

# **Japanese Style Scrambling Russian**

## **Myth and reality**

Zhanna Glushan

MPhil-avhandling  
Det humanistiske fakultet  
Universitetet i Tromsø  
Mai 2006

## **Acknowledgements**

First of all, I would like to thank my supervisor Klaus Abels for his invaluable help, insightful comments, interesting discussions, his patience and numerous hours that he spent reading through my thesis.

I am also grateful to all the teachers and students at the Department of Linguistics in Tromsø for creating a perfect atmosphere for studying linguistics and making these two years so interesting and rewarding for me.

Also I would like to thank my informants for their time and helpful comments :Gillian Ramchand, Bella Crespin, Kaori Takamine, Naoyuki Yamato, Takahiro Tamura, Aysa Arylova, Ženya Markovskaja, Ženya Romanova.

Special thanks goes to Ad Neeleman for giving me valuable suggestions for my work.

Finally, I am very thankful to my family for their love, optimism, patience and faith in me.

# Contents

## 1 Introduction

## 2 Types of scrambling constructions and properties

2.1 Scrambling constructions in Japanese.....	3
2.1.1 Configurationality parameter.....	3
2.1.2 Scrambling in the 80's. A/A'-movement discussions.....	4
2.1.3 The discovery of the undoing property. Saito (1989),(1992).....	8
2.1.4 Copy and Deletion analysis of scrambling. Saito (2001),(2003).....	10
2.1.5 Principle of Unambigious Binding. Müller and Sternefeld (1996).....	14
2.2 Scrambling constructions in Russian.....	17
2.2.1 Classics: theme –rtheme distinction.....	17
2.2.2 Discourse-driven accounts. Junghans and Zybatow (1997).....	19
2.2.3 Mysterious asymmetries of scrambling and wh-movement. Müller and Sternefeld (1993).....	26

## 3 Movement or base generation?

3.1 The LF-lowering analysis of scrambling.....	30
3.1.1 Why lowering?.....	31
3.1.2 Relativized Minimality and LF-lowering.....	33
3.2 Movement approach to scrambling.....	35
3.2.1 Derivational schema of scrambling behaviour. Bailyn (2002).....	37

## 4 The debate about scrambling

4.1.1 Theoretical issues.....	39
4.1.2 Predictions and data.....	42
4.1.3 Against semantically vacuous scrambling in Japanese.....	45
4.2 Elusive scrambling and lack of evidence.....	47

4.2.1	The problems with Bailyn's account.....	47
4.2.2.	The problems with Bošković's (2004) account.....	49
<b>5</b>	<b>Long distance scrambling in Russian. New properties revealed</b>	
5.1	Japanese style scrambling. Radical reconstruction?.....	52
5.2	Semantic effects.....	52
5.3	Ambiguity of 'kak' and two possible interpretations.....	55
5.4	Relativized minimality.....	61
5.5	Reconstruction and binding.....	65
<b>6</b>	<b>More on properties of long distance scrambling in Russian. English parallels</b>	
6.1	Locality.....	72
6.2	Heavy-NP shift and scrambling.....	80
6.3	Long wh-extraction and scrambling.....	84
<b>7</b>	<b>Conclusion.....</b>	89

## 1 Introduction

Russian is one of the languages that is claimed to have scrambling. A Certain flexibility of word order can be observed both within and across clauses in this language.

- (1) a. U Orlovyh rodilas dočka.

By Orlovs<sub>Gen</sub> born daughter<sub>Nom</sub>

‘Orlovs had a daughter’

- b. Dočka rodilas u Orlovyh.

Daughter<sub>Nom</sub> born by Orlovs<sub>Gen</sub>

- c. Rodilas u Orlovyh dočka.

Born by Orlovs<sub>Gen</sub> daughter<sub>Nom</sub>

- (2) a. Vy [dočka] slyšali čto u Orlovyh rodilas?

You<sub>Nom</sub> daughter<sub>Nom</sub> heard that by Orlovs<sub>Gen</sub> born

‘Did you hear that Orlovs had a daughter’

- b. My slyšali [dočka] čto u Orlovyh rodilas.

We heard daughter<sub>Nom</sub> that by Orlovs<sub>Gen</sub> born

‘We heard that Orlovs had a daughter’

- c. [Dočka] vy slyshali čto u Orlovyh rodilas?

Daughter you heard that by Orlovs born

‘A Daughter, did you hear that Orlovs had?’

In this thesis, I address the case of long distance scrambling, as introduced in (2). Long distance scrambling in Russian is not a well-understood phenomenon and a proper description of it is absent from the current literature. Nevertheless, there exist three distinct ideas with respect to the nature of these constructions in the literature.

Müller & Sternefeld (1993) observe that the application of long distance scrambling in Russian is less restricted than wh-movement and suggest that scrambling should be classified beyond the general A -A'-movement distinction.

Bošković and Takahashi (1998) draw a parallel between the properties of long distance scrambling in Japanese and in Russian. They assume that, similar to Japanese, scrambling in Russian is base generated and is semantically vacuous.

Bailyn (2001) argues that long distance scrambling in Russian is parallel to wh-movement and should be analyzed as an instance of discourse-driven movement.

These hypotheses contribute very different properties to scrambling in Russian.

This thesis is an attempt to investigate the properties of long distances scrambling in Russian and to compare the facts to the properties of Japanese style scrambling à la Bošković and Takahashi (1998) and to other displacement phenomena.

In this study, I argue that Russian does not have semantically vacuous scrambling and give arguments that long distance scrambling in Russian involves contrastive focus. I show that scrambling in Russian is sensitive to relativized minimality and to other locality constraints. I observe that the structures most transparent for scrambling are clausal complements of perception verbs. In other syntactic contexts, scrambling is more restricted than it is generally assumed.

I will also explore the possibilities of parallels between scrambling and wh-movement in English. I present evidence that scrambling and wh- movement pattern together with respect to locality constraints, informational backgrounds and successive cyclicity. I also show that there is no straightforward parallel between scrambling in Russian and heavy- NP shift in English.

The thesis is organized as follows: in Chapter 2, I introduce the types of scrambling constructions in Japanese, their properties and the history of research on this issue. I will also discuss the types of scrambling constructions available in Russian and cover the previous analyses of them. In Chapter 3, I summarize the two distinct approaches to scrambling: base generation approach advocated by Bošković and Takahashi (1998) and movement approach by Bailyn (2001). In Chapter 4, I address the weak points of the approaches and determine the workspace of my research. In Chapter 5, I investigate the properties of long distance dislocation in Russian and argue against the assumption of semantically vacuous scrambling. In Chapter 6, I draw parallels between scrambling in Russian, wh- movement, and heavy-NP shift in English. Chapter 7 summarizes the main findings.

## 2 Types of scrambling constructions and their properties

### 2.1 Scrambling constructions in Japanese

It is generally observed that considerable word order flexibility is allowed in Japanese. The placement of the verb is restricted to the clause final position, but the order of the other elements appears to be free. The flexibility of word order can be observed both locally and long distance.

(3)

- a. Taroo-ga ano mise-de hon-o katta (koto) (Nemoto, 1999:121)  
Taroo<sub>Nom</sub> that store-at book<sub>Acc</sub> bought  
'Taroo bought a book at that store'
- b. Hon-o Taroo-ga ano mise-de katta (koto)  
book<sub>Acc</sub> Taroo<sub>Nom</sub> that store-at bought
- c. Ano mise-de Taroo-ga hon-o katta (koto)  
that store-at Taroo<sub>Nom</sub> book<sub>Acc</sub> bought
- d. Hon-o ano mise-de Taroo-ga katta (koto)  
book<sub>Acc</sub> that store-at Taroo<sub>Nom</sub> bought
- e. Taroo-ga hon-o ano mise-de katta (koto)  
Taroo<sub>Nom</sub> book<sub>Acc</sub> that store-at bought
- f. Ano mise-de hon-o Taroo-ga katta (koto)  
that store-at book<sub>Acc</sub> Taroo<sub>Nom</sub> bought

- (4) [Mary-ga John-ni okutta telegami]-o kare-ga [dareka-ga t nususmiyomisista  
Mary<sub>Nom</sub> John –to sent letter<sub>Acc</sub> he<sub>Nom</sub> someone<sub>Nom</sub> took a peek at  
to ] omotteiru (koto)  
that thinking  
'The letter Mary wrote to John, he thinks someone took a peek at'  
(Nemoto, 1999:130)

#### 2.1.1 The Configurationality parameter

In the late 70's the flexibility of word order in Japanese was attributed to the configurationality parameter. Hale (1980) proposed that languages that have rigid word

order are configurational or have hierarchical structure, while languages with flexible word order are non-configurational or have flat structure. The flexible word order is created by the phrase structure rule as in (3) with no movement involved in the derivation.

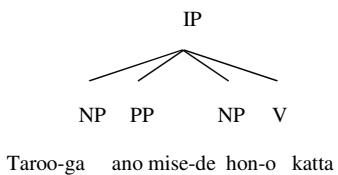
(5)

$x' \rightarrow x'^* x$  (Nemoto, 1999: 121)

The formula in (5) shows that Japanese is head final. [x]denotes a head, [ $x'^*$ ] stands for any number of xs. Correspondingly, [ $x'$ ] is the level higher than [ $x'^* x$ ].

Hence, the configurationality parameter associates a flat structure in (6) to the example in (3)a.

(6) (Nemoto, 1999:123)



However, in the 80's Saito and Hoji (1983), Saito (1985) argued against the non-configurational structure for Japanese. They introduced evidence that Japanese clause structure is not flat and analyzed the flexibility of word order observed in (3) and (4) in terms of movement. The basic word order assumed for Japanese is SOV, thus, Saito (1985) proposed that the rest of available word orders are created by an optional rule that adjoins a phrase to [IP] or [VP].

(7) Adjoin- $\alpha$ , where  $\alpha$  is  $X^{\max}$

Saito (1985) assumed that both clause internal and long distance scrambling are uniformly instances of A'-movement. The transformation was called "scrambling" in accord with Ross's (1967) terminology.

### 2.1.2 Scrambling in the 80's. A or A'- movement discussions

There are two questions that straightforwardly follow from the assumption that the origin of scrambling is movement:

- What type of movement does scrambling involve?
- What is the motivation for this movement?

Starting from mid- 80's these two questions gave rise to a mass of discussion in the literature. I begin with the latter question first.

Saito (1992), following work on Hindi done by Mahajan (1990), showed that scrambling in Japanese can not uniformly involve adjunction, since local scrambling in Japanese, like in Hindi, reveals both A and A'-properties.

Saito (1992) reports that in Japanese the dislocated element in clause internal scrambling can bind an anaphor it didn't previously c-command.

(Saito, 1992:74-5)

(8) a. ?\*Otagai<sub>i</sub> –no sensei karera-o<sub>i</sub> hihansita (koto)

each others teachers<sub>Nom</sub> them<sub>Acc</sub> criticized fact

'Each other's teachers criticized them'

b.? Karera-o<sub>i</sub> [[ otagai<sub>i</sub> –no sensei ] [ t<sub>i</sub> hihansita]]] (koto)

Them-Acc each others teachers-Nom criticized fact

'Them<sub>i</sub> [each other's teachers] criticized t<sub>i</sub>'

As we can see in (8)a the anaphor *otagai* is not bound, which leads to a violation of condition A of the binding theory. However, when the pronoun *karera* is scrambled to a position c- commanding the anaphor, binding is possible, saving the ungrammaticality of (8)b. The possibility of binding in (8)b shows that clause internal scrambling can target an A-position.

Another argument by Saito (1992) is that clause internal scrambling can save weak cross over violations. He shows that scrambling of the wh-phrase to a position c-commanding the empty pronoun, improves the grammaticality. This gives Saito (1992) additional evidence that local scrambling exhibits A-properties. See Saito (1992:71) for the relevant examples.

An important fact discovered by Mahajan (1990) for Hindi and Saito (1992) for Japaese is that clause internal scrambling is not uniformly A- movement, but, in fact, it can be both.

(9)[<sub>TP</sub>Zibunzisin-o<sub>i</sub> [Taroo-ga t<sub>i</sub> semeta ]] (koto) (Saito, 1992: 76)

self<sub>Acc</sub> Taroo<sub>Nom</sub> blamed fact

'Himself<sub>i</sub>, Taroo blamed t<sub>i</sub>'

Under the assumption that clause internal scrambling targets an A- position, the example in (9) should induce a condition C violation. Given the grammaticality of (9), Saito concludes that the landing site targeted by scrambling in (9) is in an A'- position.

The next observation with respect to local scrambling was made by Tada (1990, 1993). He showed that clause internal scrambling can change scope relations.

- (10) a. Dareka-ga daremo-o aisite iru (Nemoto, 1999:141)  
           someone<sub>Nom</sub> everyone<sub>Acc</sub> love  
           ‘Someone loves everyone’  
            $\exists > \forall * \forall > \exists$

b. Daremo<sub>i</sub>-o[ dareka-ga t<sub>i</sub> aisite iru ]  
           everyone<sub>Acc</sub> someone<sub>Nom</sub> love  
           ‘Everyone, someone loves’  
            $\exists > \forall \forall > \exists$

c. LF [Dareka<sub>j</sub>-ga [daremo<sub>i</sub>-o[t<sub>j</sub> t<sub>i</sub> aisite iru ]]]

The example in (10)a is not ambiguous with respect to scope, where the subject QP *dareka* takes wide scope. Conversely, in (10)b the scrambled QP can take scope over the subject QP. In (10)b, we can observe the ambiguity effect: on the one hand, the scrambled QP can take wide scope over the subject QP, which gives evidence that the scrambled QP can surface in the landing position at LF; on the other hand, it is still possible for the subject QP to scope over the subject QP, which shows that the scrambled QP is also subject to reconstruction. The scope ambiguity in (10)b has been viewed as an evidence that local scrambling has an effect on the interpretation.<sup>1</sup>

Another important issue about Japanese, which will become important in the course of my further discussion, is the fact that Japanese allows multiple subject constructions.

- (11) [IP Nagano-ga [IP yamato-ga[IP ki-ga kire-da]]] (Nemoto, 1999:134)  
           Nagano<sub>Nom</sub>   mountain<sub>Nom</sub> tree<sub>Nom</sub> beautiful-is  
           ‘It’s Nagano, where in mountains trees are beautiful’

To sum up, local scrambling in Japanese is reported to have mixed properties. On the one hand, the fact that a scrambled element can antecede an anaphor, the lack of WCO violations, the change of scope point into the direction that scrambling in Japanese involves A-movement. However, there are cases like (9) and (10)b<sup>2</sup>, where local scrambling exhibits reconstruction, which allows us to conclude that clause internal scrambling can also involve A'-movement.

<sup>1</sup> I will address the properties of long distance scrambling with respect to scope later.

<sup>2</sup> Here I mean the possibility of the subject QP to take scope over the dislocated QP.

The question of motivation for clause internal scrambling in Japanese will be addressed in the next subsection.

I turn to long distance scrambling in Japanese and its properties. One of the most discussed issues about scrambling in Japanese in the 80's were the differences between clause internal and long distance scrambling.

A first and very important property of long distance scrambling is that long distance subject scrambling is disallowed. Examples from (Saito, 1985:192) cited from (Nemoto, 1999:131)

- (12) a. Mary-ga John-ni [kono giron-ga okasii to] itta

Mary<sub>Nom</sub> John<sub>Dat</sub> this argument<sub>Nom</sub> strange that told  
'Mary told John that this argument is strange'

- b.\* [Kono giron]<sub>i</sub>-ga [Mary-ga John-ni [t<sub>i</sub> okasii to] itta]

this argument<sub>Nom</sub> Mary<sub>Nom</sub> John<sub>Dat</sub> strange that told  
'This argument, Mary told John that is strange'

In (12)b the embedded subject is scrambled into the matrix clause and it results in ungrammaticality.

Another important characteristic of long distance scrambling in Japanese is the fact that phrases scrambled long distance can not serve as A-binders for anaphors. Examples taken from (Saito, 1992:75-6).

- (13) a.? Karera-o<sub>i</sub> [[ otagai<sub>i</sub> -no sensei ] [ t<sub>i</sub> hihansita]] (koto)

Them-Acc each others teachers-Nom criticized fact  
'Them<sub>i</sub> [each other's teachers] criticized t<sub>i</sub>'

- b.\* [TP Karera-o<sub>i</sub> [[ otagai-no sensei ]-ga [ CP [TP Tanaka-ga t<sub>i</sub> hihansita

they -Acc each other-Gen teacher-Nom Nom criticized  
to]itta]](koto)  
that said fact .

'Them<sub>i</sub> [each other's teachers] said that Tanaka criticized t<sub>i</sub>'

In example (13)b, the object is scrambled long distance, but unlike in (13)a, the local case, in (13)b the fronted pronoun can not bind an anaphor. The impossibility of binding in (13)b is considered to be evidence that long distance scrambling in Japanese targets an A'-position.

A similar effect was observed for scope interpretation by Tada (1993). He pointed out that, unlike local scrambling, long distance scrambling doesn't affect scope. Examples from (Tada, 1993) cited from (Nemoto, 1999:142).

- (14) a. Dareka-ga [John-ga daremo-o aisiteiru to]itta  
           someone<sub>Nom</sub> John<sub>Nom</sub> everyone<sub>Acc</sub> love that said  
           ‘Someone said John loves everyone’  
            $\exists > \forall * \forall > \exists$
- b. Daremo<sub>i</sub>-o[ dareka-ga[John-ga t<sub>i</sub> aisiteiru to ]itta]  
       everyone<sub>Acc</sub> someone<sub>Nom</sub> John    love    that said  
       ‘Everyone, someone said that John loves’  
        $\exists > \forall * \forall > \exists$

Tada (1993) observes that neither (14)a nor (14)b is ambiguous with respect to scope. The scrambled QP *daremo* can not take scope over the subject QP in (14)b. This observation allows Tada (1993) to conclude that long distance scrambling behaves differently from local scrambling with respect to scope.

### 2.1.3 Discovery of the undoing property. Saito (1989), (1992)

In the late 80's Saito (1989), (1992) argued for the undoing property of long distance scrambling. Given this property, long distance scrambling is claimed to be distinct from a regular A'-movement, namely, wh- movement and topicalization.

(Saito and Fukui, 1998:441)

- (15) a. \* [IP [CP Dare-ga [IP John-ga sono hon-o katta ka]siritagatteiru] koto]  
           who<sub>Nom</sub>    John<sub>Nom</sub> that book<sub>Acc</sub> bought Q want to know fact  
           ‘the fact that who wants to know john bought that book’
- (Saito, 1992:84)
- b. [Masao-ga [CP [IP Hanako-ga dono hon-o tosyokan-kara    karidasita]  
       Masao<sub>Nom</sub>      Hanako<sub>Nom</sub> which book<sub>Acc</sub> library-from checked  
       ka] siritagatteiru]] (koto)  
       out Q want to know fact  
       ‘the fact that Masao wants to know[Q[Hanako checked out which  
       book from the library]]’  
       [...[CP [C' [IP ...wh...]Q]...]
- c.? [Dono hon-o<sub>i</sub> [Masao-ga [CP [IP Hahako-ga t<sub>i</sub> tosyokan-kara  
       Which book<sub>Acc</sub> Masao<sub>Nom</sub>    Hanako<sub>Nom</sub>    library-from  
       karidasita] ka] siritagatteiru]] (koto)  
       checked out Q want to know fact

'[the fact that which book, Masao wants to know [Q[ Hanako checked out t from the library]]]

wh<sub>i</sub> [...[CP [C<sup>-</sup> [IP t<sub>i</sub> ...]Q]]...]

It was previously observed in the literature that wh- phrases in Japanese must be contained within the [CP] headed by a [Q]- element at LF. Following Harada (1971), this condition is called wh-Q constraint.

However, as we can see in (15)c, a wh-phrase can be extracted out of the [CP] containing a [Q]- element without violating the wh-Q constraint. What follows from this observation is that wh- phrase is not represented in the scrambled position at LF, but reconstructs to its original position within the lower [CP]. Given this property, Saito (1989),(1992), Saito and Fukui (1998) conclude that long distance scrambling can be undone at LF and, hence, can be semantically vacuous.

The prediction that follows from the undoing property is that it is expected that, unlike wh-movement /topicalization in Japanese, multiple application of scrambling should be allowed. This property is reported by Saito and Fukui (1998:443). Multiple application of scrambling is grammatical both for clause internal and long distance scrambling.

- (16) a. [Sono hon<sub>i</sub>-o [John-ni<sub>j</sub> [Mary-ga t<sub>j</sub> t<sub>i</sub> watasita]]]

that book<sub>Acc</sub> John-to Mary<sub>Nom</sub> handed

'Mary handed that book to John'

- b. [IP Sono hon<sub>i</sub>-o [John-ni<sub>j</sub> [ Bill-ga [CP [IP Mary-ga t<sub>j</sub> t<sub>i</sub> watasita] to] itta]]]

that book<sub>Acc</sub> John-to Bill<sub>Nom</sub> Mary<sub>Nom</sub> handed that said

'Bill said that Mary handed that book to John'

(Saito and Fukui, 1998:441)

Both (16)a and (16)b are well-formed, thus, Saito & Fukui (1998) conclude that scrambling is different from a regular A'-movement.

Another property of long distance scrambling reported in the literature is the fact that adverbials can not scramble long distance. Examples from (Saito, 1985) cited from (Bošković & Takahashi, 1998:355).

- (17) a. Mary-ga [John-ga riyuu-mo naku sono setu-o sinziteiru to]

Mary<sub>Nom</sub> John<sub>Nom</sub> reason-even without that theory<sub>Acc</sub> believes that  
omotteiru.

thinks.

'Mary thinks that John believes in that theory without any reason.'

- b.\* Riyuu-mo naku<sub>i</sub> Mary-ga [CP John-ga t<sub>i</sub> sono setu-o sinziteiru to]  
omoitteiru.

As we can see in (17)b, extraction of adverbial *riyuu-mo naku* into the matrix clause leads to ungrammaticality.

To sum up the discussion above, Saito's (1992) discovery of the indoing property of scrambling showed that long distance scrambling in Japanese is distinct from a regular A'-movement and that a separate account of its' properties is required.

Given the fact that scrambling in Japanese exhibits mixed properties, a number of analyses made an attempt to explain the distinct properties observed in scrambling transformation within one unified account.

Several researchers (Saito (1992), Miyagawa (1996), (2001)) provide an account for the mixed properties of local scrambling, but since the focus of this study is long distance scrambling, I will not persue this issue further and move to the discussion of the analyses proposed for long distance scrambling.

#### 2.1.4 Copy and Deletion analysis of scrambling. Saito (2001), (2003)

Now we turn to the discussion of Saito's (2001), (2003) analyses, where he attempts to unify the properties of local and long distance scrambling with the help of the Copy and Deletion theory of movement.

Saito (2001) proposes to adopt Chomsky's (1993) Copy and Deletion theory of movement in order to explain the undoing properties of both local and long distance scrambling in Japanese.

He develops the ideas proposed in Tada (1990) and Kitahara (2000) and argues that scrambling is a uniform operation whether it is clause-internal or long-distance. He provides an analysis of the radical reconstruction property of scrambling and gives his account for A-/A'- distinction between long distance and clause internal scrambling.

