to economy-of-representation principles intended to ban unnecessary projections (see, e.g., Bošković 1997b, Chomsky 1995, Grimshaw 1993, Radford 1995, Safir 1993, Speas 1994). Interestingly, in actual practice they are all applied only to functional elements; that is, they are used to ban only unnecessary functional structure. We can make this “accident” more principled by taking my (1997b) position that only pure Merge of functional elements is subject to Last Resort. Let us assume, then, that functional heads are indeed merged into the structure only if there is a reason for it. As discussed in Bošković 1997b, the functional/lexical category distinction makes sense given that lexical elements determine what we want or choose to say, and functional elements merely help us build legitimate grammatical structures. In Bošković 1997b I appeal to the natural assumption that the latter (building legitimate grammatical structures), but not the former (what we want or choose to say), is subject to economy principles to justify subjecting only pure Merge of functional elements to Last Resort. Functional elements are then inserted into the structure only to the extent that they are necessary to build legitimate structures. Another way to approach this issue would be to assume that only functional categories are selected, a natural consequence of which would be to require only pure Merge of functional elements to be motivated by selectional requirements. The upshot of the above discussion is that pure Merge of a functional projection, but not pure Merge of a lexical projection, must have independent motivation. Given that the traditional NP is DP in nonscrambiling languages, and NP in scrambling languages, pure-Merging the traditional NP with X, with X projecting, will have to have independent motivation in nonscrambiling languages, but not in scrambling languages. Since scrambling is pure Merge under BT’s analysis (see also Saito and Fukui 1998 for a different perspective on this assumption), we thus derive the costlessness aspect of scrambling and capture the scrambling/NP correlation, deducing the generalization (45). To illustrate (assuming that scrambling involves non-feature checking adjunction to IP), DP (traditional NP in nonscrambiling languages) cannot be pure-Merged adjoined to IP without violating Last Resort, while
NP (traditional NP in scrambling languages) can be. A DP can still be pure-Merged in its 0-position given that such merger involves 0-feature checking.\(^\text{27}\)

To sum up, I have argued that Bailyn’s (2001) main empirical arguments from Russian against BT’s (1998) analysis of scrambling do not involve scrambling at all. I have also shown that the correlation between the absence of DP and the availability of scrambling can be accounted for under BT’s analysis of scrambling if pure Merge of functional but not lexical elements is subject to Last Resort, as argued in Bošković 1997b.

**Footnotes**

For helpful comments and discussion, I thank the participants of the Nanzan Workshop on Scrambling and my Fall 2002 syntax seminar at the University of Connecticut, where portions of the material in section 2 were presented; two anonymous *Linguistic Inquiry* reviewers, and Sandra Stjepanović. I am also grateful to Lydia Grebenyova, Natalia Rakhlin, and Arthur Stepanov for help with judgments.

1. The problem also arises under the topic movement analysis.

2. Admittedly, Bailyn does give some evidence for discourse sensitivity of scrambling.

3. Bailyn does discuss extraction out of scrambled elements, but does not discuss the problem that BT show such extraction raises regarding general conditions on extraction.

   In this respect, it is worth noting that BT only discuss cases in which a scrambled phrase follows the element extracted/scrambled out of it, not cases like (i), where a scrambled phrase precedes a phrase that scrambles out of it, which Bailyn claims raise a problem for their analysis.

   (i) * [\text{IP} \{\text{CP} \text{Mary-ga} \ e_i \ katta \ \text{to}_j \} \ [\text{IP} \text{sono} \ \text{hon-o}_i \ [\text{IP} \text{John-ga} \ e_j \ \text{itta}]\}] \ (\text{koto})

   Mary-NOM bought that that book-ACC John-NOM said fact

   ‘Lit.: That Mary bought, that book, John said.’

Bailyn argues that (i) can be accounted for under the overt movement analysis, given the Proper Binding Condition (PBC). (\(e\), violates the PBC.) However, Saito (1989, 1992), who proposed the PBC analysis of (i), points out that under this analysis it is crucial to apply the PBC at S-Structure, since after scrambling is undone in LF, (i) no longer violates it. The PBC analysis is therefore incompatible with the Minimalist Program, which has no place for S-Structure conditions.
Furthermore, assuming that the PBC applies at S-Structure is empirically untenable owing to well-known counterexamples like remnant topicalization in German.

(ii) \([_{cp} \ [vp \ t_i \ Gelesen \_], \ [c \ hat \ das \ Buch_i \ keiner \ t_j]]\)

\begin{verbatim}
read has the book no-one
\end{verbatim}

‘Read the book, no one has.’

