Against the Scrambling anti-Movement Movement*

John Frederick Bailyn
Stony Brook University

1 Introduction: Free Word Order and Scrambling

This article argues against two recent non-movement accounts of free word order in Russian—van Gelderen (2003) (hereafter VG) and Bošković (2004) (hereafter B) and in favor of Scrambling-as-Movement. Both VG and B claim (a) that the (re)ordering of major constituents in Russian results from a process that is not movement, and (b) that (most) Russian (re)orderings result from a process distinct from that of Japanese. In sections 2 and 3 of this article I present the VG and B approaches and argue against them. In the final section, I argue that Japanese and Russian do not differ in the manner described by either author, and that a unified discourse-driven account of Scrambling as Last Resort movement is both theoretically more desirable and empirically more successful.

The issue at hand is the proper account of alternative word orders for identical major constituents in so-called “free” word order languages. Typical cases are given in (1) (local) and (2) (long-distance) for Japanese and (3) for Russian (the bold element is the constituent separated from canonical (thematic) position):

(1) a. Mary-ga sono hon-o yonda    (Japanese)
    MaryNom that bookAcc read
    ‘Mary read that book.’

    b. sono hon-o Mary-ga ___ yonda
       that bookAcc MaryNom ___ read
       ‘That book Mary read ___.’

(2) sono hon-o John-ga [ Mary-ga ___ katta to] itta]
    that book JohnNom MaryNom bought that thinks
    ‘That book John thinks that Mary bought ___.’

(3) a. Mal'čiki čitajut knigi.      SVO
    boysNom read booksAcc

    b. Mal'čiki knigi čitajut __.     SOV

---

*Ideas in this article have been presented at Utrecht, Tilburg and Yale Universities, and in seminars in St. Petersburg, Moscow and Novi Sad. Thanks to those audiences for discussion, as well as to Andrei Antonenko, Boban Arsić, Hans Broekhuis, Jim Lavine, Nataša Milečević, Andrew Nevins, Øysten Nilsen, Henk van Riemsdijk and two FASL reviewers and editors. All mistakes remain my own.
Standard accounts since Ross (1967), especially Saito (1989, 1992), have assumed or argued for a “scrambling” transformation, which derives (2) as shown in (4):\(^1\)

\[(4) \text{sono hon-o John-ga [Mary-ga katta to] itta} \]
\[\text{that book John NOM bought that thinks} \]
\[\text{movement (Scrambling)}\]
\[\text{Reconstruction (if A'-mvt)}\]

Motivation for a movement account of scrambling is given in Saito (1989), based on the contrast between Japanese (5a) and (5b):

\[(5) \text{a. [sono hon-o, [John-ga [CP [Mary-ga e1 katta to]2
\[\text{that book John NOM bought that}
\[\text{[Bill-ga e2 itta] to] omotteiru]}\]
\[\text{Bill NOM said that think}
\[\text{‘That book, John thinks that [that Mary bought e1]2, Bill said e2.’} \]
\[\text{b. *[[Mary-ga e1 katta to]2 [John-ga [CP sono hon-o,
\[\text{Mary NOM bought that John NOM that book-}
\[\text{[TP [Bill-ga e2 itta] to] omotteiru]}\]
\[\text{Bill NOM said that think}
\[\text{‘[that Mary bought e1]2, John [that book], thinks that Bill said e2.’} \]

In (5a), CP2 is scrambled out of an embedded clause, and NP1 is then scrambled out of CP2. All moved elements c-command their traces, and the derivation is fine. In (5b), however, NP1 is moved first, followed by its containing CP2. The resulting structure violates the Proper Binding Condition, because within CP2, the contained NP trace e is not c-commanded by its antecedent, sono hon-o, now stranded. Thus, “traces created by scrambling and those created by WH-movement in English behave in exactly the same way with respect to the Proper Binding Condition (PBC)” (Saito 1989, p. 190)