As it was already pointed above, local scrambling in Japanese reveals mixed properties. There are cases like (8)b, where the scrambled phrase can antecede an anaphor it didn't previously c- command violating condition A of the binding theory. Apart from that, there are cases like (9), where an anphor is scrambled to a position c-commanding an R-expression leading to no condition C viloation.

To account for these facts, Saito suggests, following Lebeaux (1988), that condition A can be satisfied in the course of the derivation, which characterizes condition A as an ‘anywhere condition’. Hence, condition A is satisfied prior to the undoing of scrambling. As for condition C, the effect in (9) Saito reasons, that condition C is an LF condition. Thus, scrambling is subject to reconstruction and condition C is satisfied at LF in (9).

Saito (2001) proposes to apply the mechanism of the derivational interpretation of chains to scrambling. Adopting Chomsky’s (1993) idea that operator variable chains are created by copy and deletion, Saito adapts the relevant mechanism to the derivation of scrambling. He notes that his analysis is consistent with his previous assumption that scrambling is not feature- driven.

The copy deletion account of wh-movement proposed by Chomsky (1993) involves the [O](operator)- feature, which triggers wh- movement. Saito dispenses with the [O]-feature for scrambling.

- (18) a. [CP who [ John saw who]] (Saito, 2001:298)  
 $\{P,O,D\}$        $\{P,O,D\}$   
 b. [CP who [ John saw who]]  
 $\{P,O,\mathbf{D}\}$        $\{D\}$

According to Chomsky (1993), the derivation of the tree structure proceeds bottom up. Each of the features (P, O, D) should be retained at one position. The [P]- feature remains at the head of the chain, since it is PF- related. The [O]- feature is assumed to be the trigger for wh- movement, hence it is retained in [Spec, CP]. The [D]- feature, LF- related, is selected in the object position and thus should be deleted at [Spec, CP].

A similar mechanism is applied to scrambling.

- (19) [TP Sono hon-o<sub>i</sub> [ Yamada-ga t<sub>i</sub> yonda ]] (koto) (Saito, 2001:299).  
 that book<sub>Acc</sub> Yamada<sub>Nom</sub> read fact  
 [TP Sono hon-o... [... sono hon-o...]]  
 $\{P,D\}$        $\{P,D\}$   
 [TP Sono hon-o... [... sono hon-o...]]  
 $\{P,\}$        $\{,D\}$

In example (19), the derivation of clause bound scrambling is exhibited. Both [P]- and [D]- features are copied to [Spec, TP]. The lower copy of the [P]- feature gets

deleted, since the fronted object is pronounced in the higher position. The [D]- feature (LF-related) should be retained low, since it is selected in the object position.

- (20) [TP Sono hon-o<sub>i</sub> [ Tanaka-ga [CP Yamada-ga t<sub>i</sub> yonda to ] omoitteiru] (koto)  
 that book<sub>Acc</sub> Tanaka<sub>Nom</sub> Yamada<sub>Nom</sub> read that think fact  
 [TP Sono hon-o...[ [CP... sono hon-o...[ TP sono hon-o]]]  
 {P,D} {P,D} {P,D}  
 [TP Sono hon-o...[ [CP... sono hon-o...[ TP sono hon-o]]]  
 {P,} {P,} {,D}  
 [TP Sono hon-o...[ [CP...[ TP sono hon-o]]]  
 {P,} {,D} (Saito, 2003:493)

The derivation in (20) is an instance of long distance scrambling. Given the assumption that movement proceeds successive cyclically, the features are copied twice, with the intermediate copy deleted on the identity basis with the head of the chain. Thus, what follows is that scrambling doesn't have any semantic import because [D]- feature is retained in its low position both in local and long distance scrambling.

Saito's hypothesis allows to account for the different properties of local and long distance scrambling with respect to binding.

- (21) a.? [TP Karera-o<sub>i</sub> [[ otagai-no sensei ]-ga t<sub>i</sub> hihansita]](koto)  
 they<sub>Acc</sub> each other<sub>Gen</sub> teacher<sub>Nom</sub> ctiticized fact  
 'Them<sub>i</sub> [each other's teachers] criticized t<sub>i</sub>'  
 {P,D} {P,D}→{P} {D}  
 b.\* [TP Karera-o<sub>i</sub> [[ otagai-no sensei ]-ga [ CP [TP Tanaka-ga t<sub>i</sub> hihansita  
 they<sub>Acc</sub> each other<sub>Gen</sub> teacher<sub>Nom</sub> Tanaka<sub>Nom</sub> criticized  
 to] itta]](koto)  
 that said fact.  
 'Them<sub>i</sub> [each other's teachers] said that Tanaka criticized t<sub>i</sub>'  
 [TP karera-o [otagai...[CP karera-o [TP...karera-o]]]]  
 {P}{D} {P}{D}  
 {P} {P} {D} (Saito, 2001:299-301)

Given the assumption that condition A is an 'anywhere condition', Saito supposes that the effect in (21) a is due to the fact that the [D]- feature c-commands the anaphor before it gets deleted. The same is not possible in (21) b. Recall that Saito

(2001), following Chomsky (1993), assumes that the derivation proceeds bottom up and the interpretation takes place as a chain is created. In (21)b movement successively cyclically goes through an additional position, the [D]- feature is retained low, while the [P]- feature is copied into the higher position. The [D]- feature does not c-command the anaphor *otagai* at any stage of the derivation; hence the binding effect in (21)b is impossible.

In his (2003) paper, Saito extends some of the issues already discussed in his (2001) work. He suggests that if an anaphor contains a [+ anaphor feature], this feature can move separately from the [D]- feature. Consider the following examples:

(22)

- a.\* Karera-ga [CP Hanako-ga [CP Ziroo-ga *otagai* -o sonkeisiteotru to itta  
They<sub>Nom</sub> Hanako<sub>Nom</sub> Ziroo<sub>Nom</sub> each other<sub>Acc</sub> respect that said  
to]omoitteiru  
that think (koto)
- b. Karera-ga [CP *otagai* -o Hanako-ga [CP Ziroo-ga sonkeisiteotru to itta  
They<sub>Nom</sub> each other<sub>Acc</sub> Hanako<sub>Nom</sub> Ziroo<sub>Nom</sub> respect that said  
to]omoitteiru (koto)  
that think fact  
'They think [that each other, Hanako said [that Ziroo respects t ]]'  
(Saito, 2003:509)

In example (22)a, the anaphor can not be bound by the matrix antecedent. Apparently, when the anaphor is fronted to the clause initial position as in (33)b, the possibility of binding emerges.

Given the derivational analysis by Saito (2003), such a possibility should not occur. The features retained at the head of the chain are [P]- features, which in principle cannot correlate with binding. As a solution to the effect observed in (22)b Saito introduces the idea that the [+ anaphoric] feature can move along with the [P]- features or can be retained at any position of the chain where it can be bound.

According to Saito (2003), the derivation proceeds as follows: *otagai-o* starts out from the embedded object position in the most embedded clause bearing three features {P,D,A}(see (23a). The features get copied in the next [C]- projection. The [D]-feature is retained low. Then {P,A} features are copied further as in (23b) and the lower copy of {P,A} is deleted under identity. After that, *karera-ga* is introduced into the

subject position with features {P,D} (see (23c). The phonetic features are retained at the head of the chain, the [D]- feature can antecede the [A]- feature.

- (23) [<sub>TP</sub> Karera-ga...[<sub>CP</sub> [<sub>TP</sub> otagai-o...[<sub>CP</sub> otagai-o[<sub>TP</sub>...otagai-o]]]]]
- |    |         |            |     |
|----|---------|------------|-----|
| a. | {P}{A}  | {P} {D}{A} |     |
| b. | {P} {A} | {P} {A}    | {D} |
| c. | {P}{D}  | {P} {A}    |     |
- (Saito, 2003:510)

According to Saito (2003), this analysis can be extended to the following examples from English.

- (24) John wonders [<sub>CP</sub> [which picture of himself]<sub>i</sub> [<sub>TP</sub> Mary thinks[<sub>CP</sub> [<sub>TP</sub> Susan liked <sub>t<sub>i</sub></sub>]]]]]
- |     |       |     |                   |
|-----|-------|-----|-------------------|
| {D} | {O,P} | {A} | {O,P,}            |
|     |       |     |                   |
| {D} |       |     | (Saito, 2003:510) |

If movement of the [A]- feature along with [P]- features is allowed, then in (24), the [A]- feature of *himself* is bound by the [D]- feature of the subject *John*, similar to the effect in (23).

To conclude, Saito's (2001), (2003) main assumption is that local scrambling and long distance scrambling are uniform. The distinction of the properties they exhibit, as well as the radical reconstruction property can be derived with the use of the Copy and Deletion theory of movement and naturally follow from the analyses Saito (2001), (2003) provided.

### **2.1.5 Müller and Sternefeld (1996). Principle of Unambiguous Binding**

Another attempt to propose an account for the properties of scrambling, which allows to capture some aspects of cross-linguistic variation as well is offered by Müller & Sternefeld (1993, 1996).

They propose an account of the restrictions on overt movement of wh-phrases, which contributes to the understanding of scrambling phenomena.

The main idea they execute is that the principle of economy of derivation proposed by Chomsky (1994), Chomsky and Lasnik (1993) should be reformulated in a more specific way. They argue that the reformulation of this principle allows to account for certain problematic cases of wh-movement in German and English, as well as to

contribute to the understanding of mixed properties of scrambling in Japanese, Korean and Russian.(Chomsky, 1994) cited from ( Müller and Sternefeld, 1996:480)

## (25) Economy of Derivation

If two derivations  $D_1$  and  $D_2$  are in the same reference set, and  $D_1$  involves fewer operations than  $D_2$  then  $D_1$  is to be preferred over  $D_2$ .

(Müller and Sternefeld, 1993:461)

#### (26) Principle of Unambiguous Binding (PUB)

A variable  $\alpha$  that is  $\alpha$ -bound must be  $\beta$ -free in the domain of the head of its chain, where  $\alpha$  and  $\beta$  refer to different types of positions.

From (26) it follows that movement to the type of position  $\alpha$  should not be followed by the movement to a type of position  $\beta$ . Otherwise, the initial variable will be bound ambiguously and induce a PUB violation.

The principle Müller and Sternefeld (1993) suggest strongly depends on the classification of the landing sites of movement.

They assume the following landing sites for the corresponding types of movement: (Müller and Sternefeld, 1996:496)

- a. Wh- movement at S- structure is substitution in [Spec, CP].
  - b. Wh-movement at LF is either substitution in [Spec, CP] or, if [Spec, CP] is already filled, right adjunction to [Spec, CP].
  - c. Topicalization in the Germanic languages is substitution in [Spec, TP].
  - d. Scrambling is left-adjunction to XP ([VP] or [IP] in German and Korean)

The claim Müller and Sternefeld make is that PUB can offer a unified analysis of the asymmetries between movement types.

Consider an example of wh-topicalization from German. It is generally accepted that wh- elements unlike [-wh ] element can not undergo topicalization. (Müller & Sternefeld, 1996:482)



The ungrammaticality in (27) follows from the violation of the PUB. In (27)a, ‘wen’ occupies the [Spec, TP] position ( $\alpha$ ) at S-structure, at LF it undergoes movement to the adjoined [Spec, CP] position in the matrix clause ( $\beta$ ), incurring a violation of PUB.

The prediction most relevant for the present study involves examples of wh-scrambling in Korean. The phenomenon of wh- scrambling in Korean is problematic for

the Economy approach, but according to Müller and Sternefeld, can be easily accounted for by the PUB.

- (28)a. Cholsu-ka [VP muos-ul [VP Sunhi-eke  $t_i$  chu-oss-ni ]]?  
 Cholsu<sub>Nom</sub> what-<sub>Acc</sub> Sunhi<sub>Dat</sub> gave Q  
 ‘What did Cholsu give to Sunhi’
- b. LF [CP [Spec,CP muos-ul<sub>i</sub>] Cholsu-ka [VP ( $t'_i$ ) [ VP Sunhi-eke  $t_i$   
 β α  
 chu-oss- ni]]. (Müller & Sternefeld, 1996:501)

The grammaticality of (39) is unexpected under the PUB. The wh- element first moves from the [VP]- internal position to the [VP]- adjoined position ( $\alpha$ - position). As the next step it, supposedly, undergoes movement to the [Spec, CP] position at LF ( $\beta$ -position).

However, Müller and Sternefeld (1996), refer to the assumption introduced in their (1993) paper that languages differ with respect to the domain of the PUB application. They suggest that for languages like English and German PUB is applied at S-structure and at LF. In other languages like Japanese, Korean and Russian “the S-structure movement is forgotten at LF; thus, the S- structure part of the chain will not be checked again in the course of LF movement”(Müller& Sternefeld, 1996: 501).

According to Müller and Sternefeld, the LF  $\beta$ - movement (see (28)), does not incur a PUB violation, because the overt wh- movement is not checked at LF.

The above solution suggested in Müller and Sternefeld (1993) is reformulated in their 1996 work as a part of the definition of Form Chain. According to Chomsky (1993, 1994), the standard operation Form Chain consists of two sub parts: Move and Insert Traces.

Müller and Sternefeld propose that a Chain Linking operation should be added as a subcomponent, and the final step is Check PUB.

(29) Move >>Insert Traces>> Apply Chain Linking>> Check PUB  
 The sequence in (29), according to Müller and Sternefeld, occur in languages like English and German. As for scrambling in languages like Korean, Japanese, Russian, they posit, that the order of the operations should be the following. (Müller & Sternefeld, 1996:502)

- (30) Move >>Insert Traces>> Check PUB>>Apply Chain Linking

If we compare (29) and (30), the last two operations are switched. Now consider the example (28), the LF representation of (28) is repeated in (31).

The LF- movement to [Spec, C] applies due to Move operation. The trace  $t'_i$  is inserted according to the Insert Trace operation. At this point of the derivation, PUB incurs no violation because chain linking has not applied and the S-structure trace is not taken into account. The last step is the application of Chain Linking, which adds an S-structure trace  $t_i$ .

(31) [CP [Spec,C muos-ul<sub>i</sub>] Cholsu-ka [VP(t'<sub>i</sub>) [ VP Sunhi-ake t<sub>i</sub> chu-oss-ni]]]

$\alpha$                                      $\alpha$

Hence, the above solution captures the fact that wh- scrambling in Korean/Japanese does not violate the PUB condition. Müller and Sternefeld claim that the PUB constitutes a plausible account for the asymmetry of movement types, as well as sheds light on how the properties of scrambling are derived.

To sum up, I have discussed a number of assumptions about scrambling in Japanese that exist in the literature. There are three main issues about scrambling that most of the scholars are trying to solve

- The type of movement involved in scrambling
- the motivation for scrambling

Now I turn to the discussion of Russian scrambling and the existing literature about it.

## 2.2 Scrambling constructions in Russian

It is generally assumed that Russian is one of the languages that allow flexibility of word order. The question what is causing the flexibility of word order in Russian and how it is different or similar to the flexibility of word order observed in other languages has been widely discussed in the recent literature.

### 2.2.1 Classics: theme –rheme distinction

According to the classical assumption, the flexibility of word order in Russian is attributed to the theme-rheme distinction. For example, Krylova and Khavronina (1976) argue that “word order depends on the speaker’s aim and on the new significant information he wishes to convey in his utterance” (Krylova and Khavronina, 1976:17).

The theme denotes the subject of the message, while the rheme contains the message itself.

(32)- Kto u vas byl v gostjah? (Krylova and Khavronina, 1976:17)

‘Who visited you?’

- U nas v gostjah byl Victor.

Theme              rheme

By us<sub>Gen</sub> in visit<sub>Instr</sub> was Victor <sub>Nom</sub>

‘We had Victor visiting us’

As we can see in (32), the theme is represented in both the question and the answer, hence, the theme bears old information, while rheme conveys a piece of new information.

Krylova and Khavronina (1976) make a distinction between objective word order, which appears in non-emotive speech, and subjective word order / inversion, which occurs in emotive speech. In non-emotive speech, the theme has to precede the rheme, while in emotive speech the theme- rheme sequence can be violated.

The direct word order is assumed to conform to the sequence theme-rheme and rendered as stylistically neutral. The indirect order, in turn, is considered to be stylistically marked with the theme- rheme sequence violated. (Krylova and Khavronina, 1976:135)

(33) a. Ne ponravilos Fedoru eto predloženie.

Rheme

Not like <sub>Past</sub> Fedor <sub>Dat</sub> this offer <sub>Nom</sub>

‘Fedor didn’t like this offer’

b. Fedoru eto predloženie ne ponravilos.

Theme              rheme

However, Krylova and Khavronina (1976) note that the inverted word order does not determine the theme-rheme sequence by itself. Special intonation helps to emphasize the rheme and disambiguate the theme- rheme sequence.

(34) a. Vremya **bylo horošee.** (Krylova and Khavronina, 1976:138)

Time was good

‘It was a good time’

b. **Horošee bylo vremya.**

Good was time

‘It was a good time’

To conclude, in general in the Prague school, the variation of word order in Russian was attributed to the theme-rheme division. The direct word order was considered to be unmarked with the theme- preceding- rheme sequence. Inverted order was viewed as stylistically marked and violating the theme-rheme sequence.

### 2.2.2 Discourse-driven accounts. Junghans & Zybatow (1997)

Junghans & Zybatow (1997) advocate a different approach to the derivation of scrambling. According to them, scrambling is a regular A'- movement caused by information structure. In their paper Junghans and Zybatow discuss only cases of clause internal scrambling.

They argue that clause internal variation of word order in Russian is connected with discourse. They assume that the neutral word order for Russian is: subject, verb, dative object, accusative object. This order of the constituents appears to be unmarked because only this combination allows to get an indefinite reading of the nominal expressions. (Junghans & Zybatow, 1997:295)

- (35) a. Odna ženschina podarila malčiku jabloko.  
ok ‘A woman gave a boy an apple’  
ok ‘A woman gave the boy an apple’

The rest of the possible orders are assumed to be derived by overt movement.

- (36) a. Anton celuet Mašu.(unmarked)

- Anton<sub>Nom</sub> kiss Mary<sub>Acc</sub>  
b. Celuet Anton Mašu.  
c. Mašu celuet Anton.  
d. Anton Mašu celuet  
e. Celuet Mašu Anton.  
f. Mašu Anton celuet.

According to Junghans and Zybatow, any of the constituents in the examples (36) can be assigned a topic or a focus feature based on the communicative situation. I will discuss one of the examples from the list in (36). According to Junghans and Zybatow the subject in (36)f is assigned a topic feature, and the constituent that is assigned a topic feature should move to the topic position. Since it is assumed that this position should be able to host subjects, as well as objects and other elements, Junghans and Zybatow suggest that this position is an adjoined position to [Agr, SP].

Correspondingly, each of the orders, that deviate from the unmarked order in (36) are assumed to be derived by movement to a unique topic/focus position, where the inverted constituent checks the topic/focus feature it is assigned in accordance with the informational background. Junghans and Zybatow also note that, given the availability of various discourse movements in Russian, it is possible that Russian, as opposed to languages like English, has a richer system of discourse driven movements and apart from contrastive focus and new-information focus, there may be other types of focus.

If scrambling is analyzed as overt movement, it should be viewed as driven by strong features. According to Chomsky (1992), strong features in the lexical projections cause lexical material to raise overtly.

Junghans and Zybatow argue that the surface order in Russian is not created by movement driven by feature checking. Given the optional nature of overt movement in Russian, they think that it is implausible that relevant features change their value, being weak in some cases and strong in others.

Junghans and Zybatow claim that Russian has weak grammatical features. Due to rich overt morphology, which disambiguates the surface order of the constituents, all syntactic constituents in Russian can, in principle, stay in situ. When movement occurs, it is the information structure that requires it.

The hypothesis proposed by Junghans and Zybatow predicts the following typological division for languages. Languages differ with respect to the realization of their surface syntactic functions.

Languages like English with poor overt morphology move their syntactic constituents around for case requirements in order to establish a non-ambiguous configuration of the constituents. Languages with rich overt morphology like Russian mark their constituents on the surface, verb agreement is realized by means of suffixes, thus overt movement is not required. When movement occurs in Russian to reorder the constituents, it is required exclusively by information structure.

An important observation to point out with respect to Junghans and Zybatow's analysis is that they do not provide reconstruction tests to support their idea that local scrambling in Russian involves A'-movement. The type of evidence they provide mostly involves examples of the context, but the explicit testing of the inverted constructions is not available. Thus, it is unclear if it always the case that local scrambling in Russian is discourse driven, or it can be an instance of A-movement.

To sum up, Junghans and Zybatow view the clause internal flexibility of word order as a result of the topicalization/focalization movements. Thus, scrambling is uniformly A'-movement, that creates inverted structures in accord with the information structure.

Nevertheless, later research on clause- internal scrambling in Russian gives evidence that local instances of scrambling in Russian, parallel to Japanese, can show A-properties.

Bailyn (2002) exemplifies the effects characteristic of A-movement process in local scrambling constructions in Russian.

Bailyn shows that the scrambled phrase can bind an anaphor it did not c-command before the movement.

- (37) a.? Svoja<sub>i</sub> rabota nravitsya Maše<sub>i</sub>. (Bailyn, 2002:10)

self's work<sub>Nom</sub> pleases Maša<sub>Dat</sub>

'Maša likes her work'

- b. Maše<sub>i</sub> nravitsya svoja<sub>i</sub> rabota.

Maša<sub>Dar</sub> pleases self work<sub>Nom</sub>

'Maša likes her work'

Bailyn attributes the marginal grammaticality of (37)a to the unbound anaphor. When the antecedent is moved to the position c-commanding the anaphor, the binding becomes available. Bailyn treats the possibility of binding in (37)b as an evidence that local scrambling targets an A-position.

Another piece of evidence that Bailyn introduces comes from the anti-reconstruction test. He demonstrates that the inverted constituent can bind an R-expression it didn't previously c-command, triggering a principle C violation.(Bailyn, 2002:11)

- (38)a. [Novye znakomye Ivana<sub>i</sub>] predstavili ego<sub>i</sub> predsedatelju.

New friends Ivan<sub>Gen</sub> introduced him<sub>Acc</sub> to the chairman<sub>Dat</sub>

'Ivan's<sub>i</sub> new friends introduced him<sub>i</sub> to the chairman.'

- b.\* Ego<sub>i</sub> predstavili [novye znakomye Ivana<sub>i</sub>] predsedatelju.

Him<sub>Acc</sub> introduced new friends Ivan<sub>Gen</sub> to the chairman<sub>Dat</sub>

'He<sub>i</sub> was introduced to the chairman by Ivan's<sub>i</sub> new friends'

Given the assumption that local scrambling involves A'- movement, the ungrammaticality of (38) should not be expected. Bailyn views the effect in (38)b as evidence in favour of the assumption that local scrambling involves A- movement.

More evidence, that Bailyn introduces, comes from weak-cross over effects in Russian. As is shown in (39) local scrambling in Russian does not cause weak cross over violations. This is also a general characteristic of an EPP-driven movement. Examples based on (Bailyn, 2002:12)<sup>3</sup>

- (39) a.\* Ee<sub>i</sub> sobaka kusaet každuju devočku<sub>i</sub>.

Her dog<sub>Nom</sub> bites every girl<sub>Acc</sub>

‘Her<sub>i</sub> dog bites every girl<sub>i</sub>.’