I conclude therefore that the PBC analysis of (i) is untenable both theoretically and empirically. (There are also accounts of (i), such as that proposed by Kitahara (1994, 1997) (see also Sauerland 1999, Müller 1998), that assume that scrambling involves feature checking. The assumption, however, faces very serious problems, discussed in Fukui 1993, Saito and Fukui 1998, and Saito 2000 (see also above for arguments against the focus movement instantiation of the feature-checking analysis).)

It is worth noting that in an early draft, BT gave an account of (i) based on Barss’s (1986) claim that sideward movement is disallowed (i.e., the assumption that there must be a c-command relation between positions associated by movement regardless of whether we are dealing with raising or lowering), motivated by the unavailability of narrow scope for someone in How likely to be sick is someone? (someone would have to move sideward when undergoing quantifier lowering to get inside the scope of likely) and the assumption that the cycle applies in both overt syntax and LF (for arguments to this effect, see Bures 1993, Branigan and Collins 1993, Jonas and Bobaljik 1993, Watanabe 1995). Under the BT’s analysis, (i) is base-generated without traces as it is. The scrambled NP and the clause must move to their \(\theta\)-positions in LF. The scrambled NP cannot be moved into the scrambled clause since this would involve sideward movement. Rather, first the clause must move to its \(\theta\)-position, and then the scrambled NP can move to its \(\theta\)-position within the clause. This derivation, however, violates the cycle since the domain of the first operation properly contains that of the second.

BT also point out that this analysis accounts for (ii), a serious problem for the PBC analysis. The difference between (i) and (ii) is that the higher fronted constituent in German is not scrambled (i.e., base-generated in its S-Structure position). Rather, it undergoes overt movement to Spec,CP, leaving behind a copy. BT suggest that in LF the scrambled phrase das Buch ‘the book’ moves to its \(\theta\)-position within the copy. The cycle violation that occurred in (i) then does not arise in (ii).
(However, see section 3, where the assumption that German has scrambling is questioned. If \textit{das Buch} undergoes real overt movement (rather than scrambling) in (ii), we need to assume either that the PBC does not hold at all or that it can be satisfied during the derivation, both of which invalidate the PBC account of (i).)

4. As explicitly noted in Bošković 2002b:360, this holds (for example) for what I there called scrambling of \textit{wh}-phrases in Slavic, which Bailyn discusses briefly.

5. As Saito (1985) discusses, \textit{wa}-marked elements can also undergo scrambling, that is, such elements can either be topicalized or undergo scrambling.

6. Here, I am ignoring the \textit{li}-focus construction.

7. As BT discuss, the scrambling derivation on which \textit{kto/čto} ‘who’/’what’ in (10) are generated in Spec,CP and lower to their \( \theta \)-position in LF is ruled out by the well-known ban on LF movement of phrases located in operator positions overtly (see Epstein 1992, Lasnik and Uriagereka 1988, Lasnik and Saito 1992, Bošković 1997a, 2003b). \textit{Kto/Čto} then have to undergo overt \textit{wh}-movement in (10).

   It is worth noting here that Stjepanović (1999a) observes that (10a-b) raise a serious problem for the overt movement analysis of scrambling. In particular, the derivation on which the \textit{wh}-phrase undergoes overt scrambling out of the \textit{wh}-island prior to \textit{wh}-movement incorrectly rules in (10a-b) given that scrambling is not subject to the \textit{Wh}-Island Constraint (see (11)).

8. As BT discuss, assuming that (16) is ruled out because it involves \( \text{Ā} \)-movement across an \( \text{Ā} \) element, (17) provides another argument against the overt movement analysis of scrambling.

9. Bailyn uses this to argue against BT’s analysis of scrambling, again not offering an account of the corresponding Japanese fact discussed by BT—in other words, not offering an account of the putative contrast between Russian and Japanese.

10. It is of course possible that there is some speaker variation, speakers who reject multiple dislocation constructions not having JSS.

