Remarks and Replies

On Scrambling: A Reply to Bošković and Takahashi

John Frederick Bailyn

In this article I argue against Bošković and Takahashi's (1998) analysis of scrambling as base generation (with lowering for Ā-cases). I present evidence from Russian of scope and antireconstruction effects and scrambling/*wh*-movement parallels, all implicating a "classical" overt movement account of Ā-scrambling. I then discuss theoretical issues unresolved by the base generation/lowering account. Having shown that Ā-scrambling is (upward) movement, I argue that the account of A-scrambled arguments as base-generated also loses its force. In conclusion I suggest an alternative way to eliminate the apparent optionality associated with scrambling, while maintaining the classical analysis of scrambling as upward movement.

Keywords: scrambling, reconstruction, optionality, Ā-movement, constraints on movement, word order, Russian syntax

Bošković and Takahashi (1998) (hereafter B&T) claim that scrambled elements are base-generated in their surface (adjunct) position and undergo a postsyntactic process of lowering into θ-position, in Ā-cases, under a view where θ-roles are formal features, in the sense of Chomsky 1995. By eliminating scrambling itself, replacing it with base generation, this approach eliminates the optionality associated with scrambling, reversing the standard analysis that scrambling is optional, semantically vacuous movement (Ross 1967, Saito 1989, 1992).[1]

However, as I will show, B&T's approach runs into various difficulties: it leaves restrictions on scrambling unexplained; makes certain predictions that are falsified by the data; loses significant generalizations; and does not fully solve the problem that it sets out to address, namely, that of the apparent optionality of scrambling. Given these difficulties, I argue that the "classical" view of scrambling as upward movement must be maintained.

The article is structured as follows. Section 1 presents the classical view of scrambling and

I would like to thank Olga Arnaudova, Leonard Babby, John Bowers, Barbara Citko, Marcel den Dikken, Daniel Finer, Richard Larson, James Lavine, Gil Rappaport, Mark Volpe, John Whitman; audiences at WCCFL XIII, the Arizona International Workshop on Word Order and Scrambling, and at Princeton, Leipzig, and Potsdam Universities; and anonymous *LI* reviewers for helpful discussion. All mistakes remain my own. This research was partly supported by IREX and the Social Science Research Council.

[1] The following works propose movement accounts of scrambling: for German, see Webelhuth 1989; for Japanese, see Saito 1989, 1992, Miyagawa 1997, forthcoming; for Russian, see Yadroff 1992, Müller and Sternefeld 1993, Bailyn 1995, forthcoming; for Hindi, see Mahajan 1990; for Korean, see Lee and Santorini 1994. For nonmovement accounts, see Bayer and Kornfilt 1994, Neeleman 1994.

Linguistic Inquiry, Volume 32, Number 4, Fall 2001
635–658

B&T's alternative, identifying predictions that can be used to test them. Section 2 presents evidence from Russian against the predictions made by B&T's approach. Section 3 raises some theoretical issues that are problematic for their approach and returns to the original question of optionality. Section 4 discusses A-scrambling under B&T's account. An alternative way to eliminate the apparent optionality of scrambling while maintaining upward movement is presented in section 5.

1 Two Views of Scrambling

Let us begin by looking at the "classical" analysis of long-distance scrambling, then at B&T's alternative. The scrambling phenomenon is observed in sentence pairs like (1a–b) from Japanese (B&T's (3a–b)). Members of such pairs differ in word order, but are identical in grammatical relations, truth conditions, and morphology.

(1) a. John-ga [Mary-ga *sono hon-o* katta to] omotteiru.
John-NOM Mary-NOM that book-ACC bought that thinks
'John thinks that Mary bought the book.'
b. *Sono hon-o* John-ga [Mary-ga *t* katta to] omotteiru.

Starting with Ross (1967), many researchers have taken pairs like (1a–b) to be related by a "scrambling" transformation that derives (1b) from (1a) by reordering constituents.[2] The contrast between the classical analysis and B&T's alternative concerns whether or not (1a) is the transformational source sentence for (1b). The classical view, which I argue for in this article, claims that it is; B&T's alternative argues that the instances of *sono hon-o* in (1a) and (1b) occupy distinct base-generated positions.

1.1 The Classical Analysis of Scrambling

The classical view takes the structural relationship between a predicate and its arguments and modifiers in scrambling languages to resemble that found in nonscrambling languages. Specifically, arguments and modifiers begin in a local relation with their associated predicates; only later are they scrambled away. On this view (1a) must be the source structure in the pair (1a–b) since only in (1a) does *sono hon-o* 'that book' stand in a local structural relation with its predicate *katta* 'bought'. Thus, (1b) must derive from (1a) by scrambling *sono hon-o* away from its base position through raising, as in (2).

(2) *Sono hon-o* John-ga [Mary-ga *t* katta to] omotteiru.
that book-ACC John-NOM Mary-NOM bought that thinks
↑ *movement* ⌋

[2] Ross's original account limited the process to the local clause and would not have accounted for long-distance cases such as (1b). However, the term *scrambling* continued to be used in this derivational sense throughout the Government-Binding Theory period (see, among many others, Saito 1992). Under minimalist assumptions, on the classical view (1a) and (1b) share the same numeration. I return to this important assumption in section 3.

Within Government-Binding Theory the generally accepted view was that scrambling is an instance of Move α, constrained by universal principles such as Subjacency, the Empty Category Principle (ECP), and the Condition on Extraction Domain (CED). Debates have involved the kind of movement (A- or Ā-) and the correct analysis of cases like (3a–b) (B&T's (6a–b)), where a scrambled element is interpreted in its base position, despite surface displacement.

(3) a. *Nani-o*$_i$ John-ga [Mary-ga e_i katta ka] sitteiru.
what-ACC John-NOM Mary-NOM bought Q knows
'John knows what Mary bought.'
b. [*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.'

In (3a) the scrambled *wh*-object takes embedded scope, despite being in the main clause. Similarly for (3b): there an entire CP containing a *wh*-phrase is scrambled, but the *wh*-phrase contained in it is interpreted in the lower position.

Under this approach, then, (long-distance) scrambling is a form of Ā-movement that *reconstructs* at LF to base position, where interpretation occurs.[3] This approach is illustrated in (4).

(4) *Sono hon-o* John-ga [Mary-ga *t* katta to] omotteiru.
that book-ACC John-NOM Mary-NOM bought that thinks
movement ↑
reconstruction ↑

Reconstruction effects can be viewed as preliminary evidence in favor of the raising/movement analysis of scrambling insofar as they appear parallel to reconstruction effects found in *wh*-movement and other instances of Ā-movement. We will return to Russian *wh*-movement and scrambling parallels below.

1.2 Bošković and Takahashi's (1998) Analysis of Scrambling

B&T invert the classical picture of (1a–b), claiming that the position of *sono hon-o* 'that book' in (1b) is in fact base-generated. On their proposal (1a–b) are not derivationally related to each other; however, at LF the scrambled (for B&T: base-generated) object undergoes an obligatory lowering operation, which derives an LF structure essentially identical to (1a).

(5) *Sono hon-o* John-ga [Mary-ga *e* katta to] omotteiru.
that book-ACC John-NOM Mary-NOM bought that thinks
LF lowering ↑

[3] In more recent frameworks (Chomsky 1995) movement and reconstruction are replaced by copying and deletion (the copy theory). Reconstruction is handled by deleting the higher copy, leaving the lower copy at LF for interpretation. Assuming that under the copy theory constraints on movement are still handled as applying to the relationship between copies of a single element (effectively, a condition on chains), the copy theory either would render licit all cases of

Thus, in place of raising and LF reconstruction, there is *only* LF lowering. The differences between the two views and relevant definitions are summarized in (6).

(6) *Differences between classical (4) and B&T's (5) approaches to scrambling*
   a. i. In (4) the derivation involves *raising.* In (5) it does not.
      ii. In (4) the derivation involves *reconstruction.* In (5) it does not.
      iii. In (5) the derivation involves *obligatory lowering.* In (4) it does not.
   b. i. *Raising:* (standard) upward movement, forming a chain, leaving a trace, subject to constraints (Subjacency, ECP, CED, etc.)[4]
      ii. *Reconstruction:* LF movement of an element back to its base-generated spot (or, in the copy theory, deletion of the higher copy) for interpretive purposes
      iii. *Obligatory lowering:* feature-driven downward movement (as proposed by B&T)

The motivation for this proposal arises from the particular theory B&T assume, and from specific interpretive properties of scrambling. B&T frame their account within the Minimalist Program (Chomsky 1995), according to which there should be no optional movement, for reasons of economy. Although I differ with B&T on the driving force of scrambling and its mechanics, I share their determination to eliminate pure optionality from the grammar of natural language. This shared assumption serves as the backdrop for the discussion to follow.[5] As B&T observe, classical scrambling, on its movement analysis, appears to be optional. It is not motivated by Case considerations or the kind of feature checking that characterizes *wh*-movement.