- b. [Každuju devočku]<sub>k</sub> kusaet ee<sub>i</sub> sobaka t<sub>k</sub>.

[Every girl]<sub>i Acc</sub> bites [her<sub>i</sub> dog]<sub>Nom</sub>

‘Every girl is bitten by her dog’

As we can see in (39)b, movement of the object across the co-referent pronoun does not incur a violation. Bailyn assumes that the effect in (39) supports the idea that local scrambling in Russian is A-movement.

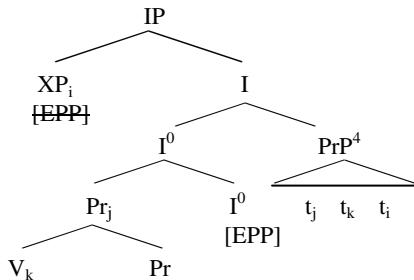
Bailyn suggests to analyze these instances of local scrambling as triggered by the external projection principle (EPP). He assumes that in Russian the EPP is an overtness requirement on the [Spec, IP] position that can be satisfied by both nominative and non-nominative [XP]. In case the [Spec, IP] position is filled by a nominative [XP], the SVO order emerges. If the [Spec, IP] position is filled by a non-nominative [XP], the verb has to raise to [I].

Bailyn adopts Holmberg and Platzack’s (1995) idea of the overt tense condition. According to the overt tense condition, an uninterpretable feature [+T] must be checked by overt movement. For Russian, Bailyn assumes, that this feature is generated in the [IP] domain and must be checked overtly, hence, providing the motivation for the verb to raise.

---

<sup>3</sup> All Bailyn’s examples involve psych verbs. To exclude this interfering factor, I constructed examples with a non-psych verb.

(40) (Bailyn, 2002: 9)



It is important to point out that Bailyn (2002) does not investigate the issue whether the inverted order versus un-inverted word order in local scrambling constructions that he points out has any influence on the interpretation. Given his assumption that clause internal scrambling is an A-movement, interpretive difference should not be expected. It is generally assumed that instances of A-movement are not driven by considerations of discourse structure. This assumption would be contradicting the discourse -driven movement approach to local scrambling by Junghans and Zybatow (1997) summarized above.

To conclude, there are two types of approaches to local scrambling in Russian. One type of approach is the discourse driven movement approach that defines local scrambling in Russian as an instance of discourse driven movement. The other type of approach treats local scrambling as an instance of A-movement. The conclusion to draw is that there are two options that can be considered:

- local scrambling in Russian has mixed properties, similar to Japanese and can exhibit both A- and A'-properties.
- there is a considerable lack of evidence that does not allow to estimate the full paradigm of properties.

I turn to the discussion of long distance scrambling constructions in Russian and the properties they exhibit.

There are two main types of approaches with respect to long distance scrambling in Russian. The first type treats long distance scrambling as an instance of A'-movement. Representatives of this approach point out that the properties of long distance scrambling are distinct from A-scrambling and discuss the parallel behaviour of wh-movement and long distance scrambling.

---

<sup>4</sup> PrP stands for predicate phrase.

The second type of existing approaches points out the asymmetry between long distance scrambling and wh- movement in Russian. They argue that long distance scrambling appears to be more unconstrained than wh-movement because scrambling tends to escape the type of constraints that normally rule out wh-movement. Hence, they treat long distance scrambling as an instance of the third type of movement, that can not be classified within the A/A'-movement distinction.

I will address the former type of approaches first and discuss their main arguments. Bailyn (2002) argues that long distance scrambling of the phrase containing an anaphor does not interfere with the binding relation.

- (41) a. Ya hoču, čtoby studenty<sub>i</sub> pročitali [knigi drug o druge<sub>i</sub>]  
I want that students read books<sub>Acc</sub> about each other  
'I want the students to read the books about each other.'
  - b. [Knigi drug o druge<sub>i</sub>] ya hoču, čtoby studenty<sub>i</sub> pročitali.  
books<sub>Acc</sub> about each other I want that students read  
'The books about each other, I want the students to read'
- (Bailyn, 2002:4)

As we can see in the example (41)b, the reciprocal is scrambled long distance to the position c-commanding its antecedent. However, the sentence is grammatical due to reconstruction. Given the evidence that the dislocated phrase reconstructs, Bailyn concludes that long distance scrambling is A'-movement.

More evidence comes from weak –cross- over effects. Bailyn points out that long distance scrambling causes weak-cross over violations and this fact can serve as evidence that long distance scrambling targets an A'- position.Examples from (Bailyn, 2002:4)

- (42) a.\* Ja hoču čtoby ee sobaka poljubila každuju devočku.  
I want that [its<sub>i</sub> dog]<sub>Nom</sub> loves [ every girl]<sub>i Acc</sub> t<sub>i</sub>  
'I want her dog to love every girl'
- b.\* Každuju devočku ja hoču čtoby ee sobaka poljubila.  
[every girl]<sub>i Acc</sub> I want that [its<sub>i</sub> dog]<sub>Nom</sub> loves  
'I want her dog to love every girl'

Hence, the facts listed above allow Bailyn to conclude that long distance scrambling in Russian is an instance of A'- movement.

A conclusion similar to Bailyn's (2002) with respect to long distance scrambling and Junghans and Zybatow's (1997) with respect to clause-internal scrambling is

reached by Strahov (2000). Based on the observation that wh-fronting behaviour is parallel to scrambling, she argues that scrambling both local and long distance has properties of A'- movement.

Strahov points out that wh-fronting as well as scrambling can be clause- internal and long distance.

- (43)a. [Èti ovošči] <sub>i</sub> on ljubit t<sub>i</sub>. (Strahov, 2000:300)

[these vegetables] <sub>Acc</sub> he <sub>Nom</sub> likes

'He likes these vegetables'

- b. Čto <sub>i</sub> ty ljubiš t<sub>i</sub>?

What <sub>Acc</sub> you <sub>Nom</sub> like

'What do you like?'

- (44) a. Vy [posylku] <sub>i</sub> videli [CP kak zapakovali t<sub>i</sub>].

You<sub>Nom</sub> parcel<sub>Acc</sub> saw how packed

'You saw how they packed the parcel'

- b. Kogo Dima prosit Svetu priglasit' t?

Who<sub>Acc</sub>Dima<sub>Nom</sub> ask Sveta invite

'Who does Dima ask Sveta to invite?'

Scrambling similar to wh-fronting does not necessarily target the clause initial position.

- (45) a. Ja [èti ovošči] <sub>i</sub> ljublju t<sub>i</sub>. (Strahov, 2000:300-1)

I <sub>Nom</sub> these vegetables <sub>Acc</sub> like

'I like these vegetables'

- b. Ty kuda<sub>i</sub> sejčas ideš t<sub>i</sub>?

You<sub>Nom</sub> where now go

'Where are you going now?'

Strahov demonstrates that both multiple scrambling and multiple wh-fronting exist in Russian and co-occur.

- (46) a. [Eto plat'e] [mne] šila podruga. (Strahov, 2000:301)

This dress<sub>Acc</sub> me<sub>Dat</sub> sewed friend<sub>Nom</sub>

'A friend sewed this dress for me'

- b. [Komu] [čto] Dima prines ?

Whom<sub>Dat</sub> what<sub>Acc</sub> Dima<sub>Nom</sub> brought

'What Dima brought to who?'

Strahov also refers to works by Bailyn (1995) and Sekerina (1997) where it is shown that wh-fronting and scrambling are restricted by the same island constraints: sentential subject constraint, adjunct constraint, complex NP- constraint, but both wh-fronting and scrambling allow extraction from a wh-island.<sup>5</sup>

Hence, Strahov concludes that scrambling and wh-fronting reveal parallel behaviour because they are both derived by the same type of movement, namely, discourse driven movement.

### 2.2.3 Mysterious asymmetry of scrambling and wh-movement

Müller & Sternefeld (1993)

Müller and Sternefeld (1993) in their discussion of long distance scrambling observe that although there is a noticeable correlation between wh-movement and scrambling, there are number of constraints that appear to restrict wh-movement but not scrambling.

For example, Müller and Sternefeld (1993) point out that in Russian wh-movement can proceed only across a subjunctive complementizer, but not across an indicative one<sup>6</sup>. (Müller and Sternefeld, 1993:466-7).

- (47)a.\* [Kakuju knigu]<sub>i</sub> ty      dumaeš čto Petr pročítal t<sub>i</sub> ?  
                   which book<sub>Acc</sub> you<sub>Nom</sub> believe that<sub>Ind</sub> Peter<sub>Nom</sub> read

b. [Kakuju knigu]; ty      dumaeš čtoby Petr pročítal t<sub>i</sub> ?  
                   which book<sub>Acc</sub> you<sub>Nom</sub> believe that<sub>Subj</sub> Peter<sub>Nom</sub> read  
                   'Which book do you think that Peter should read?'

However, this constraint does not happen to restrict scrambling. (Müller and Sternefeld, 1993:467)

- (48) On skazal [CP čto [IP noski [IP on rad [CP čto kupil t<sub>i</sub>]]]]  
 he said that<sub>Ind</sub> socks<sub>Acc</sub> he<sub>Nom</sub>glad that bought  
 'He said that he is glad that he bought the socks'.

As we can see in (48), movement of an [NP] across the indicative complementizer does not invoke a violation of the constraint. Hence, Müller and

<sup>5</sup> I will come back to this issue in detail in section 6.

<sup>6</sup> I disagree with the judgement in (47). I think the contrast reported by Mülle and Sternefeld is too strong. I think (47)a is degraded but not ungrammatical.

Sternefeld suggest that scrambling has a less restricted nature rather than wh-movement.

Müller and Sternefeld observe a similar effect with respect to the subject condition, which states that extraction out of subjects is disallowed. Subject clauses create strict islands for wh-movement, whereas, scrambling does not incur subject condition violations.

- (49) a. Mne Katju<sub>i</sub> kažetsja [CP čto [IP otpustit' t<sub>i</sub> odnu tak pozdno]] bylo by<sup>7</sup>  
Me<sub>Dat</sub> Katja<sub>Acc</sub> seems that let-go alone so late be would  
bezumiem.  
insanity<sub>Inst</sub>  
'It seems to me that it would be insane to allow Katja to go alone so late at night'.  
b.\* Kogo<sub>i</sub> tebe kažetsja [CP čto [IP otpustit' t<sub>i</sub> odnogo tak pozdno]] bylo by  
Who<sub>Acc</sub> you<sub>Dat</sub> seems that let-go alone so late be would  
bezumiem?  
insanity<sub>Inst</sub> (Müller and Sternefeld, 1993:467)

Moreover, Müller and Sternefeld introduce evidence that scrambling, as opposed to wh-movement, can escape wh-islands.

- (50) a. \* Kto ty videl kogda pod'ezžal? (Müller and Sternefeld, 1993:467-8)  
Who you saw when came  
b.?\* Čto vy videli kak zapakovali?  
What you<sub>Pl</sub> saw how (they) did up  
(51) a. Ty doktor videl kogda pod'ezžal t<sub>i</sub>?  
You<sub>Nom</sub> doctor<sub>Nom</sub> saw when arrived  
'Did you see when the doctor arrived?'  
b. Vy posylku videli kak zapakovali t<sub>i</sub>.  
You<sub>Nom</sub> parcel<sub>Acc</sub> saw how packed  
'You saw how they packed the parcel'

Given the evidence presented above, Müller and Sternefeld conclude that there exists an asymmetry between wh- movement and long distance scrambling with respect to their properties in Russian.

---

<sup>7</sup> I agree on the judgement in (49).

The observation that scrambling is less restricted than wh-movement, or regular A'-movement, leads to the idea that the scrambling transformation can, in fact, be the type of movement not captured within A- and A'-type classification.

The type of account that Müller and Sternefeld propose for the properties of long distance scrambling is based on their hypothesis of principle of unambiguous binding (PUB) discussed in the previous subsection in more detail. Müller and Sternefeld propose that besides the usual adjunction sites associated with scrambling, namely, [VP] and [IP]-adjunction, Russian has an additional position-left adjunction to [CP]. In this way scrambling becomes insensitive to wh-islands. Consider the scheme representing (52)a below: (Müller and Sternefeld, 1993:469)

(52)

$\begin{array}{c} [\text{VP doktor}_i [\text{VP} \dots [\text{CP} t [\text{CP kogda C} [\text{IP} t_i \dots]]]] \\ \alpha \qquad \qquad \qquad \alpha \end{array}$

Given the assumption that Russian has an additional adjunction to [CP]-position, Müller and Sternefeld assume that in (52), scrambling proceeds through two positions: adjunction- to- [CP] and then adjunction- to- [VP] in the matrix, thus, involving only adjunction positions ( $\alpha$ -type positions). Wh-movement, in turn, would have to proceed through the [Spec, CP] and that would create a landing position of a different type ( $\beta$ -type), which would lead to a violation of PUB. Thus, scrambling can escape a violation of PUB and wh-movement can't.

Müller and Sternefeld note that, given the possibility of [CP]-adjoined position for scrambling in Russian, it is expected that there are cases of overt adjunction to [CP] in Russian but not in German. The prediction is borne out as we can see in the example (53) below. (Müller and Sternefeld, 1993:469)

- (53)a. Ja byl [CP [CP novuju školu] i [gde strojat t<sub>i</sub>]]  
          I<sub>Nom</sub> was new school<sub>Acc</sub> where they- build  
          'I have been where they ar building the new school.'
- b. \*Ich war (dort)[ CP [NP die neue Schule] [CP wo sie t<sub>i</sub> bauen]].  
          I was there              new school            where they build

Hence, according to Müller and Sternefeld (1993), the additional properties of long distance scrambling in Russian are the result of the availability of a [CP]-adjoined position for scrambling.

To sum up the discussion above, it is appropriate to say that long distance scrambling in Russian has mixed properties. On the one hand, a number of analyses

observe certain parallelism between long distance scrambling and discourse driven movement and account for A'- properties of scrambling by assuming that scrambling is an instance of topicalization /focalization processes.

On the other hand, there are number of analyses that single out some properties of long distance scrambling that allow to classify scrambling as a less restricted type of transformation as compared to a regular A'-movement.

Unanimous agreement on this subject has not been reached in the literature and the issue open.

### *Summary*

In this section I discussed two types of scrambling constructions existing in Japanese and Russian: clause internal scrambling and long distance scrambling and the properties they are reported to possess. Scrambling in Japanese is assumed to have mixed properties: clause internal scrambling can target both A- and A'-positions; long distance scrambling is reported to have either A'- properties or can be subject to radical reconstruction, which allows to treat some instances of long distance scrambling in Japanese as a semantically vacuous transformation.

In Russian, the tendency appears to be the same: there are approaches that analyze clause internal scrambling in Russian as an instance of A-movement, others view all instances of scrambling (clause- internal and long-distance) as A'-movement. For long distance scrambling in Russian, there are approaches that treat scrambling as an instance of topicalization /focalization processes and also there are attempts to classify long distance scrambling in Russian as a type of movement distinct from a regular A'-movement.

The observation I would like to point to is that although the tendency with the treatment of Japanese and Russian scrambling in the literature is the same, but the number of analyses and the amount of tested data is more extensive and detailed for studies on Japanese than on Russian. For Russian studies there is a noticeable shortage of tested data and absence of long-time research. This might be the reason for certain disagreements on the properties of scrambling in Russian.

### 3 Movement or base generation?

#### 3.1 The LF- lowering analysis of scrambling

The standard analysis of scrambling in Japanese faces a problem of violating the last resort principle. Bošković & Takahashi (1998) (B&T) propose an alternative to the standard movement analysis in the form of an LF- lowering analysis, which, as they claim, conforms to the last resort principle and is conceptually and empirically superior to the standard idea of optional overt movement.

According to B&T, the scrambled phrases are base generated in their surface positions and undergo LF movement (lowering in this case) to the position where they are assigned  $\theta$ -roles. (Bošković & Takahashi, 1998:350)

- (54) PF [<sub>IP</sub>Sono hon-o<sub>i</sub> [<sub>IP</sub>John-ga[<sub>CP</sub> Mary-ga [<sub>VP</sub>e<sub>i</sub>[<sub>v</sub>katta]]]to omotteiru]]<sup>8</sup>  
That book<sub>Acc</sub> John<sub>Nom</sub> Mary<sub>Nom</sub> bought that think  
'That book, John thinks that Mary bought.'

LF [<sub>IP</sub>John-ga[<sub>CP</sub> Mary-ga [<sub>VP</sub>sono hon-o [<sub>v</sub>katta]]]to omotteiru]]  
John<sub>Nom</sub> Mary<sub>Nom</sub> that book<sub>Acc</sub> bought that think

As is illustrated in (78), the object *sono hon-o* is generated in the [IP]-adjoined position and undergoes lowering to its  $\theta$ -position at LF.

B&T's proposal involves a lowering operation, which is generally assumed to be unnatural. There is no specific constraint against lowering operation in the current theory, but lowering is usually ruled out by other independent principles (proper binding condition, ban against vacuous quantification). In other words, the ban on lowering doesn't exist because it would redundant with respect to other constraints. Given the absence of a constraint prohibiting lowering, B&T assume that lowering exists and that it is licit.

The lowering operation, according to B&T, takes place covertly. B&T adopt Lasnik and Saito's (1992) assumption that movement doesn't have to leave a trace when no principle requires it and, hence, argue that an LF movement doesn't have to leave a trace because of no explicit principle demanding it.

B&T point out that  $\theta$ - theory is not challenged by their analysis. Given the fact that the minimalist system can satisfy thematic requirements at LF, the scrambled

---

<sup>8</sup> 'e' stands for a  $\theta$ -position.

element being located at its canonical position at LF doesn't violate any of the principles of θ- theory. The novel assumption with respect to θ-theory that B&T introduce is that θ- roles are formal features, which can trigger movement. In this way, scrambling is motivated by feature licensing and, hence, B&T's analysis of scrambling is consistent with the last resort principle.

B&T hypothesize that θ-features can differ in strength, which allows them to capture a parametric difference between languages. The idea is that Japanese has weak θ-features and that's why they can be checked covertly. For English, B&T argue that θ-features are strong and need to be filled in overt syntax. In this way, B&T (1998) account for the fact that scrambling is not available in English, but can be found in Japanese.

### 3.1.1 Why lowering?

In order to show that the LF- lowering analysis is empirically superior to the previous analyses, B&T devote a whole section to the discussion of the problematic issues with respect to scrambling that can be naturally resolved under the LF- lowering account. I will shortly summarize the main arguments.

The first argument in favor of the LF analysis of scrambling is the conformity to last resort principle. Movement is driven by θ- features. If movement didn't apply, the scrambled phrases would fail to receive their θ- roles, which would lead to ungrammaticality.

The second argument deals with Saito's (1989) data, where he observes that scrambling being an A'- movement is different from wh-movement and topicalization in the sense that, unlike the latter two, scrambling can be undone at LF.(see examples in (15) for relevant facts)

The otherwise puzzling observation by Saito (1989) can be accounted for by the LF- lowering account. Scrambling has no semantic import and doesn't establish operator variable relation because scrambled phrases are base generated in the [IP] - adjoined position and they do not surface in that position at LF, contributing no semantic import to scrambling.

B&T's third argument is that LF lowering allows to eliminate the reference to two types of syntactic positions A- and A'- in the analysis of scrambling. The standard

assumption treats long distance scrambling as A'- movement on the basis of the fact that the scrambled [NP] cannot antecede anaphors in scrambling constructions.

B&T (1998) account for the absence of binding by the absence of c-command. Since the scrambled [NP] is obligatory lowered at LF, it doesn't c-command the anaphor, hence, binding cannot be established. If LF lowering had to leave a trace then the proper binding condition would be violated. Since B&T assume that LF lowering doesn't leave a trace, the proper binding condition is left inapplicable.

Another issue that B&T's theory can resolve is Saito's (1985) observation that adverbials can not be scrambled across a [CP] in Japanese. If scrambling is analyzed as an instance of optional movement, then it is puzzling why this optional movement can not apply to adverbials.(see examples in (17) for relevant facts)

B&T argue that, under the LF- lowering analysis, the effects described in Saito (1985) are predicted by their theory. Since adjuncts are licensed by being adjoined to categories, and, according to the base- generation hypothesis, they are generated in an adjoined position, adjuncts do not undergo LF- lowering. Adjuncts have neither a Case feature nor a θ- role, thus, the lowering is not motivated and doesn't occur.

Another important observation that B&T discuss is the one pointed out by Kikuchi (1987): scrambling out of scrambled phrases is allowed in Japanese. (Bošković & Takahashi, 1998:357)

- (55) [ Sono hon-o<sub>j</sub> [John-ga [CP[IP [CPMary-ga t<sub>j</sub> katta to]<sub>i</sub> [Bill-ga t<sub>i</sub> Itta]]to]omotteiru]].  
That book-Acc John-Nom Mary-Nom bought that Bill-Nom said  
that think

'That book<sub>j</sub>, John thinks that [that Mary bought t<sub>j</sub>] Bill said t<sub>i</sub>.'

In (55), the object is scrambled out of the embedded [CP], The [CP] [CPMary-ga sono hon-o katta to] was scrambled prior to the object. The result of the transformation is grammatical.

Under the assumption that long distance scrambling involves adjunction, the possibility of extraction out of an adjunct is puzzling. It is generally assumed that extraction out of adjuncts is disallowed. Under the LF lowering account, no extraction takes place, since the scrambled phrases are initially generated in those positions at S-structure, hence, the grammaticality of (55) is expected.

An important distinction between the LF- lowering account and overt movement that B&T point out is that these accounts have different predictions with respect to island effects.

Under the movement account, long distance scrambling is A'-movement and it should be sensitive to wh- islands, while under the LF- lowering assumption, scrambling doesn't involve movement, hence, scrambling should be able to escape wh-islands.

(Bošković & Takahashi, 1998:359)

(56) a.?\* John<sub>i</sub> you wonder whether Mary kissed t<sub>i</sub>.

b.?\* To John<sub>j</sub> that book<sub>i</sub> (Bill said that) Mary handed t<sub>i</sub> t<sub>j</sub>.

(57) a. Sono hon-o<sub>i</sub> John-ga [Mary-ga t<sub>i</sub>. yonda ka dooka] siritagatteiru.

That book<sub>Acc</sub> John<sub>Nom</sub> Mary<sub>Nom</sub> read whether wants to know  
'That book, John wants to know whether Mary read'

b. Sono hon-o<sub>i</sub> John-ni<sub>j</sub> [Bill-ga [Mary-ga t<sub>j</sub> t<sub>i</sub>. watasita to] itta].

That book<sub>Acc</sub> John<sub>Dat</sub> Bill<sub>Nom</sub> Mary<sub>Nom</sub> handed that said  
'That book, to John, Bill said that Mary handed'

Given the classical assumption that long distance scrambling is an instance of A'-movement, it is expected that the ungrammaticality in (56), which is an example of topicalization in English (A'-movement), should be as ungrammatical as Japanese examples as in (57). Thus, B&T predict that unlike topicalization in English, Japanese scrambling is not sensitive to relativized minimality.

### 3.1.2 Relativized minimality and LF-lowering

The fact that LF- lowering is exempt from relativized minimality is crucial for B&T theory, because relativized minimality facts give the main evidence for the semantically vacuous type of scrambling in Japanese.