Furthermore, Saito (1989, 1992), Webelhuth (1989), Mahajan (1990), and Bailyn (1995, 2001, 2003) show with a number of syntactic tests for Japanese, German, Hindi and Russian respectively, that standard WH-movement constraints also apply to the derivation of free word order, implicating movement. Such constraints include the Coordinate Structure Constraint, Subjacency, Weak Crossover, the Adjunct Condition and others. Finally, Saito (1989) shows that Scrambling licenses Parasitic Gaps in Japanese. We therefore approach the problem from the perspective that theories advocating non-movement should at very least maintain the level of descriptive

---

\(^1\) In a Copy Theory, Reconstruction involves pronouncing the higher copy and interpreting the lower one.
more recently, however, the Minimalist Program of Chomsky (1995) and later work has compelled us to ask questions of motivation about any movement transformation posited: “Is the movement in question (syntactically) obligatory?”, “Is the movement driven by interface (in this case interpretive) considerations?”, “Do features drive the movement, and if so, which?”
In the case of Scrambling, preliminary answers to these questions throw some doubt on the movement account. First of all, Scrambling appears always to be optional (that is, we are never forced to derive (1b) from (1a).) Second, Scrambling appears to be semantically inert in that some Logical Form (LF) relations, such as Quantifier Scope, are not affected by its operation in Japanese (Saito’s “radical reconstruction” property), meaning it may have no interface relevance (and hence should be superfluous, on minimalist assumptions). Finally, it is not clear what features might drive such movement. For these reasons, various alternatives to movement have been recently proposed, two of which I discuss, and ultimately reject, in what follows.

2 “Early Spell-Out”
Van Gelderen (2003) argues that Minimalism allows for the possibility that there exist “Early Spell-Out” languages (such as Russian) in which major constituents in a derivation can essentially move directly from the Numeration to Spell-Out, without passing through any syntactic component. This is possible in Russian because case is internally licensed, and assuming that case is a purely PF phenomenon, the syntactic component can be sidestepped, deriving the effect of free word order. In such cases “Scrambling is the result of the lack of merger, meaning that constituents arrive to PF unattached, which allows great freedom of linear order... This is what occurs in languages such as Russian” (p. 7). In short, “certain languages have ways of checking features that do not require Merge to occur” (p. 12). This is what allows all 6 of the possible constituent orders found in Russian (cf. (3)). Note that crucially for VG, nominals do undergo syntactic formation (to the level of DP) after which the predicate and its arguments are arranged at PF according to discourse principles without any further syntactic processes taking place. Languages of this kind are predicted by VG to have the following properties:

---

2It is not clear to me what is meant by “internal licensing”, that is, what morphological property of a language allows it to license case internally. Examples of such languages other than Russian are not presented in VG, so I will limit the discussion of this issue. Japanese, crucially, does not have this property, and is not an early spell-out language for VG.

3This simplifies VG’s particular claim for the six Russian constituent orders shown in (3). In fact, VG claims Early Spell-Out accounts for only three of the six orders, namely VSO, OSV and VOS, the so-called “unmerged” structures. The other three
(6) Properties of languages with Early Spell-Out (VG, pp. 23-25):
   i. Free Constituent Order: all word orders of major constituents are available
   ii. Islands: “every partial structure will be opaque for extraction”
   iii. Ambiguity: The relative order of two quantifiers will always be ambiguous
   iv. Adjuncts: “no difference is expected between arguments and adjuncts in Early Spell-Out structures”

2.1 Against early spell-out
First, let us consider the issue of free constituent order under VG’s system. Sub-constituents, such as argument DPs, are created by Merge in the usual way. Once V, DP₁, and DP₂ are built, then the derivation is sent off to Early Spell-Out. The result is any one of the six orders given in (3), the distinctions being determined by linear rules of discourse (information) structure. Crucially, for VG there is no process of syntactic Merge between a verb and its arguments in Russian (as opposed to Japanese, where the verb-final order results from complements merging with heads in head-final fashion.) However, in abandoning any kind of VP-internal merger, VG encounters significant problems with both selection and constituency. Let us consider each in turn.