Furthermore, scrambling appears to be inert semantically. That is, it appears to *obligatorily* reconstruct to its base position in the LF component (Saito 1989, 1992). In (3a–b), for instance, the displaced *wh*-phrase and *wh*-clause *must* be interpreted in their source site. With *wh*-movement and topicalization, by contrast, "undoing" the moved phrase is impossible; the movement is morphologically driven. B&T observe that "the fact that scrambling can be undone is puzzling under the standard assumption that, like *wh*-movement and topicalization, which cannot be undone, scrambling in the constructions under consideration involves overt Ā-movement" (p. 354).[6]

---

lowering (an unwelcome result) or would remain indistinct from a theory with traces, which is what both B&T and I assume. In this article I will use the movement terminology (*base position, traces,* etc.) proposed in classical accounts and maintained by B&T.

[4] This use of the term *raising* should not be confused with the narrower use associated with NP-movement out of the complement of *seem*-type verbs, which is not relevant to this article. I use *raising* simply to implicate *upward* movement, in contrast to the lowering proposed by B&T.

[5] Various attempts have been made to soften the absolute ban on optionality. For example, Chomsky (1995) argues that optionality is possible only when two convergent derivations involve an identical number of steps, and Saito and Fukui (1998) argue that a language's basic structure-building mechanisms allow for certain optional movements. I share with B&T the belief that a theory admitting true optionality is weaker than one without optionality. I return to the central issue of the apparent optionality of scrambling in section 3.

[6] B&T provide examples showing that neither *wh*-movement nor topicalization allows scope dislocations. However, the claim that *wh*-movement and topicalization are never "undone" appears to be too strong. Standard reconstruction effects hold of these constructions, as shown in (i)–(iii).

B&T derive this apparent nonparallelism between scrambling and *wh*-movement/topicalization using an analysis in which the former is something quite different from movement. Their view relies crucially on the assumption that θ-roles are formal features, which are parameterized in strength, like other formal features. A scrambling language (like Japanese) has weak θ-roles and arguments need not merge into θ-position; rather, they can be base-generated in scrambled position (adjoined to [Spec, IP]), as in (5), and lower into θ-position at LF. This LF movement is, of course, obligatory (if it were not, the θ-features would never be checked), and the Last Resort problem goes away. Japanese and English then differ in the strength of their formal θ-features.[7]

Such an analysis accounts for the following generalizations: (a) Scrambling obligatorily reconstructs. (This follows directly for B&T—scrambled arguments always lower into θ-position.) (b) Extraction out of scrambled elements is acceptable, as shown in (7).

(7) a. John-ga [Mary-ga *sono hon-o* katta to] omotteiru.
John-NOM Mary-NOM that book-ACC bought that thinks
'John thinks that Mary bought the book.'

b. *Sono hon-o*$_i$ [John-ga [$_{CP}$[$_{IP}$[$_{CP}$ Mary-ga e_i katta to]$_k$ [Bill-ga e_k itta]]
that book-ACC John-NOM Mary-NOM bought that Bill-NOM said
to] omotteiru]].
that thinks
'That book$_i$, John thinks that [that Mary bought e_i]$_k$, Bill said e_k.'

(7a) involves no scrambling. In (7b) the direct object appears outside its containing CP, which is also not in its selected position, causing no violation. (This is expected under B&T's account since after LF lowering, there is no violation.) (c) Scrambling contributes nothing to the semantics of the sentence. (This follows under B&T's account on the assumption that interpretive relations are established at LF—the displaced element is always in θ-position at LF.)

Given the assumptions shared in B&T 1998 and this article, B&T's approach makes the following predictions:

(8) a. *Prediction A*
There should be no interpretive effects associated with surface (scrambled) position. (''radical reconstruction'')

(i) [*That picture of himself*$_i$]$_k$, I know John$_i$ likes t_k. (OK only if reconstruction takes place: Principle A)

(ii) ??[*That story about him*$_i$]$_k$, I think John$_i$ heard t_k. (marginal only if reconstruction takes place: Principle B)

(iii) *[*That story about John*$_i$]$_k$, I think he$_i$ heard t_k. (* only if reconstruction takes place: Principle C)

Clearly, any theory of reconstruction must accommodate (i)–(iii) on the one hand and the contrast between scrambling and topicalization on the other. I return to this issue in section 2.

[7] B&T's account allows for another potential formalization of the typological difference between scrambling and nonscrambling languages, namely, the availability of a base-generated IP-adjunction site in the former but not in the latter. They themselves opt for the stronger claim that the strength of formal θ-features is parameterized, and I therefore follow the stronger line in discussing their account.

b. *Prediction B*
There should be no constraints on the configurational relationship between the surface and θ-positions of scrambled elements.
c. *Prediction C*
Only elements with θ-roles should participate in scrambling.

Certain points must be clarified to see how these predictions follow from B&T's approach. Prediction A follows from the obligatory nature of LF lowering for B&T. Scrambled elements are in θ-position at LF; they must therefore be interpreted there. This follows from the minimalist assumption that LF is the unique level of interface with the interpretive component of the grammar and therefore the only level at which interpretive information can be encoded.[8] Prediction B follows from the fact that lowering requires the absence of any trace in the base-generated position (B&T's claim); if there were such a trace, the Proper Binding Condition would be violated. This claim is central if B&T's approach is to allow the Proper Binding Condition to still rule out *overt* lowering, which remains generally unattested. "[O]vert lowering and lowering of elements that for independent reasons must leave traces are disallowed" (p. 351). Although B&T do not provide a specific formulation, the Proper Binding Condition ruling out overt lowering can be defined as in (9).

(9) *Proper Binding Condition* (applies at LF)
Traces must be bound.

Assuming standard c-command definitions of binding, (9) rules out any lowering process that leaves a trace, and it would rule out B&T's proposed lowering, were such movement to leave a trace. B&T's account specifically requires, therefore, that the lowering involved in (1b) *leave no trace*. "We follow Lasnik and Saito (1992) in assuming that movement does not have to leave a trace when no principle requires it. Then, the LF movement deriving [(1b)] does not have to leave a trace, rendering the Proper Binding Condition inapplicable" (p. 351). Importantly, on the assumption that a chain is defined as the relation between a trace and its antecedent (or between two copies of an element under the copy theory), in B&T's account the lack of a trace in turn denies the existence of any chain relating the base-generated site and the (LF) θ-position. Constraints on chain formation therefore should not apply in such constructions, and scrambling should be generally unconstrained (prediction B).[9] Prediction C is a central prediction of B&T's theory, namely,

[8] An *LI* reviewer suggests that B&T's approach can allow for several different kinds of processes at LF; but they do not propose this additional complication of the grammar, and it would require further justification.

[9] The preceding remarks apply primarily to representational accounts such as B&T's and the classical alternative. A derivational account employing lowering might be better equipped to capture constraints on scrambling through economy conditions (Shortest Move) but would still face the problem of disallowing overt lowering while allowing scrambling (as lowering) yet at the same time accounting for its constrained behavior. Furthermore, a lowering process such as that proposed by B&T is impossible under a purely derivational theory such as Epstein et al.'s (1998), under which a term can only be attracted by another c-commanding term, ruling out lowering of the kind B&T suggest.

that scrambling is a property of θ-marked elements.[10] We will see that these predictions are disconfirmed by data from Russian.

Russian is clearly a scrambling language in the original "stylistic" sense intended by Ross (1967) and picked up by Saito (1989, 1992). It also has overt *wh*-movement, as discussed by Müller and Sternefeld (1993) and B&T themselves, among many others.[11] Russian scrambling is therefore a good testing ground for B&T's account.[12]

2 Empirical Arguments against Scrambling as Lowering

2.1 Interpretive Effects of Overt Scrambling

Prediction A of B&T's analysis is that scrambled position is not relevant to LF interpretation (for Ā-scrambling). This follows from obligatory lowering, which expressly disallows any trace or copy in the scrambled position (the base-generated position for B&T). However, the prediction that no interpretive effects are associated with the scrambled position is too strong; we will now look at various cases in which scrambled position is directly relevant to interpretation.

2.1.1 Russian Scope Effects B&T do not state explicitly how scope is to be determined, which (as an *LI* reviewer points out) remains a serious unanswered question for the lowering account. However, B&T do not claim to be assuming anything other than the usual account of scope (namely, configurational LF relations), and thus scrambled orders should not produce changes in scope relations, since under B&T's account scrambled elements undergo obligatory lowering at

[10] An *LI* reviewer points out that this prediction does not appear to follow directly from B&T's analysis. The reviewer cites B&T's example (13), repeated here as (i), in which an adjunct *wh*-phrase is scrambled in Japanese, as a possible counterexample to prediction C (addressed by B&T exactly because of its potentially unexpected behavior under their account).

