Stony Brook University

Whoever doesn't HOP must be Superior:

The Russian left-periphery and the Emergence of Superiority

A Dissertation Presented

by

Tatiana V. Scott

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Linguistics

Stony Brook University

December 2012

Stony Brook University

The Graduate School

Tatiana V. Scott

We, the dissertation committee for the above candidate for the Doctor of Philosophy in Linguistics degree, hereby recommend acceptance of this dissertation.

John Frederick Bailyn, Ph.D
Associate Professor, Department of Linguistics
Stony Brook University, NY

Daniel Finer, Ph.D
Professor, Department of Linguistics
Stony Brook University, NY

Richard Larson, Ph.D
Professor, Department of Linguistics
Stony Brook University, NY

Marcel Den-Dikken, Ph.D
Professor, Department of Linguistics,
Graduate Center, CUNY, NY

This dissertation is accepted by the Graduate School

Charles Taber
Interim Dean of the Graduate School

Abstract of the Dissertation

Whoever doesn't HOP must be Superior!:

The Russian left-periphery and the Emergence of Superiority

by

Tatiana Scott

Doctor of Philosophy

in

Linguistics

Stony Brook University

2012

This dissertation maps the left-periphery of the Russian language, presenting a new geometry of Russian main and subordinate clauses in order to account for a number of phenomena: single and multiple wh-constructions, sluicing constructions, and coordinate multiple wh-constructions (CMW), as well as to predict various occurring word-orders. Interestingly, the theory for these constructions comes from a proposed unique wh-unrelated category, the Higher Operator Phrase (HOP), located higher than CP, which allows the aforementioned phenomena to fall out of the proposed structure.

The thesis claims that Russian exhibits 'true' wh-movement and is similar to Bulgarian in Rudin's 1988 typology, with a strong [+wh] feature on each wh-phrase, forcing it to be in the specifier of CP. Superiority holds, and emerges in various contexts in Russian; though its effects may be masked in the main clause by the ability of any wh to raise further (to SpecHOP).

There is an asymmetry with respect to surface Superiority in matrix vs. embedded clauses: Superiority emerges in embedded clauses, and in main clauses whenever HOP is overtly occupied by a non-wh-phrase. The asymmetry falls out of the theoretical assumption that the HOP position is available in matrix clauses only.

These conclusions give a rise to a re-analysis of sluicing in Russian as a wh-phenomenon where Superiority also emerges as it does in coordinate wh constructions (CMW). The analysis of CMW presented in this work is based on existing accounts utilizing sideward movement (Nunes 2001) and a coordination phrase (ConjP/&P) (Zoerner 1995, Gribanova 2009, Citko & Gračanin-Yuksek 2010).

Finally, the thesis addresses the issue of what can occupy the head position of HOP. This gives rise to a discussion of Topicalization properties of this (HOP) position. It is shown how a topic-marking particle "-TO" can occur in this position, and adds to the evidence for the existence of this position. As a result, it illustrates how various wh- and other word order constructions can be uniformly analyzed with this unique structure.

Table of Contents

List of Tables

List of Abbreviations

ACC/A	Accusative case
Adj	Adjective
ADJ	adjunct
Adv	Adverb
ARG	argument
AUX	auxiliary
CMW	coordinated multiple wh-phrases
ConjP/&P	coordination phrase
DAT/D	Dative case
EXCL	exclamative
FUT	future tense
HOP	High Operator Phrase
INF	infinitive/infinitival
IS	Information Structure
LBE	left branch extraction
mvt	movement
MWF	multiple wh-fronting
NOM/N	Nominative case
O	Object
PERF	perfective
PL	pair list reading
PL	plural
PREP	Prepositional case
QR	quantifier raising
REFL	reflective
S	Subject
SG	singular
SP	single pair reading
Spec	Specifier
V	Verb
VP	verb phrase
VPE	verb phrase ellipsis
WCO	weak crossover
3rd	3rd person

Acknowledgements

This moment has come: “I would like to thank the Academi-a!” I couldn’t have gotten here alone. Many people have contributed to this production. I owe my deepest gratitude to everyone who made it possible and supported me every step of the way.

First and foremost, I am forever grateful to professor John F. Bailyn, my advisor, my mentor, and my friend. It is rare that a student and an advisor share such personal closeness. Needless to say, our working styles are similar. I have been amazingly fortunate to have an advisor who gave me freedom to explore on my own, and whose encouragement guided me back when I faltered. Without him, this project would have died years ago. I also would like to send my thanks to John’s wife Dijana who allowed him to spend hours away from home having continuous discussions with me, sometimes going late into the night.

My cordial gratitude goes to my outstanding committee: professor Richard Larson, professor Daniel Finer, and professor Marcel den Dikken. Their knowledge of Liguistics and their personalities kept the project real. They always comforted me when most needed. I thank Richard for his constructive criticism and useful ideas that broaden my horizons in the linguistics field, and always kept me on my toes. I thank Dan for having so many creative ideas about my work and always being eager to discuss them, sometimes on the spur of the moment. His input into Chapter 4 of this work cannot be overstated. Without Marcel the project would not be nearly as wholesome as it stands today. Marcel’s immediate commentary on any part of the work longer than the work itself is what made me a better syntactician. The questions that he posed to me reshaped the dissertation in many ways, and I have no doubt, for the better. He is, of course, not responsible for any flaws of this work.

My gratitude also goes to professor Irina Sekerina. She was the voice of my conscience and has become a great colleague and a friend, reading countless unfinished drafts. Her working discipline has always been an example for me. I thank her for being there every step of the way.

I would like to thank professors Roumi Pancheva, Norvin Richards, and David Pesetsky for the data that they have thrown at me during different stages of this project. Amazingly, every one of them was eager to answer my desperate inquiries via Facebook the night before the defense: a rare graduate student luxury.

I should not forget to thank other faculty of the Linguistics department at Stony Brook: professor Bob Hoberman, professor Mark Aronoff (my wonderful qualifying committee), professor Lori Repetti, and professor Ellen Broselow. The amount of support that I received from professor Marie Huffman cannot be even enumerated. I am truly sorry for adding hours to your days with the constant paperwork requests. I do feel lucky that my graduate years took place with such wonderful faculty. All of us studying and working at the department would have never survived without Sandra Brennan who is our guardian angel.

I am also grateful for the opportunity to collect my data abroad with the financial support of the Fulbright – Hays Dissertation Abroad foundation.

I owe my gratitude to all those people who made my graduate school experience unforgettable, without them it probably would have been unbearable. They are not only exceptional drinking partners, they are just impressive people who I have been blessed to know: Dr. Stephanie Harves, Dr. Masha Vassilieva, and Dr. Barbara Citko. When I first met them, neither one was a Dr. yet.

Stephanie, thank you for all our conversations and for our writing hours, and for just being you.

Thank you, Barbara, for your knowledge, your ideas, professional ethics, and for your love of parrots.

I am grateful to Masha who despite her insatiable love of cats still loves me more. Thank you, Masha, for being a great friend from our first meeting. Without you, life would be different.

My gratitude goes to Dr. Carlos deCuba for his amazing sense of humor and back patting when I most needed it; Dr. Margo DelliCarpini who has been a great friend throughout the years; and Dr.-to-be Andrei Antonenko, who is probably, the smartest guy I know but who never admits to it. He is both a friend and a great critical informant for all my investigations. I always think about Dr. Marianne Boroff for just being there for me, the person innately and genuinely nice for no apparent reason. I am grateful for all the times we have spent together.

And last but not least, my heartfelt thank you goes to the dearest people in my life: my parents, my brother, and my life partner.

I thank my parents for never doubting me, giving me all the best since the first minute of my life, and for loving me unconditionally. I thank my brother, Mike, for bringing joy into my life, and for his creative and unique mind.

My eternal gratitude and love goes to Roman, without whom, I probably would have lived under the Brooklyn Bridge without any prospect for a bright future. His support in every way is invaluable. And now when I hear the fanfares of the time running out, I want to share this moment with him.

Chapter 1. HOP Introduction and General Discussion

1.0 Introduction: puzzles and questions

Syntax as a part of generative linguistics is concerned with the question of how different language structures are derived. The current dissertation tackles a number of word order puzzles in Russian. In fact, the thesis is built in a paradoxical way: on the one hand, most of its body is devoted to a discussion and analysis of wh-movement and wh-related phenomena in Russian. On the other hand, I state that the analysis of such phenomenon as wh-movement in Russian is impossible without postulating a specific structure that reflects both "flexibility of the focus structure" and the "flexibility of Syntax" in Van Valin's (1999) terms, and is not unique to wh-structures.

Complex variations of acceptable word orders within a language are always a theoretical challenge. Russian is notoriously challenging for linguists because of the ostensible variability of word order options, for which it is often hard to account uniformly, especially considering the vast data variability among speakers.[1]

In this introductory chapter I merely outline the major puzzle presented briefly in (1-4) below, and the main proposal to solve the puzzle, which I discuss throughout the subsequent chapters. Additionally, I discuss how the proposal put forward here fits into the general existing theoretical framework in comparison with other analyses of Russian word order.

Let us first consider the central puzzles of this work. The following sentences exhibit surface word order variability with respect to wh-phrase ordering:

(1) a. **Kto** **komu** darit podarki?
who$_{N}$ who$_{DAT}$ gives$_{3rdSG}$ presents

b. **Komu** **kto** darit podarki?
who$_{DAT}$ who$_{N}$ gives$_{3rdSG}$ presents
"Who gives presents to who(m)?"

[1] For the purposes of the current research and the analyses of constructions discussed in this work, I rely primarily on speaker data collected from a significant number of speakers with a very basic statistical analysis of the results. The methodology is discussed in the following sections.

In (1), either order of wh-phrases (Nom>Dat and Dat>Nom) is acceptable, this is commonly referred to as lack of Superiority. Russian has been therefore categorized as a language that does not show Superiority effects. However, note that when the same question is embedded (as in 2), the order of the wh-phrases is not free: the Dat>Nom combination worsens the utterance, as the contrast between (2a) and (2b) shows :

(2) a. Ja ne znaju, **kto** **komu** darit podarki.
I NEG know who$_{N}$ who$_{DAT}$ gives$_{3rdSG}$ presents
"I don't know who gives presents to who(m)".

b. ?*Ja ne znaju, **komu** **kto** darit podarki.
I NEG know who$_{DAT}$ who$_{N}$ gives$_{3rdSG}$ presents

Additionally, this "inflexible" order of wh-phrases is also observed when one (and only one) overt non-wh-item appears at the left periphery, preceding the multiple wh-sequence. This is shown in (3):

(3) a. Podarki **kto** **komu** darit?
presents$_{ACC}$ who$_{N}$ who$_{DAT}$ gives$_{3rdSG}$
"As for presents[2], who gives [them] to who?"

b. *Podarki **komu** **kto** darit?
presents who$_{DAT}$ who gives$_{3rdSG}$
"As for presents, who gives [them] to who?"

Interestingly, the structures presented in (3) with a clause –initial non-wh-element are impossible to embed, in either order, as shown in (4):

(4) a. *Ja ne znaju, podarki **kto** **komu** darit
I NEG know presents who$_{N}$ who$_{DAT}$ gives$_{3rdS}$

b. *Ja ne znaju, podarki **komu** **kto** darit
I NEG know presents who$_{DAT}$ who$_{N}$ gives$_{3rdS}$

[2] The translation "As for X…" is used loosely here to stress the discourse nature of such a word order.

Here, we see a clear difference in acceptability in matrix versus embedded clauses. In both cases rigid ordering of the wh-phrases exists[3].

I propose a uniform formal syntactic account for the Russian puzzles presented above. On the surface, the solution is fairly simple: I offer to extend the CP field, in the main clause. I give an explanation of such an extension below; more evidence on the workings of the proposed mechanism and its nature is provided throughout this work. I put forward a new category to account for several Russian word order constructions: starting with the derivation of wh-constructions. However, a closer look at such an account gives a rise to a number of crucial general, theoretical, language specific, and implicational questions, which are discussed along the way.

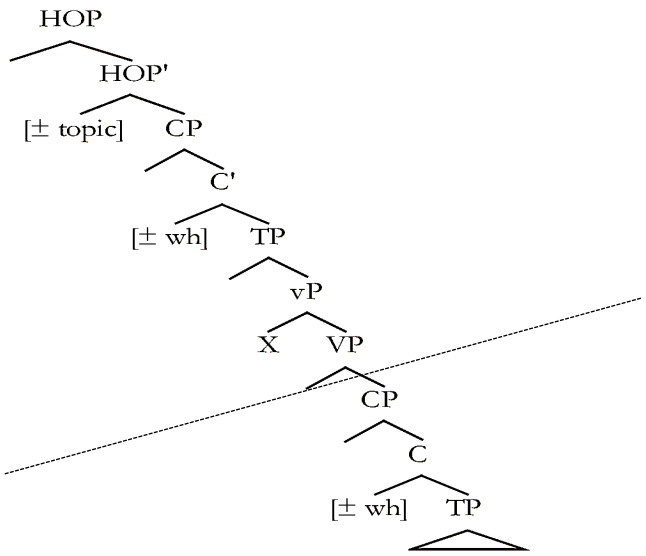
Chapter 2 is devoted solely to wh-behavior in Russian. However, I believe that the approach proposed here for Russian could be extended and allow us to reconsider the typology of wh-questions in general (i.e. in Serbian/Croatian); an extension of these ideas are discussed in Chapter 5. The following Chapters tackle wh-related constructions, i.e. sluicing and multiple coordinated wh-constructions (Chapter 3) and constructions that are discourse related, i.e. using an overt discourse particle (Chapter 4).