11. See Stjepanović 1999a for an analysis of the cases where topicalization and \textit{wh}-movement do appear to reconstruct, which maintains the above account of the contrast between scrambling and topicalization/\textit{wh}-movement regarding the undoing effect.
12. Suppose *koga* ‘whom’ could lower to the most embedded Spec,CP in LF. (This actually might not be an option in Reinhart’s (1995) and Tsai’s (1994) systems, where only *wh*-adjuncts undergo LF *wh*-movement.) The movement would have to leave a trace (in other words, the trace cannot be deleted in LF) since the trace is in the position of the variable. The derivation in question is then ruled out by the ban on vacuous quantification and the PBC. Note that, in contrast to the derivation in question, in the case of scrambling lowering no condition of the grammar forces leaving a trace behind. BT therefore assume that scrambling lowering does not leave a trace (alternatively, the trace can be deleted), which makes the PBC irrelevant. (In this respect, BT’s analysis of scrambling is similar to May’s (1977, 1985) quantifier lowering.) It is worth noting that, as BT discuss, we have here a conceptual argument against positing a ban on lowering given that the ban would redundantly rule out the *koga* lowering derivation for (29). In other words, BT observe that positing a condition specifically banning lowering would be vastly redundant given that almost all instances of lowering are ruled out by independently needed mechanisms (in fact, as BT discuss, this holds for *all lowering in overt syntax* and *all lowering of operators*, or, more precisely, elements that are forced to leave traces by independent principles of the grammar).

13. Speakers differ regarding the Russian counterpart of (30), *Skol’ko deneg potratit’ Ivan znaet kto xožet?*, some of them accepting it on the relevant reading. I tentatively attribute this to a variation in the exact formulation of the clausemate requirement, which interferes with conducting Saito’s test in Russian.

14. It is obviously more difficult to show that scrambling not only can but must be undone for Slavic than for Japanese owing to the availability of the topicalization/focalization option (as discussed in Stjepanović 1999b, SC also has topicalization and focalization). Recall that (3) provides evidence that JSS must be undone. The interfering factor with the corresponding Russian example in (7) is the availability of the topicalization/focalization derivation, on which the fronted quantifier can take wide scope.

15. Note, however, that there are some differences between Russian (more generally, Slavic) scrambling and JSS. For example, it is well known that elements undergoing short-distance scrambling in Japanese can bind anaphors (i). On the other hand, such elements cannot bind anaphors
in Russian (ii). (The topicalization/focalization derivation is irrelevant, since topicalized/focalized elements cannot serve as A-binders.)

(i) [Mary to Pam]-ni [otagai-no hahaoya]-ga e, atta.
Mary and Pam-DAT each other-GEN mother-NOM met
‘Mary and Pam, each other’s mothers met.’

(ii) *[Larisu i Tanju], [materi drug druga,] vstretili e.
Larisa-ACC and Tanja-ACC mothers-NOM each-other-GEN met
‘Larisa and Tanja, each other’s mothers met.’

For an account of this difference between Russian and Japanese, see BT 1998. Under BT’s analysis, short-distance scrambled elements can stay in their base-generated S-Structure position in LF in Japanese, but not in Russian, which gives us a straightforward account of the contrast between (i) and (ii) (the difference between Russian and Japanese is tied to a difference between the two languages regarding the availability of the multiple subject construction, Japanese, but not Russian, allowing it).

It is often assumed that there is a difference between Russian and Japanese scrambling regarding scope. Specifically, while a short-distance scrambled element in Japanese can take either wide or narrow scope with respect to elements that c-command its 0-position, it is often assumed that in Russian, the scrambled element must take wide scope in such a configuration. Before we look at short-distance scrambling, it is worth noting that Bailyn claims that even the long-distance dislocated element in (7) must take wide scope, a behavior in stark contrast with what we find in JSS. (Recall that Japanese long-distance scrambling does not affect scope at all.) My informants who accept (7) (one does not accept it at all), however, find it ambiguous: either quantifier can take wide scope. In short-distance dislocation, the dislocated quantifier indeed must take wide scope in (iii). However, this is not the case in (iv), which is ambiguous. ((iii) is more natural with kazdago celoveka ‘every person’ regardless of the reading. Also see Ionin, in press, for discussion of scope in Russian and Stjepanović 1999a for relevant discussion of SC.)

(iii) Kazdago kto-to ljubit.

everyone-ACC someone-NOM loves
‘Everyone, someone loves.’
(iv) Každgo (čeloveka) dva studenta ljubjat.

   everyone person two students love

   ‘Everyone/Every person, two students love.’

(iv) is well behaved: the topicalization/focalization option must be responsible for wide scope of the object given that the scrambling option can only yield narrow scope. (Recall that even short-distance scrambling must be undone in Russian, unlike in Japanese, as (i)-(ii) show.) On the other hand, the lack of ambiguity in (iii) is puzzling. I leave it unresolved here, merely noting that if for some reason focalization were the only option for the dislocated quantifier in (iii), the example’s lack of ambiguity could be explained given that focus facilitates wide scope.