(i) ?Naze Mary-ga [John-ga sono setu-o sinziteiru ka] sitteiru.
why Mary-NOM John-NOM that theory-ACC believes Q knows
'Mary knows why John believes in that theory.'

However, in discussing this example, B&T acknowledge that without piggybacking on the *wh*-feature forcing movement in (i), the phrase in question would have no motivation for lowering if base-generated, since it is an adjunct, and not θ-marked. More precisely, then, prediction C might read: Only elements with θ-roles should participate in scrambling, unless they are associated with *independent features* that force movement. In that case the objection might continue that Russian adverbs, which scramble freely, might well be associated with independent features that force their lowering from scrambled position, leaving B&T's base generation account intact. Recall, however, that B&T claim that the typological difference between scrambling and nonscrambling languages lies in the strength of θ-features. Therefore, the effects of the parameter should not be found on any elements other than θ-marked elements. Prediction C might then read: Only elements with θ-roles should behave differently in scrambling languages and nonscrambling languages. (Thanks to Richard Larson for discussion of this issue.)

[11] Some recent accounts of Slavic *wh*-fronting have concluded that various apparent instances of *wh*-movement are in fact instances of scrambling or its discourse equivalent, focus movement. See Stepanov 1997, Bošković 1997, 1998, Boeckx and Stjepanović 2000, and Strakhov 2001. Such analyses rely crucially on the (classical) analysis of scrambling as upward movement; the resulting generalizations are lost if scrambling does not involve movement. Thus, it appears difficult to maintain both this type of analysis and B&T's.

[12] Scrambling judgments vary considerably across Russian dialects. Scrambling is accepted freely in some (see, e.g., Yadroff 1994) and is far more restricted in others. The range of possibilities in itself argues for a more fine-grained notion of scrambling, one that is possible with stricter and milder constraints on overt raising, rather than a binary parameter of weak versus strong θ-features.

LF into θ-position to check θ-features. At LF they therefore occupy canonical argument position. The scrambled position is simply not available as a position for scope interpretation. Prediction A of B&T's analysis (namely, that no interpretive effects should be associated with the surface position of scrambled elements) should thus apply to scrambling and scope in Russian. And yet, as the following contrast shows, scrambled orders do differ from nonscrambled orders with respect to scope in Russian.[13]

(10) a. Kto-to xočet, čtoby Boris uvidel každogo mal'čika.
someone-NOM wants that Boris saw [every boy]-ACC
'Someone wants Boris to see every boy.'
(i) ∃x ∀y (ii) *∀y ∃x
b. [*Každogo mal'čika*]$_i$ kto-to xočet, čtoby Boris uvidel t_i.
[every boy]-ACC someone-NOM wants that Boris saw
'Every boy someone wants Boris to see.'
(i) *∃x ∀y (ii) ∀y ∃x

(10a) and (10b) are both unambiguous, but with different relative scope. The lack of ambiguity in (10a) is expected under standard versions of Quantifier Raising (QR) as a clause-bounded process. The possibility of wide scope for *every boy* in (10b), however, goes against the prediction made by the lowering hypothesis. This fact is a problem for an account that requires obligatory lowering into θ-position at the same level where scope relations are determined. Both sentences should have the same interpretation.[14]

2.1.2 Russian Antireconstruction Effects It is well known that reconstruction with *wh*-movement differs for different kinds of phrases, as shown in (11). (The same holds for English topicalization. See Huang 1993 and Heycock 1995 for discussion.)

(11) a. *[*How proud of John*$_i$]$_k$ do you think *he*$_i$ should be t_k?
b. [*Which question that Gore*$_i$ *got during the debate*]$_k$ do you think *he*$_i$ messed up on t_k the worst?

In (11a) reconstruction occurs, creating an LF configuration in which Principle C is violated. (11b), however, does not reconstruct; Principle C is not violated. Both Huang (1993) and Heycock (1995) argue that the distinction reduces to a difference in LF position, although they differ on the nature of the distinction. Antireconstruction effects thus provide a diagnostic for whether reconstruction has or has not occurred. Thus, for B&T antireconstruction effects such as the one

[13] Following Comrie (1973), Müller and Sternefeld (1993), and many others since, I assume that extraction in Russian is generally acceptable only out of *čtoby* (subjunctive) clauses, and not out of *čto* (indicative) clauses. Thus, Russian examples of long-distance scrambling will be given in such structures. How the *čtoby* versus *čto* distinction is to be handled is not crucial to the present discussion (but see Avrutin and Babyonyshev 1994 for an interesting account).

[14] An *LI* reviewer raises the question of a distinct process of (overt) QR for (10b). However, covert QR is required for Russian, where scope ambiguities are common (and overt QR is rare). And because QR is generally taken to be clause-bounded, overt QR could not account for (10b). Therefore, (10b) involves scrambling (lowering for B&T), and it should manifest embedded scope, contrary to speaker judgments.

illustrated in (11b) should disappear with scrambling. However, the same contrast shown in (11) obtains with Russian scrambling.

(12) a. On [*dovol'noj rabotoj Marii*$_i$]$_k$ sčitaet ee$_{j/*i}$ t_k davno.
he-TOP satisfied-FEM work-INSTR Mary-GEN considers her long since.
'He has considered her$_i$ satisfied with Mary's$_i$ work long since.'

b. On [*sluxi o Marii*$_i$]$_k$ xočet, čtoby ona$_i$ uslyšala t_k.
he-TOP rumors about Mary wants that she heard
'He wants her$_i$ to hear rumors about Mary$_i$.'

c. [*Nekotorye voprosy Goru*$_i$]$_j$ ja xoču, čtoby on$_i$ srazu zabyl t_j.
some questions Gore-DAT I want that he soon forgot
'Some questions to Gore$_i$ I want him$_i$ to immediately forget.'

The predicate in (12a) reconstructs, triggering a Principle C violation.[15] The arguments in (12b–c), on the other hand, are interpreted in their surface position; there is no Principle C violation. This is the opposite of what we would expect under B&T's proposal. Given obligatory lowering we should not expect any antireconstruction effects with scrambling (prediction A) since all scrambled arguments *obligatorily* lower for θ-checking.[16] However, such effects exist in Russian, as shown in (12). On the other hand, a theory that allows raising to scrambled positions (followed by reconstruction) has the flexibility to handle these data. We now turn to reconstruction in raising accounts.

2.1.3 "Sensitive" Reconstruction Heycock (1995) discusses the reconstruction facts in (11) and concludes that reconstruction is obligatory for "nonreferential" phrases and optional or impossible for others (primarily arguments). This is reminiscent of "optional" movement processes in Germanic (especially object shift), which are related to semantic effects by Diesing and Jelinek (1996) and others, whereby the LF "type" of an optionally shifted element determines its LF position (inside or outside its home VP, for example). If LF reconstruction after movement is "sensitive" to the semantic nature of the element, as such accounts maintain, we can allow for the distinctions found above while maintaining a classical raising account. It is unclear how B&T's account would accommodate semantic sensitivity of this kind since all arguments (and only arguments) must lower at LF to θ-position. Clearly, some are not interpreted in θ-position.[17]

[15] Notice that the existence of a Principle C violation here is a problem under prediction C as well, since the constituent involved is not an argument and yet appears to reconstruct/lower.

[16] Lebeaux (1988), Heycock (1995), Safir (1999), and many others discuss the possibility of a process of "late insertion" for adjuncts that would allow them to be located high in the structure even in cases where the argument they adjoin to is interpreted in lower position. Late insertion will not be enough, however, to explain away all antireconstruction effects, even if the relevant constituents in (12) are shown to be true adjuncts, something that is not immediately clear. Moreover, Safir shows that there are cases in which *both* the upper and lower copy are required for interpretation, something B&T's account does not appear to predict to be possible.

[17] Here and elsewhere B&T would have to resort to introducing another, different kind of LF level (Huang's LF′, for example) to allow both lowering and location in a different position at the time the binding theory applies (a possibility B&T do not discuss). (The same holds for scope, as shown in section 2.1.1.) In effect, this would amount to reintroducing the scrambled position as relevant to the interface, and thus would undermine the spirit of B&T's own analysis.

This accords with a general direction in the literature on LF going back to Chomsky 1976, 1981, and also found in Diesing 1992, Huang 1993, Collins and Thráinsson 1996, and elsewhere. The existence of "sensitive" reconstruction appears to be incompatible with B&T's LF lowering approach without further elaboration of the accompanying semantics.

2.2 Scrambling as Raising: Syntactic Constraints

Let us now turn to prediction B—namely, that under B&T's account no raising is involved in the derivation of alternative word orders.