My approach to the analysis of the constructions in (1-4) involves a unique category, the High Operator Phrase (HOP), located higher than CP in matrix clauses, which can on the one hand extend the direct mechanisms of wh-movement, and on the other, can host discourse related material. HOP has a dichotomy of features; let us assign it a [±topic] feature for the time being. (I return to question of what can legitimately be hosted in HOP throughout the chapters.) The linear and the tree structure in (5) and (6) respectively illustrate the proposal.

(5) $[\text{HOP}_{[\pm \text{topic}]} \ldots [\text{CP}_{[\pm \text{wh}]} \ldots [\text{TP}]]]$

[3] Often, such variation in acceptability of Russian judgments is assigned to register differences (see Zolotova 2000, Yanko 2001, Svedova 1980, Bonnot 1990, Dyakonova 2009 among many others). Speaker judgments are often inconsistent (Cohen & Cohen 1975, Sprouse and Almeida 2010) and the extent to which they can be inconsistent in Russian appears overwhelming in many cases. Interestingly, some of the inconsistency comes exactly out of the former problem: the existence of registers. Several terms have been used for this: formal/informal (Yanko 2001); written/spoken; standard/colloquial (Dyakonova 2009). Independent of the term, the idea is the same: while in general registers exhibit similar grammatical relations, there are certain differences that separate registers, and in some cases this creates enough data confusion for analysis, especially formal syntactic analysis. Thus, in the course of this work, I assume an analysis of word order that reflects discourse requirements, however, I refrain from discussing the difference in registers and use findings that are irrelevant to the discussion of registers, a practice that has been followed in most recent works (for instance, Dyakonova 2009).

(6) HOP Structure (with relevant features)



(7) Questions about (6):

i. Does such a structure that expands the CP-domain fit cartographic assumptions?
ii. Is this structure able to formalize discourse related notions into syntactic theory? Can the structure that is introduced in (6) be universally applied?
iii. What is the nature of HOP? How can labeling it Operator be supported?
iv. What elements can legitimately appear in HOP?

In the current chapter, I discuss the answer to questions (i), and in part to (ii). Here, I just touch upon the idea of formalizing discourse related categories into a formal syntactic theory. The more profound question of the structure's universality, as well as how such a category can be universally applied (questions ii and iii) are addressed in the following chapters and in detail in Chapter 5. The empirical answer to (iii) is given in each chapter subsequently, while the theoretical grounds for hosting these elements (*whs* and topics) are discussed in Chapter 5.

The objective of this chapter, then, is to set the theoretical field for the current work. Its main purpose is to evaluate the primary proposal given here from several theoretical perspectives each discussed in its turn: an existing Cartographic analysis of Russian word-order (Dyakonova

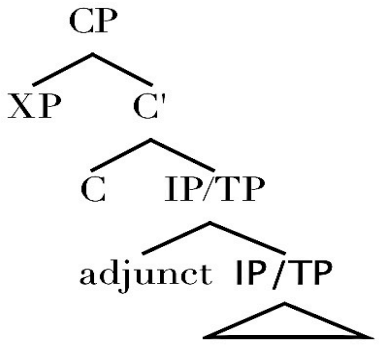
2009); Cartographic views in general as they apply to Information Structure phenomena (Neeleman, de Koot, Titova, Vermeulen 2009, Wagner 2009, Williams 2009), and a brief overview of alternative approaches to Information Structure (Erteshik-Shir 2007). I start by discussing the theoretical approach that underlies the proposal of expanding the CP domain in the first part of this chapter. Then, I move to the discussion of how Information Structure (IS) elements can be formally analyzed, since I assume a topic feature is present in HOP.

1.1 Cartography and its Alternatives

Several theoretic approaches have been suggested to analyze languages that exhibit word order variations that are in part dependent on their discourse properties. Information Structure (IS) is consistently considered a driving force for scrambling-type movements to reflect properties related to discourse. Hence, since IS context is not always present, derivations that would reflect discourse properties are commonly optional, that is to say that if a certain IS context is present, certain reordering rules apply, if not then no additional derivations and/or scrambling take place. Any kind of optionality is a challenge to the Minimalist Program (Chomsky 1995). But this is in fact how some of the behavior in discourse configurational languages looks on the surface: optional. Kiss (1995) defines discourse configurational languages "in intuitive terms" as having: (A) an expression of a particular structural relation (in other words, a specific structural position) for Topics, which "serve to foreground a specific individual that something will be predicated about", and (B) a particular structural relation, or movement into a particular structural position, for Focus. The categories A and B often co-occur but are independent, according to Kiss.

It has been a complex task to formalize such languages' processes under current linguistic theory where optionality is difficult to account for: additional mechanisms and categories are often proposed (as in Kiss 1995, for instance). The fact that the Minimalist Program (Chomsky 1995) has reduced the inventory of possible processes to a minimum compared to earlier versions of the Government and Binding (GB) theory, creates problems for such analyses. The traditional view of the CP-domain is illustrated in (8) where no space for additional categories exists.

(8) CP-domain: traditional approach (Chomsky, 1995)



Some have supported the introduction of a more fine-grained approach to the left-periphery (within, outside, or inside the CP-domain), where additional functional categories (i.e. TopP, FocP) are proposed (Kiss 1995, Horvath 1995b, Liptàk 2011, and Rizzi 1997, 2004).

Since the current thesis offers an extension of the CP-domain[4], the approach developed here might at first sight be perceived as a Cartographic one. In this section, I discuss the basic premises that Cartographic approaches postulate in order to establish whether the proposed account does or does not fit Cartographic views.

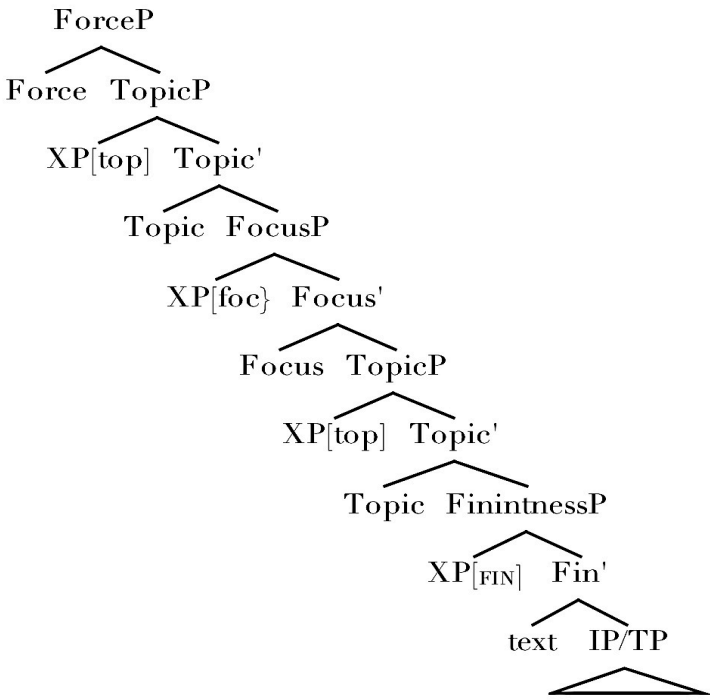
Cartographic investigations developed in parallel with Minimalist investigations. Zwart (2007) states: "It is possible that the two approaches represent two sides of a single coin." The Cartographic research project proposes a systematic matching between morphosyntactic and semantic features and functional projections, providing "as precise and detailed maps as possible of syntactic configurations" (Cinque and Rizzi 2008: 43). Cartography extends the inventory of the functional categories from the original: DP, IP, CP, and vP. Rizzi (1997) proposed to "disassemble" the 'left-periphery' (CP-domain) into ForceP, TopicP, FocusP, and Fin(itness)P.

More specifically, Rizzi (1997), assuming Relativized Minimality and the principle of movement as "last resort", identifies three syntactic layers: (i) the lexical layer headed by the verb in which theta assignment takes place; (ii) the inflectional layer, headed by functional heads corresponding to concrete or abstract morphological specifications on the verb, and responsible for licensing of argument features such as case and agreement and (iii) the complementizer layer, typically headed by a free functional morpheme, and hosting topics and various operator-like

[4] Such an extension might be easier understood as either a Split CP structure, or a CP-shell-like structure (see DeCuba 2007). More detailed explanation of the nature of the CP-extension of matrix clauses only is discussed in the consequent chapters (see also Scott 2010 for preliminary ideas on the extension for Russian).

elements such as interrogatives and relative pronouns, focalized elements, etc. Such a layered system gives a rise to Force-Finiteness System that is also often called a "split" CP structure. Each layer corresponds to a certain function. The Force layer type-clauses the sentence: interrogative, declarative, exclamatory, relative, comparative, etc. The Fin layer determines tense specification, which matches the one expressed on the lower inflectional system. Lastly, the inner domain (lower layer) hosts the content of IP embedded in it. Thus, for Rizzi, the C system is fundamentally distinct from the I system, that is to say that it is not an extension of IP and therefore, of the V system. Additionally, there are optional layers for the Topic – Focus System (Top-Foc), where topic is defined as 'knowledge that the speaker presupposes to be shared with the hearer ("Your book, you should give to Paul *t*")'. Topic and Focus movement are then brought in line with the view that movement (or, in more neutral terms, the construction of non-trivial chains) is "last resort", and must be triggered by the satisfaction of a "criterion" (or feature checking, in Chomsky's 1993 terminology). In its turn the Top-Foc system is integrated into the Force-Finiteness System. The Top-Foc System is optional, but if present must be "sandwiched" between Force and Fin (Rizzi 1997). The structure, then, looks as follows.

(9) CP-domain: Cartography (Rizzi 1997)



In essence, the latest development in Cartographic research proposes what can be termed the "one feature one head" hypothesis (OFOH) (Cinque, Belletti 2004). More specifically, the order proposed in such an approach is rigid, universal, and all clauses are claimed to be identical in structure. Such ordering is entailed by the WH-Criterion (Rizzi 1991): "a Wh operator and a head endowed with the Wh feature must be in Spec/head configuration at S-structure; if the Wh feature is generated under T in main questions, I to C movement must apply to bring the feature to the C system, where the Criterion is satisfied." (Note that here, Rizzi does not specify in which layer of C the wh-feature is housed.) The last important conclusion that Rizzi makes is that there are several Topical elements and only one Focus available per clause. Rizzi thus proposes the following structure for the Complementizer layers.

(10) ForceP >TopP*>FocP>TopP*>Fin>IP

The important generalizations are: there exists an iterative Topic position, a unique Focus position, and no adjunction is possible under the Cartographic one-feature one head (OFOH) rule.

These ideas of Rizzi were widely adopted by Dyakonova (2009) in application to Russian. (In later sections, I discuss how this approach might be problematic, following Neeleman et al. 2009). Before I turn to the discussion of these approaches of accounts of IS word order, let us first define the IS primitives, Topic and Focus. These definitions are crucial for this work.

1.1.1 Discourse primitives

In order to discuss the pragmatic effects derived by various word orders, it is essential to define Information Structure primitives, namely Topic and Focus. The definition of Topic and Focus depends on whether the notions are understood syntactically, semantically, and/or pragmatically. Several statements should be true about IS: the general assumption is that focal and topical information are necessary components of utterance interpretations, all languages have certain means of encoding focal and topical information, and such encoding varies from language to language.

Rizzi (1997) states that "the topic is a presupposed element characteristically set off from the rest of the clause by comma intonation ("Your book$_i$, you should give t_i to Paul") and normally expressing old information, somehow available and salient from previous discourse;

the comment is a kind of complex predicate, an open sentence predicated of the topic and introducing new information." While some of the definitions state that Topic is "old", "assumed", "given" information, which is (somehow) salient from the discourse, others state that Topics are what the utterance is about ("aboutness" topics). The theory of Informatics developed by Vallduví (1992) initially extracts Focus, the informative part of the utterance. The Ground acts as a "vehicular frame for the informative focus", in other words it guarantees the appropriate entry of the information (Valluduví 1992:46).

(11) IS division of the utterance (Vallduvi 1992 :46)

Sentence = {Focus, Ground}
Ground = {Link, Tail}

Erteshik-Shir (2007) provides descriptions and measures of topichood: all specific DPs that provide the hearer with "old" or "given" information and may be dislocated to the left-periphery (on the example of Danish). Vallduví (1990) discusses two types of old information in Catalan: defined as Links, which are left-detached (or topicalized) and Tails, which appear at the right. Topics can also be "Stage" topics, those that have entered the discourse as being already set-up. This is easy to understand because every utterance is produced in a specific place at a specific time. Thus, several researchers claim that topicalized material often occurs at the left-edge of the clause, or pre-verbally.

An example of a Stage topic is given below in (12):

(12) Q: What happened?
A: John has washed *the dishes*.

(13) It's dark. *The moon* has disappeared.

Erteshik-Shir also defines "permanently available topics", such as 'the president', 'the train', 'the moon': these do not need introduction in previous discourse (as in (13)). The distinction between different types of topics is essential for this research, and is discussed in more detail in Chapters 4 and 5.

Several important questions arise from the variety of possible constituents that can be defined as Topics. Consider the following definition from Strawson (1964), who has identified three central properties of topics:

(14) Strawson's (1964) definitions of Topic:

(a) The topic is what a statement is about.

(b) The topic is used to invoke "knowledge in the possession of an audience."

(c) "The statement is assessed as putative information about its topic."

The property in (a) is similar to the definition of Links from Vallduví (1990), the property in (b) is nothing else but "old" or 'given'. The property in (c) however, talks about the assessment of the topic referent and the truth value of the sentence. If there is nothing to refer the topic to in a sentence, then the sentence gaps in truth value, in Strawson's terms "a statement which, by hypothesis, is about something is really about nothing."