According to B&T, the explanation why the LF- lowering is exempt from relativized minimality follows from the definition of relativized minimality. According to Rizzi (1990), W can induce relativized minimality effect for movement from X to Y only if W c-commands X and doesn't c-command Y.

(58)

[Y ←<sub>c-command</sub> [W<sub>c-command</sub> → X]

In case of lowering from X to Y, W doesn't induce a relativized minimality effect because W doesn't c-command X and c-commands Y.

(59)

[X $\leftarrow_{\text{c-command}}$  [W $\rightarrow_{\text{c-command}}$  Y]

Due to these reasons LF-raising is subject to relativized minimality but LF-lowering is not.

The next important claim that B&T make is that scrambling in Russian is also exempt from relativized minimality effects and, hence, is also semantically vacuous parallel to Japanese.

According to B&T, Russian meets the requirements to be tested for the availability of semantically vacuous scrambling, since it uncontroversially has overt wh-movement and allows scrambling.

- (60) a. \*Kto ty videl kogda pod'ezžal ? (Bošković & Takahashi, 1998:359)  
Who you saw when came  
'Who did you see when came?'  
b. Ty doktor videl kogda pod'ezžal  $t_i$ ?  
You<sub>Nom</sub> doctor<sub>Nom</sub> saw when arrived  
'Did you see when the doctor arrived?'

The example in (60)a demonstrates that wh-extraction across a [CP] out of embedded question is ungrammatical. However, as we can see in (60)b, subject dislocation from the relative clause into the matrix clause across the interrogative [CP] is acceptable. The subject dislocation is not sensitive to weak islandhood as would be expected if the dislocation involved focus movement.

Based on the facts above, B&T conclude that the Russian has a semantically vacuous type of scrambling, similar to Japanese.

In general, the idea that long distance scrambling in Russian is semantically vacuous has never been introduced in the literature before. As was already discussed in the previous section, there have been attempts to argue that scrambling in Russian is interrelated to discourse (Junghans Zybatow (1997), Bailyn (2002)) and that scrambling has to do with the theme- rheme distinction (Krylova & Khavronina(1976)). Thus, the claim that long distance dislocation in Russian is semantically vacuous contradicts a number of previous assumptions.

However, long distance scrambling in Russian has never been given considerable attention to in the literature and there are number of questions left unexplained with respect to the examples in (60).

B&T do not comment upon why long distance subject dislocation in Russian is grammatical, while the most famous characteristic property of long distance scrambling in Japanese is that it does not apply to subjects. Moreover, it is unclear why semantically vacuous dislocation in Russian has to proceed to a clause- medial position, while for Japanese it is clause- initial position that is occupied by the scrambled element.

These questions as well as the general characteristics of long distance scrambling in Russian will be addressed in this paper later.

To sum up the discussion above, B&T (1998) propose a base generation analysis of long distance scrambling in Japanese. On the one hand, their analysis is superior to the movement analysis because it allows to account for the number of properties that can not be resolved under the movement approach. On the other hand, their analysis involves LF lowering, which is a conceptually new assumption both with respect to Japanese and Russian.

The question remains if it is justified to assume an LF operation to account for the flexibility of word order and whether this analysis can be extended to analyze a free word order phenomenon cross-linguistically. I will point to more problems with B&T's (1998) analysis later in this section.

Now I turn to the discussion of the movement approach to scrambling.

### **3.2 Movement approach to scrambling**

The classical analysis of scrambling involves movement. Since this paper mainly addresses the properties of scrambling in Russian, I am going to discuss one of the movement approaches to scrambling that is based on the data from Russian. By far, the most extensive generative literature on this issue is written by Bailyn (2001), (2004).

Bailyn (2001) makes an attempt to create a universal derivational system of scrambling that would be able to capture the distinct properties of local and long distance scrambling both in Russian and Japanese, as well as to solve the problem of optional movement for scrambling in general.

The derivational system proposed by Bailyn (2001) is based on three distinct assumptions independently proposed in the literature.

(Bailyn, 2001:2)

- Local Scrambling occurs to satisfy the EPP. EPP is an overtness condition on the [Spec, IP].
- Long distance scrambling is discourse- driven. It is often assumed to be adjunction to [IP] but can also be implemented as movement to the [Spec]-position of a high functional category.
- The notion of a purely derivational system where all conditions on linguistic expressions apply derivationally (this assumption is adopted from Epstein et al (1998)).

The subcomponents relevant to the analysis are the “multiple- spell- out hypothesis” and the notion of feature splitting in the sense of Saito (2001), whereby only the formal features attracted by the head move and others do not move or get deleted.

The main problem Bailyn (2001) addresses in his analysis is how to unify the notions of A- and A'- scrambling. According to Bailyn (2001), the distinct properties of these processes should naturally follow from a uniform derivational system.

Among the general properties of long distance scrambling discussed in Bailyn (2001) are:

- Long distance scrambling does not change binding relations as it was shown in (41).
- It causes weak- cross over violations as we have seen in (42).
- It is parallel to wh-movement with respect to subjacency.

(61)a.\* Kogo ty pozvonil agentu kotoryj ljubit? (Bailyn, 2002:5)

Whom<sub>Acc</sub> you phone spy<sub>Dat</sub> who loves  
 ‘Whom did you phone a spy who loves?’

b.\* Boris<sub>i</sub> ty pozvonil agentu kotoryj ljubit t<sub>i</sub> !

Boris<sub>Acc</sub> you phone a spy who loves  
 ‘It’s Boris you phoned a spy who loves!’

For local scrambling, the properties assumed are radically distinct:

- The scrambled constituent can bind an anaphor it didn’t previously c-command. See example (37) for the relevant facts

- Local scrambling does not cause weak-cross over effects. See example (39) for the relevant facts.

The question Bailyn (2002) raises is why the two processes are called ‘scrambling’, given the fact that they possess distinct properties.

Bailyn proposes that the two types of scrambling should be treated separately with respect to the triggers of the movement. Local scrambling should be treated as an instance of EPP- driven substitution into the [Spec, IP] position, while long distance scrambling should be treated as an instance of discourse- driven movement targeting the left edge of the clause.

### 3.2.1 Derivational schema of scrambling behaviour. Bailyn (2002)

Now I turn to the discussion of Bailyn’s (2002) derivational schema itself.

The derivational schema that Bailyn (2002) assumes is based on the idea that interpretations are built up derivationally. He adopts this assumption from Saito (2001) as well as the following ideas about the mechanism of derivation. (Bailyn, 2001:15-16)

- Assume copy theory of movement
- Assume XP arguments have these features: [P](PF-relevant), [D], [Op] (LF-relevant)
- Assume wh-movement and long distance scrambling are driven by [Op]-feature
- Assume inversion is driven by [D]-feature

Bailyn proposes that [NP]s are interpreted and enter into binding relations at any point of the derivation where their [D]-features are active.

#### (62) Derivational schema of scrambling behaviour

##### a. EPP-driven scrambling: (local,A)

[<sub>IP</sub> **XP<sub>i</sub>** [<sub>D</sub>,<sub>P</sub>] [<sub>r...t<sub>i</sub></sub> [<sub>D</sub>,<sub>P</sub>] ...]])

##### b. Discourse-driven scrambling: (long,A’)

[<sub>IP</sub> **XP<sub>i</sub>** [<sub>P</sub>,<sub>OP</sub>] [<sub>IP...t<sub>i</sub></sub> [<sub>D</sub>,<sub>P</sub>][<sub>OP</sub>] ...]]) (Bailyn, 2001: 17)

As was already mentioned above, Bailyn argues that local scrambling is EPP-driven. The EPP, in turn, is triggered by the strong nominal [D]- feature. For cases of local scrambling, Bailyn assumes that the [D]- feature is active in the landing site, hence, the interpretive component interprets the moved [XP] high (see (62)a). This factor allows to account for the A-properties of local scrambling.

As for long distance scrambling, Bailyn assumes that it is discourse- driven. Similar to wh-movement, long distance scrambling is triggered by an [Op]-feature. The nominal [D]-feature, crucially, does not have to raise and the nominal [XP] is interpreted low, deriving A'-properties of long distance scrambling.

In this way, Bailyn derives the distinct properties of local and long distance scrambling within one derivational schema.

However, there are number of issues that are unclear with Bailyn's approach. For example, Bailyn claims that his schema can unify the properties of scrambling in Russian as well as in Japanese. However, he does not discuss cases where long distance scrambling in Japanese is argued to be semantically vacuous. His schema presupposes that all cases of long distance scrambling are discourse-driven, thus, it is puzzling how the semantically vacuous cases (if they exist) can be incorporated in his schema.

Moreover, there are cases in Japanese where local scrambling exhibits A'-properties. These cases are also left unexplained. I come back to these issues as well as to other problems with Bailyn's account in the next section.

To conclude the section, I want to point out that both the base generation and the movement approach have their strengths and weaknesses. At first sight, it looks as though the base generation analysis of scrambling is superior to the classical movement analysis with respect to the amount of issues resolved. However, the lowering operation it assumes is quite radical and, thus, it requires solid theoretical argumentation as well as evidence cross-linguistically. The question whether the LF- lowering analysis of scrambling can be assumed for other languages than Japanese remains open. Bailyn's derivational approach to scrambling has the opposite problems. The amount of data it can account for is limited to the explanation of a number of effects for Russian as well as for Japanese. The claim that Bailyn's derivational schema of scrambling can be applied cross- linguistically appears to be too strong. Thus, at the end neither of the approaches is superior to the other.

## 4 The debate about scrambling

In the previous section, I have discussed two radically different approaches to the derivation scrambling: the base generation approach by B&T (1998) and the overt movement approach by Bailyn (2001). These two accounts represent two opposing stands on the issue. The debate between these two opposite approaches concerns not only the derivation of scrambling, but also its properties cross-linguistically as well as the definition of scrambling itself.

In this section, I address the points of disagreement between the overt movement approach and the base generation approach represented by Bailyn (2001) and B&T(1998) respectively. I will discuss the arguments and data they introduce in support of their hypotheses and make an attempt to assess both of the accounts critically, revealing their weak and strong points.

### 4.1.1 Theoretical issues

The main and very important point of disagreement between Bailyn (2001) and Bošković (2004) concerns the definition of scrambling.

For Bailyn (2001), long distance scrambling in Russian is a discourse-driven movement. He assumes that Russian is “a clearly scrambling language in the original stylistic sense” (Bailyn, 2001:641) and he posits that the properties of scrambling found in Russian should also be assumed for Japanese.

Bošković (2004) doesn’t define the term scrambling explicitly, but what follows from his argumentation is that scrambling in its pure form is a type of dislocation to the left that doesn’t have any semantic import, does not change the scope relations and applies freely, namely, is not feature or discourse- driven. This kind of scrambling, Bošković assumes, is available in Japanese and in Russian.

According to Bošković (2004), the interfering factor in Russian is the absence of morphology associated with topicalization /focalization, as opposed to Japanese where the topicalized elements bear an overt topic marker–*wa*. Thus, Bošković posits that there are two main theoretical possibilities for Russian<sup>9</sup> “the freedom of word order is

---

<sup>9</sup> SC also, but I will not discuss SC in this paper in any detail.

the result of topicalizing/focalizing movements, coupled with some optionality regarding subject and object A- raising” (Bošković, 2004: 619) or Russian has focalization, topicalization and Japanese style scrambling (JSS). JSS in Russian would have to be very illusive, because discourse- driven movement and scrambling look similar. However, Bošković argues that it is possible to tease the two apart and the availability of JSS in Russian will become obvious. Focalization, topicalization options can be eliminated if the construction is tested with respect to islandhood, namely, relativized minimality.

A test that could unambiguously reveal JSS is Saito’s (1992) test (see (15) for the relevant facts). However, Bošković (2004) claims that this testing is not available for Russian. The interfering factor in Russian is the clause mate requirement on wh-phrases in Russian: “wh- phrases must be clause mates in overt syntax with the +wh [C] heading the [CP] where they are interpreted” (Bošković, 2004:625). Russian is a multiple wh-fronting language, where all wh- phrases front and get interpreted in the fronted position. Thus, the structures similar Saito’s (1992) would violate the clause mate requirement for wh-phrases in Russian. Therefore, Saito’s (1992) test can not be applied to scrambling in Russian

An important theoretical issue that Bailyn discusses is the problem of optionality. B&T claim that they dispense with the problem of optionality of scrambling because under their analysis no movement operation is involved. Merge as opposed to Move is not subject to last resort. However, Balyin points out that B&T’s theory allows more than one position at Merge. According to Bailyn, B&T’s account does not solve the problem of optionality, but transfers it to the base structure level.

Bošković, in turn, points out that he makes a distinction between Merge of functional and lexical elements, reasoning that lexical Merge (what we choose or want to say) is not subject to last resort, while functional Merge (building of legitimate grammatical structure ) is restricted by the principle of last resort.

Bošković discusses several assumptions concerning lexical insertion. Chomsky (1995) assumes that no aspect of lexical insertion, including pure Merge, is subject to last resort. Chomsky (2000), on the contrary, assumes that pure Merge is subject to last resort, thus, enriching the theory of selection with this assumption. Bošković (2004) takes a middle stand on this issue assuming that only pure Merge of functional elements is subject to last resort. (Bošković, 2004:632)

Another theoretical consequence that, according to Bailyn, follows from B&T's theory is that it becomes licit to generate arguments in an adjoined position. Given the legitimacy of base generation in adjoined positions, it is expected that multiple scrambling construction should be commonplace, with no preference for the constructions where arguments appear in their  $\Theta$ -positions. However, Bailyn points out that multiple scrambling constructions in Russian are rare and often ungrammatical for many Russian speakers.

- (63)a. Ivan hočet čtoby Boris peredal kassetu Saše .

Ivan<sub>Nom</sub> wants that Boris gave cassette<sub>Acc</sub> Saša<sub>Dat</sub>

'Ivan wants Boris to give the cassette to Saša.'

- b.\* Ivan Saše<sub>i</sub> kassetu<sub>j</sub> hočet čtoby Boris peredal t<sub>i</sub> t<sub>j</sub>.

Ivan<sub>Nom</sub> Saša<sub>Dat</sub> cassette<sub>Acc</sub> wants that Boris gave (Bailyn, 2001:649)

Under the LF lowering account both (63)a and (63)b should be treated as equally grammatical, while under the movement account (63)b is ruled out by relativized minimality. Hence, the effect in (63) is not predicted by B&T's theory, while given the classical assumption that scrambling is A'-movement, the effect is expected.

According to Bošković, the fact that multiple dislocation is disallowed in Russian can indicate that Russian has only topicalization / focalization option, with JSS not available. However, his informants find the example (63) grammatical in case the subject *Ivan* in the matrix clause is substituted by a pronoun. (Bošković, 2004:621)

- (64) On Saše<sub>i</sub> kassetu<sub>j</sub> hočet čtoby Boris peredal t<sub>i</sub> t<sub>j</sub>.

He<sub>Nom</sub> Saša<sub>Dat</sub> cassette<sub>Acc</sub> wants that Boris gave

'He wants Boris to give the cassette to Saša.'

Given the grammaticality of (64) Bošković reasons, that the example (63) pointed out by Bailyn is an instance of topicalization, while the example in (64), in fact, involves scrambling.

As a Russian speaker, I think that (63)b is degraded, but not ungrammatical as Bailyn (2001) reports. However, I prefer (64) over (63). I think the contrast between the two is not predicted neither by Bailyn (2001) nor Bošković (2004). Therefore, the facts can not be viewed as favoring one assumption over the other.

Moreover, it is unclear why the change of the category in the subject position is so crucial for grammaticality of structures in (63)-(64). If Bošković's reasoning is right and (64) is an instance of JSS, it becomes totally unexpected that scrambling can not apply across a nominal subject as opposed to pronominal one.

#### 4.1.2 Predictions and data

In his reply to B&T (1998) paper, Bailyn (2001) points out three predictions that the LF lowering analysis makes and shows that all of the three predictions are falsified.

The first prediction is that scrambling should have no interpretive effect. He argues that the claim about the lack of interpretive effects associated with the scrambled position is not true for Russian. Bailyn introduces evidence that the scrambled order of elements contrasts with the unscrambled one in scope interpretation.

- (65) a. Kto-to hočet, čtoby Boris uvidel každого malčika. (Bailyn, 2001:642)

Someone-Nom wants that Boris saw [every boy]<sub>Acc</sub>

‘Someone wants Boris to see every boy’

- (i)  $\exists x > \forall y$  (ii)  $*\forall y > \exists x$

- b. Každого malčika kto-to hočet, čtoby Boris uvidel.

[Every boy]<sub>Acc</sub> someone<sub>Nom</sub> wants that Boris saw

‘Every boy, someone wants Boris to see’

- (i)  $*\exists x > \forall y$  (ii)  $\forall y > \exists x$

The effect in (65) is unexpected under B&T’s account. Under the LF-lowering assumption, the scrambled order should not produce changes in scope relations, since the dislocated element undergoes obligatory lowering into the  $\theta$ -position. The scrambled position is not interpreted at LF, hence, it cannot be available for scope interpretation.

Contrary to that, Bošković (2004) points out that in Japanese, unlike the facts reported for Russian in (65), the quantifier can not take wide scope in the scrambled position. The relevant example is repeated below.

- (66) Daremo-ni<sub>i</sub> dareka-ga [Mary-ga t<sub>i</sub> atta to] omotteiru.

Everyone<sub>DAt</sub> someone<sub>Nom</sub> Mary<sub>Nom</sub> met that thinks.

‘Everyone, someone thinks that Mary met.’

- (i)  $\exists x > \forall y$  (ii)  $*\forall y > \exists x$  (Bošković, 2004:614)

Bošković notes that, given the comparison of (65) and (66), neither the base generation can approach, nor the overt movement approach provide a unifying account for the observed properties. Thus, he suggests that the distinction between the effects in (65) and (66) should not be attributed to the inadequacy of the analysis but to the possibility that the phenomena are different in nature. Bošković reasons that, given the

absence of the undoing effect in (65), it is likely that the structure presented by Bailyn is not an instance of scrambling, but, in fact, involves topicalization.

As a Russian speaker, I agree on the judgement reported in (65). However, I would hesitate to interpret the above data as an argument in favour of any of the approaches. It has been previously discussed in the literature that the tests that appeal to quantifier/ existential scope effects are very problematic for Russian. Due to the absence of overt articles, one has to use other means (like the indefinite pronoun *kto-to* in this case) to create an effect parallel to that in English, which interferes with the naturalness of the reading and confuses the informant.

The next prediction of B&T's (1998) analysis that Bailyn discusses is that scrambling should be generally unconstrained in terms of locality.

Bailyn claims that this prediction is not born out for Russian. He argues that scrambling patterns with wh-movement with respect to locality constraints. Consider an example below. (Bailyn, 2001:646)

- (67) a.\* Kogo<sub>i</sub> Marina znaet [čto[ Ivan ljubit t<sub>i</sub> ]]?

Who<sub>Acc</sub> Marina<sub>Nom</sub> knows that Ivan<sub>Nom</sub> loves

'Who does Marina know that Ivan loves?'

- b.\* Borisa<sub>i</sub> Marina znaet [ čto [Ivan ljubit t<sub>i</sub> ]]?

Boris<sub>Acc</sub> Marina<sub>Nom</sub> knows that Ivan<sub>Nom</sub> loves

'Marina knows that Ivan loves Boris?'

Bailyn assumes that (67)b is ruled out by relativized minimality similar to (67)a. He also discusses a correlation between other constraints on movement and scrambling and argues that both wh-movement and scrambling are subject to subjacency (see examples in (61)) and coordinate structure constraint (see (68) below). These facts are not expected under the base generation assumption and Bailyn views these facts as evidence in favor of movement analysis.

- (68) a.\* Kogo ty hočeš, čtoby Ivan videl [t<sub>i</sub> i Mašu]? (Bailyn, 2001:647)

Whom<sub>Acc</sub> you want that Ivan saw and Maša<sub>Acc</sub>

- b.\* Borisa ty hočeš, čtoby Ivan videl [t<sub>i</sub> i Mašu]?

Boris<sub>Acc</sub> you want that Ivan saw and Maša<sub>Acc</sub>

Bošković (2004), in his turn, notices that, in the discussion of island effects, B&T (1998) mostly concentrate on the relativized minimality effects, which are considered to be well understood in the current theory. Bošković argues that such island effects as the adjunct condition are not understood well enough and can not be used for

teasing the movement and the base generation approaches apart. Due to the same reasoning, Bošković does not treat (68) as a strong argument, because the nature of the coordinate structure constraint itself is a disputable point. He refers to Munn's (1993) analysis where it is shown that the coordinate structure constraint in fact doesn't correlate with the syntactic locality, but it is a constraint on semantic interpretation unrelated to movement.

Bošković also refers to other analyses like Müller and Sternefeld (1993), Müller&Sternefeld (1996), where they introduce evidence that scrambling behaves differently from wh-movement with respect to a number of locality constraints. For example, wh- movement can take place out of subjunctive, but not out of indicative clauses, while scrambling can do both.(see (47),(48) for the relevant facts)

The third prediction of B&T's theory that the LF lowering analysis makes is that only elements that can be assigned a  $\Theta$ -role can participate in scrambling. Since covert movement is driven by  $\Theta$ -checking, it follows that the elements that do not bear a  $\Theta$ -role can not be base generated in the scrambled position. Bailyn argues that this prediction is not correct for Russian because non- arguments can scramble in Russian.

- (69) a. Ja *bystro<sub>i</sub>* hoču, [čtoby oni *t<sub>i</sub>* dopisali kursovy]. (Bailyn, 2001:648)

I quickly want that they wrote papers

'I want them to write their papers quickly'

- b. Ja *zelenuju<sub>i</sub>* hoču, [ čtoby ona kupila *t<sub>i</sub>* knigu].

I green <sub>Acc</sub> want that she bought book<sub>Acc</sub>

'I want her to buy the green book'

According to Bailyn, in (69)a the adverb 'bystro' modifies the embedded verb, though it is dislocated into the main clause. Given the fact that the adverb is interpreted low, and since it is generally assumed that adjuncts do not bear a  $\Theta$ -role, the motivation for LF lowering of the adverb is unclear.

Bošković treats the example in (69) as irrelevant, because for him it shows nothing but the fact that adverbs can be topicalized /focalized in Russian. For Bošković, the very possibility of adverb dislocation in Russian already eliminates the possibility that scrambling might be involved in this type of construction.

I think that in (69)a the adverb is not modifying the embedded verb, as Bailyn reports, but it modifies the verb in the matrix clause with the scope available only in the matrix clause. This fact shows that the adverb is base generated in the matrix clause and is topicalized, which can be considered as an argument in favour of Bošković's (2004)

hypothesis. Interestingly, the only possible interpretation available for (69)b, is the one where the adjective scopes in the lower clause. Since it is unclear whether (69)b involves a reconstruction effect observed with A'-movement or it is an instance of radical reconstruction, this observation does not give evidence in favour of any of the approaches.

#### 4.1.3 Against semantically vacuous scrambling in Japanese

Bailyn (2001) elaborates on the analysis that he suggests is plausible for scrambling in Russian as well as in Japanese. He makes the following generalizations with respect to the motivation of scrambling

(Bailyn (2001:654).