2.2.1 Scrambling out of Scrambled Phrases B&T argue that scrambling out of scrambled phrases is possible because scrambling involves a trivial chain, out of which further movement is acceptable. Recall (7), repeated as (13).

(13) a. John-ga [Mary-ga *sono hon-o* katta to] omotteiru.
John-NOM Mary-NOM that book-ACC bought that thinks
'John thinks that Mary bought the book.'

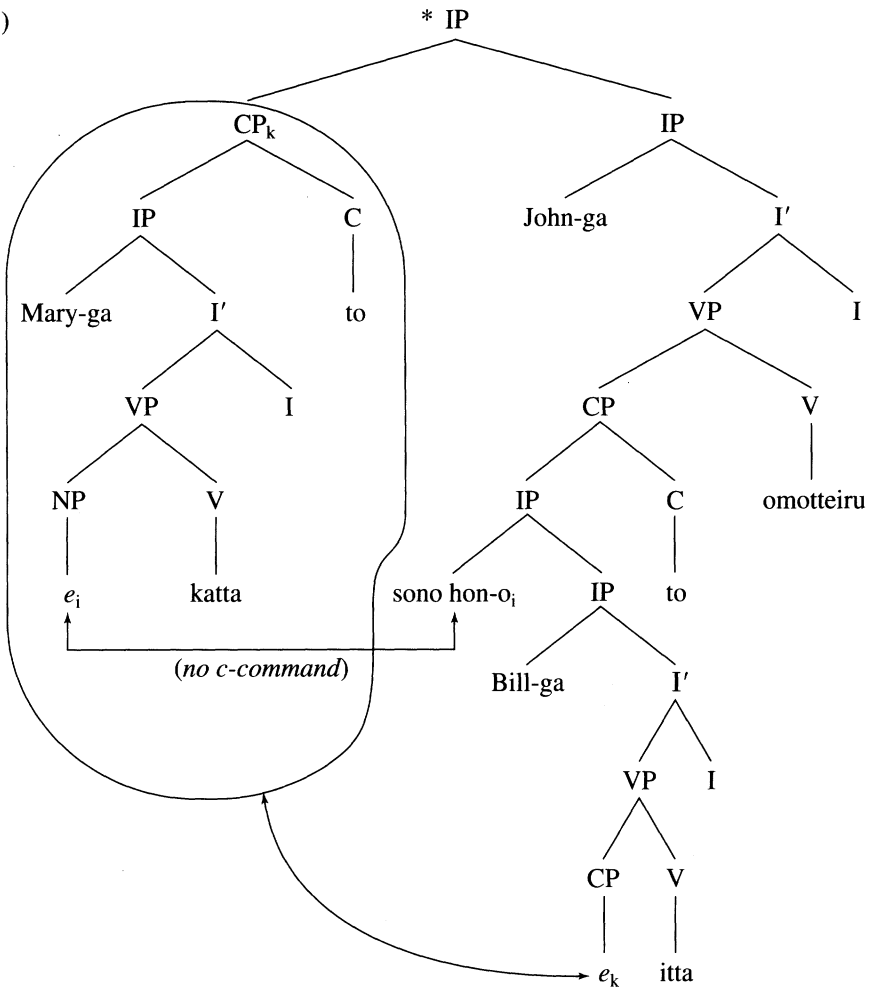
b. [*Sono hono-o*$_i$ [John-ga [$_{CP}$[$_{IP}$[$_{CP}$ Mary-ga e_i katta to]$_k$ [Bill-ga e_k itta]]
that book-ACC John-NOM Mary-NOM bought that Bill-NOM said
to] omotteiru]].
that thinks
'That book$_i$, John thinks that [that Mary bought e_i]$_k$, Bill said e_k.'

B&T's account is straightforward: (13b) involves two instances of base generation followed by lowering. Both the smaller (embedded NP) and larger (CP) components are base-generated in adjoined position and undergo LF lowering. Because there are no traces at all, the derivation is licit. "In [(13b)] . . . no relevant extraction takes place, since both 'scrambled' phrases are base-generated where they are at S-Structure" (p. 357). Notice, however, that a raising approach fares no worse here. The chains formed by overt movement deriving (13b) violate no derivational principles in a raising account either. In particular, e_k is c-commanded by its antecedent, as is e_i. Where the two accounts differ is with respect to a different derivation: one in which the two displaced constituents, CP_k and NP_i, appear *in the opposite order.* In this case the lowering analysis predicts grammaticality (in fact, no order should be worse than any other), whereas the raising analysis predicts a violation. The latter prediction is confirmed by the ungrammaticality of (14).

(14) *[[$_{CP}$ Mary-ga e_i katta to]$_k$ [John-ga [$_{CP}$ *sono hon-o*$_i$ [$_{IP}$ [Bill-ga e_k itta]]
Mary-NOM bought that John-NOM that book-ACC Bill-NOM said
to] omotteiru]].
that thinks
'[That Mary bought e_i]$_k$, John [that book$_i$] thinks that, Bill said e_k.'

Under a raising analysis the ungrammaticality of (14) is the expected result because there is an improper chain between [$_{NP}$ that book] and its trace, ruled out by the Proper Binding Condition, as shown in (15).

(15)



Under the lowering analysis the order found in (14) is predicted to be grammatical because the CP and NP lower into argument position, and no Proper Binding Condition violation is possible (on B&T's assumption that this kind of lowering leaves no trace). The sentence should fare no worse than (13b). It is not clear how B&T's proposals can account for the contrast between (13b) and (14).[18]

[18] (14), though ungrammatical, is reminiscent of German remnant movement constructions, which are acceptable. However, the acceptability of such constructions does not weaken the difficulty (14) presents for B&T's account (indeed, their account predicts such cases to be the norm); that is, it predicts any case of scrambling of a larger XP over a more locally scrambled smaller piece from within XP to be acceptable, contrary to fact, and contrary to the prediction made by the Proper Binding Condition. Furthermore, German remnant movement may be restricted to VP remnants, which would not fall under B&T's account in any event, since VPs are not θ-marked constituents. See Epstein et al. 1998 for further discussion.

2.2.2 Wh-*Movement and Scrambling Parallels* B&T's lowering account of long-distance scrambling denies the existence of a "record" in base-generated (adjoined) position; in this way LF lowering is allowed. Thus, we do not expect (typical) movement constraints to hold of long-distance scrambling under B&T's analysis (prediction B), assuming either that constraints hold of chains formed between traces and their antecedents or that they apply to the structural position of traces left by movement. And yet there is evidence that such constraints do apply to scrambling, and with considerable regularity, providing additional support for the classical claim that long-distance scrambling is overt Ā-movement. Consider the Russian sentences in (16).

(16) a. **Kogo*$_i$ Marina znaet [čto [Ivan ljubit t_i]]?
who-ACC Marina-NOM knows that Ivan-NOM loves
'Who does Marina know that Ivan loves?'
b. **Borisa*$_i$ Marina znaet [čto [Ivan ljubit t_i]].
Boris-ACC Marina-NOM thinks that Ivan-NOM loves
'Marina thinks that Ivan loves *Boris*.'

(16a) shows that *wh*-movement is ungrammatical out of an indicative embedded clause. (16b) shows that the same restriction applies to scrambling.[19]

Let us consider possible approaches to such parallels. First, it is possible that (16b) *does not* involve scrambling. If it involves no movement, the (ungrammaticality) parallel with *wh*-movement is unexpected.[20] If it is another kind of movement (not scrambling), as an *LI* reviewer suggests, then the data are being excluded as belonging to another class of phenomena and a significant generalization is lost (either the *wh*-movement/scrambling parallels or a unified account of scrambling). The alternative remains that (16b) *is* a case of scrambling, and therefore under B&T's account it involves lowering. Something about this lowering, therefore, causes the violation. However, B&T claim that the *only* movement in such cases is lowering (which leaves no trace), and in that case parallels such as the one in (16) are unexpected. In Russian (as in German; see Webelhuth 1989) *wh*-movement and scrambling are restricted in highly parallel fashion. A partial list of relevant constraints restricting movement is provided in (17) (see Fowler 1987 and Bailyn 1995 for discussion).[21]

[19] The exact nature of the violation with extraction is the subject of some dispute. See Comrie 1973, Avrutin and Babyonyshev 1994, and Bailyn 1995 for various views. Solving this problem is beyond the scope of this article, but I take the parallel violations in (16) as evidence that upward movement has taken place in both cases.

[20] There are constructions that appear to be grammatical versions of (16b), primarily with nominative arguments. Indeed, B&T cite such an example, (22b), from Müller and Sternefeld (1993), which goes back to Comrie 1973, to show that *wh*-movement and scrambling are not parallel. However, I show in Bailyn 1995 that these are not cases of scrambling, but cases of left-dislocation with a null resumptive pronoun, which explains why nominative morphology is strongly preferred. Without it, as in (16), left-dislocation is unavailable, scrambling remains the only possibility, and the effect of movement constraints remains.