(15) The King of France is bald.

(16) The exhibition was visited by the King of France.

Then, a logical question about truth value of topics arises: how to assess the truth values of Topicalized sentences? Some of the ideas along the lines of the solution comes from the fact that topics are commonly specific and hence are a pivot for the truth value assignment where truth value comes out of the property of specificity (Reinhart 1982 "aboutness" topics). One of the important results of the definition of a topic as the pivot for truth value assignment is that the scopal relations in the sentence depend on topic assignment: The topic will have wider scope than other elements in the sentence because the predicate of the sentence is evaluated with respect to the topic.

The second question is the availability of "multiple" topics within one utterance. That topic-focus assignment is recursive has been argued by Krifka (1992), Lambrecht (1994), Erteschik-Shir (1997) (among others). Let us consider the following question-answer pair (17):

(17) Q: Did *you* see Peter yesterday?
A: Yes, *I* saw *him* yesterday at school.

It is apparent that there are two pronouns that can be considered Topics in the answer statement. If the truth value is based on the notion of topic, as Erteschik-Shir (2007) believes, having two elements that are topics in the evaluation the sentence will become ambiguous. Hence, she states there should be just one "main" topic – what the sentence is truly about. The rest of the topics if

present are "subordinate". However, such definitions are more within a semantic domain do not assign specific structural positions to either Topic or Focus. Still, one might assume that there should be structural difference between "main" and "subordinate" (in Erteschik-Shir terms) topicalization.

According to Kiss (2007), both Topic and Focus are defined by syntactic, semantic, and prosodic properties. She takes this assumption even further for Hungarian in stating that Focus and Topics are structurally defined in Hungarian – are assigned specific structural positions, and thus, in the definition of each she defines the structural position of both:

(18) **FOCUS** (Hungarian, Kiss 2007)

> The focus is an immediately preverbal constituent, expressing exhaustive identification, bearing a pitch accent.

"Syntactically", - she notes, - "the Hungarian focus is an XP occupying an invariant A-bar position, identified by Brody (1990) as the specifier of a FocP. The finite V, which follows the verbal particle in neutral sentences (10a), is left-adjacent to the focus (10b), which may be due to V movement across the particle – into the head of a Non-NeutralP according to Olsvay (2004). FocP is subsumed by TopP"(Kiss 2007: 74)

(19) **TOPIC** (Hungarian, Kiss 2007: 70)

> The topic is an XP extracted from the functionally extended verb phrase into the left periphery of the sentence. It precedes the pitch accent that marks the left edge of the functionally extended verb phrase in Hungarian. It is interpreted as the logical subject of predication.

Kiss (2007: 70) states that

> On the syntactic level, the topic is an argument preposed from the maximally extended verb phrase into clause-initial position, with a trace/copy in the vP. Sentence adverbials base-generated external to the maximal verbal projection are not topics. Referential locative and temporal adverbials, however, can be analyzed not only as sentence adverbials but also as optional arguments binding traces in the vP, hence they can function as topics in the left periphery. The landing site of topics is the specifier of the functional projection TopP. In the case of multiple topicalization, the iteration of TopP is assumed. The relative order of topics and sentence adverbials is free.

Defining Topic and Focus on the basis of their structural position seems to be a non-generalized but a logical approach (not everyone defines these notions through their position in a structure). However, in order to appear in a certain position structurally, Topics and Focus have to maintain the properties of topicality and focusability. Thus, there are certain elements that can appear in a focused position: wh-phrases, certain (not universal) quantifiers, phrases marked by *only* can appear in Focus position (at least in Hungarian).

For Russian, Dyakonova (2009) gives a definition of Focus and Topic without referring to specific positions but rather defines them as "new" for Focus and "shared" for Topics. One of the reasons she avoids defining these notions structurally is that Focus (rather than Topic) in Russian does not always appear in a specified position (it might be post- or pre-verbal unlike Hungarian). It is believed cross-linguistically that Focus appears in the left-periphery; however, in Russian (and in Spanish (Dominguez 2004)) focused constituents may appear sentence finally (at the edge of the right periphery), in the left-periphery, or *in-situ*. Thus, a purely structural definition on a par with the one given for Hungarian does not work for Russian at least when it comes to Focus in Dyakonova's (2009) analysis.

(20) **Focus** (Dyakonova 2009):

> **Foc** refers to material or expression that is *prosodically or syntactically prominent* because it introduces "new" information.

(21) **Topic** (Dyakonova 2009):

> **Top** refers to material or expression that is generally "old", "known" or "commonly shared".

To summarize this section, we can state that Topic and Focus are often defined in terms that reflect the discourse information of a given language, which can be identified by stress, contrast, and such. It is not nevertheless immediately clear how all this prosodic, semantic and pragmatic information is encoded into a syntactic derivation. Under some theories, it is expressed at the level of S-structure in a Spec-Head relation, or under FocP or TopP configurations, and it is interpreted at LF or even later (*Assertion Structure*: Zubizaretta 1998, *Functional Form*: Bailyn 2012). The features that distinguish between these two categories are often argued about. In this

sense defining these two notions structurally might be an attractive approach when they constrain specific word ordering requirements.

Thus, for the purposes of the current work, I assume Dyakonova's general definitions of Focus and Topic in (20) and (21), though I refrain from her structural analysis of how such sentences are derived. Let us see how the analysis of Russian by Dyakonova (2009) works. I briefly evaluate her account in the following section: her approach uses Rizzi's (1997) structure (10) as one of its main premises.

1.1.2 Dyakonova's 2009 "Phase-based Cartographic Approach" to Russian WO

Dyakonova (2009) outlines four questions related to word order variation in Russian. First, following the empirical evidence that topical and focal information in Russian is "edge-oriented" (Topic is sentence initial and Focus is usually sentence-final), the question arises whether a word order permutation phenomenon is just a linear PF requirement and is not related to any deep syntactic motivation. Secondly, if Syntax plays a role in encoding IS, what is the nature of the syntactic operations: are they accessing IS as a feature-checking operation or as an independent IS-related derivation? Third, assuming Chomsky's (2001) approach to computation as cyclic, or by phase, can it be that the IS-related information is relevant at the phasal rather than at the clausal level? And fourth, since word order requirements are often treated as optional in Russian, "is it possible to reconcile the syntax of Russian with the recently developed theory whereby any syntactic operation is motivated by feature-checking?" (Dyakonova 2009)

In order to approach any kind of word-order permutations (optional or not), it is important to understand underlying basic Russian word order (the myth of the freedom of word order has long been dispelled in the literature, see King 1995, Bailyn 1995, 2003, 2004, Junghanns and Zybatow 1997, Sekerina 1997, Rodionova 2001, van Gelderen 2003, Pereltsvaig 2004, Slioussar 2007, among others). The general agreement is that Russian exhibits underlying SVO order, and Dyakonova (2009) proposes a Basic Word Order Rule for Russian presented in (22).

(22) Basic Word Order Rule for Russian (Dyakonova 2009):

> In neutral contexts the external argument always precedes the verb and the internal argument always follows it, unless the internal argument is a pronoun of any kind, in the sentence final position. In the latter case the internal argument tends to precede the verb.

In the Basic WO Rule "the external argument" refers to Subject and "internal arguments" refers to Objects (both direct and indirect). Hence, any deviation from the general rule produces and reflects certain pragmatic effects, in Dyakonova's own terms: specific context-related interpretations. While Dyakonova does not state what these effects are, one can assume that topical, focal, and question constructions would be the first ones to deviate from the Basic WO Rule (since these constructions require additional derivations and movement).

While a substantial part of Dyakonova's thesis is devoted to the discussion of the hierarchy of internal arguments, I will leave out this discussion here, and return to it when necessary in the discussion of the empirical issues in the relevant chapters of the current work. Here, I examine her theoretical assumptions in detail.

The premises that are used in her thesis are based on a number of theoretical postulates that create the foundation of her account. On the one hand, she refers to phase-based derivations, parallel chains, and locality as it is worked out within the framework of Derivation by Phase (Chomsky 2001, 2005). On the other hand, she relies heavily on the structures proposed by Rizzi (1997) accounting for the left-periphery phenomena. Let us discuss whether the two are compatible.

Within the phase-based theory, a syntactic structure is decomposed into relatively independent chunks, or phases; usually CP and *vP are considered phases (Chomsky 2001). Dyakonova also assumes that structures are built by the operation Merge: internal or external. External Merge mainly satisfies theta-requirements. Internal Merge is what is otherwise known as Move. The derivation is assumed to proceed strictly by phase. Additionally, following Pesetsky and Torrego (2007), a feature-checking procedure is assumed. Pesetsky and Torrego (2007) propose that every syntactic feature has two independent components: interpretability and valuation, which results in four possible options of each feature: [+interpretable, +valued]; [-interpretable, -valued]; [+interpretable, - valued], and [- interpretable, + valued]. Such a description of features separates them into Probes and Goals. The process of Agree then is described in:

(23) **Agree** (Pesetsky and Torrego 2007 as presented in Dyakonova 2009: 22)

(i) an unvalued feature F (a probe) on a head H at syntactic location α (Fα) scans its c-command domain for another instance of F (a goal) at location β (Fβ) with which to agree.

(ii) Replace Fα with Fβ so that the same feature is present in both locations

Additionally, Dyakonova (p.24) states that

> One of the problems with [Topic] and [Focus] often noted before is that they do not conform to the conventional way of feature checking, whereby an uninterpretable-interpretable pair is of crucial importance. Rizzi (1997) resolves this issue by treating Topic and Focus as Criteria satisfaction rather than feature checking, because they must be interpretable on the attracting head and on the attractee and cannot be deleted. However, the revised feature-checking-as-feature-sharing approach helps to avoid the problem of LF interpretability. Features are not deleted in this model. Moreover the distinction between value and interpretability allows to treat [Topic]/[Focus] on a par with other features, say [Tense]. Let me illustrate how the system can work using [Focus].

She then uses Rizzi's (1997) Force–Fin structure presented in (9) above, which becomes crucial for her account. The application of a cartographic structure, in the cases where Minimalism assumptions limit her account for certain word order facts, forces her to create a number of rules that systematically account for any data that she presents. Her claim rests heavily on cartographic assumptions but they alone do not in fact account for the attested word orders. Using all the background theoretical postulates listed above, her proposal to account for word order variations comes down to the following (here, I only present theoretical proposals without additional discussion of the empirical notions).

In Chapter 3, stating that in every Russian sentence word order is constrained by IS, she puts forward the following general structure of the Russian utterance (reproduced here in (24) Dyakonova 2009: 55 (3)), which is not directly related to cartographic categories:

(24) *IS Ordering Rule:*

Topic > Discourse Neutral Material (DNM) > Focus

By using this rule and investigating various word orders, Dyakonova is able to draw a comprehensive typological inventory of all possible word orders in Russian, SVO, VSO, OVS, where SVO is considered the discourse neutral word order. In this she derives the general underlying Russian word order convention. Everything else has to be identified as edge-oriented Focus or Topic.

TABLE 1. CORRELATION BETWEEN WORD ORDER AND IS ARTICULATION IN RUSSIAN (Dyakonova 2009, ch.III)

Structure	IS composition
S V O	$[S]_{Topic}$ V $[O]_{Focus}$
	$[S]_{Topic}$ $[V\ O]_{Focus}$
	$[S\ V\ O]_{Focus}$
S V IO DO	$[S]_{Topic}$ V IO $[O]_{Focus}$
	$[S]_{Topic}$ $[V\ IO\ DO]_{Focus}$
	$[S\ V\ IO\ DO]_{Focus}$
Structure	IS composition
S V DO IO	$[S]_{Topic}$ V DO $[IO]_{Focus}$
S O V	$[S]_{Topic}$ O $[V]_{Focus}$
(O/Avd) V S	$[O/Adv]_{Topic}$ V $[S]_{Focus}$
	$[O/Adv]_{Topic}$ $[V\ S]_{Focus}$
O_1 V O_2 S	$[O_1]_{Topic}$ V O_2 $[S]_{Focus}$
V O S	$[V\ O]_{Topic}$ $[S]_{Focus}$
V S O	$[V\ S\ O]_{Focus}$

However, *no motivation for Topic or Focus appearing on the edges is provided*. In order to account for edge phenomena, she proposes a D-Linking Rule that she also calls the Scrambling Rule, provided in (25) below. (Note that Dyakonova assumes that IS encoding and D-linking encoding are two independent mechanisms, i.e. D-linking is not an IS primitive.)

(25) *Scrambling Rule* (Dyakonova 2009: 75 (57))

A D-linked constituent should be preposed to a position in the pre-verbal area

In her words, "this rule has to do with the referential status of constituents and demands that D-linked constituents be preposed to a preverbal area." The interaction between the two ordering principles result in "ranking" the rules with respect to each other. While the two rules can in effect produce the same results, according to Dyakonova, they should be kept apart in order to result in correct word orders. Neither is strictly Cartographic.