Within Minimalism, c-selection is replaced by feature checking, as in Adger (2003), where the requirement that a preposition take an NP complement, say, is formalized as the P head bearing an uninterpretable [+uN] feature, which must be eliminated by being checked, at Merge, by a complement bearing an interpretable [N] feature, that is, by a nominal. So P must take an NP complement. Verbs that take CP complements, (indicative, interrogative, subjunctive, etc), small clause complements, and so on, are similarly marked. The featural requirements that constitute c-selection are satisfied when Merge with the appropriate category occurs.

It should be immediately apparent that the Early Spell-Out system, which expressly denies a merger process, will not as it stands be able to handle selectional relationships. Selectional restrictions cannot be captured at the level that determines linear order (PF) due to lack of adjacency. Nor are they able to be satisfied earlier in the derivation under VG. This leaves the LF component as the only possibility; and this is the level where VG assumes such relations are handled. But what exactly is the process of ‘checking’ like at LF? Is it configurational? Does it involve features? Does it require adjacency? It is generally assumed that categorical requirements cannot be fully reduced to semantics (ask the time vs. *wonder the

(SVO, SOV, and OVS) result from Merge in the usual way. Space considerations prevent me from arguing against this classification of Russian word order patterns. Instead, I will concentrate on the general approach and its empirical and theoretical weaknesses, for which the “unmerged” orders are enough to make the case.
time), so it would not appear that selection can be somehow checked at LF, despite VG’s assumption that it can. Any kind of combinatorial approach to selection is unavailable to VG, as the account bypasses syntactic combination. If selection is handled through feature checking, as in Adger 2003, the uninterpretable categorial features driving selection must be eliminated before LF. Some other device is required. Without further elaboration on how selection is to be handled, the system has weakened the grammar far more than it has strengthened it.4

The next problem for Early Spell-Out involves constituency. The Early Spell-Out system denies constituency of both VP/nP and TP. In some respects, it essentially restates Hale’s (1973) non-configurationality parameter. However, substantial evidence for VP and TP constituency exists for Russian—ellipsis and sluicing (Grebenyova, this volume), coordination, and other standard constituency tests demonstrate the necessity of the VP and TP groups. A third problem concerns the claim that embedded clauses are fully opaque to scrambling. It is generally known that extraction is possible at least from subjunctive čtoby clauses (see Bailyn 1995 and elsewhere). It is not clear how Early Spell-Out can derive surface order in these cases but not allow separation from the argument clause in indicatives. Fourth, changes in word order directly affect scope, as shown in Ionin (2001), an aspect of free word order that Early Spell-Out denies the possibility of (recall that LF relations have no connection to PF orders in this system). VG claims all double quantifier structures will be ambiguous, contrary to the well-observed fact that surface scope is highly preferred to inverse scope by most speakers in both derived word orders.5 Thus the VG system loses empirical coverage in its effort to answer some of the questions about Scrambling raised under Minimalism.6

4 Similarly, in a diathesis system, such as Babby (forthcoming), selectional properties are captured as a two-tiered representation of argument structure. Selection is thus lexical, though the system assumes a rigid order of hierarchical combination, which leads to strong empirical predictions about word order. Early Spell-Out weakens the predictive force of both a Bare Phrase Structure and a diathetic approach and would need additional machinery to account for constituency and derived word orders.

5 The same is apparently not true of Japanese (see Bošković & Takahashi 1998) among many others. This distinction is what leads Bošković (2004) to argue that Russian does in fact have overt movement in such instances, although he labels it Topic/Focus movement. We return to this proposal in Section 3 of this article. For now suffice it to say that Russian facts speak against Early Spell-Out and in favor of movement, as Bošković shows.