- (70) a. A'-scrambled and non scrambled orders are always associated with different discourse /informational interpretations.  
b. The movement deriving scrambled orders is motivated by discourse /informational considerations.

As I already discussed, according to Bailyn, these generalizations are true for Japanese. Consider the examples below. Bailyn, 2001:653)

- (71) a. John-ga [CP Mary-ga sono hon-o katta to] omotteiru.  
          John<sub>Nom</sub>     Mary<sub>Nom</sub> that book<sub>Acc</sub> bought that think  
          ‘John thinks that Mary bought that book’

b. [Sono hon-o<sub>i</sub> John-ga [CP Mary-ga t<sub>i</sub> katta to] omotteiru.  
          That book<sub>Acc</sub> John<sub>Nom</sub>     Mary<sub>Nom</sub> bought that think  
          ‘That book<sub>i</sub>, John thinks that [that Mary bought t<sub>i</sub>]’

Bailyn argues that (71)b is appropriate in the context, where ‘the book’ is part of the informational background. Bailyn claims that in Japanese as well as in Russian, the scrambled and non scrambled orders have distinct discourse representation and introduces the following examples in (72) that allow him to conclude that scrambling in Japanese is, in fact, interconnected with discourse and theme- rheme distinction.

- (72) a. John-wa dou shiteiru no? (Bailyn, 2001:653-4)  
           John<sub>Top</sub> how doing Q  
           ‘How is John doing?’

b. John-ga [CP Mary-ga sono hon-o katta to] omotteiru.  
       John<sub>Nom</sub> Mary<sub>Nom</sub> that book<sub>Acc</sub> bought that think

‘John thinks that Mary bought that book’

- c. \* [Sono hon-o<sub>i</sub> John-ga [CPMary-ga t<sub>i</sub> katta to]omotteiru.

That book<sub>Acc</sub> John<sub>Nom</sub> Mary<sub>Nom</sub> bought that think

‘That book<sub>i</sub>, John thinks that [that Mary bought t<sub>i</sub>]’

- (73) a. Sono hon ni-kanshite nani-ka atta no?

that book about something happened Q

‘Did anything happen to that book?’

- b. \* John-ga [CPMary-ga sono hon-o katta to] omotteiru.

John<sub>Nom</sub> Mary<sub>Nom</sub> that book<sub>Acc</sub> bought that think

‘John thinks that Mary bought that book’

- c. [Sono hon-o<sub>i</sub> John-ga [CPMary-ga t<sub>i</sub> katta to]omotteiru.

That book<sub>Acc</sub> John<sub>Nom</sub> Mary<sub>Nom</sub> bought that think

‘That book<sub>i</sub>, John thinks that [that Mary bought t<sub>i</sub>]’

In the example (72) ‘John’ is a theme of the discourse and the appropriate response does not need to have the structure with ‘that book’ element fronted (see the contrast between (72)b and (72)c). Conversely, the example in (73) introduces the element ‘that book’ as a theme, thus the structure of the answer is more appropriate with the scrambled order fronting the theme rather than without.

However, the contrasts that Bailyn presents with respect to Japanese are not very informative. My Japanese informants think that the question in (72)a is incompatible with both of the answers in (72). They prefer to change the question in (72)a *John-wa dou shiteiru no?* (How John is doing?) to the one more compatible with the answers in (72)b,c: *John -wa dou omotteiru?* (What does John think?). If the question is changed, they choose the answer in (72)c.

Moreover, some Japanese speakers point out that asking the question with -wa morpheme requires the use of a wa-morpheme in the answer. This is exactly what Bailyn is trying to avoid, because wa- morpheme is a topicalization marker in Japanese and he is trying to make an argument about scrambling in Japanese, excluding topicalization option.

As for the contrast in (73), the Japanese speakers unanimously prefer (73)b, contrary to what Bailyn informs. Moreover, some of my informants pointed out that they would even choose to drop the object in (73)c. So what we can observe is that Japanese allows the argument drop.

For the reasons listed above, I assume that the conclusion Bailyn draws with respect to scrambling in Japanese based on these particular examples is premature.

## 4.2 Elusive scrambling and the lack of evidence

In this subsection, I will make an attempt to evaluate Bailyn's (2001) and Bošković's (2004) approaches. I will address the weak points of the accounts and establish the reasons for my own investigation of the problem.

### 4.2.1 The problems with Bailyn's (2002), (2001) accounts

The first issue I would like address is the definition of scrambling. In the debate summarized above, we can see that two distinct definitions of scrambling clash. For Bailyn, scrambling is a discourse- driven movement, for Bošković (2004), scrambling is a derivation principally different from discourse- driven movement, but, nevertheless, the debate proceeds and as expected the discussion does not come to any conclusion. An important factor to be noted is that Bailyn supports his arguments by examples from Russian with minor evidence from Japanese, while Bošković (2004) relies on the analysis by B&T, which is mostly based on data from scrambling in Japanese. The reasonable conclusion that one could draw is that their definitions are distinct because the nature of the phenomena they describe is different. However, none of the debaters want to accept this and proceed with their argumentation.

Coming back to Bailyn's definition of scrambling, it appears to be unclear why he, in fact, needs the term. If, according to him, long distance scrambling is discourse-driven movement and local scrambling is EPP- driven movement, then the term scrambling appears to be redundant. Discourse- driven movement is generally labeled as topicalization/focalization processes and EPP- driven movement is raising. Suppose that by referring to scrambling, Bailyn means something more than the usual discourse driven movement, but then the additional properties should be clearly defined and supported by the data. Unfortunately, I could not find this clarification in Bailyn's papers and the usage of the term 'scrambling' was left ambiguous for me.

When it comes to the properties of scrambling, Bailyn is inconsistent in his argumentation. In his (2002) paper he draws a parallel between scrambling in Japanese and Russian, arguing that they have similar properties. For example, he argues that for local dislocations both in Russian and Japanese the scrambled element can bind an

anaphor it didn't previously c-command (see (37) for the relevant facts, also (Bailyn, 2002:5).

In contrast to that, in his reply to B&T's paper, Bailyn (2001) argues that scrambling in Russian has the properties distinct from the ones pointed out for Japanese. For example, Bailyn argues that unlike in Japanese, non-arguments can scramble in Russian (see (69)).

Thus, it is unclear why, in his reply to B&T's paper, Bailyn is so convinced that Russian scrambling has the properties distinct from the Japanese counterpart, while in his previous work, namely, *Derivational Approach to Russian scrambling*(2002), he draws numerous parallels between Russian and Japanese with respect to both local and long distance scrambling.

The argument about the possibility of scrambling adjuncts in Russian also turns out to be ambiguous. It is not clear whether Bailyn means that within his understanding of scrambling as discourse driven movement adjuncts can scramble, or that Russian has the type of scrambling described by B&T but with the additional characteristic of having the possibility of adjunct scrambling.

It seems that Bailyn completely disregards the properties of Japanese scrambling as described by B&T and insistently tries to unify these two phenomena, ignoring the major distinctions in their properties, which he himself points out in his reply paper.

Bailyn denies the existence of JSS not only for Russian but also for Japanese. But his evidence from Japanese is too poor to make this claim. Moreover, as I already discussed earlier, some Bailyn's examples, namely, examples of discourse effects, are poorly constructed and, therefore, are not informative on the issue.

Bailyn doesn't introduce any other piece of evidence that could unambiguously show that, contrary to B&T, JSS is non existent in Japanese, so the claim he makes about discourse- driven nature of scrambling in Japanese is vacuous.

In his derivational approach to Russian scrambling Bailyn follows Junghans & Zybatow (1997) in the assumption that long distance scrambling is discourse-related movement that is semantically vacuous only with respect to binding effects. It is a generally accepted fact that topicalized /focalized elements cannot antecede an anaphor, but there is evidence from English that they can be bound.

(74) John<sub>i</sub> wondered [which pictures of himself<sub>i/k</sub> [ Bill<sub>k</sub> likes t]].

Given Bailyn's assumption that scrambling is A'-movement, it is expected that the effect similar to (74) should be available in Russian.<sup>10</sup>

What also follows from Bailyn's derivational schema of scrambling is that long distance scrambling in Russian should always reconstruct. The [D]-feature is active in the launching site, and, given the assumption that "NPs are interpreted and enter into binding relations at any point of the derivation where their [D]-features are active" (Bailyn (2001:17), the scrambled element should always be interpreted in the lower position. This appears to be true for Japanese examples, but Bailyn does not give a single example from Russian that would support his analysis. Given the absence of evidence from Russian, the claim that discourse driven movement in Russian always reconstructs appears to be premature.

To sum up, considering the arguments pointed out above, Bailyn's analysis cannot be considered complete. Bailyn's definition of scrambling is obscure at times, he tends to generalize over the distinct properties of scrambling in Russian and Japanese, and the analysis he offers leads to more confusion and uncertainty.

#### 4.2.3 The problems with Bošković's (2004) account

Now let's address the weak points within B&T's (1998), Bošković's (2004) argumentation. Contrary to Bailyn, B&T have a clear idea what kind of properties defines scrambling. Bošković (2004) points out that the use of the term 'scrambling' requires careful and justified treatment and before identifying some derivation as scrambling, one needs to verify if the construction doesn't involve an instance of discourse- driven movement.

However, Bošković (2004) doesn't demonstrate this careful treatment of the phenomena with respect to his argument concerning Russian. For Russian, Bošković (2004) argues, it is problematic to tease discourse-driven movement and scrambling apart, because Russian doesn't mark the topicalization/ focalization morphologically.

Moreover, the behavior of these two constructions is similar and "the dislocated elements in Russian can do everything that both topicalized /focalized elements can do" (Bošković, 2004:621). Given the situation in Russian, where scrambling appears to be

---

<sup>10</sup> We will return to these examples in the next section.

completely ambiguous with discourse- driven movement, it is important that the claim that Russian has JSS is supported by sufficient evidence.

Though Bošković (2004) talks about the availability of tests that can unambiguously reveal JSS in Russian, the only test he applies is the relativized minimality test (see (60) for the relevant facts).

Bošković (2004) doesn't comment upon the fact why both of the examples involve dislocation to the clause medial position. It is unclear whether this position should be assumed for JSS in Russian or whether it is a pure coincidence that JSS targets this particular position in his examples.

If JSS doesn't have a fixed position in Russian, the effect is expected to remain the same for cases when NPs are dislocated to clause initial position. The relevant examples are absent from Bošković's paper.

As I have already discussed earlier, the distinctive property of Japanese scrambling is the fact that subjects can not scramble long distance in Japanese. Hence the observation that subjects scramble long distance in Russian is puzzling,

Bošković (2004) also does not elaborate on the idea if the initial unscrambled order has the same interpretation as the scrambled one. If we suppose that (60)b is an instance of JSS, then we shouldn't expect any difference in interpretation. However, B&T as well as Bošković (2004) do not provide evidence that would show that both of the scrambled and non- scrambled orders are appropriate in the same context. Again, given the absence of evidence, the uncertainty remains.

Since according to B&T, the main characteristic of JSS is the radical reconstruction property, it should be possible to apply reconstruction tests to reveal if the dislocated phrases in have this property. For Russian the use of negative polarity items (NPI) and genitive of negation test can be informative in this respect. None of the relevant tests are applied by B&T or Bošković (2004).

Bošković (2004) discusses the fact that Saito's (1992) reconstruction test (wh-phrase extraction out of [CP] in Japanese) is not applicable to Russian due to interfering factors. Thus, it seems that for Bošković (2004) there are only two reliable tests that could reveal JSS in Russian and one of them is unavailable.

To sum up, Bošković (2004) is not consistent in his argumentation, because on the one hand he talks about the availability of tests, that can unambiguously identify JSS, but on the other hand his claim that JSS exists in Russian is not supported by due evidence.

Hence, the reader is given two options: one can either trust Bošković's (2004) claim that Russian has JSS, assuming that the evidence is available, but not explicitly given, or one could test the construction introduced by Bošković (2004) as an instance of JSS and check if the properties assumed for JSS can, in fact, be confirmed for this construction.

I choose the latter and in the next section I am going to test the constructions exemplified by Bošković (2004) as JSS in order to answer the question if JSS exists in Russian or if long distance dislocation in Russian results only from topicalizing /focalizing movements.

#### *Summary*

In this section I have summarized two radically different approaches to scrambling: base generation approach by B&T (1998) and movement approach by Bailyn (2001). I have discussed the mechanisms B&T (1998) and Bailyn (2001) assume for their analyses and listed their argumentation in favor of their hypotheses.

I have pointed to certain problems with both of the analyses and concluded that neither of the two analyses can be considered as superior to the other.

The analysis proposed by Bailyn (2001) strongly relies on the assumption that scrambling both in Russian and Japanese occur for discourse consideration. His analysis tends to generalize over certain properties of scrambling in Russian and in Japanese and, thus, is not complete. The analysis suggested by Bošković (2004), in its turn, is incomplete in terms of evidence presented and leaves a lot of issues unexplained.

## 5 Long distance scrambling in Russian. New properties revealed

### 5.1 Japanese style scrambling in Russian. Radical reconstruction?

In this chapter, I will make an attempt to answer the question if Russian has Japanese style scrambling. I will test the constructions pointed out by Bošković (2004) as examples of semantically vacuous scrambling in Russian and check if the properties of long distance scrambling assumed for Japanese are found in Russian.

The examples are repeated below in (75) and (76), (Müller and Sternefeld, 1993:467-8).

(75) a\* Kto ty videl kogda pod'ezžal ?

Who you saw when came

b.?\*Čto vy videli kak zapakovali?

What you<sub>Pl</sub> saw how (they) did up

(76) a. Ty [doktor]<sub>i</sub> videl kogda pod'ezžal t<sub>i</sub>?

You<sub>Nom</sub> doctor<sub>Nom</sub> saw when arrived

'Did you see when the doctor arrived?'

b. Vy [posylku]<sub>i</sub> videli kak zapakovali t<sub>i</sub>.

You<sub>Nom</sub> parcel<sub>Acc</sub> saw how packed

'You saw how they packed the parcel'

Bošković (2004) claims that the example in (76) shows that long distance dislocation of an [NP] in Russian, as opposed to long distance wh-movement (see (75)), is not subject to relativized minimality. This fact gives evidence that a dislocated [NP] in Russian is subject to radical reconstruction and this observation is treated as an argument in favor of the existence of semantically vacuous scrambling in Russian.

It is assumed that the main property of Japanese style scrambling is a radical reconstruction property and long distance scrambling in Japanese is also not sensitive to wh-islands. Based on these facts, it is argued in Bošković (2004) scrambling in Russian and in Japanese exhibit a phenomenon of the same type, namely, semantically vacuous scrambling.

### 5.2 Semantic effects

Japanese style scrambling is claimed to have no semantic effect and the claim is that Russian long distance dislocation does not have one either. If scrambling is

semantically vacuous, then both the scrambled order and non-scrambled order should be felicitous in the same type of context.

I will make an attempt to create the context for the examples pointed out by Bošković (2004) (see examples in (76)) as an instance of JSS in Russian and check whether the scrambled order requires a change of context and, thus, has an effect on the interpretation.

- Informational background 1

*We are present at the crime investigation. A murderer disguised as a doctor managed to get into the victim's apartment, violently slaughtered the old lady and stole jewellery and money. We are questioning the neighbour about what he has seen.*

The neighbour:

*I saw a pizza boy arrive at 5, then there was a cleaning lady who comes to Johns' every Monday at 6....*

The question of the detective:

(77)

- a. Vy [doctor] videli kogda pod'ezžal?

You [doctor]<sub>Nom</sub> saw when arrived  
‘The doctor, did you see him arriving?’

- b. \*Vy videli kogda pod'ezžal [doctor]?

You saw when arrived [doctor]<sub>Nom</sub>  
‘Did you see when the doctor arrived?’

The answer:

(78)

- a. Doktora ya ne videl.

The doctor I haven't seen .

- b. \*V 7 večera ya videl doktora

At 7p.m. I have seen a doctor

- Informational background 2

*We are present at the same investigation. We found a witness that claimed that he has seen a doctor arriving on that day. We are interested in the time of the event.*

The question of the detective:

(79)

- a.\*Ty [doctor] videl kogda pod'ezžal?

You [doctor]<sub>Nom</sub> saw when arrived

b. Ty videl kogda pod'ezžal doctor?

You saw when arrived [doctor]<sub>Nom</sub>

The answer of the witness:

(80)

a. Ya dumayu videl ego okolo 7.

I think I saw him around 7p.m.

b. ??Doktora ya videla okolo 7.

The doctor I saw around 7 p.m.

There are two important observations from the informational backgrounds above and the compatibility of word orders. The fact to be noticed is that the scrambled word order is felicitous with background 1, but not acceptable in the background 2. The non-scrambled word order favours the background 2, and is not acceptable with the background 1. Thus, we can conclude that long distance scrambling in Russian does have an effect on the interpretation and the scrambled order can become infelicitous in one context and the only one possible in the other.

The second observation is that the informational background 1 needs to include the pair-list options in order to become appropriate for the scrambled word order. If we omit the pair-list enumerations, the scrambled order would no longer be favoured.

One more observation with respect to the examples is that the word order in the answers in (78) and (80) reflects the word order in the questions. It also shows that the word order has an effect on the interpretation and reordering of the constituents in the sentence reflects the change of the context the sentence is used in.

An important factor to mention is the fact that in Stepanov (1998) the pair list reading is discussed with respect to the interpretational aspects of multiple wh-fronting. For example, the example below, depending on the context where it is used, can trigger either a pair list answer or an individual answer.

(81) Kto čto kupil?

Who <sub>Nom</sub> what <sub>Acc</sub> bought

'Who bought what'

Stepanov (1998), adopts the assumption from Stepanović (1995) that wh-phrases are inherently contrastively focused and their fronting occurs for checking focus features. Thus, it is plausible that long distance scrambling in Russian is an instance of contrastive focus.

Since we have observed that long distance scrambling and multiple wh-fronting co-occur in the same type of informational background in Russian, the legitimate question to address would be whether there is a straightforward parallel between these two phenomena with respect to their properties. I want to postpone this issue for now.

### 5.3 Ambiguity of ‘kak’ and two possible interpretations

The second issue I would like to discuss with respect to Bošković’s (2004) examples is the structural ambiguity of ‘kak’ (how) in one of the structures (see (75), the relevant example is repeated below in (82)).

(82) a.?\*Čto vy videli kak zapakovali?

What you<sub>Pl</sub> saw how (they) did up

b. Vy [posylku]<sub>i</sub> videli kak zapakovali t<sub>i</sub>.

You<sub>Nom</sub> parcel<sub>Acc</sub> saw how packed

‘You saw how they packed the parcel’

I disagree with the judgement in (82)a. I consider the example in (82)a as well as the list in (83) below grammatical.

I think that the structure in (82)a is ambiguous between the one where ‘kak’ functions as a wh-element and the one where ‘kak’ is a complementizer. This is crucial in our case because the possibility of extraction out of wh-islands is a characteristic that defines scrambling. Hence, for cases where ‘kak’ is a complementizer, radical reconstruction property can not be straightforwardly postulated.

(83) a. Čto vy videli kak zapakovali ?

what<sub>Acc</sub> you saw how packed

‘What did you see them packing?’

b. Kogo ty slyšal kak otrugali?

Whom<sub>Acc</sub> you hear how scolded?

‘Whom did you hear them scolding?’

c. Kogo ty pomniš kak zabrali v armiyu?

Whom<sub>Acc</sub> you<sub>Nom</sub> remember how took-they in army?

‘Whom do you remember being taken to the army?’

In order to create a structure unambiguously containing ‘kak’ as a wh-element, ‘kak’ should be substituted by unambiguous ‘kakim obrazom’ (in which way).

The prediction is that if the unambiguous wh-element ‘kakim obrazom’ intervenes in the structure in the examples in (83), it should produce ungrammaticality. This is the case as (84) shows below.

(84) a.\* Čto vy videli kakim obrazom zapakovali ?

what<sub>Acc</sub> you saw how packed

‘What did you see them packing?’

b. \* Kogo ty slyšal kakim obrazom outrugali?

Whom<sub>Acc</sub> you hear how scolded?

‘Whom did you hear them scolding’

c. \* Kogo ty pomniš kakim obrazom zabrali v armiyu?

Whom<sub>Acc</sub> you<sub>Nom</sub> remember how took-they in army?

‘Whom do you remember being taken to the army?’

The next question to answer would be when ‘kak’ functions as a complementizer and when it appears as a wh-element. What are the two syntactic structures that intersect and create this ambiguity.

According to my observations (as we can see in (83)), ‘kak’ is ambiguous when it is embedded under verbs of perception.

As discussed in Boivin (1998), perception verbs like *feel*, *see*, *hear* can take clausal complements in French and English. She points out two possible interpretations for this type of constructions in French: a concrete and an imaginative reading. Consider the examples below.

(Boivin, 1998:104)

(85) Je ne le vois pas adopter cet enfant.

I neg him see not adopt this child

‘I don’t see him adopting this child’

The sentence in (85) can be paraphrased into: *I do not imagine him adopting this child.* (imaginative reading). The second possible interpretation is that the event of the adoption of the child can not be seen (concrete reading).

For Russian, both imaginative reading and concrete meaning are available for ‘kak’ constructions, as well as the additional reading, where ‘kak’ is interpreted as a wh-element.<sup>11</sup>

(86) Ya vižu kak Boris igraet na gitare.

---

<sup>11</sup> The reading where ‘kak’ functions as a wh-element can be available only if special intonation is used with the emphasis placed on ‘kak’.

I see how Boris<sub>Nom</sub> play on guitar

The possible reading of this sentence can be:

- I am witnessing the event of Boris playing the guitar
- I imagine Boris playing the guitar in the future<sup>12</sup>.
- I can see how Boris is playing the guitar (for ex., using too much effort)

In order to show that the concrete reading, available for ‘kak’ constructions in Russian, is one of the kind described by Boivin (1998), I will briefly discuss the properties Boivin points out with respect to this type of constructions in French and check if these properties can be revealed for ‘kak’ in Russian.

Boivin argues that stage- level predicates in the complements of *voir* (see) yield an ambiguous sentence, which can have either concrete or imaginative reading. Opposed to that, individual –level predicates in the complements of *voir* (see) can not yield concrete reading and can only be statements of imaginative view.

(Boivin, 1998:107)

(87) Je vois Jean posséder une maison.

I see John own a house

‘I see John owning a house.’(imagine)

Boivin argues that (14) is ungrammatical under the concrete reading. The tendency appears to be the same with Russian.

(88) Ya vižu kak Boris znaet otvet na vopros.

I see how Boris<sub>Nom</sub> knows answer<sub>Acc</sub> to question

‘I can see Boris knowing the answer to the question.’(imagine)

The example in (88) is completely ungrammatical under the concrete reading (witnessing the event of Boris knowing the answer), but is acceptable under the imaginative reading (imagining that Boris became an intellectual and knows the answer to the question).

Another distinctive property with respect to clausal complements of perception verbs pointed out by Boivin (1998), Felser (1998) is the observation that bare plural subjects can not be assigned a generic interpretation in the complements of perception verbs. The only interpretation of the sentence available is existential.