In her discussion of Topic and Focus in Russian, Dyakonova (2009) makes several solid empirical observations. First, she considers which positions of the left-periphery can be targeted by Topicalization, where the latter is understood as a broad term for movement. She claims that there are several such positions: a unique TopP position for a unique "strong" Topic, using her terminology, which is not recursive. Crucially, for our purposes, she states this about the

ordering between Focus and Topic and wh-phrases: Topicsa always precede Focus; however, wh may precede or may follow Topics (wh>topic, topic>wh). However, these orderings differ because there are several different topics available. The ordering depends on which kind of topic is present. Using Rizzi's (2004) proposal, she distinguishes between three types of landing sites for Topicalization: FrameP, TopP, and topP, which follow this specific hierarchy: FrameP>TopP>topP. Of these, TopP is the only possible landing site for the Topic in a traditional sense: "what the sentence is about" (or strong Topic). Discourse-oriented constituents, or weak topics, are hosted by freely generated *topPs. Moreover, she argues for the existence of a separate projection for frame-setting adverbials, FrameP. These can be either a single constituents or a domain. Hence, she proposes the structure shown in (26):

(26) [ForceP Force^{o} [FrameP Frame^{o} [InterP Inter^{o} [TopP Top^{o} [topP top^{o} [FocP Foc^{o} [topP top^{o} [FinP Fin^{o}]]]]]]]]

The structure in (26), essentially the one proposed by Rizzi (1997, 2005) is cartographic, and therefore, presumes a rigid ordering of each category within the clause, and between clauses. However, since Dyakonova postulates a phase–based approach, she immediately states that there is a question about phase edges that differ from each other. It appears that some of the projections are missing in lower and higher phases. She relates this to the way the information is encoded within domains, and that such information varies. In her words, "it means that the two phases differ with respect to how much of the "outer", discourse-related, information must be accessible with the phase when the phase is being constructed" (p.153). However, it is not clear which phases she is talking about using the structure in (26) under the assumption that phases are usually considered to be CP and *vP. This brings us to the larger issue of how phases and Cartography relate.

Gallego (2009) convincingly argues against the compatibility of Cartography and phases. Bailyn (2011) points out (emphasis is mine):

> Gallego [2009] shows that individual languages treat PHASES (CP, vP, and DP) in uniform fashion, and that the parameterized behavior of phases in a language determines significant generalizations *beyond the descriptive reach of cartography*, which distinguishes phases from other functional categories.

In other words, it must be independently determined which categories act as phase edges within the exploded CP domain, and even if IS notions are sometimes related to phase edges (which can only account for left-edge effects of course), a cartographic account alone does not suffice.

In brief, the facts that Dyakonova attempts to account for are the following: In IS-related word order, Topics appear at the left edge of the clause (see Table 1), Focus tends to appear clause-finally, but may also appear in a pre-posed position, the optional ordering of quantifiers depends on their scope, and wh-phrases vary in position with respect to Focus and Topic within the sentence. In order to account for all these word order variations, Dyakonova provides a number of mechanisms where one overrules, over-ranks, or substitutes for the other, and thus makes her analysis in principle unfalsifiable. This is a problem with Cartography generally, as Bouchard (2010) points out: the descriptive tools of Cartography can successfully derive any word order and scope. And this is consequently true for Dyakonova's account.

To summarize, Dyakonova's (2009) account is dexterous in putting together empirical facts and a considerable number of assumptions and machinery of rules to account for them. However, at a certain point, the number of mechanisms used (the two word order rules independent of Cartography: *IS* and *Scrambling*, along with Rizzi's rigid structure of the left-periphery), leads to significant confusion. When so much machinery is applied, it can of course account for a lot of facts: if one mechanism does not work, the other one would. It is not clear how these mechanisms are compatible with each other.

One of the major problems that Dyakonova's account faces is the use of the rigid cartographic structure to account for IS related structure. It is this approach that makes her analysis partially unsuccessful. The impossibility of utilization of Cartography to account for "fleeting" IS-related elements is nicely discussed in the work of Neeleman et al. (2009) as well as in the work of Abels (2009) and Wagner (2009). In the following section, I briefly outline their argumentation and alternatives as well as stating how the account proposed in this work is compatible with their suggestions.

1.2 Information Packaging: To Map or Not to Map?

1.2.1 Analyzing Focus and Topic

Cartography postulates two main principles. First, it assumes a one-to-one correspondence between a position and interpretation. Second, it decomposes syntactic structures into categories. The first principle should result in Topic and Focus being realized in designated functional projections. The second leads to the idea that if two interpretations are present, then, two positions should license each of the interpretations. Neeleman et al. (2009) show that if one considers the notion of Contrast, one can see that there should be distinct positions for Focus, Topic, and Contrast. In fact, however, Contrast can only appear as a "sub"-interpretive feature of either Focus or Topic. Molnàr (2006) gives the following definition for Contrast:

> …the notion of contrast must be regarded as a linguistically relevant phenomenon, because contrast is a necessary condition for the use of certain syntactic and phonological means - both in the field of focusing and topicality. The data from different languages show, however, that this notion is important to a different degree - depending on which language and which specific structure is considered. In those languages where formal marking of contrast is only optional and not unambiguous (as in English or German), the strict distinction of the contrastive and non-contrastive cases is empirically not so well-founded - supporting the gradient view of contrast. From a cross-linguistic perspective it seems well-motivated, though, to assume that the notion of contrast has a crucial impact on linguistic forms, in many languages demanding an obligatory formal realization. (p. 106)

Thus, if Contrast is a separate feature, there should not be any problem in treating Focus, Topic, and Contrast as different privative features that are realized in a specific structural position in accordance with the Cartographic tenet. The problem arises when either of these features can appear in different position in a language without a visible motivation for such behavior.

One of the apparent problems with Cartography is that Cartography supposes that features and functional categories should always have a one-to-one correspondence, but in reality they often do not. It is important to keep in mind that one should define syntactic categories and syntactic functions independently of one another, since one function may be served by various categories, and one category may serve more than one function. Hence, the idea of syntactic decomposition becomes impossible in cases where a one-to-one feature/category is not borne out. Similar issues appear in other attempts to discuss Information Structure categories within cartography, as pointed out by Neeleman et al. 2009 and Wagner 2009.

Neeleman et al. (2009) discuss the problems that arise with a cartographic analysis of discourse related categories such as [contrast][5]. They state that topic and focus do not have designated positions in the clause, and this claim runs counter the tenet of Cartography of OFOH. The presence of [contrast], according to them, is conditional on the presence of [topic] or [focus]. Thus, an encoding of these features as head of separate functional projections TopicP, FocusP and ContrastP cannot capture the data where occurrence of these features (a) optional, (b) does not always happen in the same designated position of the clause. The sample languages are Dutch, Japanese, and Russian. Dutch is used as an illustration for the distribution of the feature of Contrast, Japanese is used as an illustration of a Topic-fronting language, and Russian is used in the discussion of Focus.

First, let us consider the following examples from Dutch that illustrate that A'-scrambling can target various positions irrespective of whether the moving phrase is a Topic or Focus. It can land between the complementizer and the subject (27), between the subject and the indirect object (28), or in the first position in the main clause (29). Neeleman et al. (2009) claim that other landing sites are available in structures that contain adverbs that are usually freely ordered in Dutch with respect to moved topics and foci. Conisder the following examples from Dutch (Neeleman et al. 2009: 20 (6-9))

(27) a. Ik geloof dat [alleen DIT boek]$_1$ Jan Mariet$_1$ gegeven heeft.
I believe that only this book Jan Mariet given has
"I believe that Jan has given only this book to Mary."

b. Ik geloof dat [zo'n boek]$_1$ allen JAN Mariet$_1$ gegeven heeft.
I believe that such a book only JAN Mariet given has.
"I believe that only Jan has given such a book to Mariet."

[5] The notation here might deviate from the general notation of the thesis but is maintained from the original work. Different notation does not entail reference to different categories, however.

(28) a. Ik geloof dat Jan [alleen DIT boek]$_1$ Mariet$_1$ gegeven heeft.
I believe that Jan only this book Mariet given has
"I believe that Jan has given Mariet only this book."

b. Ik geloof dat JAN [**zo'n boek**]$_1$ allen Mariet$_1$ gegeven heeft.
I believe that JAN such abook only Mariet given has.
"I believe that Jan has give this book only to Mariet"

(29) a . [Alleen DIT boek]$_1$ zou Jan Mariet$_1$ geven.
only this book would Jan Mariet give
"John would give Mary only this book."

b. [**Zo'n boek**] zou alleen JAN Mariet$_1$ gegven
such a book would only JAN Mariet give
"Only Jan would give Mariet such a book."

Neeleman et al. (2009) show for Japanese that Topics do not necessarily occupy clause initial positions, and that Contrast can shift the position of the topicalized unit. In their discussion of Russian, they postulate a rule for Focus licensing in Russian (30):

(30) Focus is licensed in clause-final position. (Neeleman et al. 2009: 36)

The following examples illustrate the application of this rule in Russian.

(31) **Čto** čitaet Saša? Saša čitaet KNIGU.
what reads Sasha Sasha reads book
"What does Sasha read?" "Sasha reads a book."

(32) **Komu** Anja dala knigu? Anja dala knigu KATE
who Anja gave book Anja gave book Kate
"Who did Anja give a book to?" "Anja gave a book to Kate."

However, the Focus shifts to the front of the clause when it's contrastive.

(33) JAZZ-PIANISTA mal'čiki slušali vystuplenie (a ne jazz-gitarista).
jazz-pianist$_{GEN}$ boys$_{NOM}$ heard performance and not jazz guitarist$_{GEN}$
"The boys heard the performance of the jazz –pianist, and not the jazz –guitarist."

As a result of the analysis of separate data, the authors conclude that according to the requirements of decomposition, Contrast should be assigned a separate special functional category ContrastP, but this is not possible, because this is the feature that is not separable from Topic and Focus. However, its occurrence with either one might impact where either TopicP or FocusP would appear in a derivation, and hence, violates the rigid clausal composition required by Cartography.

To sum up, it appears that a Cartographic account is insufficient because the requirements of Cartography are either too strong (an existing position for the feature must be filled) or too weak (a rigid position for the feature cannot be moved), for such an account.

Wagner (2009) discusses the possibility of the analysis of Focus within Cartography and refers to Neeleman's et. al. (2009) account as a "templatic account" in which there is a template of the hierarchy of phrases TopP FocusP and ContrastP, in which some possibilities are ruled out and some are not. He suggests that while the tenets of Cartography are in fact a reasonable approach to analysis of some structures, he proposes not a decompositional but a compositional approach to the analysis of Focus.

Wagner states that "up to a point" the two approaches are compatible. However, some word order patterns found across languages might not fit into the "rigid theory of a universally fixed total ordering". I have shown similar impossibility of "fitting in" for Russian on the example of Dyakonova (2009) discussed in the previous sections. Any attempt to formalize the discourse categories into a rigid Rizzian structure required additional mechanisms (as seen in her account). Her account is not incompatible with a Cartographic one, since she uses the Force/Fin structure as the basis for any structure. Wagner (2009) states that since the approaches are possibly compatible, but there are obvious problems with Cartography, then perhaps, some of the assumptions have to be re-evaluated.

Neeleman et al. (2009) and Wagner (2009) show that the Cartographic approach is not sufficient, or rather not entirely sufficient, to account for Information Structure categories such as Focus, and more specifically, Contrast.

1.2.2 Cartography and Its Alternatives

In the previous section, I have shown how a cartographic approach to word-order in Russian is not sufficient to account either for IS phenomenon, or for other phenomena that are related to word order, such as wh-movement. It requires too much machinery and hence, appears to be unfalsifiable. Furthermore, I have discussed the general claims of the Cartographic program and whether it is compatible with approaches to account for the distribution of such categories as Topic, Focus, and Contrast. On the basis of the research provided by Neeleman et al. (2009) and Wagner (2009) we have seen that with the way the Cartographic tenets are applied, accounts of information packaging apear unfeasible, at least in a way the principles are postulated now. Even researchers that work within the Cartographic program have to make a number of additional assumptions (Dyakonova 2009 for Russian, see also Liptàk 2011 for Hungarian). Therefore, we have to look at other possibilities in analysis of IS phenomena. In the next section, I provide a brief overview of the basic alternatives to the analysis of IS from a syntactic and a pragmatic point of you.

1.2.3 Approaches to Information Packaging

Before the Cartographic research program appeared, the question of how to account for Information Structure phenomena, or for the distinctions between focus, presupposition, and propositional attitude toward entities in the discourse conveyed by phrasal intonation, and broadly understood as a trigger for an optional divergence from the basic word order to reflect certain pragmatic/semantic effects of the utterance, was widely discussed. Some provided purely syntactic explanations, while others have provided purely pragmatic explanations, and others had proposed mixed views, which allow semantico-pragmatic triggers to syntactic operations. Most agree, however, that the linear order of constituents has to do with notions that are related to what is contextually known and what is not, at least to some extent. The definition of these notions and their realization in syntax has been the topic of much research since they were introduced by the Prague School in the beginning of the twentieth century. Different terms have been used for IS since Halliday 1967, each accompanied by an explanatory approach: Chafe (1976) coined the term “packaging” to refer to a speaker’s use of a particular syntactic form to serve a particular pragmatic function, or, in his words, to accommodate “states of the addressee’s

mind." Prince (1998:208) puts it more succinctly: "speakers seem to form their utterances so as to structure the information they are attempting to convey, usually or perhaps always in accordance with their beliefs about the hearer: what s/he is thought to know, what s/he is expected to be thinking about." Vallduví (1990) uses the term information packaging. Other terms are topic-focus-structure, or focus structure as in Erteschik-Shir (1997, 2007). In more basic decomposition IS structure consists of Topics, Focus, and possibly contrast.

Thus, the problem arises how to account for languages that exhibit Information Structure phenomenon and how other syntactic phenomena are related to IS if it is. Erteshik-Shir adopts Van Valin's (1999) dichotomy of language types with respect to the flexibility of IS organization and the flexibility of Syntax.