[21] In a footnote B&T set aside the issue of (some of) these constraints by claiming that they are potentially not movement constraints at all: "We ignore here the Coordinate Structure Constraint, the Left Branch Condition, and the Specificity Condition, since it is not at all clear that these are movement constraints" (p. 358, fn. 17). I will not address that possibility further here until the case for the sudden exclusion of these well-known constraints is better made. Since their effects on *wh*-movement are well known, I will continue to treat them in standard fashion.

(17) *Movement constraints active in Russian*
  a. Subjacency/Complex NP Constraint
  b. Constraint on extraction out of adnominal genitives
  c. Coordinate Structure Constraint
  d. Empty Category Principle (includes *that-t* effect)
  e. Condition on Extraction Domain

Examples of Subjacency violations are shown in (18).

(18) a. *$Kogo_i$ ty pozvonil agentu, kotoryj ljubit t_i?
  whom-ACC you-NOM phone spy-DAT who loves
  'Whom did you phone a spy who loves?'
  b. *$Borisa_i$ ty pozvonil agentu, kotoryj ljubit t_i!
  Boris-ACC you-NOM phone spy-DAT who loves
  'It's BORIS you phoned a spy who loves!'

(18a) is a Subjacency violation caused by *wh*-movement. (18b) is a Subjacency violation caused by scrambling. If scrambling involves lowering leaving no trace, such a parallel is predicted not to exist. Another example involves the Coordinate Structure Constraint, which holds in strict form in Russian (and probably universally). Like Subjacency in (18), this constraint applies equally to scrambling and to *wh*-extraction.

(19) a. *$Kogo_i$ ty xočeš', čtoby Ivan videl [t_i i Mašu]?
  whom-ACC you want that Ivan saw and Masha-ACC
  '*Whom do you want Ivan to see and Masha?'
  b. *$Borisa_i$ ty xočeš', čtoby Ivan videl [t_i i Mašu].
  Boris-ACC Ivan want that Ivan saw and Masha-ACC
  '*BORIS you want Ivan to see and Masha.'

In (19a) *wh*-movement violates the Coordinate Structure Constraint. (19b) is equally ungrammatical. This parallel cannot be captured under an obligatory lowering account: in (19b) *Borisa* should be generable in IP-adjoined position and then lower into θ-position, without causing any violation, again because no trace is left and no chain is formed. In fact, however, the sentence is just as ungrammatical as the parallel *wh*-extraction case (19a). Examples of this kind abound (see Bailyn 1995 for Russian and Webelhuth 1989 for German). B&T's analysis appears to leave such parallels unexplained.

2.3 Nonarguments

B&T's theory makes another strong testable prediction, prediction C—namely, that only elements assigned a θ-role should be able to scramble.[22] If θ-role "checking" is the driving force behind the ability to appear displaced from base position, then *only* those elements that receive θ-roles

[22] See footnote 10 for discussion of whether features other than θ-roles could drive the lowering process. If they could, then B&T's typology of scrambling breaks down.

should show variation in word order. Non-θ-marked items should never be found dislocated from their immediate clause. However, Russian demonstrates scrambling of adjectives, adverbs, and other nonarguments.

(20) a. Ja xoču, čtoby oni bystro dopisali kursovye.
I want that they quickly wrote papers
'I want them to write their papers quickly.'
b. Ja *bystro* xoču, čtoby oni *t* dopisali kursovye.
I quickly want that they wrote papers
'I want them to write their papers quickly.'
c. *Ja *bystro* znaju agenta, kotoryj pišet *t*.
I quickly know agent who writes
'I quickly know an agent who writes.'
d. Ja *zelenuju*$_i$ xoču, čtoby ty kupila [t_i knigu].
I green-ACC want that you bought book-ACC
'I want you to buy a green book.'

In (20b) the embedded manner adverb *bystro* 'quickly' is located in the main clause, yet it modifies the embedded verb just as in (20a). Under B&T's theory the acceptability of (20b) is unexpected, without additional assumptions. To be interpreted in the lower clause, *bystro* must be in that position at LF. And yet there is no (feature-driven) motivation for such LF movement. Prediction C is disconfirmed. (20c), a standard Subjacency violation, provides additional evidence for the raising account of scrambling even in cases involving adverbs. (20d) is a case of so-called split scrambling whereby an attributive adjective from the embedded clause is scrambled to the main clause, a movement unexpected under B&T's account.

3 Some Theoretical Issues Raised by Lowering

In addition to empirical issues, the theory proposed by B&T raises theoretical points involving θ-theory and feature checking, as well as the original issue of optionality and economy.

3.1 θ-Theory, Feature Checking, and Multiple Scrambling

Under standard minimalist assumptions (Chomsky 1995), θ-marking occurs at Merge, but this requirement is simply stipulated. It appears, however, that any theory of θ-marking that does not associate it with Merge greatly increases the empirical burden put on theories of phrase structure, which specifically rule out movement into complement position. If thematic relations are not local at Merge, even those between a head and its complement, the traditional arguments that sisterhood is the closest syntactic relation are lost.

Indeed, in Chomsky 1995 it is made quite clear that θ-roles are not formal features: "Under any approach that takes Attract/Move to be driven by morphological features . . . *there should be no interaction between θ-theory and the theory of movement*" (p. 312, emphasis mine). Some arguments in favor of this are the following: (a) The Chain Condition defines a chain, formed

by Move, as meeting several conditions "which we take to be part of the definition of the operation [Move] itself. . . . α must c-command its trace, so that there cannot be an operation that lowers α . . . ; *movement is raising,* in the specific sense defined by c-command" (p. 253, emphasis mine). The tail of an A-chain is defined by its status as a θ-position; under B&T's theory the notion of an A-chain would then be different for scrambling and nonscrambling languages. (b) Internal arguments are merged into complement position of their predicates. However, complements are crucially *not* within the checking domain of a head. Clearly, in a theory such as B&T's θ-relatedness will (often) not result from Merge at the base at all (in Japanese-type languages, at least) but will be subject to checking in complement position. We might therefore expect arguments to be regularly generated in adjoined positions and to assume their θ-positions only at LF, predicting multiple scrambling to be the norm as there should be no preference for arguments to appear in θ-position.[23] In fact, multiple long-distance scrambling in Russian is not only rare—it is ungrammatical for many speakers, something the base generation account does not predict.

(21) a. Ivan xočet, čtoby Boris peredal *kassetu* *Saše.*
Ivan-NOM wants that Boris gave cassette-ACC Sasha-DAT
'Ivan wants Boris to give the cassette to Sasha.'
b. Ivan xočet, čtoby *kassetu*$_i$ Boris peredal t_i Saše.
c. Ivan xočet, čtoby *Saše*$_k$ Boris peredal kassetu t_k.

(21b–c), with one argument scrambled, do not distinguish between the two theories. However, any attempt to scramble both arguments leads to a violation.

(22) a. *Ivan *Saše*$_k$ *kassetu*$_i$ xočet, čtoby t'_i t'_k Boris peredal t_i t_k.
b. *Ivan *kassetu*$_i$ *Saše*$_k$ xočet, čtoby t'_i t'_k Boris peredal t_i t_k.

Under the base generation and lowering account, which involves no traces (or chain formation), such constructions should be commonplace, and certainly no less acceptable than (21); whereas under a raising account (22a–b) can be ruled out by Relativized Minimality.[24] Thus, we have both empirical and conceptual reasons not to abandon the idea that θ-role assignment is unrelated to movement.[25]

3.2. Optionality and Economy

Perhaps the most serious theoretical issue for B&T's account concerns the very problem they begin with: optionality. Recall the main motivation for B&T's analysis: to eliminate the apparent

[23] It is possible that there are languages that allow only this kind of dislocation strategy. They would not, however, have the properties of Russian and Japanese.

[24] The proper account of this violation assumes Rizzi's (1990) Relativized Minimality, under which Ā-movement over an occupied Ā-position is a violation. Such restrictions appear hard to capture in B&T's account.