6 Note also that VG does not in fact eliminate Scrambling from the grammar. Japanese, where not all six constituent orders occur, must have Scrambling as movement and (left-branching) VP constituency to assure V-final structures. Thus the VG typology posits both Early Spell-Out languages (Russian) and Scrambling languages (Japanese), hence more language types, without relevant discussion of the loss of explanatory adequacy. I will not discuss this further, other than to note the problem.
3 Base-Generation and LF Lowering

Bošković (2004) builds on Bošković & Takahashi (1998)’s account (hereafter BT) of Scrambling as a base-generated process followed by obligatory LF lowering. In this system, \( \theta \)-roles are features. Languages differ as to whether \( \theta \)-relations are ‘weak’ and can be checked at LF (Japanese) or whether they are ‘strong’ and must be checked at Merge, hence no Scrambling (English). Thus Scrambled elements are pronounced in their base-generated position and then undergo obligatory lowering to their LF position in order to check their \( \theta \)-role. There is no optional movement, and radical reconstruction effects in interpretation follow automatically after lowering. The proposal is schematized in (7):

(7) \textbf{sono hon-o John-ga [Mary-ga katta to] itta]}

\textit{that book John\textsubscript{Nom} Mary\textsubscript{Nom} bought that thinks}

\textit{(base position)} \textit{LF Lowering (LF position: \( \theta \)-checking)}

3.1 Against scrambling as base-generation and lowering

Despite the theoretical desirability of eliminating the optionality of Scrambling in this way, the original BT proposal has been challenged in the literature (see Bailyn 2001, Boeckx 2003 among others). The reader is referred to those works for detailed argumentation. Two major issues, however, require some discussion, and are as follows:

A. The BT account of Scrambling predicts the absence of surface interpretive effects associated with the high (scrambled) position. Empirically, this claim appears too strong, as shown in (8-10), where we see surface scope effects and anti-reconstruction binding effects respectively:

Russian Surface Scope Effects (see also Ionin 2001)

(8) a. Kto-to xočet, čtoby Boris uvidel každago maľčika.

\textit{Someone wants that Boris saw [every boy]}

\textit{‘Someone wants Boris to see every boy.’}

i) \( \exists x \forall y \) ii) \( ^*\forall y \exists x \) (?! for some speakers)

b. [\textbf{Každago maľčika}] kto-to xočet, čtoby Boris uvidel t.

\textit{[every boy] ACC someone wants that Boris saw}

\textit{‘Every boy, someone wants Boris to see.’}

i) \( ^*\exists x \forall y \) ii) \( \forall y \exists x \)

In (8a), we see that Russian quantifiers prefer surface scope interpretations, as discussed in Ionin (2001). When an embedded

\footnote{Needless to say, this is far from an uncontroversial assumption. See Chomsky 1995 among many others for discussion. I will assume in what follows that such a characterization of thematic relations \textit{is} possible, and my critique of BT will be limited to empirical domains, rather than taking on this larger, framing assumption.}
object is scrambled, it acquires surface scope. On theories where quantifier scope is determined at LF (the standard assumption), (8b) is incompatible with Lowering, since the scrambled element should always be interpreted in its low thematic position. The same problem occurs with anti-reconstruction effects shown in (9).

Anti-Reconstruction Effects (see Heycock 1995)

(9)  a. *[How proud of John] do you think he should be?*
    b. *[Which question [that Gore got during the debate]] do you think he should be?*

(10)  a. *Ja xoču, čtoby oní srazu zabýl
        I want that he right away forget
        nekotorye voprosy Gorů
        some questions to Gore
        ‘I want him to immediately forget some questions to Gore.’
    b. *[Nekotorye voprosy [Gorů]] ja xoču,
        some questions to Gore I want
        čtoby oní srazu zabýl t
        that he right away forget
        ‘Some questions to Gore I want him to immediately forget.’