TABLE 2. IS V. SYNTAX LANGUAGE TYPOLOGY (Erteschik-Shir 2007:82)

	Rigid Syntax	**Flexible Syntax**
Rigid Focus Structure	French, Toba Batak	English, Toura
Flexible Focus Structure	Sesotho, Italian	Russian, Polish

It is not surprising that Russian occupies the position along with Polish. In Chapters 2 and 3, I show that there are certain distinctions between word order possibilities in Russian and in Polish. This falls out of the claims put forward here: one of which is that some word orders in Russian are fairly rigid, such as the context of the emergence of Superiority, which will be discussed in detail in Chapter 2. That is why even such a dichotomy is not able to divide languages into configurational and discourse-configurational unambiguously. Languages may choose to mark only the topic or only the focus configurationally.

There are several additional questions that arise: intonation marking of specific discourse related categories, and obligatoriness of Topic/Focus marking. In English, topicalization by movement is possible, but is not often employed, in Danish it is common, but not obligatory, in Hungarian it is strongly preferred. I extend the discussion here, and state that Russian behaves on a par with Hungarian in this respect: moved topics are strongly preferred. Whenever the Focus exists, it must be marked by intonation (Erteshik-Shir 2007).

The idea of intonational focus marking comes from Selkirk (1995). Selkirk suggested that F-marked elements possess a special status in an utterance, and even if they are not marked with

Focus, they do represent "new" information. Russian seems to be different in this respect. Russian exhibits 7 basic Intonation Contour patterns, only some of which would result in a "neutral" interpretation of the utterance, and some of them can be combined, which would add pragmatic/semantic interpretation (Bryzgunova 1999). The patterns are discussed in detail in many works on Russian (Bailyn 1995, 2012, King 1995, Yanko 2001 among others). However, for the account put forward here, intonation patterns are not of primary importance.

Thus the account that I propose is a syntactic but not a Cartographic one. In the final chapter, I discuss the nature of HOP and how it fits into the existing body of theoretical findings, and whether it can be applied uniformly to other languagues, i.e. Serbian/Croatian that exhibits similar behaviors. The discussion of these issues folds out of the main claims of the following Chapter: The Emergence of Superiority Program.

1.3 Chapter Conclusions and Dissertation Organization

1.3.1 Chapter Summary

In this chapter, I have put forward a simple sketch of the main claim behind the account that is used throughout this dissertation, the existence of a main clause left-edge phrase HOP. Such an account does not amount to a commitment to a cartographic analysis. I provided a critical overview of previous cartographic work on Russian word order (Dyakonova 2009), and have argued how her account is unfalsifiable and requires numerous assumptions beyond Cartography to account for word order facts. I have also examined Cartographic approaches and determined that the principles of Cartography alone do not account for IS phenomenon. On the other hand, the options of the Minimalist Program (Chomsky 1995) compared to Principles and Parameters also limit the possibilities of approaches to IS without suggesting an alternative. Hence, I proposed a high functional category, but did not otherwise adopt a cartographic structure, providing the evidence how my account is not cartographic.

HOP provides the correct architecture of the Russian syntactic sentence structure in structures with wh-movement, wh-related, and discourse-related utterances. These are discussed in Chapters 2, 3, and 4 respectively.

1.3.2 Dissertation Organization

This dissertation consists of 6 chapters. In Chapter 1, I I have briefly presented some word order puzzles of Russian and summarized the basic ideas put forward in the current work, namely that of a single high functional category in the main clause only, HopP. This is the basic structural geometry that accounts for a range of Russian word order variations. Hop is ususally ssociated witth a ususal [Topic] function, though we shall see that other elements can also be located there, including wh-phrases. I have also discussed general theoretical problems that arise in accounting for the discourse-syntx connection in languages with variable word orders. I have shown how such word orders are claimed to be accounted for within the Cartography framework, and how cartography is not enough.

Chapter 2 is devoted to the analysis of wh-behavior in Russian, and presents us with the contexts of the Emergence of Superiority. The Emergence of Superiority is a keystone of the discussion of this work. The chapter also discusses the theoretical implication of Superiority as a phenomenon that is claimed to be unparameterized. I provide a detailed discussion of how Superiority operates, and take into consideration existing theories and facts of Superiority. I also establish the position of Russian in the wh-Slavic typology following Rudin's (1988) original discussion. I claim that Russian belongs in the class with Bulgarian and requires wh-fronting of all wh-elements into CP. I also show the environments in which Superiority emerges, relying on a statistical study of grammaticality judgments. I give an explanation for how Superiority works and why there is a misconception about Russian wh-behavior that has been analyzed as either a wh-in-situ language or as a language where wh-phrases move to a Focus position. I show how these accounts fail and provide an alternative account of wh-behavior: The HOP category allows me to account for both the Emergence of Superiority and a seemingly "unsuperior" word order when multiple wh-phrases are utilized.

Chapter 3 covers wh-related phenomena such as sluicing and coordinate multiple wh-constructions. With the findings of Chapter 2, I show that sluicing in Russian can be reduced to regular ellipsis of IP following Merchant's (2001) account. I then discuss multiple wh-movement, including coordinate multiple wh-constructions. I examine existing and possible analysis of multiple wh-coordination and how they can be applied to Russian. In order to provide an analysis of these constructions, I create a typology of coordinate wh-behavior in Russian which consists of three types of coordination: coordination of two arguments, of two

adjuncts, and of an argument and an adjunct. I analyze each type separately. I follow the analysis of coordinate wh-constructions provided by Gribanova (2009), however, I provide additional findings and account for the mechanics of a ConjunctionP analysis. I utilize Nuñes (2001) sideward movement analysis, with discussion of how ConjP is built. I also discuss the availability of the reverse sluicing analysis (Giannakidou & Merchant 1998), and its limitation. As a result, I provide a mechanism that accounts for the peculiarities of multiple wh-coordination in Russian which are discussed in this chapter in detail: Superiority, and the unavailability of certain types of coordinated multiple wh-movement.

In Chapter 4, I tackle another phenomenon of word order that can be analyzed with HOP. N these cases, a wh-phrase is not present in HOP, HOP has a [+topic] feature, and it is able to host overt topics. I discuss the Russian particle – TO, which I claim is located at the head of HOP. I discuss how–TO marked constructions work. I also show how they are incompatible with wh-movement into HOP when the [+topic] feature is present. This provides additional evidence that HOP is unique and can only host one element at time. I also show that fronted –TO-marked constructions are impossible in embedded clauses.

Chapter 5 addresses the question of the nature of the category HOP introduced here, discussing the elements that legitimately can appear in this position, and how it can be universally applied. In essence, it discusses the consequences of the current research for the syntactic theory in general.

Chapter 6 provides a summary and conclusion of all the findings of the dissertation, and possible extensions of the research undertaken here.

Chapter 2. Russian WH: HOP and The "Emergence of Superiority" Program

2.0 Introduction

The current chapter examines wh-behavior in Russian. Universal Grammar (UG) allows three distinct types of languages with regard to wh-questions (Richards 1998): languages like Bulgarian, where all wh-elements move to the front (i); languages like Japanese and Chinese, where all wh-elements stay in situ (ii); languages like Italian and English, which move only one wh-element to the front while leaving the other(s) *in-situ* (iii).

(i)	[CP wh-phrase C wh-phrase C [TP ... t ... t ...]]	(Bulgarian, …)
(ii)	[CP [TP wh-phrase . . . wh-phrase]]	(Chinese, Japanese, …)
(iii)	[CP wh-phrase C [TP . . . t . . . wh-phrase]]	(English, Italian, …)

There is little consensus on the place of Russian in this typology. It has been claimed under some analyses that Russian patterns with type (ii) such as Chinese and Japanese (Strahov 2001, Stepanov 1998). Under other assumptions, it was suggested that Russian might pattern with type (iii) where one wh-phrase adjoins to CP, and others to IP, not unlike English, with the difference that in English one wh-phrase fronts and the other remains *in-situ*, and in Russian one wh-phrase fronts to CP and another moves to IP or FocP positions.

Below, I present wh-constructions in Russian: first questions with a single wh-phrase, then questions with multiple wh-phrase through which I revisit the place of Russian in the wh-typology, and explore general theoretical consequences (i.e. Superiority Condition) that unfold from the discussion. I provide data that re-examines the role Superiority plays in Russian. The discussion of Superiority allows me to argue for two essential characteristics of Russian wh-behavior: (A) Russian is of the Bulgarian type of Slavic multiple-WH movement languages (as in (i) above). (B) Russian has an extra layer of main clause structure, a functional category I call High Operator Phrase (HOP). I propose that there is in fact not only wh-fronting but wh-movement in Russian: an obligatory fronting of the wh-phrases to the CP-domain, into SpecCP. A high structural functional category HOP above CP explains the puzzle of Russian wh in a unique manner, and exposes why many facts of Russian wh were misinterpreted before (for instance, the location of wh-phrases in a non-initial position). Both (A) and (B) are novel for the analysis of Russian syntactic behavior. (A) and (B) should become easier to make a point for if

two problems that interfered with adequate analyses of Russian wh-behavior are discussed: obedience to Superiority in Russian (and consequently, what it means theoretically since several accounts exist), and the existence of a good body of data: there has always been considerable controversy about the grammaticality judgments of certain utterances.

To address the issue of the problematic data, I provide an analysis of wh-behavior in Russian using data collected in linguistic surveys from monolingual Russian speakers. An important example of how Russian wh data have been misinterpreted is found in the works of Stepanov 1998 and Strahov 2001, who both suggest that there is no wh-movement in Russian in the traditional sense. And while authors agree on this conclusion, due to data variability they provide different reasoning for their claims. Thus, Stepanov classifies Russian together with Serbian/Croatian (type ii) with movement into a Focus Phrase position, and Strahov compares Russian to Japanese and Chinese, languages without wh-fronting (*in-situ*) (type iii). Here, however, I will argue that the proper classification for Russian is type (i), like Bulgarian.

This chapter is organized as follows. First, I briefly present the situation with wh-behavior in Slavic languages (represented by Serbian/Croatian and Bulgarian), fine graining the typology introduced in (i-iii), following Rudin's (1988) criteria. Then, I introduce the methodology for collecting data used in this work. I then move to the discussion of single wh-phrase constructions. I show how the data do not support an analysis of Russian as a single-wh fronting or wh-*in-situ*. I discuss the shortcomings of such accounts of Russian, showing how they fail to capture important. Secondly, I turn to multiple wh-constructions, and show that apparent violations of Superiority in wh-phrase ordering can be captured by a simple extension to the grammar of Russian, namely HOP. Third, I consider the behavior of wh-phrases in embedded clauses where we find matrix/subordinate asymmetries with respect to wh-movement, providing additional evidence for the account suggested in this chapter. Finally, I move toward a reconsideration of the place of Russian within the existing Slavic wh-typology, and show that it should in fact be grouped with Bulgarian as a Richards' (1998) style "CP-absorption language" with obligatory overt wh-movement of all wh-phrases (to check a +WH and not a Focus feature), with all wh-phrases belonging in CP. I also revisit aspects of Superiority. I conclude by running several wh-movement diagnostics briefly verifying wh-constituency and multiple wh-extraction possibilities. Both diagnostics support the ideas put forward in the first part of the chapter.

2.1 WH: The Basics

The basic questions that have always existed in the literature on Russian wh are: what triggers wh-fronting, and which position does this wh-fronting target (CP, IP, FocP)? (Stepanov 1998, Strahov 2001, Scott 2003, Liakin 2005, Zavitnevich 2005, Dyakonova 2009, Bailyn 2011). Since Rudin's seminal 1988 work, it has been standard to assume that Slavic languages fall into two types with respect to handling wh-constructions: multiple wh-fronting languages with multiply-filled specifiers ([+MFS]) (where all wh-phrases appear in SpecCP), such as Bulgarian; and languages where only one wh-phrase appears in SpecCP, and the rest are adjoined to IP [-MFS] – such as Serbian/Croatian and Polish (Bošković 1996, 1997; Lambova 1999; Citko 1997). Some state that Russian clusters with Serbian/Croatian (Bošković 1997, 1998; Richards 1997), others assert that Russian is a true wh-*in-situ* language, where fronting happens for other reasons, e.g. focus movement (Stepanov 1998) or as a result of focalization / topicalization (Strahov 2001).

Consider the Russian wh questions ((1)a, 0a) where the wh-phrases are fronted, and ((1)b, 0b, 0c) without any wh-movement, which is unambiguously ungrammatical under a neutral non-echo reading.

(1) a. Kogo Miša videl?
who Misha saw
"Who did Misha see?"

b. *Miša videl kogo?
Misha saw who
"Who did Misha see?"

(2) a. Čto kuda zaprjatali gnomy?
what where hid gnomes
"What did the gnomes hide where?"

b. Čto kuda zaprjatali gnomy?
what where hid gnomes
"What did the gnomes hide where?"

c. *Čto gnomy zaprjatali kuda?
what gnomes hid where
"What did the gnomes hide where?"

d. *Gnomy zaprjatali kuda čto?
gnomes hid where what
"What did the gnomes hide where?"

Examples of single (1) and multiple wh 0 questions demonstrate that leaving any of the wh-phrases *in-situ* results in ungrammaticality. In the sections below, I consider Russian single and multiple wh-behavior separately.

Researchers who have analyzed wh-behavior as *in-situ* in Russian have usually presented us with a limited number of basic examples: in some cases the data were misrepresented (Stepanov 1998); and when the facts were correct, the data were misinterpreted. In order to make sure that there is no data controversy, I have conducted a thorough collection of data about Russian wh-sentences. The methodology is described in section 2.2.