[25] Inherent Case assignment, which relates lexical Case marking to particular θ-roles, also raises serious questions for B&T's account as it stands. How can this relation be handled if inherently Case-marked elements scramble (which is possible in both Japanese and Russian)? A full account of these difficulties and their potential solution, however, requires a theory of inherent Case marking under minimalism that is beyond the scope of this article.

optionality associated with scrambling. Notice, however, that in solving the Last Resort problem through base generation and obligatory lowering, this analysis creates a new problem involving optionality. Optionality becomes unresolvable in a new way, for although the LF movement itself is motivated, a theory that allows more than one potential Merge position for arguments creates a new problem: optionality at Merge. Under B&T's theory languages with weak θ-roles have the *option* of base-generating arguments in either of two positions; optional *movement* may be eliminated, but *optionality* remains, being simply transferred to the base structure. Consider, for example, the numerations associated with (1a) and (1b). They are identical, on standard assumptions. Now consider the operations involved in deriving (1a) and (1b) from this numeration. Clearly, (1b) involves an additional step: the LF lowering, which is absent in (1a). Thus, (1b) should be blocked by (1a) in all cases by (global) economy considerations. B&T therefore adopt Collins's (1997) local economy view that prohibits look-ahead. However, if we are to maintain a strong theory of economy, we should consider the possibility that (1a) and (1b) differ in some way and hence do not have the same numeration, a conclusion that would also solve the initial optionality issue. The numeration underlying (1b) must contain certain information that forces movement. This is in the spirit of recent accounts of long-distance scrambling to which we return in section 5. It is also behind current accounts of object shift (Collins and Thráinsson 1996) as well as accounts of local scrambling (Miyagawa, forthcoming, Bailyn, forthcoming), to which we turn next.[26]

4 Local Scrambling

Thus far we have only examined cases of Ā-scrambling, which are the core cases covered by B&T. Indeed, B&T themselves focus on long-distance scrambling cases. For local (short) scrambling, B&T's base generation proposal does not involve lowering. Rather, they maintain base generation in IP-adjoined position in cases of local scrambling as well, *with no further movement.* This is consistent with the fact that local scrambling appears not to reconstruct, changes scope relations, and so on. To handle θ-checking in these cases, B&T propose that these arguments check θ-features in situ at LF, meaning they never occupy θ-position. In particular, B&T "hypothesize that when moved to I, a verb may θ-mark its object in the IP-adjoined position, allowing it to stay there at LF" (p. 360). This leads B&T into certain problems regarding apparent reconstruction effects with Principle C, which they resolve by redefining the segment created by the initial IP-adjunction.

I will not take a strong stand on this account of local scrambling, primarily because the advantage of claiming base generation in such cases rests on the lowering account of Ā-scrambling, which we have seen to be problematic. However, it is worth pointing out that many of the difficulties associated with the lowering account of Ā-scrambling remain for these cases as well.

[26] Collins (1997) considers a similar issue in the derivation of locative inversion constructions in English, concluding that Economy is a local, not global, condition. So long as Last Resort and Minimality are observed, Economy is not violated. However, since B&T's theory for the first time introduces optionality with regard to where arguments are base-generated in Japanese, the optionality problem (choosing IP-adjoined position over θ-position) remains.

First, there is the issue of rigidity of interpretation, in this case associated with the higher position, since the argument is claimed never to be in θ-position. Indeed, Brown (2000) gives evidence from Russian for reconstruction of A-chains, which provides a counterexample to B&T's account of local scrambling. Again, "sensitive" reconstruction appears to be empirically closer to the mark. Second, it is predicted again that (local) scrambling should also be limited to θ-marked elements, a prediction that appears to be too strong for Russian, where adverbials, adjectives, and other nonarguments can scramble locally, as shown in (23).

(23) a. *Bystro*$_i$ Ivan t_i čitaet knigu.
quickly Ivan reads book
'Quickly Ivan reads the book.'
b. [*O Nabokove*]$_i$ Ivan včera kupil knigu t_i.
about Nabokov Ivan yesterday bought book
'About Nabokov Ivan bought a book yesterday.'
c. Ivan *xorošuju*$_i$ kupil [t_i knigu].
Ivan good-ACC bought book-ACC
'Ivan bought a good book.'

Third, there is evidence from both Japanese and Russian that local scrambling is not distinct from movement satisfying the EPP. Thus, both Miyagawa (forthcoming) and I (Bailyn forthcoming) analyze local argument scrambling as EPP-driven movement applying to constituents other than the canonical subject. Fourth, it is unclear how Case checking can work, if locally scrambled arguments do not lower even at LF. How, in that case, would a locally scrambled, say, direct object be associated with the appropriate position where its Case is checked? Fifth, the same scope and binding difficulties discussed above will apply to local scrambling cases as well. Sixth, under a base generation account there again appears to be no reason not to left-adjoin *all* internal arguments rather than (sometimes) base-generating them in actual θ-position, since the mechanism for θ-role assignment to the higher site is available. Thus, we can see that the base generation analysis for local scrambling will also require further work before it can be generally accepted.

5 An Alternative Account of Apparent Optionality

5.1 Possible Approaches

Let us return again to the optionality involved in (1a) and (1b). Logical possibilities for analyzing such apparent optionality are given in (24).

(24) *Possible accounts of (1a) and (1b)*
a. (1a) and (1b) are not derivationally related.
i. Some languages allow nonconfigurational structures. (nonconfigurationality: Hale 1983)
ii. Obligatory lowering accounts for (1b). (B&T 1998)
b. (1a) is derived from (1b). (requires overt lowering)

c. (1b) is derived from (1a).
   i. The raising is truly optional, for independent reasons. (Saito and Fukui 1998)
   ii. The raising is not optional; instead, it is related to another part of the grammar. (Zubizarreta 1998; this article)

(24ai) is assumed to be false in much work on languages like Russian and Japanese and is also incompatible with constraints on scrambling. (24aii) is B&T's approach. (24b) would amount to a B&T-style account of English as well, under which an object in canonical object position is overtly lowered (forced by strong θ-features). However, this would require loosening constraints on overt lowering, with obvious empirical difficulties. (24ci) is Saito and Fukui's (1998) account whereby some movement is "truly optional," in the sense that it is not driven by feature checking, but is costless depending on the directionality of phrase structure, so that a left-headed language allows optional movement to the right, whereas only right-headed languages allow left-(IP)-adjunction scrambling. However, this account is clearly too narrow to accommodate right-branching languages with left-adjunction scrambling, such as Russian or Serbo-Croatian. Stjepanović (1998) demonstrates convincingly that this approach will simply not allow enough variation to account for such languages. Only (24cii), a (nonoptional) raising account, remains. Such an approach would account for the parallel behavior noted above. It would also solve many of the conceptual problems discussed earlier. The central question becomes, then, what motivates the movement that raises a constituent out of its position at Merge to a distinct surface location? Such movement does not take place for Case reasons, or for θ-reasons. There is no relevant operator-variable chain, and therefore the typical feature-driven accounts will not hold. It is to this issue that I now turn.

5.2 Behind the Optionality of Scrambling

In this section I indicate a possible direction for solving the paradox of apparent optionality of scrambling while still allowing a strong derivational approach to alternative word orders. Following Zubizarreta (1998), I assume that scrambling satisfies Last Resort in being forced by a mismatch between the discourse structure (her Assertion Structure) of the utterance and its intonational structure. Thus, for Zubizarreta, Romance subject postposing is a case of "P(rosodic)-movement" without which the eventual Assertion Structure would be underivable by the given word order and intonation. This is also behind Miyagawa's (1997) proposal that Ā-scrambling "is motivated by something like focus" (p. 21). I claim that Ā-scrambling is related to the topic-focus or information structure component of the grammar (see Chomsky 1971, Jackendoff 1973, Rochemont 1980, Vallduví 1992, Lambrecht 1994, Zubizarreta 1998, among many others, for similar ideas).

It is well known that word order variants in Russian are associated with distinct discourse interpretations. The usual description, given in Švedova 1980, says that "word order can vary, but at the same time *it is not free:* the meaning of a sentence, its *communicative goal,* differs with different word orders" (p. 191, emphasis mine). The standard account goes as follows: Varying the word order affects the "functional perspective" of the sentence, with a systematic

relationship holding between rheme and theme. A typical contrast is the one shown in (25)–(26) for Russian.

(25) a. Kto čitaet knigu?
who reads book-ACC
'Who is reading the book?'
b. Knigu čitaet Ivan.
book-ACC reads Ivan-NOM
'IVAN is reading the book.'

(26) a. Čto delaet Ivan?
what does Ivan-NOM
'What is Ivan doing?'
b. Ivan čitaet knigu.
Ivan reads book
'Ivan is reading A BOOK.'

(25a) and (26a) are questions that determine appropriate discourse relations in the answer (see Kovtunova 1976 for discussion of this question test). In particular, (25a) asks about the reader of the book. In the answer, therefore, *Ivan* takes final position (usually allotted to the rheme). In (26), however, the situation is reversed and *the book* is the rheme, hence its position at the end of the sentence. Returning to (1a–b), repeated here, we can see that similar factors are at play in Japanese as well.

(1) a. John-ga [Mary-ga *sono hon-o* katta to] omotteiru.
John-NOM Mary-NOM that book-ACC bought that thinks
'John thinks that Mary bought the book.'
b. *Sono hon-o* John-ga [Mary-ga *t* katta to] omotteiru.