(10b) shows that for Russian scrambling, just as for English wh-movement (9b), some fronted constituents containing an R-expression obviate the Principle C violation that is incurred when they are in base position as in (10a). This is then a case of “anti-reconstruction”. Note that the R-expressions in the (b) sentences are either adjuncts or within adjuncts. This has led to the proposal that adjuncts are attached late in the derivation, never being associated with the lower position, hence the anti-reconstruction effect. However, in a system where the displaced argument itself must obligatorily lower at LF, the LF representation will not have the argument and modifier in the same location, making semantic interpretation impossible. A system like that of Heycock 1995, where the reconstructability of an element depends on its referential status, is crucially not available in the BT system.

B. The BT account specifically requires that there be no trace (or copy) in scrambled (high) position. This is necessary to make the Lowering process itself syntactically legitimate. However, this also implies the lack of any locality or other syntactic constraints on Scrambling, assuming the usual accounts of such constraints as being constraints on chains or constraints on traces. However, the predicted lack of locality and other syntactic constraints on scrambling is contradicted by literature on many free word languages (see Saito 1989, 1992, Webelhuth 1989, Mahajan 1990, Bailyn 1995, 2001, 2002 among many others). A partial list is given in (11):
(11) **Known syntactic constraints on Scrambling**

a. Proper Binding Condition (Saito 1989)
b. Subjacency (Webelhuth 1989)
c. Complex NP Constraint (Webelhuth 1989)
d. The Empty Category Principle (incl. that-t effect) (Bailyn 1995)
e. No extraction out of Russian člo-clauses (Bailyn 1995)
g. Constraint on Extraction Domains (Webelhuth 1989)
h. Constraint on extraction out of Russian adnominal genitives (Bailyn 1995)

Space considerations preclude a full presentation of syntactic effects in scrambling here. However, Russian examples of Subjacency and ECP effects are given in (12) and (13-14) respectively:

(12) a. *Kogo tı pozvonil [agentu [ kotoryj ljubit tı ]]?  
   Whom you phone spy who loves   
   ‘Whom did you phone a spy who loves?’

b. *Borisa tı pozvonil [agentu [ kotoryj ljubit tı ]]
   Boris you phone spy who loves  
   ‘It’s BORIS you phoned a spy who loves!’

(13) a. Komu ty xočes’, [ čtoby Ira pozvonila tı ]?  
   who you want that Ira phoned  
   ‘Who do you want Ira to call?’

b. *Kto ty xočes’, [ čtoby tı vljubilsja v Iru ]?
   who you want that fall in love (to) Ira  
   ‘Who do you want that fall in love with Ira?’

(14) a. Ja Borisui xotel, [ čtoby Ira pozvonila tı ]
   I Boris wanted that Ira phone  
   ‘I wanted Ira to phone Boris.’

b. *Ja Borisij xotel, [ čtoby tı vljubilsja v Iru ]
   I Boris wanted that fall in love (to) Ira  
   ‘I wanted Boris to fall in love with Ira.’

Clearly, known Scrambling is sensitive to Subjacency and ECP effects. This is not expected under BT. BT do not directly address the issue of the PBC. However, they do acknowledge the problem of known syntactic constraints on Scrambling in a footnote: "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" (BT, fn 17, p. 358). No proposals are made as to how to account for the parallel effects of those constraints on Scrambling and WH-movement in a system without Scrambling-as-Movement. An important generalization is thus lost, and empirical coverage of the resulting theory, however theoretically preferable, is sacrificed without sufficient theoretical compensation, until CSC and other scrambling effects are accounted for. The empirical coverage of GB accounts has to be maintained in non-movement accounts in some other way, not provided by BT (or VG).
3.2 Movement but not scrambling?
In Bailyn 2001 I presented these and other objections to the BT account, primarily using Russian data, showing the effects discussed above (that Scrambling has interpretive effects and that locality constraints hold), and that adjuncts can Scramble (another effect not predicted by the BT account). In reply, Bošković (2004) (B) acknowledges the importance of such examples, but argues that they do not undermine the BT approach of Base-Generation and Lowering. "Russian examples that Bailyn (2001) uses to argue against Bošković & Takahashi’s (1998) analysis of scrambling are irrelevant to the analysis because they do not in fact involve scrambling (Bošković 2004: 613, emphasis JFB). In particular, B concedes that movement is involved in Russian instances which show interpretive effects and are subject to standard movement constraints, but that these instances are not in fact Scrambling. Rather, these are cases of “Topic/Focus Movement,” which is assumed by B to be a standard syntactic movement process.