---

<sup>12</sup> Imaginative reading in Russian can also be created by participial constructions:

Ya vižu Borisa igrayuschim na gitare

I see Boris<sub>Acc</sub> palying the guitar

Some Russian speakers prefer the use of ‘kak’ constructions, others prefer the use of the participle to convey the imaginative reading.

(Felser, 1998:367)

- (89) a. Dinosaurs ate kelp. EX/GEN  
b. We saw dinosaurs eat(ing) kelp. EX/\*GEN

The same observation holds for Russian examples. When a clausal complement with ‘kak’ is embedded under a stative verb, both existential and generic readings are available. However, when a clausal complement is embedded under a perception verb, the generic reading is no longer observed.

- (90) a. Ya ne znayu kak dinozavry eli zolu. EX/GEN  
I not know how dinosaurs<sub>Nom</sub> ate kelp  
'I don't know how dinosaurs were eating kelp'  
b. Ya videl kak dinozavry eli zolu. EX/\*GEN  
I saw how dinosaurs ate kelp  
'I saw dinosaurs eating kelp'

Thus, to sum up, I have shown that the properties generally discussed with respect to clausal complements of perception verbs in languages like French and English are parallel to the ones revealed in Russian ‘kak’-constructions. There are at least two<sup>13</sup> syntactic structures intersecting in the ‘kak’-type of examples pointed out by Bošković (2004), creating an ambiguity.

The question I want to address next is how we can disambiguate these structures.

A possible way to disambiguate these structures would be to embed a clausal complement with ‘kak’ under the type of verb that is not a perception verb and, importantly, one that does not normally take a declarative complement. In this way, both the concrete and imaginative reading will be eliminated and the only reading of ‘kak’ left available will be a wh- reading.

In English, the type of verb that does not take a declarative complement is the verb ‘know’ under negation with first person singular subject.

- (91) a. I know that he came.  
b.\* I don't know that he came.

I construct the same type of examples in Russian with ‘kak’ clausal complements embedded under the negated verb ‘know’. The prediction is that extraction out of these complements should be more problematic than extraction out of complements with perception verbs. If our reasoning is on the right track and long

---

<sup>13</sup> In fact, Boivin (1998) argues for two distinct structures for concrete and imaginative reading. I am not addressing this issue here, because it is not directly relevant to the question I am investigating.

distance [NP] dislocation, in fact, involves contrastive focus, then it is expected that extraction out of these complements should induce weak-island effects.

(92) a. My ne znaem kak rebjenok kupil sigarety.

We not know how child bought cigarettes

‘We don’t know how the child bought cigarettes’.

b.?? My [rebjenok]<sub>i</sub> ne znaem kak t<sub>i</sub> kupil sigarety.<sup>14</sup>

We<sub>Nom</sub> child<sub>Nom</sub> not know how bought cigarettes

‘We don’t know how child bought cigarettes’

The only possible interpretation of ‘kak’ here is ‘how, in which way’. The same grammatical judgement is obtained if we substitute ‘kak’ by ‘kakim obrazom’.

(93)?? My [rebjenok]<sub>i</sub> ne znaem kakim obrazom t<sub>i</sub> kupil sigarety.

We<sub>Nom</sub> child<sub>Nom</sub> not know how bought cigarettes

‘We don’t know how child bought cigarettes’

Therefore, we have evidence that ‘kak’ is ambiguous between a complementizer and a wh-element. When ‘kak’ is disambiguated, weak island effect is straightforward. The absence of weak island effects in the ‘kak’-constructions pointed out by Bošković (2004) is due to an ambiguity that allows long-distance extraction of focused [NP] across complementizer inducing no violation of relativized minimality.

In order to show that the ambiguity of ‘kak’ can be shown at the contextual level I construct two contexts that would reflect a complementizer / wh-element correlation.

- ‘kak’ as a complementizer

*We are present at the terrorist act investigation. A terrorist packed a bomb into a parcel at the post office and sent it to the government. We are questioning the staff about who they have seen packing at the post office on that day.*

The witness:

*I saw a boy packing his skies, then there was a lady packing her flower basin...*

The interviewer:

(94) Vy posylku videli kak /\*kakim obrazom kto-nibud’ upakovyval?

You parcel(Acc) saw how/\*in which way anybody packed?

‘Did you see anybody packing the parcel?’

The witness:

---

<sup>14</sup> The effects in (92)b and (93) can be improved if the dislocated [NP] is specified by the preceding context. This observation corresponds to one of the strategies of extraction out of weak islands described by Starke (2001). I come back to these facts later in this section.

(95) Yes/No / \*with the tape.

What we can observe in the above context is that the interviewer is interested in the fact if the event of packing was seen, but not in the manner of packing. Hence the manner of packing stated in the answer of the witness is unexpected and confusing. ‘Kak’ can not be substituted by the unambiguous wh-element ‘kakim obrazom’, hence, it gives evidence that ‘kak’ is a complementizer when used in the context described above.

- ‘kak’ as a wh-element

*Christmas fairytale: The little dwarfs are packing presents for children. They pack teddy bears with ribbons, LEGOs in wrapping paper with stars, candies into boxes.*

The interviewer:

(96)?? Vy lyži (Acc) videli kak/ kakim obrazom gnomы zapakovali?

You skies saw how/in which way dwarfs packed?

‘Did you see how the dwarfs packed the skies?’

The witness:

(97) Yes/no/ Lyžy zapakovali v dlinnuyu korobku.

Skies packed in a long box

‘The skies were packed in along box’

In the context above the manner of packing is a part of the context. As we can see in (96), ‘kak’ can be substituted by an unambiguous wh-element without inducing ungrammaticality.<sup>15</sup> The answer needs to include the manner of packing and so does the context, hence, we have sufficient evidence to conclude that ‘kak’ behaves like a wh-element in the example (96) within the context described above.

An important factor that contributes to disambiguation of structures is intonation. Thus, in example (94), the fronted [NP] but not the ‘kak’ is emphasized by intonation, which helps to parse the construction as involving a complementizer and not a wh-element. However, in the example (96), both the fronted [NP] and the ‘kak’ need to be emphasized by intonation, which helps to parse ‘kak’ as a wh-element and creates a problematic situation characteristic of a weak island effect (two elements focused at the same time).

---

Example (96) is degraded but not ungrammatical.

Nevertheless, we still have to account for the grammaticality of the other example, pointed by Bošković (2004), which does not involve ‘*kak*’, but, crucially, is grammatical. I repeat the relevant example below.

- (98) a. Vy [doctor] videli kogda pod'ezžal?

You [doctor]<sub>Nom</sub> saw when arrived

‘The doctor, did you see him arriving?’

Apart from this example, there are examples originally pointed out by Müller and Sternefeld (1993) that involve long distance dislocation of [NP] across a wh-element, but, crucially, have no perception verb in the matrix.

(Müller and Sternefeld, 1993:469)

- (99) a. Ya byl [novuju školu] gde strojat.

I was new school<sub>Acc</sub> where they build

‘I have been to the place where they build a new school’

I am going to address this question in the next subsection.

## 5.4 Relativized minimality

In the previous subsection I found evidence that long distance [NP]-dislocation has an effect on interpretation and is felicitous in a pair-list context, which shows that long distance scrambling in Russian can involve contrastive focus. However, the idea of contrastive focus contradicts the facts introduced by Müller and Sternefeld (1993) as in (99) and by Bošković (2004) as in (98). These examples reveal no weak-island effect, while it should not be the case if we analyze these examples as involving contrastive focus.

According to Bošković and Takahashi’s (1998) theory, Japanese style scrambling is exempt from relativized minimality effects due to its radical reconstruction property. Under the assumption that Russian has Japanese style scrambling it is expected that long distance scrambling in Russian is also not subject to weak-islandhood. Consider the examples below.

- (100) a.?? Vy [čelovek] <sub>i</sub> videli kuda ušel t<sub>i</sub>?

You<sub>Nom</sub> person<sub>Nom</sub> saw where went

‘Did you see where a man went?’

- b. Vy [čelovek v zelenoi kurtke] <sub>i</sub> videli kuda ušel t<sub>i</sub>?

You person in a green jacket saw where went

‘Did you see where the person in a green jacket went?’

As we can see in (100), bare subject extraction into the matrix clause leads to marginal grammaticality, while when the subject is modified by the [PP], grammaticality improves.

(101) a.? Ty [školu] znaeš gde stroyat t<sub>i</sub> ?

You school <sub>Acc</sub> know where they build

‘Do you know where they build a school?’

b. Ty [novuyu školu] znaeš gde stroyat t<sub>i</sub> ?

You new school <sub>Acc</sub> know where they build

‘Do you know where they build a new school?’

The same effect is observed in (101). The grammaticality improves when the dislocated noun is modified by an adjective. An important factor to mention is that (100)a and (101)a are ungrammatical in ‘out of the blue’ context, but became acceptable if the background of the conversation includes the item fronted in the structure. The speaker and the hearer should share a ‘common knowledge’ presupposition about the element that is fronted in the structure.

As is discussed in Starke (2001), a wh- element can not cross another wh-element. However, there are cases when wh-phrases can escape wh- islands. “Wh-phrases (Q) which can extract out of weak islands are those that have some additional property  $\beta$ ”. (Starke, 2001:10)

(102) a.\*Q<sub>i</sub> ... Q<sub>j</sub>...Q<sub>i</sub>

b. Q $\beta$ ... Q... Q $\beta$

Wh-phrases that have an additional property  $\beta$  constitute an element Q $\beta$  that can cross over an intervening Q- element because the movement of Q $\beta$ - element can be blocked only by another intervening Q $\beta$ - element. According to Starke (2001), Q $\beta$ -movement as opposed to Q-movement carries existential presupposition with wide scope.

Starke distinguishes two types of presupposition that legitimate extraction out weak islands: specificity- based and range-based. Specificity-based presupposition requires the type of context where the wh-phrase has a specific antecedent (Starke, 2001:19)

*You are a car mechanic working in a garage; one morning, as you come late, you hear that the list of cars to be repaired might have been reshuffled. To get up to date, you ask:*

(103) Which car is it now unclear whether we should repair?

Range-based presupposition requires a context, which provides a *range* for the wh-phrase.

(Starke, 2001:20)

*You are a car mechanic working in a special garage catering to customers who are picky about how their car is repaired. To this effect, the garage keeps a list of cars paired with how their owners want them repaired. One morning as you come in, you hear that the list has been made partially unreadable by water infiltration so that some cars have lost their how-to instructions. To get up to date on the situation you ask:*

(104) Which car is it now unclear how we should repair?

If we compare Starke's (2001) contexts to the ones I constructed above for 'kak' ambiguity (see *Terrorist attack* context, p.59, *Christmas fairy tale*, p.60), we can observe the similar tendencies in them.

Specificity-based presupposition requires one list: list of customers in *Terrorist* context; list of cars in Starke's *Garage* context. Range-based presupposition requires two lists: *Christmas fairytale*: a list of presents, a list of ways to pack them; *Garage context*: a list of cars and a list of repair instructions. Thus what follows is that the contexts I created to show the structural ambiguity of 'kak', in fact, allow to see the specificity-based/ range-based distinction.

However, the important difference is that both Starke's (2001) structures involve extraction out of embedded questions, while I argued that structures with 'kak' are ambiguous. Crucially, the structure that involves a presupposition similar to specificity-based presupposition is scrambling across a complementizer.(see context and example (94), p.59.)

If this reasoning is on the right track, then scrambling out of declarative clauses requires a specificity-based presupposition. Range-based presupposition is required when scrambling applies out of wh-islands.<sup>16</sup>

---

<sup>16</sup> This reasoning here is quite speculative in nature and based on intuitions. However, the similarities between the context's from Starke (2001) to the ones constructed for scrambling here are striking and informative.

Coming back to observations about (100) and (101), they can be translated into Starke's characteristics of weak-islandhood. Extraction of a bare subject in (100)a, (101)a in 'out of the blue' context equals to bare Q-extraction, which is ungrammatical.

(105)\*  $Q_i \dots Q_j \dots Q_i$

However, when the context includes the fronted item in (100)a and (101)a and, hence, the presupposition is justified, the grammaticality improves.

(106)  $\exists Q \dots Q \dots \exists Q$

The type of presupposition involved can be defined in the examples in (100)b and (101)b. Both of these examples presuppose a list of items. Example (100)b presupposes a list of people, say, standing in a line with the speaker asking about one specific person. Similarly, example (101)b is felicitous in a context where there is a small number of schools in town, with only one of them new and reference made to the new school. The type of extraction out of weak island in (100) and (101) involves the following configuration pointed out by Starke (2001).

(Starke, 2001:28)

(107)  $SQ \dots Q \dots SQ$

Hence, we have evidence to assume that, contrary to standard assumptions, long distance scrambling in Russian does induce relativized minimality effects and examples earlier pointed out as scrambling, in fact, involve focus movement. The type of extraction involved corresponds to constructions characteristic of weak-islandhood.

Now I turn to the discussion of Bošković's (2004) example repeated below.

(108) a. Vy [doctor] videli kogda pod'ezžal?

You [doctor]<sub>Nom</sub> saw when arrived

'The doctor, did you see him arriving?'

If we follow the reasoning above, (108) should be at least degraded, similarly to (100) and (101) above. But it is reported to be grammatical.

Recall, that the type of context, where (108) is felicitous should include two lists: a list of arrivees and a list of times (see informational background 1, p.53). Therefore, we have evidence to assume that one factor that improves the grammaticality of (108) is range-based presupposition that legitimates extraction out of weak island.

There is another factor that contributes to the grammaticality of (108), which is a perception verb in the matrix clause. In the next chapter, I am going to introduce

evidence that clausal complements embedded under the verbs of perception are generally more transparent for extraction both in Russian and in English.

Presumably, that the second factor by itself can not legitimate extraction out of weak islands but as an additional factor has its influence on the grammaticality.

Example (108) would be ungrammatical in ‘out of the blue’ context i.e. with no presupposition involved, independently of the fact that a perception verb is used in the matrix clause. However, if we compare an example with a perception verb and with a non-perception verb in the matrix with no presupposition (‘out of the blue’ context (see (109))), then the observation is that in order to become grammatical the non-perception verb example would require a stronger presupposition than the one with perception verb in the matrix.

- (109) a.\* Vy [doctor] videli kogda pod'ezžal?  
You [doctor]<sub>Nom</sub> saw when arrived  
'The doctor, did you see him arriving?'  
b.\* Vy [doctor] znaete kogda pod'ezžal?  
You [doctor]<sub>Nom</sub> know when arrived  
'The doctor, do you know when he arrived?'

Hence, we have evidence to show that even though on the surface it seems that (108) and other examples pointed out by Müller and Sternefeld are not subject to relativized minimality, but, in fact, they are subject to weak islandhood and the relevant examples that are grammatical correspond to cases of legitimate extraction out of weak islands.

The presence of RM effect in this construction is crucial in our case, because, according to B&T, the LF lowering operation is exempt from relativized minimality effects. This fact allows us to conclude that the construction pointed out by B&T is not an instance of Japanese style scrambling.

## 5.5 Reconstruction and binding

Given the facts discovered above, namely, that long distance dislocation in Russian involves focus movement; we should expect that reconstruction should also occur in the type of examples discussed. As is shown in Starke (2001), reconstruction effects in extraction out weak islands can be observed in English.

(Starke, 2001:65)

- (110) a. [Which picture of herself<sub>i</sub>] does Simone<sub>i</sub> wonder whether Humphrey will  
dare show to his parents <which picture of herself<sub>i</sub>>?  
b. [Which picture of herself<sub>i</sub>] does Simone<sub>i</sub> wonder when Humphrey will  
finally throw away <which picture of herself<sub>i</sub>>?

The prediction is that reconstruction should occur in the Russian examples, but dislocation should be subject to the constraints extraction out of weak islands is normally subject to.

In order to check if there is reconstruction I apply a test involving genitive of negation.

There is a type of constructions in Russian, where genitive case on the noun is licensed by clause-mate sentential negation as in (111).

- (111) a. Ya ne čitaju gazet.  
I not read newspapers<sub>Gen</sub>  
'I don't read newspapers'  
b.\* Ya čitayu gazet  
I read newspapers<sub>Gen</sub>

Given the radical reconstruction property assumed for long distance dislocation in Russian by Bošković and Takahashi (1998), the noun in the genitive case dislocated into the matrix clause should not incur ungrammaticality.<sup>17</sup>

- (112)
- a. Ty pomniš kak Ivan nikogda ne čital knig?<sup>18</sup>  
You remember how Ivan<sub>Nom</sub> never *ne* read books<sub>Gen</sub>  
'Do you remember how Ivan never read books?'
- b. \*?? Ty [knig]<sub>i</sub> pomniš kak Ivan nikogda ne čital t<sub>i</sub>?  
You [books<sub>Gen</sub>] remember how Ivan<sub>Nom</sub> never *ne* read t<sub>i</sub>  
c. ? Ty [knig po biologii]<sub>i</sub> pomniš kak Ivan nikogda ne čital t<sub>i</sub>?  
You [books<sub>Gen</sub> about biology] remember how Ivan never *ne* read t<sub>i</sub>  
<sub>Acc</sub> 'Do you remember how Ivan never read biology books?'

---

<sup>17</sup> There is still a question though at which level of derivation the genitive case on the noun is licensed.

<sup>18</sup> These examples involve 'kak', which as I argued in the previous subsection is ambiguous. However, the ambiguity of 'kak' can not interfere with radical reconstruction. It is expected that the dislocated [NP] should reconstruct both across a wh- element and a complementizer.

As we can see in (112)b the dislocation of the object in genitive case into the matrix clause leads to ungrammaticality. However, when the object is modified by a [PP] as in (112)c, the grammaticality is improved.

A similar effect is to be observed with negative polarity items (NPI). Negative polarity items in Russian, the so called *ni*-words, must be accompanied by clause mate sentential negation. The dislocation of bare negative polarity items into the matrix clause is not allowed.

- (113) \*Ty [nikto<sub>i</sub>] pomniš kak t<sub>i</sub> ne pomog tebe v trudnuju minutu?

You [no one]<sub>Nom</sub> remember how <sub>Neg</sub> helped you <sub>Dat</sub> in hard time?

‘Do you remember how you had no one to help you in hard time?’

- (114)\*Ty [nikogo<sub>i</sub>] pomniš kak on ne priglasil t<sub>i</sub> na den' roždenija?

You [no one]<sub>Acc</sub> remember how he <sub>ne<sub>Neg</sub></sub> invited on birthday?

‘Do you remember how he didn't invite anyone to his birthday party?’

- (115)\*Ty [nikomu<sub>i</sub>] pomniš kak on togda ne pomog t<sub>i</sub>?

You anyone<sub>Dat</sub> remember how he then <sub>ne<sub>Neg</sub></sub> help?

‘Do you remember how he didn't help anyone at that time?’

However, when the negative polarity item is modified by the partitive construction, the grammaticality improves.

- (116)? Ty [nikto iz druzei<sub>i</sub>] pomniš kak t<sub>i</sub> ne pomog tebe v trudnuju

You [none of the friends]<sub>Nom</sub> remember how <sub>ne<sub>Neg</sub></sub> help you <sub>Dat</sub> in hard  
minutu?

time?

‘Do you remember how none of the friends helped you in hard time?’

- (117)? Ty [nikogo iz sester<sub>i</sub>] pomniš kak on ne priglasil t<sub>i</sub> na den'

You [none of the sisters]<sub>Acc</sub> remember how he <sub>ne<sub>Neg</sub></sub> invited on birthday?  
roždenija?

‘Do you remember how he didn't invite any of the sisters to his birthday  
party?’

- (118)? Ty [nikomu iz gruppy<sub>i</sub>] pomniš kak on togda ne

You [no one from the group]<sub>Dat</sub> remember how he then <sub>ne<sub>Neg</sub></sub>  
pomog t<sub>i</sub>?  
help?

‘Do you remember how he didn't help anyone from the group at that  
time?’

Given the assumption of radical reconstruction for long distance [NP] dislocation in Russian, it is unexpected that bare noun in genitive case (see (112)b) as well as bare NPI (see (116),(117),(118)) can not be extracted into the matrix clause.

However, the observations above can be explained under a movement account. The effects above are again reminiscent of a weak island effect. Movement of a bare noun or an NPI is illicit, unless the dislocated item is modified by a property  $\beta$  (specificity or range- based presupposition). In example (112)c, the [PP] modifying the dislocated object narrows the broad reference to ‘books’ to a less broad entity of ‘biology books’. In other words, it makes the relevant entity more specific.

The same is true for the NPI examples. The partitive construction in examples (116)-(118) restricts the reference of NPI to the entity of ‘friends’, ‘sisters’ etc.

Hence, we could analyze (112)c and (116)-(118) as involving the following configuration of extraction out of weak islands.

(119) SQ....Q....SQ

However, this explanation would not be quite right, because we have postulated that ‘kak’ is ambiguous.

Hence, what the examples with genitive of negation and NPIs show are that bare overt Q- extraction is illicit both across a wh-phrase and a complementizer.

This can be considered as additional evidence to our hypothesis that scrambling always requires a presupposition: specificity based for extraction out of declarative clauses and range-based for extraction out of embedded questions.

Now I turn to binding effects, the other issue discussed with respect to radical reconstruction. According to B&T, the element scrambled long distance can not establish new binding relations at the landing site, because, given the radical reconstruction property, the scrambled [NP] does not surface in that position at LF.

However, Saito (2003) points out that when the anaphor is scrambled into the matrix clause, the matrix subject can antecede the anaphor and the grammaticality improves.

(Saito, 2003:509).

- (120) a.\*Karera-ga [CP Hanako-ga [CP Ziroo-ga otogai-o sonkeisiteiru to ] itta  
 they Nom Hanako Nom Ziroo Nom each other Acc respect that said  
 to] omotteiru (koto)  
 that think fact  
 \*‘They think [that Hanako said [that Ziroo respect each other]]’

- b. Karera-ga [CP otagai-o<sub>i</sub> Hanako-ga [CP Ziroo-ga t<sub>i</sub> sonkeisiteiru to] itta  
 they<sub>Nom</sub> each other<sub>Acc</sub> Hanako<sub>Nom</sub> Ziroo<sub>Nom</sub> respect that  
 to] omotteiru(koto)  
 that think fact

‘They think [that each other<sub>i</sub>, Hanako said [that Ziroo respect t<sub>i</sub>]’

Saito (2003) views this example as evidence supporting his assumption that condition (A) of the binding theory is an ‘anywhere condition’. He assumes that, given the radical reconstruction property, the anaphor should reconstruct to its base generated position at LF, hence, condition (A) cannot be satisfied at LF and should be an ‘anywhere condition’.

However, in the context of our discussion the example above can be viewed as evidence that there are cases in Japanese<sup>19</sup>, where the phrase scrambled long distance can be bound by the matrix subject. These examples contradict the assumption of the radical reconstruction property for Japanese.

Now consider the Russian examples. As we can see in the examples (121) below, parallel to the Japanese example in (46)b, the anaphor dislocated into the matrix clause can be bound by the matrix subject.