(1b) and (1a) differ in discourse structure in the expected way. In particular, (1b) (but not (1a)) is appropriate in contexts where *the book* is part of the preceding discourse or in the ''shared common concern'' in the sense of Yokoyama 1986. This is shown in (27)–(28).

(27) a. John-wa dou shiteiru no?
John-TOP how doing Q
'How is John doing?'
b. John-ga [Mary-ga *sono hon-o* katta to] omotteiru.
John-NOM Mary-NOM that book-ACC bought that thinks
'John thinks that Mary bought the book.'
c. #*Sono hon-o* John-ga [Mary-ga *t* katta to] omotteiru.

(28) a. Sono hon ni-kanshite nani-ka atta no?
that book about something happened Q
'Did anything happen to that book?'

b. #John-ga [Mary-ga *sono hon-o* katta to] omotteiru.
John-NOM Mary-NOM that book-ACC bought that thinks
'John thinks that Mary bought the book.'
c. *Sono hon-o* John-ga [Mary-ga *t* katta to] omotteiru.

(27) shows that a question introducing *John* as the theme of the discourse elicits (27b) as a response, and not (27c). Conversely, a question such as (28a) that introduces *the book* as thematic elicits (28c) as a response, and not (28b). Thus, the scrambled and nonscrambled orders differ regarding the discourse status of the utterance.[27] Minimalist assumptions force us to take such differences seriously, since they bear directly on the linguistic encoding of the sentence, which only reflects factors relevant to the interfaces. The following generalization regarding long-distance scrambling is a strong claim that I hope will provoke further discussion:

(29) *The scrambling generalization*
a. Ā-scrambled and nonscrambled orders are *always* associated with different discourse/informational interpretations.[28]
b. The movement deriving scrambled orders is *motivated* by discourse/informational considerations.[29]

(29a) is in keeping with traditional discourse grammars. (29b), providing a motivation for long-distance scrambling, is in full keeping with the spirit of B&T's account. Thus, under (29) long-distance scrambling is not optional. However, (29) associates its obligatoriness with discourse notions, not with θ-relations. (29) captures the discourse effects mentioned above, and maintains a raising account.[30] It also avoids the theoretical difficulties discussed in section 3.

Preliminary evidence for (29a) involves the availability of so-called "functional ambiguity," whereby more than one discourse structure is available given a certain word order. As shown in Lavine 1999, in Russian only canonical SVO orders (or orders derived with A-movement alone) allow functional ambiguity. Thus, the discourse structure of an SVO clause such as (30a) could be any of the structures given in (30b–d).

[27] Of course this distinction is nothing close to a new discovery in linguistics. It has a long tradition in Slavic linguistics, for example, going back at least to Mathesius 1939, appearing throughout Soviet and Prague School linguistics, and being picked up in various forms in many branches of pregenerative linguistics and functional approaches to syntax (see, among many others, Adamec 1966, Babby 1980, Prince 1984, Yokoyama 1986, Kuno 1987, Hajičová and Sgall 1987). Its integration with current understanding of derivational violations appears therefore to represent a step forward in our understanding of human language.

[28] (29a) is limited to Ā-scrambling because A-scrambling appears to be less discourse-related (if it is discourse-related at all). On the other hand, A-scrambling presents less of an optionality problem because of its possible association with formal features (such as the EPP). See Miyagawa, forthcoming, and Bailyn, forthcoming, for discussion.

[29] The question arises here whether or not discourse-related movement is driven by formal features of Topic and Focus, or whether other factors are at play. This issue is beyond the scope of this article; but for the optionality problem to be solved, the numerations underlying (1a) and (1b) must be distinct, presumably differing by discourse-related information in one form or another. This issue is also left open in Zubizarreta 1998.

[30] Of course, some discourse effects, such as those resulting from left-dislocation, are indeed base-generated (see footnote 20). Others clearly are not. The claim here is unidirectional—namely, that Ā-scrambling entails a discourse structure distinct from nonscrambled orders, but not that all such distinct discourse structures are derived by Ā-scrambling.

(30) a. čto Ivan čitaet knigu
that Ivan reads book
'that Ivan reads/is reading a book'
b. Ivan čitaet [knigu] (subject + verb = topic; NP object = comment)
c. Ivan [čitaet knigu] (subject = topic; VP = comment)
d. [Ivan čitaet knigu] (no topic; all comment)

However, once long-distance scrambling has occurred, discourse structure is no longer ambiguous.[31] This is shown in (31).

(31) a. čto $knigu_i$ [Ivan čitaet t_i]
that book Ivan reads
'that the book, Ivan reads'
b. [knigu] Ivan čitaet (scrambled object = topic; subject + V = comment)
c. [knigu Ivan] čitaet ???(scrambled object + subject = topic; verb = comment)
d. [knigu Ivan čitaet] *(no topic; all comment)

How might a theory such as B&T's represent discourse relations? Given that the scrambled element is associated with the higher position only before LF lowering, discourse relations would have to be represented at the level of the initial phrase marker—that is, at Merge (if such a level could be isolated, contra general assumptions of minimalism). Relations established at this level would be "undone" by obligatory lowering and thus would not be represented at any unified interface level. This difficulty serves as an additional theoretical argument against B&T's approach, on the assumption that Zubizarreta (1998), Heycock and Kroch (1999), and others are right in associating discourse relations with the interpretive component of the grammar.

6 Conclusion

We have seen that B&T's claim that scrambling involves base generation followed by obligatory LF lowering creates more problems than it solves, empirically and theoretically. Their claim is both too weak and too strong. It is too weak in that it does not allow for constraints on scrambling and does not account for the parallel behavior of *wh*-movement and scrambling, whereas a raising account does. On the other hand, it is too strong in terms of the nature of the LF process involved, which, if morphologically driven, cannot be sensitive to semantic considerations. And yet scrambling in fact seems to pattern with *wh*-movement (and topicalization) in undergoing (semantically sensitive) reconstruction, an unexpected result for B&T's approach. Their claim also leads us to expect more scrambling than is actually observed and introduces new issues of optionality and economy. Base generation also does not allow for uniform representation of discourse relations at an interface level.

[31] Intonation also plays an important role in determining discourse status in spoken language. However, the relation between intonation and discourse structure is a complicated issue that falls mostly outside the scope of this article. For this example, I assume neutral intonation that follows some version of the Nuclear Stress Rule. For discussion, see Bryzgunova 1971, 1981, Rochemont 1980, Selkirk 1984, and Zubizarreta 1998.

References

Adamec, Přemysl. 1966. *Porjadok slov v sovremennom russkom jazyke.* Prague: Academia.

Avrutin, Sergei, and Maria Babyonyshev. 1994. Obviation in subjunctive clauses and AGR: Evidence from Russian. In *Formal Approaches to Slavic Linguistics: The Ann Arbor Meeting 1992,* 1–24. Michigan Slavic Publications, University of Michigan, Ann Arbor.

Babby, Leonard H. 1980. *Existential sentences and negation in Russian.* Ann Arbor, Mich.: Karoma Publishers.

Bailyn, John Frederick. 1995. A configurational approach to Russian "free" word order. Doctoral dissertation, Cornell University, Ithaca, N.Y.

Bailyn, John Frederick. Forthcoming. Generalized inversion. *Natural Language & Linguistic Theory.*

Bayer, Josef, and Jaklin Kornfilt. 1994. Against scrambling as an instance of Move-alpha. In *Studies on scrambling: Movement and non-movement approaches to free word-order phenomena,* ed. Norbert Corver and Henk van Riemsdijk, 17–60. Berlin: Mouton de Gruyter.

Boeckx, Cedric, and Sandra Stjepanović. 2000. The *wh*/clitic-connection in Slavic: Theoretical implications. In *Formal Approaches to Slavic Linguistics: The Philadelphia Meeting 1999,* 22–40. Michigan Slavic Publications, University of Michigan, Ann Arbor.

Bošković, Željko. 1997. Fronting *wh*-phrases in Serbo-Croatian. In *Formal Approaches to Slavic Linguistics: The Indiana Meeting 1996,* 86–107. Michigan Slavic Publications, University of Michigan, Ann Arbor.

Bošković, Željko. 1998. *Wh*-phrases and *wh*-movement in Slavic. Position paper presented at the Comparative Slavic Morphosyntax Workshop, Bloomington, Ind. Available at: http://www.indiana.edu/~slavconf/linguistics/download.html.

Bošković, Željko, and Daiko Takahashi. 1998. Scrambling and Last Resort. *Linguistic Inquiry* 29:347–366.

Brown, Sue. 2000. Evidence for A-chain reconstruction. Poster presented at the 9th Annual Workshop on Formal Approaches to Slavic Linguistics, Indiana University, Bloomington, February 2000.

Bryzgunova, Elena. 1971. O smyslorazličitel'nyx vozmožnostjax russkoj intonacii. *Voprosy jazykoznanija* 4:42–51.