16. Bailyn mentions several ways of implementing the ban, all of which are based on conceptually problematic, arbitrary stipulations. Chomsky’s most recent way of implementing (a part of) the ban, namely, his (2000) principle that allows arguments to be merged only in 0-positions (BT’s analysis is incompatible with it), is also obviously problematic conceptually because of its stipulatory nature. Moreover, the principle is massively redundant (Epstein and Seely 1999, Bošković 2002a). Thus, it rules out all constructions containing an argument that never receives a 0-role, which are ruled out independently by Full Interpretation, such “arguments” being uninterpretable. In other words, they are ruled out by the part of the 0-Criterion (an argument must bear a 0-role) that follows from independently needed mechanisms (see Bošković 1994 and Brody 1993, where it is shown that everything other than the biuniqueness requirement of the traditional 0-Criterion, argued against in Bošković 1994, follows from general considerations of interpretation). The principle in question, as well as most other ways of implementing the ban on movement into 0-positions, strikes me as an attempt to impose D-structure on a system that has no natural place for it, for well-known reasons discussed in Chomsky 1995.

17. See Bošković 1994 for arguments that querer ‘to want’ indeed 0-marks Juan in (32) and for additional arguments supporting the analysis summarized in the text.

18. See Roehrs (2002) for similar examples from Icelandic. One example, originally due to Andrews (1990), is given in (i)—note that the matrix subject is assigned inherent Case and a 0-role by ganga ‘to go’.
(i) Drennum segist ganga vel við vinnuna.

the boy-dat says to-go well at work

‘The boy says he is doing well at work.’

Also see Bošković (1994) for an explanation of why this type of construction is not found with all verbs taking infinitival complements.

19. Notice also that at the point of lexical insertion, John and Bill are featural non-distinct with respect to all potentially relevant features in Chomsky’s (1999, 2000) system.

20. It is worth noting here that, unlike in earlier work (Chomsky 1995), Chomsky (2000, 2001) subjects even lexical insertion (more precisely, pure Merge) to Last Resort, the underlying assumption being that pure Merge is driven by selectional requirements. The move is conceptually rather unappealing since it enormously complicates the theory of syntactic selection, which previously had been essentially eliminated. (The dominant line of research since Pesetsky 1982, and a very productive one, has been to show that all selectional requirements follow from semantic properties (i.e., meaning) of relevant lexical items. Chomsky’s (2000, 2001) system, on the other hand, crucially needs a very rich theory of syntactic selection owing to the above-mentioned assumption concerning lexical insertion, which ends up trivializing the notion of Last Resort.) Chomsky’s (2000, 2001) system also crucially relies on rather freely allowing optionally selected additional Merge (see, e.g., his account of successive-cyclic movement), which can be easily used to implement the gist of the BT analysis in this system, characterized by a very rich selectional component. Thus, “scrambled” elements could be merged via optional additional Merge requirements on I that could be available in both English and Japanese, with 0-properties still playing the crucial role in making the relevant distinction between English and Japanese, as proposed in BT.

21. How do non-Slavic/Romance languages fare with respect to (35)? Japanese, Korean, Turkish, Hindi, Chukchi, and Warlpiri all fit the generalization in that they have scrambling and no articles—that is to say, they have no independent lexical items functioning as articles. (I am putting aside here the controversial question of whether the numeral *one* functions as an indefinite article in some languages.) These languages confirm that there is at least a strong tendency among scrambling languages to lack articles. Note also that taking (35) seriously leads to the conclusion that
German, which clearly has DP, does not have scrambling in the relevant sense of the term. German is traditionally considered to have scrambling. However, its “scrambling” differs in a number of respects from scrambling in, say, Japanese: It does not have the hallmark case of scrambling, long-distance scrambling out of finite clauses (see Ross 1986); it shows no evidence of the undoing effect associated with scrambling (in fact, its scrambling is claimed to always have semantic effects; see e.g., Diesing 1992, Lenerz 1977, Moltmann 1991, Sauerland 1999); and it does not allow scrambling of *wh*-phrases (see, e.g., Fanselow 1990, Grewendorf and Sabel 1999, Müller and Sternefeld 1993).

I assume, therefore, that German does not have scrambling in the sense of the term used here.

22. See these works for the details of the analyses. See also Ticío 2003 for relevant discussion of Romance. Ticío shows that potential counterexamples from Romance to the ban on adjunct extraction out of NPs involve argument rather than adjunct extraction.

23. As discussed in Bošković 2003a, we are not dealing here with a two-way correlation since the lack of D is not the only prerequisite for adjectival left-branch extraction and extraction of adjuncts out of NP.