- (121) a. Ty videl kuda on<sub>i</sub> postavil [svoyu<sub>i</sub> mašinu]?  
 You<sub>Nom</sub> saw where he<sub>Nom</sub> parked self car<sub>Acc</sub>?  
 ‘Did you see where he parked his car?’
- b.? Ty<sub>j</sub> [svoyu<sub>i/j</sub> mašinu] videl kuda on<sub>i</sub> postavil e?<sup>20</sup>  
 You<sub>Nom</sub> [self car<sub>Acc</sub>] saw where he<sub>Nom</sub> parked?
- c. [Svoju mašinu<sub>i</sub>] ty videl kuda on<sub>i</sub> postavil t<sub>i</sub>?  
 [self car<sub>Acc</sub>] you<sub>Nom</sub> saw where he<sub>Nom</sub> parked?  
 ‘Did you see where he parked his car?’
- (122) a. Ty videl kuda ya<sub>i</sub> položila [svoi<sub>i</sub> kl'uči]?  
 You<sub>Nom</sub> saw where I<sub>Nom</sub> put self keys<sub>Acc</sub>?  
 ‘Did you see where I put my keys?’
- b. ? Ty<sub>j</sub> [svoi<sub>i/j</sub> kl'uči] videl kuda ya<sub>i</sub> položila t<sub>i</sub>?  
 You self keys<sub>Acc</sub> saw where I<sub>Nom</sub> put?

<sup>19</sup> For more evidence of this type for Japanese see Dejima (1999).

<sup>20</sup> There is a certain variation with grammaticality judgements. Some speakers do not allow binding in the matrix clause. However the contrast to Japanese example in (120)a reported as ungrammatical is still informative.

- ‘Did you see where I put my keys?’
- c. [Svoi kl’uči<sub>i</sub>] ty videl kuda ya<sub>i</sub> položila t<sub>i</sub>?  
 Self keys Acc saw where I Nom put?  
 ‘Did you see where I put my keys?’

The examples (121) b and (122) b are ambiguous. The pronoun ‘svoi’ can be bound both in the matrix and in the subordinate clause. The possibility of binding in Russian examples as well as in Japanese examples is unpredicted by B&T’s theory.

Interestingly, the effects in (120)b, (121)b and (122)b also cannot be captured by Bailyn’s derivational schema of scrambling either. Recall, that for Bailyn (2001) long distance scrambling is a discourse related A'- movement, driven by operator type of features similar to wh- movement. For discourse driven scrambling, he assumes that it is the [Op]- feature that triggers the dislocation and the strong nominal [D]- feature doesn’t have to raise for economy reasons. In order to capture the effect described above the nominal [D]- feature must be active in the matrix clause to allow new binding, but Bailyn’s derivational schema prescribes for the [D]-feature to become active low for long distance dislocations and the possibility of binding in the matrix clause is thus unexpected.

Apparently, subject dislocation does not give rise to ambiguity in contrast to (121)b and (122)b.

- (123)\*Ty [svoi nachal’nik otdela<sub>i</sub>] pomniš kak oskorbilsya t<sub>i</sub>?  
 You [your head of the department (Nom)] remember how got insulted  
 ‘Do you remember how the head of the department got insulted?’

The subject in the matrix clause cannot antecede the scrambled phrase in its landing site and there is no possible antecedent for the dislocated phrase in the subordinate clause, which leads to ungrammaticality in (123).

Hence, the generalization is that binding facts both for Russian and in Japanese show that the dislocated anaphor can be bound in the matrix clause. The possibility of binding contradicts the radical reconstruction hypothesis for Japanese and Russian.

However, the binding effects described above for Russian are expected under the assumption that long distance [NP] dislocation in Russian involves contrastive focus.

The cases where a wh-element is bound after movement have been discussed in the literature with respect to English.

(Reinhart and Reuland, 1993:683)

- (124) a. Which pictures of himself/herself does Max think that Lucie likes?

b. Max knows which picture of himself/herself Lucie likes.

Thus, it is plausible to assume that the binding effects observed in Russian are of the same kind pointed out for English. The focused element can not antecede the anaphor, but can be bound. This generalization is not true for subject dislocation in Russian. I leave this issue for future research.

The facts discovered by Saito (2003) with respect to possibilities of binding in Japanese create a serious problem for B&T's (1998) base generation analysis of long distance scrambling in Japanese and require further investigation.

### *Summary*

In this section I have presented evidence against semantically vacuous scrambling for Russian. The following facts were discovered:

- The scrambled versus non-scrambled order requires different types of context
- ‘kak’ in Russian is ambiguous between a complementizer and a wh-element when it is embedded under verbs of perception
- Long distance [NP] dislocation in Russian is sensitive to weak-islands. The cases where it looks that it is not, are cases of legitimate extraction out of weak islands as described in Starke (2001)
- Long distance scrambling in Russian is focus movement.
- The possibilities of binding in the matrix clause can be pointed out both for Russian and Japanese (see Saito (2003)). Binding effects contradict the base generation approach of B&T (1998).

In the next section I am going to describe the properties of long distance scrambling in Russian further. I will explore how scrambling in Russian behaves with respect to other island constraints. I will also address the question to what extend scrambling in Russian parallels heavy-NP shift and long wh-movement in English.

## 6 More on properties of long distance scrambling in Russian. English parallels?

In the previous section I have argued that long distance scrambling in Russian is sensitive to relativized minimality. I have also pointed to the fact that the constructions with a perception verb in the matrix clause turn out to be more transparent for scrambling than constructions with other types of verb.

In this section, I am going address the following questions:

- What properties does long distance scrambling in Russian have in general?
- Are there any parallels between heavy- NP shift in English and scrambling in Russian?

### 6.1 Long distance scrambling in Russian. Locality

We have seen in the previous section that long distance scrambling gives a more grammatical result when it proceeds from a clause embedded under a verb of perception. I am going to elaborate more on this issue and make an attempt to show how the choice of verb in the matrix clause correlates with the possibility of scrambling.

For verbs of perception, the observation is that both subject dislocation and object dislocation are grammatical out of finite clauses.

- (125) a. Ya [naš sosed]i slyšal čto t<sub>i</sub> novuju mašinu kupil.

I<sub>Nom</sub> our neighbour<sub>Nom</sub> heard that new car<sub>Acc</sub> bought

'I heard that our neighbour bought a new car'

- b. Ya [portfel'] i videla kak<sup>21</sup> Ivan v prihožei ostavil t<sub>i</sub>.

I bag<sub>Acc</sub> saw that Ivan in hall left

'I saw Ivan leaving his bag in the hall'

However, when scrambling applies out of embedded questions with verbs of perception in the matrix, subject dislocation is slightly degraded.

- (126) a.?Vy [doktor] i videli kogda t<sub>i</sub> pod'ezжал?

You<sub>Nom</sub> doctor<sub>Nom</sub> saw when arrived

'Did you see when the doctor arrived?'

- b. Vy [konverty] i videli kuda mama položila t<sub>i</sub>?

---

<sup>21</sup> I assume that 'kak' is a complementizer here.

You<sub>Nom</sub> envelopes<sub>Acc</sub> saw where mother left  
 ‘Did you see where mother left the envelopes?’

Interestingly, clausal complements of bridge verbs are less transparent for scrambling than the ones with perception verbs.

(127)a.? Maša [doktor]<sub>i</sub> skazala čto skoro pod’edet t<sub>i</sub>.

Maša<sub>Nom</sub> doctor<sub>Nom</sub> said that soon arrive  
 ‘Maša said that the doctor will arrive soon.’

b. Maša [pis’mo]<sub>i</sub> dumaet čto poslala t<sub>i</sub>.

Maša<sub>Nom</sub> letter<sub>Acc</sub> think that sent  
 ‘Maša thinks that she sent the letter’

As we can see in (127) the subject dislocation becomes degraded out of decalrative clauses and a similar effect is observed for embedded questions as in (128).

(128)a. ? Maša [doktor]<sub>i</sub> skazala kogda pod’edet t<sub>i</sub>?

Maša<sub>Nom</sub> doctor<sub>Nom</sub> said when arrive  
 ‘Did Maša say when the doctor will arrive?’

b. Maša [pis’mo]<sub>i</sub> skazala kto ei poslal t<sub>i</sub>?

Maša<sub>Nom</sub> letter<sub>Acc</sub> think who her-to sent  
 ‘Did Maša say who sent her the letter?’

The next type of verbs to consider are stative verbs. The observation is that clausal complements embedded under stative verbs are even less available for extraction than perception verbs and bridge verbs.

(129)a.\*?? Ya [volny]<sub>i</sub> lyublyu čto v štorm podnimayutsya t<sub>i</sub>.

I<sub>Nom</sub> waves<sub>Nom</sub> love that in storm rise  
 ‘I love that the waves rise in the storm.’

b. ?? Ya [zavtrak] nenuvižu čto pozdno nakryvayut.

I<sub>Nom</sub> breakfast<sub>Acc</sub> hate that late serve  
 ‘I hate that they serve breakfast late.’

As we can see in (129)a, subject dislocation from a clausal complement embedded under a stative verb is ungrammatical. Object dislocation is slightly better, but also marginal. Hence, compared to examples in (125) and (127), the difference in grammaticality is significant. This fact allows us to conclude that scrambling applies more freely in constructions with verbs of perception and bridge verbs, than with stative verbs.

The next contrast to consider is scrambling out of questions embedded under stative verbs.

(Zemskaja, 1973:399)

- (130) a.??Ty [volny] <sub>i</sub> lyubiš kogda bolšie t<sub>i</sub> ?

You<sub>Nom</sub> waves<sub>Nom</sub> like when big

‘Do you like when there are big waves?’

- b.?? Ty [zavtrak] <sub>i</sub> ljubiš kogda muž v postel prinosit t<sub>i</sub>?

You<sub>Nom</sub> breakfast<sub>Acc</sub> like when husband in bed brings

‘Do you like it when your husband brings you a breakfast in bed?’

Hence, as we can see in (130), subject and object scrambling out of embedded questions under stative verbs is equally marginal.

As for clausal complements of other types of verbs, the observation is that they pattern with bridge verbs with respect to the degree of transparency. Note, that I can only consider the verbs that take clausal complements. Consider the examples below:

- (131) implicative verbs<sup>22</sup>

- a.?Ty [Pavel]<sub>i</sub> zabył čto t<sub>i</sub> konfety uže kupil?

You<sub>Nom</sub> Pavel<sub>Nom</sub> forgot that candy already bought

‘Did you forget that Pavel already bought the candy?’

- b. Ty [konfety] zabył čto Pavel uže kupil t<sub>i</sub>?

You<sub>Nom</sub> candies<sub>Acc</sub> forgot that Pavel<sub>Nom</sub> already bought

‘Did you forget that Pavel already bought the candy?’

- c.? Ty [Pavel]<sub>i</sub> pomniš gde t<sub>i</sub> konfety kupil?

You<sub>Nom</sub> Pavel<sub>Nom</sub> remember where candy bought

‘Do you remember where Pavel bought the candy?’

- d. Ty [konfety]<sub>i</sub> pomniš gde Pavel kupil t<sub>i</sub>?

You<sub>Nom</sub> candy<sub>Acc</sub> remember where Pavel<sub>Nom</sub> bought

‘Do you remember where Pavel bought the candy?’

- (132) factive verbs<sup>23</sup>

- a.? Ty [syn] žaleeš čto t<sub>i</sub> v akademiyu ne postupil?

You<sub>Nom</sub> son<sub>Nom</sub> regret that in academy not get admitted

‘Do you regret that your son didn’t get admitted to the academy?’

---

<sup>22</sup> The verb classification is taken from Landau (2000).

<sup>23</sup> Factive verbs, as well as the verbs belonging to the class of propositional and desideratives in Russian do not take embedded questions as their complements.

- b. Ty [botinki] žaleeš čto kupil t<sub>i</sub>?  
 You shoes regret that bought  
 ‘Do you regret that you bought these shoes?’

(133) propositional

- a.? Ya [professor] <sub>i</sub> zayavlyayu čto t<sub>i</sub> segodnya ne pridet.  
 I<sub>Nom</sub> professor<sub>Nom</sub> claim that today not come  
 ‘I claim that the professor wont come today’
- b. Ya [lekciju] <sub>i</sub> predpolagaju čto segodnya otmenili t<sub>i</sub>.  
 I<sub>Nom</sub> lecture<sub>Acc</sub> suppose that today they-cancelled  
 ‘I suppose that they cancelled the lecture today’.

(134) desideratives

- a.? Ya [syn] <sub>i</sub> nadeyus čto t<sub>i</sub> postupit v universitet.  
 I son hope that enter in university  
 ‘I hope my son will be admitted to the university’.
- b. Ya [konfety] <sub>i</sub> nadeyus čto Pavel kupil t<sub>i</sub>.  
 I<sub>Nom</sub> candy<sub>Acc</sub> hope that Pavel bought  
 ‘I hope that Pavel bought the candy’.

(135) interrogatives

- a. Ya [Pavel] <sub>i</sub> ugadal čto t<sub>i</sub> kupil konfety .  
 I<sub>Nom</sub> Pavel<sub>Nom</sub> guessed that bought candy  
 ‘I guessed that Pavel bought the candy’
- b. Ya [konfety] <sub>i</sub> ugadal čto Pavel kupil t<sub>i</sub>.  
 I candy guesses that Pavel bought  
 ‘I guessed that Pavel bought the candy’
- c.?? Ty [Pavel] <sub>i</sub> sprosil kuda t<sub>i</sub> položil konfety?  
 You<sub>Nom</sub> Pavel<sub>Nom</sub> asked where put the candy  
 ‘Did you ask where Pavel left the candy?’
- d. Ty [konfety] <sub>i</sub> sprosil gde Pavel ostavil t<sub>i</sub>?  
 You<sub>Nom</sub> candy<sub>Acc</sub> asked where Pavel left  
 ‘Did you ask where Pavel left the candy’

Hence, as we can see from the examples above, the transparency effects with the types of verb listed in (131)-(135) are quite stable and parallel to the effects pointed out earlier for bridge verbs as in (127) and (128).

Thus, summarizing the facts above, verbal complements can be divided into three classes based on the transparency effects with respect to long distance scrambling. We have seen that there exists a dependency between the type of verb used in the matrix clause and the transparency of its clausal complement for scrambling. The results are given in (136).

(136)

Scrambling out of finite complements		
Perception verbs	Bridge verbs	Stative verbs
subject ✓	subject ?	Subject *
object ✓	object ✓	Object ??
Scrambling out of embedded questions		
Perception verbs	Bridge verbs	Stative verbs
subject ?	subject ?	subject ??
object ✓	Object ✓	object ??

The most transparent structure is the one with the verbs of perception. It gives the most grammatical result with respect to application of scrambling. The least transparent structures for scrambling are complements of stative verbs.

Interestingly, the examples used by B&T (1998) as an argument for the existence of Japanese style scrambling in Russian involve only constructions with verbs of perception.

I have shown that the change of the matrix verb can affect the availability of scrambling. Given the radical reconstruction hypothesis for Russian scrambling, the effects described above are unexpected. Thus, it is plausible that part of the reason for the grammaticality of B&T's (1998) examples lies in the general transparency of the structure that they used.

The next issue I am going to investigate is the behaviour of long distance scrambling with respect to islands. We have seen in the previous section that scrambling can extract out of weak islands. I argued in the previous section that the strategy of extraction for scrambling is very similar to the one generally assumed for wh-movement. I have shown that, similar to wh- extraction out of weak islands, scrambling requires a presupposition.

Given this parallel between wh- movement and scrambling, it is expected that long distance scrambling should be sensitive to the adjunct constraint. Consider the examples below:

(Adger, 2003:399)

- (137) a. Hephaestus had run away, before the executioner murdered Hera.  
 b.\* Who had run away, before the executioner murdered <who>?

It is shown in (137) that wh- movement in English is sensitive to the adjunct constraint. Now consider the examples with scrambling in Russian.

- (138) a. Moi drug ušel prezde čem ya svarila kofe.

My friend<sub>Nom</sub> left before I<sub>Nom</sub> made coffee

‘My friend left before I made coffee’

- b.\* Moi drug [kofe]<sub>i</sub> ušel prezde čem ya svarila t<sub>i</sub>.

My friend<sub>Nom</sub> coffee<sub>Acc</sub> left before I made

‘Coffee, my friend left before I made’

As we can see in (138), scrambling out of an adjunct is banned. Hence, in this respect, scrambling also parallels wh-movement.

The next type of islands, that wh- extraction is considered to be sensitive to are relative clauses. Consider the examples from English:

(Ross, 1967:69)

- (139) a. Phineas knows a girl who is jealous of Maxime.

b. \*Who<sub>i</sub> does Phineas know a girl who is jealous of t<sub>i</sub>?

Scrambling out of a relative clause is also ungrammatical, similar to wh-movement.

- (140)a. Oni pozvali nas v derevnyu, v kotoroi kupili dom.

They<sub>Nom</sub> called us to the village in which they- bought a house<sub>Acc</sub>.

‘They invited us to the village where they bought a house’.

- b.\* Oni [house]<sub>i</sub> pozvali nas v derevnyu, v kotoroi kupili t<sub>i</sub>.

They<sub>Nom</sub> house<sub>Acc</sub> invited us to the village in which bought

‘They invited us to the village where they bought a house’

In this connection, I want to discuss one of the examples from Müller and Sternefeld (1993), where they point out that scrambling in Russian has distinct properties from wh- movement. The example is repeated below.

(Müller and Sternefeld, 1993:469)

- (141) a. Ya byl [novuju školu] gde strojat.<sup>24</sup>

I was new school<sub>Acc</sub> where they build

'I have been to the place where they build a new school'.

Müller and Sternefeld argue that (141) involves a dislocation across an embedded question, which allows them to show that scrambling does not induce a weak-island effect.

However, the observation is that the clausal complement in (141) can not get a question interpretation.

If we try to force the question interpretations on the 'be' verb complement in (141) with the use of Russian 'li' particle, the result is ungrammatical.<sup>25</sup>

- (142) a.\*Ty byl novuju školu strojat li?

You<sub>Nom</sub> were new school<sub>Acc</sub> they build Q?

- b. Ty znaeš novuju školu strojat li?

You know new school they- build Q?

'Do you know if they build a new school?'

As we can see in (18) the complement of the verb 'be' as opposed to the compliment of the verb 'know' can not take an embedded question as its complement.

Moreover, the complement of the verb 'be' in (17) can be substituted by a [PP], while this is not the case for the verb 'know'.

- (143) a. Ya byl [PP na stroike novoi školy].

I<sub>Nom</sub> was on construction area new school<sub>Gen</sub>

'I have been in the construction area of the new school'

- b.\* Ya znayu na stroike novoi školy.

I<sub>Nom</sub> know on construction<sub>Prep</sub> area new school

Thus, we can see that the complement of the verb 'be' in (141) occurs in the position where a [PP] can occur. The conclusion that follows from the above is that (141) does not involve an embedded question, but a free relative clause.

Additional evidence comes from the observation, that the formal head noun ('tam') can be introduced in the structure with 'be' verb complement, but, crucially, not with 'know' verb complement.

- (144) a. Ya byl tam gde strojat novhju školy.

I<sub>Nom</sub> was there where they build a new school

---

<sup>24</sup> Some Russian speakers, including the author, think that (141) is slightly degraded.

<sup>25</sup> The particle 'li' is considered to be an overt realization of Q-element in Russian.

‘I have been at the place where they build a new school’.

b.\* Ya znaju tam gde strojyat novuju školu.

I<sub>Nom</sub> know there where they build new school

Thus, we have reasons to assume that the examples originally introduced as an instance of scrambling across an embedded question, in fact, do not involve a dislocation across an embedded question, but scrambling out of a free relative clause.

However, it is generally assumed that free relative clauses induce islands effects. The same observation is true for Russian. Consider the examples below.

(145)a.\* Kto ty byl gde pohoronen?

Who<sub>Nom</sub> you<sub>Nom</sub> was where buried

b.?? Čto ty byl gde stroyat?

What you<sub>Nom</sub> was where build-they

Now consider the examples with subject and bare object scrambling in Russian.

(146)a.\* Ya [drug detstva]<sub>i</sub> byl gde pohoronen t<sub>i</sub>.

I friend of childhood was where buried

‘I have been to the place where my friend form childhood is buried’

b. ?? Ya [školu]<sub>i</sub> byl gde stroyat t<sub>i</sub>.

I school was where build

‘I have been to the place where they build a school.’

Thus, we observe that the effects in (146) pattern with the effects of wh-extraction out of a relative clause. Therefore, the argument made by Müller and Sternefeld (1993) in (141) with respect to scrambling in Russian does not hold for cases of subject extraction as well as unmodified object extraction. In other words, we see more evidence that long distance scrambling gives rise to islands effects.

To summarize the discussion above, I have pointed out the following properties of long distance scrambling with respect to locality:

- Long distance scrambling applies freely out of structures where finite complements are embedded under verbs of perception. For other type of structures, scrambling transformation in Russian is more restricted.
- Long distance scrambling is sensitive to relativized minimality
- Long distance scrambling can not apply out of an adjunct
- Long distance scrambling can not apply out of a relative clause

The conclusion that can be drawn so far is that scrambling in Russian is more restricted than it is generally assumed in the literature (Müller & Sternefeld (1993), B&T (1998) Bošković (2004)).

In terms of locality, it looks like scrambling in Russian is blocked by the same type of constraints like wh- movement. Hence, we have evidence to assume that long distance [NP]- dislocation is a transformation of the type similar to wh-movement.

B&T's (1998) hypothesis of a semantically vacuous dislocation is not well supported and, hence, can not be assumed for Russian.

## 6.2 Heavy- NP shift and scrambling

In this subsection, I want to look at heavy- NP shift in English in comparison to scrambling. Both of these transformations involve discourse-driven dislocation of [NP]s and there have been several attempts to unify these two phenomena in the literature. One of them, namely, Saito & Fukui's (1998) analysis was already discussed in sec 2.

Rochemont and Culicover (1990) describe heavy-NP shift as the type of transformation derived by the rightward application of Move  $\alpha$  to the [NP] that appears sentence finally.

(Rochemont and Culicover, 1990:116)

- (147) a. John bought a painting that he liked for his mother.

b. John bought  $t_i$  for his mother [ a painting that he liked]<sub>i</sub>

Rochemont (1978) points out two main requirements of heavy –NP shift: first, the constituents it applies to are heavy [NP]s and, second, the shifted [NP]s should be focused.

For scrambling in Russian, I have shown that the dislocated [NP] is focused, which could be seen from the informational background the dislocation is felicitous in. I have also pointed to the fact that scrambling of the heavier constituents is preferred in Russian.

However, crucially, heavy- NP shift, unlike long distance focus movement in Russian applies locally. This property was first pointed out by Ross (1967) and is usually referred to as the ‘right roof constraint’.

- (148) No element is moved rightward by a transformation, may be moved out of the next node S.

An example ‘right roof constraint’ violation is given below.

(Postal, 1974) cited from (Tanaka, 2004:6)

- (149) \*I have expected [CP that [<sub>TP</sub> I would find <sub>t<sub>i</sub></sub>] since 1939 [<sub>NP</sub> the treasure said to have been buried on the island]<sub>i</sub>

Thus, as we can see in (149) rightward movement out of finite clauses is blocked by the ‘right roof’ constraint.

This property of heavy-NP shift is distinct from scrambling. As we know, scrambling in Russian can apply long distance. Moreover, as we could see in the table (136), [NP]- extraction out of declarative clausal complements gives a more grammatical result than extraction out of embedded questions.