Bryzgunova, Elena. 1981. *Zvuki i intonacija russkoj reči.* Moscow: Russkij Jazyk.

Chomsky, Noam. 1971. Deep Structure, Surface Structure, and semantic interpretation. In *Semantics,* ed. Danny Steinberg and Leon Jakobovits, 183–216. Cambridge: Cambridge University Press.

Chomsky, Noam. 1976. Conditions on rules of grammar. *Linguistic Analysis* 2:303–351.

Chomsky, Noam. 1981. *Lectures on government and binding.* Dordrecht: Foris.

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

Collins, Chris. 1997. *Local economy.* Cambridge, Mass.: MIT Press.

Collins, Chris, and Höskuldur Thráinsson. 1996. VP-internal structure and object shift in Icelandic. *Linguistic Inquiry* 27:391–444.

Comrie, Bernard. 1973. Clause structure and movement constraints in Russian. In *You take the high node and I'll take the low node: Papers from the Ninth Regional Meeting, Chicago Linguistic Society,* 291–304. Chicago Linguistic Society, University of Chicago, Chicago, Ill.

Diesing, Molly. 1992. *Indefinites.* Cambridge, Mass.: MIT Press.

Diesing, Molly, and Eloise Jelinek. 1996. Distributing arguments. *Natural Language Semantics* 2:1–54.

Epstein, Samuel David, Erich M. Groat, Ruriko Kawashima, and Hisatsugu Kitahara. 1998. *A derivational approach to syntactic relations.* New York: Oxford University Press.

Erteschik-Shir, Nomi. 1997. *The dynamics of focus structure.* Cambridge: Cambridge University Press.

Fowler, George. 1987. The grammatical relevance of theme/rheme partition. In *CLS 23,* 93–104. Chicago Linguistic Society, University of Chicago, Chicago, Ill.

Hajičová, Eva, and Petr Sgall. 1987. The Ordering Principle. *Journal of Pragmatics* 11:435–454.

Hale, Ken. 1983. Warlpiri and the grammar of non-configurational languages. *Natural Language & Linguistic Theory* 1:5–47.

Heycock, Caroline. 1995. Asymmetries in reconstruction. *Linguistic Inquiry* 26:547–570.

Heycock, Caroline, and Anthony Kroch. 1999. Pseudocleft connectedness: Implications for the LF interface level. *Linguistic Inquiry* 30:365–397.

Huang, C.-T. James. 1993. Reconstruction and the VP-Internal Subject Hypothesis. *Linguistic Inquiry* 24: 103–138.

Jackendoff, Ray. 1973. *Semantic interpretation in generative grammar.* Cambridge, Mass.: MIT Press.

Junghanns, Uwe, and Gerhilde Zybatow. 1995. Syntax and information structure of Russian clauses. Paper presented at Formal Approaches to Slavic Linguistics 4, Cornell University, Ithaca, N.Y.

Kovtunova, Irina. 1976. *Porjadok slov i aktual'noe členenie predloženija.* Moscow: Prosvečšenie.

Kuno, Susumu. 1987. *Functional syntax: Anaphora, discourse, and empathy.* Chicago: University of Chicago Press.

Lambrecht, Knut. 1994. *Information structure and sentence form.* Cambridge: Cambridge University Press.

Lavine, James. 1999. Subject properties and ergativity in North Russian and Lithuanian. In *Formal Approaches to Slavic Linguistics: The Seattle Meeting 1998,* 307–328. Michigan Slavic Publications, University of Michigan, Ann Arbor.

Lebeaux, David. 1988. Language acquisition and the form of grammar. Doctoral dissertation, University of Massachusetts, Amherst.

Lee, Y.-S., and Beatrice Santorini. 1994. Towards resolving Webelhuth's paradox: Evidence from German and Korean. In *Studies on scrambling: Movement and non-movement approaches to free word-order phenomena,* ed. Norbert Corver and Henk van Riemsdijk, 257–300. Berlin: Mouton de Gruyter.

Mahajan, Anoop. 1990. The A/A′ distinction and movement theory. Doctoral dissertation, MIT, Cambridge, Mass.

Mathesius, Vilém. 1939. O tak zvaném aktualním členění věty. In *Slovo a slovesnost: Ročník V,* ed. Vilém Mathesius and Bohuslav Havránek, 171–174. Prague: Melantricha A.S.

Miyagawa, Shigeru. 1997. Against optional scrambling. *Linguistic Inquiry* 28:1–26.

Miyagawa, Shigeru. Forthcoming. EPP, scrambling, and *wh* in situ. In *Scrambling and word order,* ed. Simin Karimi. Oxford: Blackwell.

Müller, Gereon, and Wolfgang Sternefeld. 1993. Improper movement and unambiguous binding. *Linguistic Inquiry* 24:461–509.

Neeleman, Ad. 1994. Scrambling as a D-Structure phenomenon. In *Studies on scrambling: Movement and non-movement approaches to free word-order phenomena,* ed. Norbert Corver and Henk van Riemsdijk, 387–430. Berlin: Mouton de Gruyter.

Prince, Ellen. 1984. Topicalization and left-dislocation: A functional analysis. In *Discourses in reading and linguistics,* ed. Sheila J. White and Virginia Teller, 213–225. Annals of the NY Academy of Sciences 443. New York: NY Academy of Sciences.

Rizzi, Luigi. 1990. *Relativized Minimality.* Cambridge, Mass.: MIT Press.

Rochemont, Michael. 1980. *Focus in generative grammar.* Philadelphia: John Benjamins.

Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT, Cambridge, Mass.

Safir, Ken. 1999. Vehicle change and reconstruction in Ā-chains. *Linguistic Inquiry* 30:587–620.

Saito, Mamoru. 1989. Scrambling as semantically vacuous Ā-movement. In *Alternative conceptions of phrase structure,* ed. Mark Baltin and Anthony Kroch, 182–200. Chicago: University of Chicago Press.

Saito, Mamoru. 1992. Long distance scrambling in Japanese. *Journal of East Asian Linguistics* 1:69–119.

Saito, Mamoru, and Naoki Fukui. 1998. Order in phrase structure and movement. *Linguistic Inquiry* 29: 439–474.

Selkirk, Elisabeth. 1984. *Phonology and syntax: The relation between sound and structure.* Cambridge, Mass.: MIT Press.

Stepanov, Artur. 1998. On *wh*-fronting in Russian. In *NELS 28,* 453–467. GLSA, University of Massachusetts, Amherst.
Stjepanović, Sandra. 1998. Scrambling in Serbo-Croatian. Ms., University of Connecticut, Storrs.
Strakhov, Natalia. 2001. A scrambling analysis of *wh*-questions in Russian. In *Formal Approaches to Slavic Linguistics: The Bloomington Meeting 2000.* Michigan Slavic Publications, University of Michigan, Ann Arbor.
Švedova, Natalia, ed. 1980. *Russkaja grammatika.* Vol. 2. Moscow: Nauka.
Vallduví, Enric. 1992. *The informational component.* New York: Garland.
Webelhuth, Gert. 1989. Syntactic saturation phenomena and the modern Germanic languages. Doctoral dissertation, University of Massachusetts, Amherst.
Yadroff, Mikhail. 1992. The syntactic properties of adjunction (scrambling in Russian). Ms., Indiana University, Bloomington.
Yokoyama, Olga. 1986. *Discourse and word order.* Philadelphia: John Benjamins.
Zubizarreta, Maria Luisa. 1998. *Prosody, focus, and word order.* Cambridge, Mass.: MIT Press.

Department of Linguistics
Social and Behavioral Sciences
SUNY at Stony Brook
Stony Brook, New York 11794-4376
jbailyn@notes.cc.sunysb.edu

Zero Derivations

Luigi Burzio

Rubach (2000) proposes a modified version of Optimality Theory (OT) that features derivations. While Prince and Smolensky's (1993) original formulation requires some modification, I argue here that, rather than reintroducing derivations, the correct approach is to take fuller advantage of OT's inherent parallelism. I propose that outputs must be related not only to inputs, but to other, ''neighboring'' representations as well—a feature that is shared by both the output-to-output faithfulness approach and the theory of targeted constraints developed by Wilson (2000, to appear). I show that all the cases cited by Rubach that seem to support derivations are in fact handled by the latter two related theories, and that both of these have significant advantages over derivations.

Keywords: derivations, opacity, output-to-output faithfulness, targeted constraints

1 Introduction

Rubach (2000) proposes a version of Optimality Theory (OT) that features derivational levels or stages: Derivational OT (DOT), a development also advocated by Kiparsky (1998). In this article

I am grateful to Bob Ladd, Colin Wilson, and two anonymous reviewers for insightful comments and suggestions.

Linguistic Inquiry, Volume 32, Number 4, Fall 2001
658–677