2.2 Method

Participants were 76 monolingual Russian speakers in survey, men and women, age range, 18 to 65 years.[6] They performed a grammaticality judgment task for targeted sentences in the form of a questionnaire. 80% of participants were administered the questionnaire face-to-face (60 speakers) whereas 16 participants filled it out in the written (pencil-and-paper) form.[7] The study was conducted in three different regions of Russia: Moscow, St.-Petersburg, and Rostov-on-Don.

Each questionnaire consisted of 90 target and 10 filler Russian language items written in the Cyrillic alphabet. The instructions to the questionnaire asked the participants to rate each sentence's grammaticality on a scale from 1 to 5, where 1 was the lowest score of "fineness" and 5 was the highest score of "fineness". A 5-point scale is very familiar to Russians from the grading system used in Russian schools, and therefore, it assured clarity in judging the sentences.

---

[6] This research was in part supported by the Fulbright-Hays Doctoral Dissertation Abroad Award in 2003-2004 in the Russian Federation, PI – Scott, Tatiana (The research acquired Stony Brook University IRB approval).

[7] *n=76:* survey for chapter 2. *n* =35 monolingual speakers for survey 2 (used in chapter 3 predominantly). Procedures are the same, number of items is comparable.

All scores were translated into the following accepted linguistic notation shown in parenthesis below.

1– the sentence is impossible (*****- in linguistic notation);

2–the sentence is nearly impossible (**?*** - in linguistic notation)

3–the sentence is strange/unclear (coincided with commentaries such as: "I can understand it, but I wouldn't say it, I can imagine someone saying this") – (**??** – in linguistic notation);

4–the sentence is nearly acceptable, almost fine, but is in contrast with 'fine/perfect' (**?** – in linguistic notation).

5–the sentence is fully acceptable and grammatical.

The statistical analysis included calculation of proportional averages per relevant syntactic grouping of targets across the group and within one-speaker judgments. The latter controlled consistency of judgments of one speaker for the same type of targets. Statistically significant differences were calculated across the group for different types of targets.

Throughout the chapter, the numbers of participants responding to the relevant sentences in the questionnaire are indicated in parentheses (e.g., *n* out of 76 speakers and the percentage it constitutes rounded to the second decimal point). For example, if 74 speakers find target example grammatical (scoring at "5"), it is indicated by the following notation: *n*(5) = 74/76; 97.4%). If it is ungrammatical and the same percentage agree with that judgment, it is indicated as follows: *n*(1)=74/76; 97.4%). Thus, a high percentage does not mean a high percentage find it grammatical – it means a high percentage support the given judgment.

In addition to the target items, I have recorded natural discourse exchanges with a digital recorder, and extracted sentences relevant to the topics of discussion in this thesis. Thus, there are actual examples that are taken from real conversations with people: I will indicate these in text as "(from recording)".

2.3 Single WH

Examples (3) - (5) below present Russian information (non-echo) questions with a single wh-phrase. Notice that while examples (a) where a wh-word is moved to the left edge of the sentence are grammatical, their (b) counterparts where a wh-word is left in-situ are bad.

(3) a. Čto studenty izučajut?[8] (fronting: *n*(5)=76/76, 100%)
what$_{ACC}$ students$_{NOM}$ study
"What do the students study?"

b. Studenty izučajut čto? (*in-situ : *n*(1)=76/76, 100%)
students$_{NOM}$ study what$_{ACC}$
"What do the students study?"

(4) a. Kogo deti razbudili? (fronting: *n*(5)=76/76, 100%)
who$_{ACC}$ children woke-up
"Who did the children wake up?"

b. *Deti razbudili kogo?
children woke up who$_{ACC}$
"Who did the children wake up?"

(5) a. Čem$_i$ Marina rasstroena t_i? (fronting: *n*(5)=75/76; 98.6%)
what$_{INSTR}$ Marina$_{NOM}$ upset
"What is upsetting Marina?"

b. *Marina rasstroena čem$_i$? (*in-situ : *n*(1)=74/76; 97.3%)
Marina upset what$_{INSTR}$
"What is upsetting Marina?"

Thus far, the data allow us to establish the following fact about Russian: wh-phrases must be fronted: in-situ wh results in ungrammaticality. When it comes to single wh-phrases Russian exhibits behavior similar to wh-behavior in English, at least on the surface.

2.3.1 Single WH: A Puzzle

Despite the clear data given above, many leading views on Russian wh claim that Russian is not a wh-movement language. One account states that Russian wh is the result of wh-fronting into a Focus position (Stepanov 1998), while others state that Russian wh is a result of a discourse-driven phenomenon that is achieved by scrambling (Strahov 2001, Dyakonova 2009). What both views have in common that they claim that wh-fronting happens for reasons other than true wh-

[8] There is a possibility (often preferred) to say (i) Čto izučajut studenty?
what$_{ACC}$ study students$_{NOM}$
For a detailed discussion of this word order, see Bailyn (1995, 2004).

movement, in other words it does not involve movement to check off a [+wh] feature. What led researchers to believe that Russian patterns with wh-*in-situ* languages? I believe that it is the fact that Russian presents us with a variety of wh-structures, one of which looks as if wh-phrases do not front, or at least do not move as high as the CP-domain (or do not move at all). Here are some examples of non-initial single-WH in Russian:

(6) a. Miša gde rabotaet?
Misha where works
"Where does Misha work?"

b. Miša kogo videl?
Misha who_{ACC} saw
"Who did Misha see?"

Because of sentences such as those given in (6), Stepanov (1998) and Strahov (2001) claim that Russian patterns with Japanese and Chinese respectively. Stepanov (1998), in his analysis of wh-fronting in Russian, suggests that wh-movement in the traditional sense does not take place in Russian. According to Stepanov (1998), wh-phrases in Russian are fronted to check their focus feature. This suggests that wh-phrases have a strong focus feature [+Foc] that must be checked off. Such an assumption is attractive because it allows an explanation of cases like (6)a-b.

Stepanov (1998) analyzes (6) as an echo-question, which has a fronted wh-phrase for the reasons of Focus. In this, he follows Bošković (1999) who states that in SC all wh-phrases must be fronted and cannot be left *in-situ* even in an echo-reading situation. Bošković argues that there is an inherent [+Focus] feature on whs that triggers the movement. Stepanov adopts this approach for Russian. However, for Russian, the claim does not work: in Russian echo-questions, wh-phrases must remain *in-situ* (which they are not in (6)), while in information questions they cannot. The two differ in intonation significantly. This is shown in (7).

(7) a. Ivan kupil **ČTO**? (echo-reading)
Ivan bought what
"Ivan bought WHAT?

b. *Ivan kupil **čto**? (non-echo-reading)
Ivan bought what
"Ivan bought WHAT?

c. *Ivan rasskazal o **kom**? (non-echo-reading)
Ivan told about whom$_{PREP}$
"Who did Ivan told a story about?"

c'. Ivan rasskazal o **KOM**? (echo-reading)
Ivan t old about whom$_{PREP}$
"Ivan told a story about WHO?"

d. *Ivan udaril **kogo**? (non-echo-reading)
Ivan hit who$_{ACC}$
"Who did Ivan hit?"

d'. Ivan udaril **KOGO**? (echo-reading)
Ivan hit who$_{ACC}$
"Ivan hit WHO?"

e. *Ivan uexal **kogda**? (non-echo-reading)
Ivan left when
"When did Ivan leave?"

e'. Ivan uexal **KOGDA**? (echo-reading)
Ivan left when
"Ivan left WHEN?"

The successful echo-questions in (7) (the basic "echo"-order is reflected in (8)) do not have the same word order as in (6), that is, they cannot be fronted at all and remain *in-situ* if the question has an echo interpretation.

(8) Kto kupil **ČTO**? (echo-reading)
who bought what

However, a discussion of the nature of echo-questions is not the main goal of the current investigation. The fact remains that in Russian echo-questions, wh-phrases do not move. However, the questions in (6) above and in (9) below are not echo-questions. More evidence that the Focus analysis is not correct is forthcoming.

Strahov (2001) presents several instances for wh-question formation in Russian: non-movement cases (which we have seen to be ungrammatical), fronting of the whs to the beginning of the clause, as seen above, and, crucially, cases such as (6), with the wh in the second position.

(9) a. Ty **kuda** sejčas ideš'? (Strahov)
you where now go
"Where are you going?"

b. Marija **kogo** vstretila na ulice? (Strahov)
Maria whom met on street
"Who did Maria meet outside?

Structures where a wh-phrase is not *in* situ, and appears either in initial position or in any non-initial position, such as the examples in (9), are analyzed as instances of scrambling: Thus Strahov (2001) analyzes Russian as an *in-situ* language with scrambling, because wh-phrases do not always move to the far left, and can appear in a non-initial position. The discussion of the shortcomings of such an approach is coming up.

First, let us consider the Russian wh-data more carefully. Questions where a wh-phrase appears in the second position are indeed very productive in colloquial Russian: in (10) - (14) I provide examples with various wh-phrases appearing in the second position: including subjects, objects, prepositional phrases, and adjuncts. All of these were collected from natural speaker conversations. I also compare them to the wh-*in-situ* constructions (as modeled in (10)).

(10) a. Dima **čto** budet zakazyvat'? (from recording)
Dima what will order
"What will Dima order?"

b. *Dima budet zakazyvat' **čto**? (*in-situ : *n*(1)=75/76=98.7%)
Dima will order what
"What will Dima order?"

Example (11) shows a wh-object in second position, and the example in (12) shows a wh-subject in the second position.

(11) a. Rodionu **kto** èto rasskazal? (from recording)
Rodion$_{DAT}$ who this told
"Who told this to Rodion?"

b. **Kto** èto rasskazal Rodionu? (fronting: *n*(5)=76/76, 100%)
who this told Rodion$_{DAT}$
"Who told this to Rodion?"

(12) a. Oni **o** **kom** pereživajut? (from recording)
they about who$_{PREP}$ worry
"Who do they worry about?"

b. *Oni pereživajut o kom? (*in-situ : *n*(1)=75/76, 98.7%)
they worry about who$_{PREP}$
"Who do they worry about?"

c. **O** **kom** oni pereživajut? (fronting: *n*(5)=75/76, 98.7%)
about who$_{PREP}$ they worry
"Who do they worry about?"

(13) a. Ty **v** **kotorom** **času** planirueš' ix priglasit'? (from recording)
you in which hour plan them invite
"When do you plan to invite them?"

b. *?Ty planirueš' ix priglasit' **v** **kotorom** **času**?
you plan them invite in which hour
"When do you plan to invite them?" (**in-situ* : *n*(2)=74/76=97.4%)

c. **V** **kotorom** **času** ty planirueš ix priglasit'?
in which hour you plan them invite
"When do you plan to invite them?" (fronting: *n*(5)=76/76, 100%)

(14) a. Ty **gde** sejčas rabotaeš'? (from recording)
you where now work
"Where do you work now?"

b. *Ty sejčas rabotaeš' **gde**? (**in-situ* : *n*(1)=75/76=98.7%)
you now work where
"Where do you work now?

c. **Gde** ty sejčas rabotaeš'? (fronting: *n*(5)=76/76, 100%)
where you now work
"Where do you work now?"

These data (summarized in Table 1) indicate three facts about Russian: wh-*in-situ* is ungrammatical (the b examples): wh-phrases can appear at the left-edge (the c examples), and a wh phrase in a second position is possible (the a examples).

TABLE 1. SINGLE WH DISTRIBUTION DATA

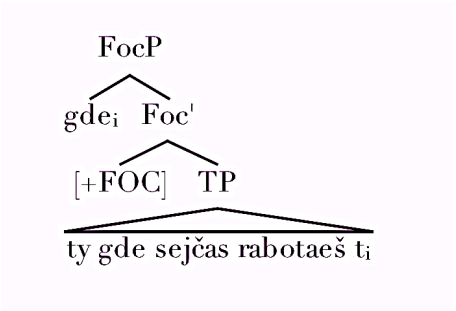
	Position of WH		
Type of Question	**in-situ**	**Left Edge**	**Second Position**
Information questions	n(1)=75: 98.7%	n(5)=76:100%	n(5)=73: 96.7%
Echo – questions	n(5)=74: 97.4%	n(1) =76: 100%	n (2)=69: 90.1%

The occurrence of the wh-word in the second position is, however, puzzling. It is these constructions that have led to the Focus and Scrambling analyses of the Russian wh.

2.3.1.1 Stepanov's Approach: Wh-fronting as Focus Movement

In Stepanov's (1998) analysis, wh-phrases occupy a SpecFoc position lower than CP as schematized (in (15)).

(15) Structure of (14) (focus-driven)



The driving force behind an account of wh-phrase fronting into the Focus position is the fact that wh-fronted multiple wh-phrases do not obey the Superiority Condition. However, Bailyn (2012) provides several reasons as to why the focus-driven account of wh-fronting is not reasonable: The most important of these is that it is not clear what it means to state that wh-phrases are "inherently focused": the answer to the question is a Focus as identified by the question – answer test eliciting the new information (Hajičová & Sgall 1975).

(16) a. 'Who wrote this song?' (= 'For which X is it true that X wrote this song?')
b. JOHN wrote this song. (= ' X= John')

Bailyn states that the evidence that JOHN here is a Focus is both "interpretive and phonological". If the answer to the question is the Focus, then under the Bošković/Stepanov's claim that the wh-phrase itself also fulfills the role of Focus we should expect a Focus feature clash. This would happen for two reasons: the location of Focus in Russian (Yokoyama 1986, King 1994, Bailyn 1995, Dyakonova 2009) and the fact that only one Focus per clause is generally allowed (Kiss 1998, for instance). It is evident that other elements than the wh-phrase can be focused in Russian wh-questions, as in (17):

(17) Gde rabotaet OLGA?
where works OLGA
"Where does OLGA (and not Daria) work?"