In the final section of this article I address the general issue of whether or not there is a significant distinction between the derivation of free word order variation in Japanese and Russian, as argued in both VG and B, concluding that the difference between the two languages is more superficial than claimed in those accounts. Before that discussion is possible, however, it is critical to examine the nature of B’s proposed distinction between the two language types.

B’s primary claim is this: Russian has (morphologically unmarked) Topic/Focus movement, which is assumed to be standard (upward) A’-movement. The high (scrambled) position determines interpretive effects (scope). The usual locality and other constraints apply as the movement is carried out, and adjuncts can participate.

Let us examine the motivation for claiming that the two languages are so different: B’s primary argument is that those Russian examples that do not have the radical reconstruction property can not result from Scrambling. “The undoing property is taken to be

---

8 Surface scope effects mask the unusual nature of B’s assumption about (standard) A’-movements (WH-mvt and TOP), namely that surface position alone determines interpretative effects (as vs. “Japanese-style” Scrambling). This contradicts the general assumption that A’-movement can reconstruct (for binding) (Fox 1999, a.o.) ([Which pictures of himself] does John hate t?). A’-reconstruction for binding is mentioned only in a footnote, in which B appeals to a derivational Binding Theory, following Epstein et al (1998), whereby binding relations are established in the course of the derivation (before wh-movement). Derivational approaches are indeed promising for Principle A (Grewendorf & Sabel 1999). However, Principle C appears to apply only at LF (Saito 2003), meaning that another aspect of B’s proposed distinction between standard A’-movements and Scrambling is weakened, in that anti-reconstruction facts with WH-movement and Scrambling cannot be treated in parallel fashion under the non-movement approach to Scrambling.

9 B later adds that Russian also has Japanese-style Scrambling, a claim to which I return below. For now what matters is the concession that (most) Russian free word order results from overt movement which is distinct from Lowering and thus not contraexemplary to the claims made about Japanese.
the defining and most interesting property of Japanese-style Scrambling (JSS)” (B, p. 618). The crucial difference between the two languages concerning interpretation and word order involves quantifier scope, and is illustrated in (16-17):

(16) [Daremo-ni] dareka-ga [Mary-ga t_1 atta to ] ometeiru
everyone everyone Mary met that thinks
‘Someone thinks that Mary met everyone.’

\[ \exists x \forall y \]
\[ \forall y \exists x \]

(17) [Každogo mal’čika]i kto-to xočet, čtoby Boris uvidel t_1
[every boy] ACC someone wants that Boris saw
‘Every boy someone wants Boris to see.’

\[ \forall y \exists x \]
\[ \exists x \forall y \]

In (16), the scrambled embedded object does not acquire surface scope in Japanese. In (17), on the other hand, the surface order determines scope, as we have seen. This difference is significant, and will be discussed below. Otherwise, the primary distinction seems to be related to the fact that Russian has no overt Topic/Focus devices other than word order (and intonation), whereas Japanese regularly marks Topics with *wa*. However, the overt nature of this particle does not entail that non-*wa*-marked word order variation in Japanese might not also serve a discourse function. Indeed there are significant examples of the discourse relevance of Japanese free word order (Miyagawa 1997 and Bailyn 2001). At the same time, Japanese Scrambling obeys the Proper Binding Condition, as in (5), as well as Subjacency and other movement constraints, Saito (1989, 1992, 2003). These effects are not accounted for in BT or B.10

---

10B responds to the issue of the PBC in an extended footnote, where it is claimed that PBC effects in Scrambling are irrelevant for two reasons: first, because the PBC does not hold in German remnant movement instances, where a fronted infinitive can contain the trace of a lower element as in (i):

(i) [t Gelesen] [hat das Buch keiner]
read has the book no one
‘Read the book, no one has.’