Another characteristic of heavy-NP shift discussed in the literature is that heavy-NP shift is allowed out of control complements in English.

(Tanaka, 2004:6)

- (150) I have expected [<sub>TP</sub> PRO to find <sub>t<sub>i</sub></sub>] since 1939 [<sub>NP</sub> the treasure said to have been buried on that island]<sub>i</sub>.<sup>26</sup>

Control complements in Russian are also transparent for scrambling. Consider the examples below.

- (151) a. Ya poprosila ego naučit' menja čitat' [stihi Puškina].

I<sub>Nom</sub> asked him<sub>Acc</sub> to teach me to read poems Puškina<sub>Gen</sub>  
'I asked him to teach me how to read Puškin's poems'.

- b. Ya [stihi Puškina]<sub>i</sub> poprosila ego naučit' menja čitat' <sub>t<sub>i</sub></sub>.

I<sub>Nom</sub> poems Puškina<sub>Gen</sub> asked him<sub>Acc</sub> to teach me to read  
'I asked him to teach me how to read Puškin's poems'

The fact that both heavy-NP shift and scrambling can apply out of control complements is not surprising. It has been previously noted in the literature that non-finite clauses are very transparent in general for various grammatical processes like extraction, binding, etc.

For example, extraction from a finite wh-island is more problematic than from an infinitival wh-island.

(Coopman & Stevenson, 1991:359)

- (152) a. \*What did you wonder [whether [he fixed t]]  
b. ? What did you wonder [whether to fix t]]

---

<sup>26</sup> Tanaka (2004) argues, contrary to Chomsky's (1999) analysis that control complements are bare [TP]s.

Among the properties of heavy-NP shift discussed in Tanaka (2004) is the fact that extraction out of ECM constructions in English leads to ungrammaticality.

- (153)\*? I have expected Mary [CP [TP  $t_j$  to find  $t_i$ ]] since 1939 [ the treasure said to have been buried on that island]<sub>i</sub>.

It is generally assumed in the literature that there is no syntactic counterpart to English ECM constructions in Russian. Hence, we can not compare heavy –NP shift and scrambling in this respect. However, in the previous subsection, I have pointed to the fact that clausal complements of perception verbs are more transparent for scrambling in Russian. Perception verbs can also be used with ECM constructions in English and my hypothesis is that heavy-NP shift should apply more freely out of ECM constructions with verbs of perception than with other types of verb. Consider the following examples.

- (154) a. [I desire  $t_i$  with all my heart] [all his linguistics books that he chose to sell]<sub>i</sub>.  
b. [I saw  $t_i$  with my own eyes] [all his old linguistics books that he was chose to sell ]<sub>i</sub>.

As we can see in (154), it is possible to heavy- NP shift locally across the modifiers of the verbs ‘want’ and ‘see’. The next step would be to apply heavy-NP shift out of ECM clause with a perception verb and a volitional verb across the same type of modifiers as in (154). Since the modifiers are generated in the matrix clause, we can make sure that heavy-NP shift applies long distance.

- (155) a.\* [I wanted John to sell  $t_i$  with all my heart] [all his old linguistics books with comments in them]<sub>i</sub>.  
b. [I saw John sell  $t_i$  with my own eyes] [all his old linguistics books with comments in them]<sub>i</sub>.

As we have already seen above in (153) heavy-NP shift is ungrammatical out of ECM clauses and this is true for (155)a as well. However, the grammaticality of (155)b is noticeably better than (155)a, hence, our prediction is borne out. As expected, we can observe that in English extraction out of ECM constructions with verbs of perception is considerably more transparent than extraction out ECM constructions with other types of verb.

If our reasoning is on the right track, then we have evidence to assume that perception verbs predicates behave differently with respect to extraction both in Russian and in English.

Another interesting point about the contrast in (155) is that it contradicts Tanaka's (2004) theory. He argues that raising-to-object involves raising out of a [CP]. His reasoning is the following: heavy- NP shift is blocked out of finite clauses by the 'right roof constraint', which prevents rightward movement out of a [CP]. Heavy-NP shift is ungrammatical out of ECM clauses, which, he posits, is due to the fact that raising-to object complements are also [CP]s and, hence, extraction out of them is ruled out by the 'right roof constraint' similar to finite clauses.

However, contrary to Tanaka's (2004) claim, we have seen in (155)b that extraction out of ECM complements with verbs of perception is allowed. This fact creates a problem for Tanaka's (2004) analysis of raising-to-object constructions.

The reason why perception verb complements are the most transparent can be explained by the fact that there is no [TP] projection in case of a perception predicate. This idea has already been suggested in the literature before. For example, Castillo (2001) argues that perception verb predicates lack [TP]s. One of the arguments she introduces is that perception verb predicates do not permit a negative complement.

(Castillo, 2001:127)

(156) a.\* He saw Mary not open the door.

b. I consider him not to be able to win the elections.

The same effect can be observed in Russian 'kak'- constructions.<sup>27</sup>

(157)\* Ya videl kak doktor ne priejal.

I saw            doctor not arrive

'I saw the doctor not arrive'

Castillo (2001) discusses the recent approaches by Haegeman (1995) and Zanuttini (1996), where the choice of imperative/subjunctive is in interrelation with [NegP] and [TP]. From the interconnection of [NegP] and [TP] and given the fact that negative complement is not allowed with perception verbs, Castillo concludes that perception verb predicates are bare [VP]s. This idea is plausible for English, given the observation that complements of perception verbs in English are bare infinitives, while complements of other verbs always occur with 'to'. The infinitival particle 'to' in English is assumed to be generated in [TP].

(158) a. I wanted him to dance.

---

<sup>27</sup> Russian does not have ECM constructions , but I have argued in section 4 that Russian constructions with 'kak' share some properties with English ECM constructions with perception verbs and can be considered as their semantical counterpart.

- b. I saw him dance.
- c.\*I saw him to dance.

Thus, it is possible to assume that the transparency of perception verb complements with respect to heavy –NP shift as well as scrambling is due to the absence of [TP] projection in perception verb predicates both in English and in Russian.

Given the facts discussed above, we reach the following conclusion:

Even though both heavy-NP shift and scrambling in Russian are instances of discourse driven movement, it appears to be implausible to talk about a straightforward parallel between these two processes. Such crucial distinctions like long distance application of scrambling and clause boundedness of heavy- NP shift point to the fact that these two types of dislocation should be treated separately. More over, we have seen that heavy NP shift can not proceed out of finite clauses in general, while scrambling gives a more grammatical result out of declarative clausal complements than out of embedded questions.

The comparison that seems to be more promising so far is the comparison of long distance scrambling to wh-movement. This is the issue I want address this issue next.

### **6.3 Long wh- movement and scrambling**

In this subsection I want to elaborate on the parallel of long distance scrambling in Russian and long wh-movement in English.

We have already pointed to the following properties of scrambling that coincide with the properties of long wh-movement:

- Sensitivity to relativized minimality
- Sensitivity to adjunct constraint
- Scrambling can not apply out of a relative clause

The issue I want to discuss next is successive cyclicity of scrambling in Russian. It is a well known fact that long wh-movement in English applies successive cyclically.

(Adger, 2003:361)

- (159) a. What did you say [<sub>CP</sub> that the poet had written]?
- b. What did John tell [<sub>CP</sub> that Mary think [<sub>CP</sub> that Fred saw]]?

It is assumed in the literature that long wh- movement applies in several steps: first, the wh-phrase moves to the embedded [Spec, CP] position, second, it proceeds to the matrix [Spec, CP] position.

Given the similarity of properties between scrambling and wh- movement, we should expect that long distance scrambling in Russian can also apply successive cyclically and several clauses up. Consider the examples below:

- (160) a.? Maša [pismo]<sub>i</sub> dumala [CP čto Petr znaet [CP čto Olya polučila t<sub>i</sub>]].

Maša<sub>Nom</sub> letter<sub>Acc</sub> think that Petr knows that Olya get  
 ‘Maša thinks that Peter knows that Olya got the letter’.

- b.?? Maša [doktor]<sub>i</sub> dumala [CPčto Petr znaet [CPčto t<sub>i</sub> pod’edet pozže]].

Maša<sub>Nom</sub> doctor<sub>Nom</sub> thinks that Petr knows that arrive late  
 ‘Maša thinks that Peter knows that the doctor will arrive later’

The contrast to observe is that application of scrambling across two clauses is possible with object dislocation, but is marginal with subject dislocation.

The application of scrambling across three clauses is ungrammatical not only with subjects but also with objects.

- (161)a.\* Maša [doktor]<sub>i</sub> dumala [CP čto Petr znaet [CP čto Olya ponyala [CP  
 Maša<sub>Nom</sub> doctor<sub>Nom</sub> thought that Petr knows that Olya understood  
 čto priedet pozže t<sub>i</sub>.]]

that arrive late

‘Maša thinks that Peter knows that Olya understood that the doctor will  
 arrive late’

- b.\* Maša [pismo]<sub>i</sub> dumala [CP čto Petr znaet [CP čto Olya videla [CP  
 Maša<sub>Nom</sub> letter<sub>Acc</sub> think that Petr knows that Olya saw  
 kak Vanya polučil t<sub>i</sub>.]]

how Vanya get

‘Maša thinks that Peter knows that Olya saw Vanya get the letter’

Thus, we can see that object scrambling and subject scrambling marginally is allowed across two clauses. Given the fact that scrambling can apply long distance, and since we have already seen certain similarities of properties between scrambling and wh-movement, I hypothesize that scrambling transformation applies successive cyclically similar to long wh-movement.

However, in the context of this paper it will remain just a hypothesis, due to the fact that the evidence that long distance scrambling applies through several intermediate positions is not available.

In the previous subsection I have pointed to the three classes of verbs with respect to the availability of scrambling out of their complements. The issue I want to address now is whether wh-movement can be parallel to scrambling in transparency effects with respect to constructions with the verbs discussed above.

- perception verbs

(162) a. What did you hear that John say?

b.? Who did you see (\*that) hit the child?

c.\* What did you hear when John say?

d.\* Who did you see how cooked pasta?

- bridge verbs

(163) a.What do you think that John will cook tonight?

b.? Who do you think (\*that) will cook tonight?

c.\* What do you think how John will cook?

d.\* Who do you think how will cook?

- stative verbs

(164) a. What do you hate that your girlfriend says?

b.\*Who do you like that run away from home?

c.\* What do you hate when your girlfriend says?

d.\*Who do you like when cooks spaghetti?

The facts above are summarized in table (165).

(165)

(136)

Wh-extraction out of declarative clausal complements			Scrambling out of declarative clausal complements		
Perception verbs	Bridge verbs	Stative verbs	Perception verbs	Bridge verbs	Stative verbs
subject ?/*	subject ?/*	Subject *	subject ✓	subject ?	Subject *
object ✓	object ✓	Object ✓	object ✓	object ✓	Object ??
Wh-extraction out of embedded questions			Scrambling out of embedded questions		
Perception verbs	Bridge verbs	Stative verbs	Perception verbs	Bridge verbs	Stative verbs
subject *	subject *	subject *	subject ?	subject ?	subject ??
object *	Object *	object *	object ✓	Object ✓	object ??

If we compare the table in (165) to the one in (136) repeated above, the following observations can be made. Wh- extraction reveals a stronger subject –object asymmetry, than scrambling. Subject wh- extraction is blocked for all of the cases considered while subject scrambling is ungrammatical only with stative verbs in declarative complements. What concerns object extraction, there is a parallel for wh-extraction and scrambling in declarative complements. There is a noticeable distinction for object scrambling out stative verb complements though.

If we look at embedded questions, the effects are very distinct. From the results observed for embedded questions, we can conclude that wh-movement is more sensitive to wh- islands than scrambling for the types of verb discussed.

However, there are certain cases of wh-extraction out weak islands in English that allow wh-object extraction. But these cases do not parallel with scrambling with respect to the choice of verb.

(166) a.? What do you wonder whether John will cook?

b.? What is it unclear whether we should repair? (Starke, 2001:10, 19)

It has been argued in the literature that the examples as in (166) are very specific and need to be clearly contextualized to give grammaticality effects reported above.

Hence, the following conclusion is reached: there are number of properties which parallel for long wh-movement in English and long distance scrambling in Russian. Among them are: sensitivity to weak islands, sensitivity to adjunct islands, sensitivity to complex NP- constraint, successive cyclicity, interconnection to informational background.

Nevertheless, as we have seen from the comparison in (165)-(136), it is inappropriate to talk about a direct parallel between scrambling and wh-movement with respect to the degree of transparency in particular structures. Even though we have seen that both scrambling and wh- movement are sensitive to weak islands, the effects in comparison show that the degree of sensitivity to weak islands differ for scrambling and wh- movement.

Why this is so, I leave this issue for future research.

To sum up the section, we have seen that the properties of long distance scrambling in Russian are such that it is more appropriate to compare it to wh-movement, than heavy NP-shift. However, the parallel between scrambling and wh-movement is not direct. Scrambling gives a relatively more grammatical result out of weak islands, than wh- movement.

### *Summary*

In this section, I have looked through the properties of long distance scrambling in Russian in general. I have shown that the structures most transparent for scrambling are clausal complements of perception verbs. The structures with other types of verb are less available for scrambling.

I have compared the properties of heavy NP-shift in English and scrambling in Russian. The conclusion I came to is that these two transformations are not parallel with respect to their properties.

I have observed that scrambling is sensitive to a number of constraints that wh-movement is sensitive to. I have pointed to the fact that long distance scrambling can be successive cyclic.

However, what concerns sensitivity to weak island effects, I have discovered that the degree of sensitivity as well as the type of transparent structures are distinct for scrambling and wh- movement.

## 7 Conclusion

This study was mainly concerned with long distance scrambling transformation in Russian. The following questions have been addressed:

- What properties does long distance scrambling have with respect to locality, binding and semantic interpretation?
- Can we talk about Japanese style scrambling in Russian?
- What other phenomena cross- linguistically is Russian scrambling parallel to?

The conclusion reached is as follows:

Long distance scrambling in Russian occurs for discourse considerations and involves contrastive focus. The dislocation is felicitous in a pair list context and is licit only with specificity based /range-based presupposition. Within the absence of presupposition, the dislocation is not allowed.

No evidence for the semantically vacuous type of dislocation can be found. Thus, Bošković and Takahashi's (1998) hypothesis that Japanese style scrambling is available in Russian is not confirmed.

Scrambling in Russian parallels to wh-movement in English in many respects:

It is subject to relativized minimality, adjunct constraint, complex [NP] constraint and can apply successive cyclically. Similar to wh- movement, object scrambling gives a more grammatical result than subject scrambling.

However, certain distinctions between scrambling and wh-movement have been pointed out and require further investigation. Unlike wh- movement, scrambling distinguishes three types of verb complements with respect to transparency: perception

verb, bridge verb and stative verb complements. Same observation is not true for wh-movement.

Another puzzling fact is that the degree of sensitivity to relativized minimality islands is smaller for scrambling than for wh- extraction. I leave these issues for future research.

## References

- Adger, D. 2003. Core Syntax. A Minimalist Approach Oxford University Press: Oxford.
- Bailyn, J. 2001. On scrambling: A reply to Bošković and Takahashi. *Linguistic Inquiry* 32: 635-658.
- Bailyn, J. 2002. A (Purely) Derivational Approach to Russian Scrambling. In FASL 11: the Amherst Meeting, Michigan Slavic Publications, Ann Arbor, Michigan.
- Bailyn, J. 2004. Generalized Inversion. *Natural Language and Linguistic Theory* 22:1-49.
- Boivin, M. 1998. Complementation and Interpretation: The Concrete and Imaginative Readings of Visual Perception Verbs. *MIT Working Papers in Linguistics* 25:103-123.
- Bošković, Ž. 2002. On multiple wh-fronting. *Linguistic Inquiry* 33:351-383.
- Bošković, Ž. 2003. On left branch extraction. In *Investigations into formal Slavic Linguistics: Contributions of the Forth European Conference On Formal Description of Slavic Languages - FDSL IV*, ed. by Peter Kosta, Joanna Błaszcak, Jens Fasek, Ljudmila Geist and Marzena Zygis, 543-577. Frankfurt am Main: Peter Lang.
- Bošković, Ž. 2004. Topicalization, Focalization, Lexical Insertion, and Scrambling. *Linguistic Inquiry* 35:613-638.
- Bošković, Ž. and Takahashi, D. 1998. Scrambling and Last Resort. *Linguistics Inquiry* 29:347-366.
- Castillo, C. 2001. The configuration of ECM Structures. *Studia Linguistica* 55(2):113-139.
- Chomsky, N. 1993. A Minimalist Program for linguistic theory. In Ken Hale and S. Jay Keyser (eds.), *The view from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Cambridge, Mass.: MIT Press.
- Chomsky, N. 1994. Bare Phrase Structure. *MIT Occasional Papers in Linguistics* 5. MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass. Published in *Government and Binding Theory and the Minimalist Program*. Webelhuth (ed.), 383-439. Oxford: Blackwell. 1995.
- Chomsky, N. 1995. The Minimalist Program. Cambridge, Mass.: MIT Press.

- Chomsky, N. 2000. Minimalist inquiries. In *Step by Step: Essays on minimalism in honor of Howard Lasnik*, Roger Martin, David Michaels and Juan Uriagereka (eds.), 89-155. Cambridge, Mass.: MIT Press.
- Chomsky, N. and Lasnik, H. 1993. The theory of principles and parameters. In *Syntax: An International handbook of contemporary research*. Armin von Stechow, Wolfgang Sternefeld and Theo Vennemann (eds.), 506-569. Berlin: Walter de Gruyter.
- Coopman, P. and Stevenson, S. 1991. How Extraction from Finite and Infinitival Complements: A Surprising Asymmetry. *Linguistic Inquiry*. 22(2): 359-367
- Felser, C. 1998. Perception and Control: a Minimalist analysis of English direct perception complements. *Journal of Linguistics* 34:351-385.
- Haegeman, L. 1995. *The syntax of negation*. Cambridge: Cambridge U. Press.
- Hale, K. 1980. Remarks on Japanese Phrase Structure: Comments on the Papers on Japanese Syntax. *MIT Working Papers in Linguistics* 2:185-203.
- Harada, Kazuko I. 1972. Constraints on wh-Q binding. *Studies in Descriptive and Applied Linguistics* 5:180-206. Tokyo: International Christian University, Division of Languages.
- Holmberg, A. and Platzack, C. 1991. On the Role of Inflection in Scandinavian Syntax. In Werker, Kosmeijer, Reuland (eds.). *Issues in Germanic Syntax*. Berlin, Mouton de Gruyter
- Junghans, U. and Zybatow, G. 1997. Syntax and Information Structure of Russian Clauses. *Annual Workshop to Formal Approaches to Slavic Linguistics 4. The Cornell Meeting 1995*. Ann Arbor, Michigan Slavic Publications: 289-319.
- Kitahara, H. 2000. *Case and Scrambling: A Derivational view*. Ms., Keio University.
- Krylova, O. and Khavronina, S. 1976. Word order in Russian Sentences. Russian Language Publishers, Moscow.
- Kuroda, S.-Y. 1988. Whether we agree or not: A comparative syntax of English and Japanese. In William Poser (ed.), *Papers from the Second International Workshop on syntax of English and Japanese*, Stanford: CSLI Publications.
- Landau, I. 2000. *Elements of control. Structure and meaning in infinitival constructions*. Dordrecht, The Netherlands: Kluwer.
- Lebeaux, David. 1988. Language acquisition and the form of grammar. Doctoral Dissertation, University of Massachusetts, Amherst.
- Mahajan, A. 1990. The A/A-bar distinctions and movement theory. Doctoral dissertation, MIT. Cambridge, Mass.

- Miyagawa, S. 1996. Word order restrictions and nonconfigurality. *MIT Working Papers in Linguistics* 29:171-141.
- Miyagawa, S. 2001. EPP, Scrambling, and Wh-in-situ. In *Ken Hale: A Life in Language*, ed. Michael Kenstowicz, 293-338. Cambridge, MA: MIT Press
- Munn, Alan. 1993. *Topics in the syntax and semantics of coordinate structure*. Doctoral Dissertation, University of Maryland, College Park.
- Müller, G. and Sternefeld, W. 1993. Improper Movement and Unambiguous Binding. *Linguistic Inquiry* 24:461-507.
- Müller, G. and Sternefeld, W. 1996. A-Chain Formation and Economy of Derivation. *Linguistic Inquiry* 27:480-511.
- Nemoto, N. 1999. Scrambling. In Tsujimura (ed.). *The handbook of Japanese Linguistics*. Oxford: Blackwell.
- Postal, P. 1974. *On raising*. Cambridge, Mass.: MIT Press.
- Rochemont, M. and Culicover, P. 1990. *English focus constructions and the theory of grammar*. Cambridge University Press, Cambridge.
- Saito, M. 1985. *Some Asymmetries in Japanese and their theoretical Implications*, PhD dissertation, MIT.
- Saito, M. 1989. Scrambling as semantically vacuous A'-movement. In *Alternative conceptions of phrase structure*. Mark Baltin and Anthony Kroch (eds.) 182-200. Chicago: University of Chicago Press.
- Saito, M. 1992. Long Distance scrambling in Japanese. *Journal of East Asian Linguistics* 1:69-118.
- Saito, M. 2001. Toward the Reunification of Japanese Scramblings. *MIT Working Papers in Linguistics* 41, 287-307.
- Saito, M. 2003. A derivational approach to the interpretation of scrambling chains. *Lingua* 113:481-518.
- Saito, M. and Fukui, N. 1998. Order in Phrase Structure and Movement. *Linguistic Inquiry* 29:439-74.
- Saito, M. and Hoji, H. 1983. Weak cross over and move  $\alpha$  in Japanese. *Natural Language and Linguistic Theory* 1:245-259.
- Stepanov, A. 1998. On Wh-fronting in Russian. In Tamanji and Kusumoto (eds.), *NELS* 28: 453-467.
- Strahov, N. 2000. A Scrambling analysis of Russian wh-Questions. In Steven Franks, Tracy Holloway King, Michael Yadroff (eds.) *Annual Workshop of Formal Approaches*

- to Slavic Linguistics: The Bloomington Meeting, 2000.* Ann Arbor: Michigan Slavic Publications. 293-310.
- Reinhart, T. and Reuland, E. 1993. Reflexivity. *Linguistic Inquiry*, 24(4):657-720.
- Ross, J. *Constraints on variables in Syntax*, PhD dissertation, MIT.
- Starke, M. 2001. *Move dissolves into Merge: A Theory of Locality*. Doctoral Dissertation, University of Geneva.
- Tada, H. 1993. *A/A' Partition in derivation*. Doctoral Dissertation. MIT, Cambridge, Mass.
- Tada, H. 1990. *Scrambling(s)*. Workshop on Japanese Syntax, Ohio State University.
- Tanaka, H. 2004. On the Categorial Status of Raising Complements. *York Papers in Linguistics* 2(1):213-222.
- Rizzi, L. 1990. *Relativized Minimality*. Cambridge, Mass.: MIT Press.
- Zemskaja, E. 1973. *Russkaja razgovornaja reč*. Moskva: Nauka.
- Zanuttini, R. 1996. On the relevance of tense for sentential negation. In Beletti and Rizzi (eds.) *Parameters and Functional heads*. Oxford: Oxford U. Press.