> … even if the claim can be maintained that [+WH] elements are inherently focused, this would not predict the required fronting found in all standard WH-movement examples, because Russian appears to not have a fixed position for Focused elements, as shown in Bailyn (1995) and elsewhere. That is, Focused elements *may* be fronted through A'-movement (in some accounts this is known as Focalization, in others simply as Scrambling). But clearly, *they do not have to be*; various intonation and other devices are available to mark the focused elements as such. In fact, as shown in

> King (1994), and argued for in a more modern framework by Neeleman et al. (2009), if there is a neutral focus position in the Russian sentence it is *on the right edge of the sentence* (Bailyn, 2012, p. 99).

This point can be illustrated by the following example:

(18) Question: Who went to Leningrad?
Answer:

a. V Leningrad poexal Ivan. (neutral intonation possible)
to Leningrad went Ivan
"It was Ivan who went to Leningrad."

b. IVAN poexal v Leningrad. (Focus intonation only)
IVAN went to Leningrad
"It was Ivan who went to Leningrad"

The wh-elements do not share the features of the focused elements: they do not allow heavy stress (the same as in 18b), and they do not prefer a sentence-final position. Thus, the fronting of the wh-phrases cannot be motivated by focus at least in Russian where Focus is generally assumed to be at the right-edge (Bailyn 1995, Dyakonova 2009). Bailyn (2012) also demonstrates that wh-phrases and Focus occupy different positions in Russian, and that there exist true wh-islands.

2.3.1.2 Strahov's Approach: Wh-fronting as Scrambling

Strahov (2000) proposes a scrambling analysis of wh-behavior in Russian.[9] However, this approach is also incorrect given that wh-movement and scrambling have different properties in Russian: while "scrambling in Russian appears to operate in a rather unconstrained manner (Zemskaja 1973, Yadroff 1991), wh-movement is heavily restricted" (Müller and Sternefeld 1993). This is shown in the contrasts below between (20) and (20), and (21) and (22) respectively.

[9] Dyakonova (2009) applies a scrambling analysis only to the matrix clause.

(19) Mne Katju$_i$ kažetsja [$_{CP}$čto [$_{IP}$otpustit' t$_i$ odnu tak pozdno]]
me$_{DAT}$ Katja$_{ACC}$ seems that to-let-go alone so late

bylo by bezumiem.
be would insanity$_{INSTR}$

"It seems to me that it would be insane to allow Katja to go alone so late at night."

(20) ?***Kogo*** tebe kažetsja [$_{CP}$čto [$_{IP}$otpustit' t$_i$ odnu tak pozdno]] bylo
who$_{ACC}$ you$_{DAT}$ seems that to-let-go alone so late be

by bezumiem
would insanity$_{INSTR}$

Specifically, Müller and Sternefeld (1993) show that long-distance scrambling is possible in the environments where wh-islands create a barrier for movement.[10]

(21) a. Ty [$_{VP}$ doktor$_i$ [$_{VP}$ videl [$_{CP}$ **kogda** [$_{IP}$ t$_i$ pod"ezžal]]]]?
you doctor saw when came

"Did you see when the doctor came?"

b. Vy [$_{VP}$ posylku videli [$_{CP}$ **kak** zapakovali t$_i$?]]
you$_{PL}$ parcel$_i$ saw how (they)-packed

"Did you see how they packed the parcel?"

(22) a. ?*Kto$_i$ [$_{VP}$ ty [$_{VP}$ videl [$_{CP}$ **kogda** [$_{IP}$ t$_i$ pod"ezžal]]]]?
who you saw when came

"Who did you seen when he came?"

b. ?*Čto vy [$_{VP}$ videli [$_{CP}$ **kak** zapakovali t$_i$?]]
what you$_{PL}$ saw how (they)-packed

"What did you see how they packed?"

[10] According to the data collected in my study, examples (23a) and (23b) received a score of 2 (?*) by 67 and 65 speakers respectively (out of 76). On he other hand, examples (22a) and (22b) are judged as acceptable at (4) by 13 speakers and (5) by 63 speakers. So the judgments are reversed. Here, then, Russian exhibits the behavior that Muller and Sternefeld assign to German, where scrambling is more restricted than wh-movement. In either case, the distribution of the wh-phrases is not the same as scrambling.

Such an asymmetry between scrambling and wh-movement allows the authors to state that wh-movement and scrambling have different properties, and different restrictions, i.e. wh-islands. Indeed, Müller and Sternefeld (1993)'s Principle of Unambiguous Binding is based on the empirical generalization that Scrambling and wh-movement have distinct properties. Clearly Strahov's analysis will not easily be able to capture such differences.

Thus far, we have seen that wh *in-situ* wh results in ungrammaticality; wh-*in-situ* analyses such as Stepanov 1998, Strahov 2001 and Dyakonova 2009 are clearly not on the right track. In fact, some fronting to some left periphery position *must* occur. Still, it is possible for wh-words to occur in the second position. This is the single-WH puzzle.

2.3.2 Single WH in Embedded Clauses

When it comes to embedded clauses, two important issues have to be discussed: the unavailability of the second position wh-phrase construction.

Dyakonova (2009) noticed that second position wh-phrases are not available in embedded clauses, which led her to analyze Russian as French, namely as a wh-in-situ language when it comes to matrix clauses and a wh-movement language when it comes to embedded clauses. This proposal was initiated by Bošković (2001) for Serbian/Croatian.

When it comes to Russian embedded clauses, a single wh-phrase in questions appears at the edge of the embedded clause, and any variation in this ordering results in ungrammaticality: a wh-phrase left in-situ , or a wh that is not at the edge of the clause (second position wh), in direct contrast to the main clause constructions we have seen. This is demonstrated in the examples below.

(23) a. Oni zainteresovalis', **kto** èto rasskazal Rodionu. (*n*(5)=76/76, 100%)
they interested$_{PERF}$ who this told Rodion
"They got interested who told this to Rodion?"

b. *Oni zainteresovalis', Rodionu **kto** èto rasskazal. (*n*(1)=69/76, 90.8)
they interested$_{PERF}$ Rodion$_{DAT}$ who this told
"They got interested who told this to Rodion?"

(24) a. Boris menja sprosil, **o** **kom** oni pereživajut.
Boris me asked about who they worry

"Boris asked me who they are worried about?"
(*n*(5)=75/76, 98.7%)

b. *Boris menja sprosil, oni **o** **kom** pereživajut.
Boris me asked they about who worry

"Boris asked me who they are worried about?"
(*n*(1)=75/76, 98.7%)

c. *Boris menja sprosil oni pereživajut **o** **kom**.
Boris me asked they worry about who_{DAT}

"Boris asked me who they are worried about?"
(*n*(1)=76/76, 100%)

(25) a. Vsem interesno, **v** **kotorom** **času** ty planirueš ix priglasit'.
all_{DAT} interested in what hour you plan them invite

"Everyone is interested what time you plan to invite them?"
(*n*(5)=76/76; 100%)

b. *Vsem interesno, ty **v** **kotorom** **času** planirueš ix priglasit'.
all interested you in what hour plan them invite

"Everyone is interested what time you plan to invite them?"
(*n*(1)=76/76, 100%)

c. *Vsem interesno, ty planirueš ix priglasit' **v** **kotorom** **času**.
all inteterested you plan them invite in what hour

"Everyone is interested what time you plan to invite them?"
(*n*(1)=75/76, 98.7%)

(26) a. Ja ne uveren, **gde** ty sejčas rabotaeš'.
I not sure where you now work
"I am not sure where you work now."

b. *Ja ne uveren, ty **gde** sejčas rabotaeš'.
I not sure you where now work
"I am not sure where you work now."

c. *Ja ne uveren, ty sejčas rabotaeš' **gde**.
I not sure you now work where
"I am not sure where you work now?"

The data in (23) – (26) are embedded questions of (10) - (14), and the speaker judgments of these constructions are summarized in the following table. Please keep in mind that the percentages in this table do not just reflect the percentages of the examples above but of all the relevant target items judgment proportions (mean value of the scores of 5 respectively).

TAB LE 2. SINGE WH ACCEPTABILITY IN MATRIX AND EMBEDDED CLAUSES

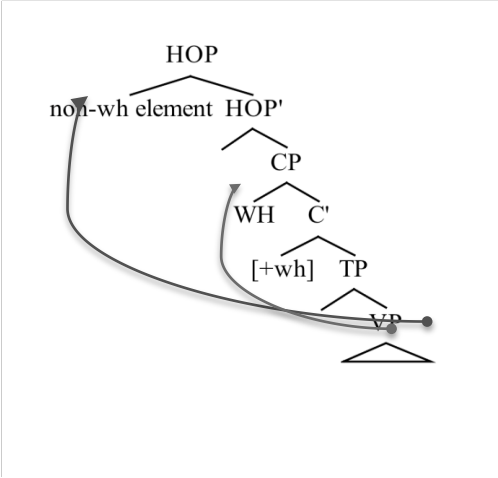
	Position of the wh-phrase		
CLAUSE	Left Edge	in-situ	Second position
Matrix	99.3% *n*(5)=75.5	0.9% *n*(5)=0.07	**96.7%** *n*(5)=74
Embedded	100% *n*(5)=76	0.4% *n*(5)=0.03	**1.3%** *n*(5)=0.9

The data in Table 2 show that embedded wh-fronting in Russian is not only preferred, it is obligatory. Furthermore, the numbers in this table provide a striking contrast between the availability of the second position wh-phrases in matrix vs. embedded clauses. Any account of Russian wh-movement must solve the puzzle identified above in main clauses, as well as account for the main/embedded asymmetries shown here. In the next section, I propose an approach that accounts for both of these phenomena in the Russian language.

2.4 Single WH: WH-fronting is WH-movement with a HOP

The fact that wh-phrases must front to the edge of the subordinate clause yields the result that Russian should be analyzed as a wh-movement language, especially taken together with the fact that the Focus movement and scrambling accounts fail to cover the basic facts of wh-behavior in Russian. The structure of a single wh in second position then looks as demonstrated in (27):

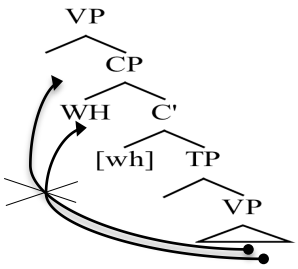
(27) Single WH structure: Main Clause



I claim that at least one wh-phrase in Russian must satisfy the wh-requirement and check off the [+wh] feature. The reason a non-wh element can appear before a wh-phrase (a wh-phrase in the second position) is the existence of the HOP phrase. The non-wh element appears in the Spec of HOP position, while the wh-word appears in SpecCP. However, the HOP position is not available in embedded clauses[11], hence, the asymmetry between the two. The structure of the embedded clause is shown below in (28):

[11] The reasons for HOP unavailability in embedded clauses are discussed in chapter 5.

(28) **Single WH: Embedded Clause**: HOP is unavailable



Now we have to test if this analysis is satisfactory.

First, the claim that wh-movement takes place would give rise to wh-movement constraints.

2.4.1 Constraints on Single WH

If wh-movement is postulated we would expect that common constraints on wh-movement that hold in other languages, hold in Russian as well.

2.4.1.1 Island Constraints

In this section, I consider Russian wh-islands. Wh-islands are ambedded wh-constructions out of which a wh-word cannot be extracted: as is well-known, wh-movement out of an embedded clause to the main clause requires an intermediate stop in the SpecCP position. Therefore, when SpecCP of the embedded clause is occupied by another wh, such movement becomes unavailable.

(29) a. How did you say [(that) everyone knows [(that) Masha fixed the car t_i]]?
b. *How did you ask [**why** everyone knows [that Masha fixed the car t_i]]?
c. *How did you say [(that) everyone can't understand [**why** Masha fixed the car t_i]]?

The wh-phrase '**why'** located in the SpecCP of the lower clause blocks the movement of '**how'** into the SpecCP of the matrix clause. In Russian it works in the following way. Both (30) and (31) are independently possible, the question in (32) where a lower CP is occupied by a wh-phrase is ungrammatical on the lower rading of *kak ('how')*.

(30) **Kak** Ivan počinil mašinu?
how Ivan fixed car
"How did Ivan fix the car?"

(31) **Komu** Ivan počinil mašinu?
who$_{DAT}$ Ivan fixed car
"Who did Ivan fix the car for?"

(32) ***Kak$_i$** on sprosil [**komu$_j$** Ivan počinil mašinu t_i t_j]?
how he asked who Ivan fixed car
"How did he ask who Ivan fixed the car for __ ?"

Analogous to English (29), the Russian questions in (32) and (33) create wh-islands that block wh-movement.

(33) a. ***Komu** ty sprosil, **kogda** Ivan pozvonil?
who$_{DAT}$ you ask when Ivan called
*"Who did you ask when Ivan called?"

b. ***Kogda** ty sprosil, **komu** Ivan pozvonil?
when you asked who$_{DAT}$ Ivan phoned
*"When did you ask who Ivan called?" (out on lower clause reading of *when*)

Thus, the existence of wh-islands gives additional syntactic evidence that wh-phrases occupy the SpecCP position rather than any lower position.

2.4.1.2 The Complex NP Constraint

The complex NP constraint, also known as Subjacency, is formulated as follows in (34):

(34) *wh$_i$ [$_{DP}$...[$_{CP}$ ...t_i ...]...]

(35) Bill made [$_{DP}$ the claim [$_{CP}$ that he read something in the syntax book]]

(36) *What$_i$ did Bill make [$_{DP}$ the claim [$_{CP}$ that he read t_i in the syntax book]]?