However, Van Riemsdijk (pc) points out that cases such as (i) mask the more general situation of the PBC applying exactly as expected in cases like (ii), where the presence of *was-für* split clearly implicates movement.

(ii) *[t Für Bücher gelesen] [weiss ich nicht was erhat]
for books read know I not what he has
‘I don’t know what for books he has read.’

Thus something like the PBC holds in case of German remnant movement cases, although bare infinitival constructions like (i) may involve base-generation. Cases such as (ii), as well as English WH-movement, Japanese LF-WH movement, and Japanese Scrambling can be united in this regard only by assuming movement constrained by the PBC applies.

Second, B rejects a PBC account because “it is crucial to apply the PBC at S-Structure... [and] is therefore incompatible with the Minimalist program, which has no place for S-structure conditions” (p. 617). Instead a sideways movement account of (5) is proposed. However, this account cannot extend to the other known PBC
One more aspect of B’s account requires discussion. B shows that not all instances of Russian word order variation involve Topic/Focus movement. Japanese-style Scrambling (JSS) also occurs in Russian. This is important for B because scrambling in Russian is (sometimes) able to escape WH-islands:

(18) Ty doktor; videl kogda [ip t pod’ezžal ]? you doctorNom saw when came ‘The doctor did you see when (he) came?’

(19) *Kto, ty videl kogda [ip t pod’ezžal ] (WH-island) who you saw when came ‘Who did you see when came?’ (Müller & Sternefeld 1993)

Because (18) appears to escape island constraints, it cannot for B be an instance of Topic/Focus movement, which (always) obeys islands. B therefore claims that (18) is an instance of JSS (Base-generation and Lowering), as vs. (19), which shows the effects of a WH-island violation. However, Lowering cannot be the correct account of (18). For if (18) is Japanese style-scrambling, then the Lowering account immediately predicts Japanese-style low quantifier scope in such constructions, since the undoing property always characterizes JSS on B’s assumptions. This prediction is not borne out:

(20) Ty každuju devušku; videl kogda you [every girl]Acc saw when [kakoj-to mal’čik celoval ti]? some boy-Nom kissed ‘Did you see when some boy kissed every girl?’

In (20), an embedded quantifier escapes a WH-island, but has surface scope. If the undoing property (Scope) is the primary diagnostic for JSS, then (20) must be overt movement, since it has surface scope. If escaping islands is the diagnostic, then (20) must be non-movement (JSS). B’s account has achieved a paradox.11

To sum up thus far: the claim that Russian free word order is (usually) driven by movement whereas in Japanese it is not encounters significant problems. Despite the proposed distinction, both languages obey movement constraints in Scrambling, both use
surface word order to encode discourse relations, (Miyagawa 1997, Bailyn 2001), and Scrambling, constrained in its operation like other movement processes, is the best candidate for how they are derived.

As for the nature of the contrast in (18), I follow Müller & Sternefeld (1993) in accounting for this contrast as coming from the nature of the two movements involved. In Minimalist terms, different features trigger WH-movement and Scrambling. The former cause a Relativized Minimality violation in WH-island contexts (19), the latter do not (18). Many speakers do not find (18) perfect, exactly as expected if there is a mild Subjacency violation but no Relativized Minimality violation. Thus it appears clear that Russian in fact has one mechanism for deriving free word order—Movement. In the next section we turn to the issue of whether there remains good cause to claim that Japanese is really any different.