24. See also Fukui 1986 for arguments that Japanese has no DP.

25. This is the effect of my (1997b) analysis. I actually assumed that all pure Merge is subject to Last Resort and provided a loophole to avoid requiring independent motivation for insertion of lexical elements. Note that here, I am generalizing the position I took with respect to lexical insertion to pure Merge in general.

26. Note that I assumed in Bošković 1997b that functional elements are not present in the numeration.

27. See also footnote 7 regarding pure Merge in the interrogative Spec,CP.

A number of issues arise here that I cannot go into in this article but that I addressed in Bošković 2002c. For example, regarding clausal scrambling, I suggested following Stepanov (2001) that there is a DP/NP on top of CP, as a result of which CP scrambling works like DP/NP scrambling. As for PP scrambling, we can assume either that there is a parallel functional structure on top of PP in nonscrambling languages (which would not be surprising in light of a number of PP/CP parallelisms noted in Bošković, in press), or that PPs are actually NPs/DPs, Ps being Case markers/particles (see BT 1998: 351, Kang 2002). Note that in Bošković 2002c, I gave an alternative
deduction of (45) that maintains the idea that an argument DP but not an argument NP must establish a θ-relation immediately, and that deduces another generalization regarding scrambling going back to Sapír 1921 (see also Alexander 1990), namely, that only overt case-marking languages may have scrambling. (Note also that in languages that have both overtly case-marked and non-case-marked NPs, such as Choctaw and Japanese, only the former can scramble.)

In the spirit of Cheng’s (1997) clausal typing requirement, according to which all clauses must be typed at S-Structure (a clause being typed as interrogative either with a question particle or by placing a wh-phrase in Spec,CP—that is, interrogative position—at S-Structure), I proposed the Argument Identification Requirement, according to which an argument must be identified at S-Structure, that is, in overt syntax. Argument identification is done either through overt case marking (the underlying assumption here is that overt case marking does have some semantic import, as in many traditional grammars (e.g., Stevanović 1969 for SC) and in recent work by Sigurðsson (2002), Uriagereka (2002), Butt and King (in press), Svenonius (in press), and Stjepanović (in preparation)), or by placing an argument in a θ-position. Given a further assumption that N, and not D, is the actual source of case (in some languages, D can get Case through low-level morphological case agreement/spreading, which is irrelevant for our purposes), we then capture both the NP/scrambling correlation and the overt case marking/scrambling correlation. In DP languages, an argument can be identified only by placing it in a θ-position in overt syntax. In NP languages, an argument can also be identified through case marking; hence, it does not need to be placed in a θ-position in overt syntax. Under BT’s analysis of scrambling, it follows that only NP languages can have scrambling. Among other things, I show that the analysis in question explains why Japanese subjects cannot scramble, given Saito’s (1985) claim that they do not bear "regular" case (in Saito’s terms, ga is not the phonetic realization of an abstract nominative Case or, in my terms, ga cannot identify an argument).

References


Academic Press.


Corver, Norbert. 1992. On deriving left branch extraction asymmetries: A case study in parametric
syntax. In NELS 22, ed. by K Broderick, 67-84. Amherst: University of Massachusetts, GLSA.

Culicover, Peter, and Michael Rochemont. 1992. Adjunct extraction from NP and the ECP. 


Ionin, Tania. In press. The one girl who was kissed by every boy: Scope, scrambling, and discourse function in Russian. Proceedings of ConSole X.


Leko, Nedžad. 1999. Functional categories and the DP in Bosnian. In Topics in South Slavic syntax and semantics, ed. by Mila Dimitrova-Vulchanova and Lars Hellan, 229-252. Amsterdam:
John Benjamins.
Bibliography:


Zlatić, Larisa. 1998. Slavic noun phrases are NPs, not DPs. Paper presented at the Comparative Slavic Morphosyntax Workshop, Spencer Creek, Ind.

Department of Linguistics, U-1145
University of Connecticut
Storrs, CT 06269
zeljko.boskovic@uconn.edu
Abstract: In this reply, I show that Russian examples that Bailyn (2001) uses to argue against Bošković and Takahashi’s (1998) analysis of scrambling are irrelevant to the analysis because they in fact do not involve scrambling. I also establish a crosslinguistic correlation between lack of articles and availability of scrambling and provide an account of the correlation under Bošković and Takahashi’s approach to scrambling.

Keywords: DP, focalization, Japanese, Last Resort, lexical insertion, Russian, scrambling, topicalization