3 On Supposed Differences between Japanese and Russian

VG and B each propose radically distinct grammars for Russian and Japanese. For VG Japanese has Scrambling (movement), but Russian has Early Spell-Out (non-Movement). Both are discourse-driven. For B, following BT, Japanese has Base-Generation and Lowering (non-movement), whereas Russian primarily has Topic-Focus Movement (directly driven by discourse). Thus both admit the need for movement in describing free word order and both acknowledge the well-known connection between free word order and discourse effects, about which there exists a significant literature (Adamec 1966, Kovtunova 1976, Yokoyama 1986 and many others). In this situation, then, it is only natural for research attention to be focused on one primary question: is it possible to limit the derivation of free word order to a single device, whose motivation is discourse-driven, and which is driven by an interpretive component of the grammar?12

We have seen that a Movement account is preferable for Russian. The next question is this: is there evidence that Japanese word order variation is also derived by movement? And of course there is the original evidence in Saito 1989, and further evidence in Kawamura 2004 and Saito 2003. Further, there is the question of surface interpretive effects. We have seen that scope does not appear to change with word order variation. But anti-reconstruction effects like those in (14) for Russian also obtain in Japanese, as shown in the contrast between (21a) and (21b):

12I assume that there is a distinct level of information structure, as proposed in many places (Rochemont 1980, Vallduví 1992, Lambrecht 1994, Bailyn 1995, Zubizarreta 1998 a.o.). However, nothing here in making the case for movement requires that discourse relations be an independent linguistic level. The question of how discourse relations are encoded does not bear on the issue of whether syntactic movement is (always) involved, for which the empirical evidence presented here remains the strongest argument.
(21) a. *[kare-wa [Mary-ga [John-ga tukutta jodan-o],
   tukatta to] omotteiru].
   used C thinks
   ‘He thinks that Mary used the joke that John made.’

b. [John-ga tukutta jodan-o], [kare-wa [Mary-ga
   tukatta to] omotteiru].
   used C thinks
   ‘The joke that John made, he thinks that Mary used.’

(21a) shows a Principle C effect, with an R-expression inside an
embedded object bound by a higher pronoun. (21b) shows that
Scrambling of the object bleeds the effect, that is, it does not behave
as if it radically reconstructs, contrary to what the Lowering account
predicts. The general picture is thus emerging that Japanese too uses
overt movement to derive free word order.

Of course there remain two outstanding questions concerning the
claim that Japanese and Russian both have discourse-driven overt
Scrambling, subject to essentially the same constraints. First, why do
the two languages differ in the interaction of scope and word order, if
not by movement vs. non-movement? Second, why can adjuncts not
scramble in Japanese?

Let us take the second question first. I assume that the restriction
in question is not on non-arguments per se, but rather on adjuncts in
particular in Japanese. This is confirmed by the fact that adjuncts also
are not acceptable in wa topic constructions, which are certainly not
theta-driven. As for scope, Japanese appears to have an scope
principle distinct from other languages in that it is interpreted without
regard to movement (Hungarian is a well-known example of the
opposite, see Kiss 1986). This is of course true for Japanese WH
phrases, which can also be scrambled without scope changes (the
WH-Q effect). Let us call this the Scope Locality Effect:

(22) The Scope Locality Effect (Japanese):
   A quantifier must be interpreted in its local argument domain

With (22) in place, we can maintain a strong derivational system of
free word order in which overt A’-movement, driven by discourse-
considerations, derives alternative word orders in all languages.
Reconstruction applies in the usual way (interpreting of a lower copy
where relevant). Scope differences derive from (22), and we are left
with two kinds of languages: Scrambling languages (Japanese and
Russian) and non-Scrambling languages (English), and one kind of
Scrambling: movement.
References


Grebenyova, Lydia. this volume. Sluicing puzzles in Russian. 156-170.

É. Kiss, Katalin. 1986. The order and scope of operators. In Topic, Focus, and Configurationality, eds. Werner Abraham, and Sjaak...
de Meij. Amsterdam: John Benjamins.

jbailyn@notes.cc.sunysb.edu