(37) *[Which cake]$_i$ did you see [$_{DP}$ the man [$_{CP}$ who baked ti]]?

In Russian, the following examples (38)b and (39)b violate the complex NP constraint and are ungrammatical:

(38) a. Ty znaeš' agenta, **kotory$_j$** ljubit Marinu
you know agent that loves Marina
"You know the agent that loves Marina."

b. ***Kogo** ty znaeš' agenta, **kotory$_j$** ljubit?
who$_{ACC}$ you know agent which loves
*"Who do you know an agent who loves?"

(39) a. Ja znaju studenta, **kotoryj$_j$** uvidel dvornika
I know student who saw janitor
"I know a student who saw the janitor."

b. ***Kogo** ty znaeš' studenta, **kotoryj$_i$** uvidel?
who$_{ACC}$ you saw student which saw
*"Who do you know a student who saw?"

Thus, the extraction of a wh-phrase out of a complex NP also is impossible as expected in the wh-movement environment.

2.4.1.3 Coordinate Structure Constraint

The coordinate structure constraint is formulated as follows: extraction out of a coordinate structure is prohibited unless the extraction affects all conjuncts equally; schematically it is expressed in (40):

(40) *wh$_i$ ...[$_{XP}$ [$_{XP}$ ...t$_i$ ...] *conj* [$_{XP}$ ...]]
or *wh$_i$ ...[$_{XP}$ [$_{XP}$ ...] *conj* [$_{XP}$ ...t$_i$ ...]]
or *wh$_i$ ...[$_{XP}$ t$_i$ *conj* [$_{XP}$ ...]]
or *wh$_i$ ...[$_{XP}$ [$_{XP}$ ...] *conj* t$_i$]

The examples in (41) and (44) show a coordinate structure, and (42) where the extraction happens to both conjuncts equally, while in (43) and (45) only one part of the coordinate structure is extracted and that results in ungrammatically.

(41) Sveta ljubit Borisa i Petju
Sveta loves Boris$_{ACC}$ and Petja$_{ACC}$
"Sveta loves Boris and Peter."

(42) **Kogo** ljubit Sveta?
who$_{ACC}$ loves Sveta
"Who does Sveta love?"

(43) ***Kogo** ljubit Sveta t_i i Petju
who$_{ACC}$ loves Sveta$_{NOM}$ and Petja$_{ACC}$
"*Who does Sveta love and Petya?"

(44) Oni posetili Pariž i Venu
they visited Paris and Vienna
"They visited Paris and Vienna."

(45) ***Čto** oni posetili Pariž i t_i?
what$_{ACC}$ they visited Paris and
"*What did they visit Paris and?"

Thus, violations of the traditional constraints on wh-movement, such as extraction out of coordinate structures and out of wh-islands, and violation of subjacency, give us ungrammatical results when we are dealing with a single wh-phrase in Russian. Thus, the idea that Russian wh-fronting is, in fact, movement, must be on the right track.

2.5 Summary: Single WH

In the previous section, I have shown that wh-movement takes place in Russian when it comes to single wh-phrases. There is a high operator phrase HOP that explains the asymmetry between the matrix and the embedded clauses with respect to certain S-structures for questions in Russian, i.e. second position wh. I have shown why *in-situ* analyses do not make the correct predictions, including the existence of wh-islands and the impossibility of wh-movement constraint violations, the existence of which would be impossible if any other position is considered as a landing site of a wh-phrase.

Now I turn to multiple wh-movement.

2.6 Multiple WH

The consideration of multiple wh-constructions traditionally gives syntacticians a more precise answer to the question where a language belongs in the wh-typology. Multiple wh-construction behavior works as a probe into the nature of their derivation: whether a language is an *in-situ*, a

wh-fronting, a wh-movement, or partial wh-movement language; whether the movement is overt or covert. Here, I would concentrate on the findings of Rudin (1988), Bošković (1996, 2004) and in part use Richards' (1997) taxonomy to establish where Russian belongs. I have already claimed that in the case of a single wh-phrase present in the structure, mandatory wh-movement takes place. However, this is not enough to determine where Russian fits in the typology if we consider the criteria suggested by Rudin (1988) for two kinds of Slavic languages, exemplified by Bulgarian and Serbian/Croatian.

2.6.1 Rudin's criteria

When it comes to multiple wh-phrases, Rudin (1988) identifies the differences between Bulgarian and Serbian/Croatian based on the differences in availability of ordering of the fronted wh-phrases as shown in (46) for Bulgarian (BG) and (47) for Serbian/Croatian (SC), where (b) is impossible in BG and fine in SC.

(46) a. **Koj** **kakvo** kupuva? (BG)
who what bought
"Who bought what?"

b. ***Kakvo** **koj** kupuva?[12]
what who bought
"Who bought what?"

(47) a. **Ko** **šta** vidi? (SC)
who what sees
"Who sees what?"

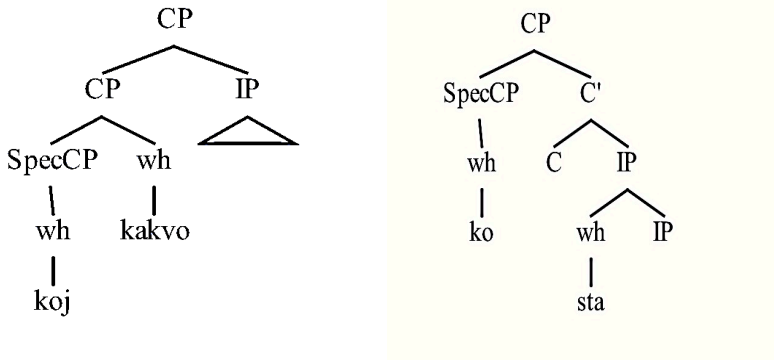
b. **Šta** **ko** vidi?
what who sees
"Who sees what?"

The suggested analysis of Bulgarian and Serbian/Croatian respectively runs as follows: Bulgarian is a multiple wh-fronting language where all wh-phrases are fronted in order to check

[12] The ungrammaticality of the Bulgarian example in (46)b is introduced in Rudin's original (1988) article. However, in Chapter 5, I revisit this issue, and provide examples from Rudin (1985) and Jaeger (2004) where Superiority violations in Bulgarian are sometimes acceptable. I discuss the contexts in which Bulgarian allows no Superiority.

off wh-features. According to Rudin (1988), they move and multiply fill SpecCP. Rudin labels these languages [+MFS].[13] The Bulgarian structure is shown in ((48)a). SC-type languages have been analyzed as languages where one wh-word undergoes wh-movement and the rest of the wh-words move top a TP-adjoined position for other purposes (Rudin 1988, Citko 1996). Rudin calls such languages [-MFS]. Their structure is presented in ((48)b).

(48) a. Bulgarian [+MFS] b. Serbian/Croatian [-MFS]



The basis for Rudin's investigations comes from the freedom of NOM/ACC word order of wh-phrases in question constructions (the Superiority Condition). She also inspects whether multiple extraction from a clause, wh-island violations, and various parentheticals after the first wh-phrase are possible. While Rudin (1988) used the tests below to describe Bulgarian as [+MFS]; Bošković's (1997) applies these criteria to draw out the differences between BG and SC.

2.6.2 WH Ordering

The first test involves NOM/ACC word order. The picture is straightforward in Bulgarian – NOM>ACC word order is preferred over ACC>NOM word order, while SC shows no difference in preferences (at least in matrix clauses). This Bulgarian restriction is commonly analyzed as Superiority.

Several accounts of Superiority exist. An early statement of Superiority comes from an observation for English of Kuno and Robinson (1972) that state that a wh-phrase cannot be

[13] Bošković (1996) suggested an account of Bulgarian where only one *wh*-phrase moves up for Q-checking, while the rest undergo focus movement.

preposed by crossing over another wh-phrase. Chomsky (1973) argued that (46) and (47) are explained by the Superiority Condition as formulated in (49):

(49) *Superiority Condition* (Chomsky 1973; Bošković 1997)

(a) No rule can involve X, Y in the structure … X … […Z…WYV…]…where the rule applies ambiguously to Z and Y, and Z is superior to Y.

(b) The category A is superior to category B if every major category dominating A dominates B as well but not conversely.

Under more recent theoretical assumptions, the Superiority Condition follows from the requirement that the [+wh] feature should be checked in the most economical way. Specifically, using Pesetsky's (1987) interpretation, the wh-phrase that is the *closest* is the one that moves first to SpecCP. Richards (1997) proposes that when the second wh-phrase undergoes movement, Shortest Move forces it to move to a lower specifier ("tucking in"); thus, crossing fewer nodes than it would on the way to the higher specifier.

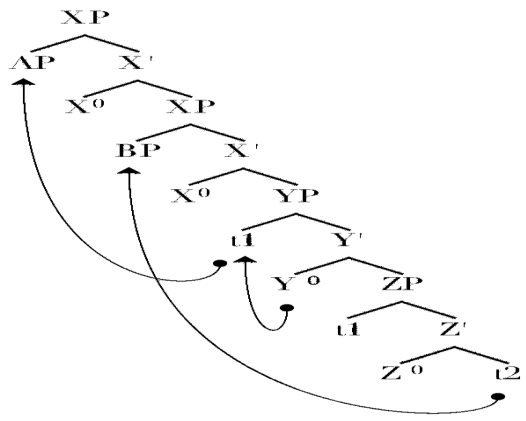
For the time being, I adopt Richards' account of Superiority that essentially consists of three components. (i) a derivational theory of the Superiority effects of English using Attract Closest (Kitahara 1994, 1997, building on Kuno and Robinson 1972, Chomsky 1973, 1993, 1995). (ii) Elimination of the strict cycle condition in favor of "featural cyclicity" (Chomsky 1995), and (iii) tucking in (required by Shortest Move). However, later, I will revisit this approach to Superiority since "tucking in" is not an ideal way of handling Superiority, and no exact mechanism exists.

Following Chomsky (1995) Richards derives cyclicity from (50):

(50) A strong feature must be checked as soon as possible after being introduced into the derivation, before any new structure is merged.

Richards demonstrates featural theory of cyclicity, together with a certain definition of Shortest Move results in the following derivation of phrase ordering in (51):

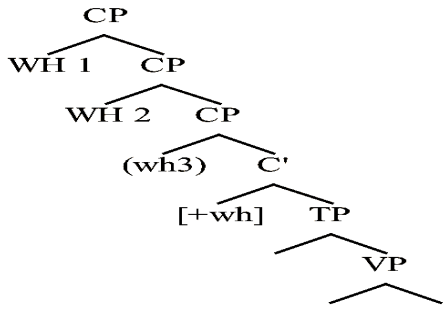
(51) Superiority derivation (Richards 1997)



Richards (1997) divides languages into CP-absorption and IP-absorption languages. For him, such a derivation (51) explains Superiority in CP-absorption languages such as Bulgarian, and lack thereof in IP-absorption languages such as Serbian/Croatian.

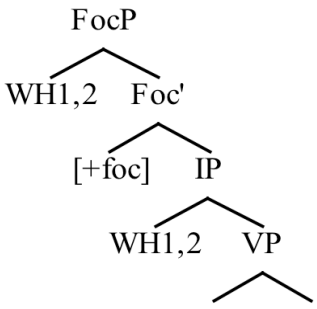
Bulgarian wh-phrases obey Superiority, while SC wh-phrases do not, and the structures in (48) reflect this, according to Rudin (1988). Thus, for Bulgarian-type languages Superiority works in the following way: the highest wh-phrase is attracted to SpecCP by [wh] on C^0; the lower wh-phrase is also attracted to C^0 and either right adjoins to the first wh-phrase (as in Rudin 1988) or the lower wh-phrase is also attracted to C^0 and "tucks in" to a lower SpecCP (as in Richards 1997). In both accounts underlying word order is preserved. This structure of BG with Richards' modification is presented in (52).

(52) BG multiple Specifiers of CP (Lambova 2004, Bošković 1996, Richards 1997)



Bošković's (1996-1998) analyzed the lack of Superiority in SC not in terms of CP- vs. IP-absorption, but rather claimed that wh-movement in SC phrases undergo focus movement since all whs are inherently focused. For Bošković, the order of the focused wh elements is not constrained by Superiority.

(53) Focus Approach to WH in SC main clause (Bošković 1996-1998)



When we examine Russian multiple wh-constructions, we see that in fact they *appear* to pattern with Serbian/Croatian: that is, they appear not to obey Superiority (both NOM > ACC and ACC > NOM orders are fine. This is shown in examples (54) - (58) below.

(54) a. **Kto** **čto** podaril Ivanu? (NOM>ACC: *n*(5)=76/76, 100%)
who$_{NOM}$ what gave Ivan$_{DAT}$
"Who gave what to Ivan?"

b. **Čto** **kto** podaril Ivanu? (ACC>NOM: *n*(5)=70/76, 92.0%)
what who$_{NOM}$ gave Ivan$_{DAT}$
"Who gave what to Ivan?"

(55) a. **Kto** **kogo** ugovoril ostat'sja? (NOM>ACC: *n*(5)=76/76, 100%)
who$_{NOM}$ who(m)$_{ACC}$ convinced to stay
"Who convinced whom to stay?"

b. **Kogo** **kto** ugovoril ostat'sja? (ACC>NOM: *n*(5)=74/76, 97.4%)
who(m)$_{ACC}$ who$_{NOM}$ convinced to stay
"Who convinced whom to stay?"

(56) a. **Kto** **čto** posovetoval Darii? (NOM>ACC: *n*(5)=76/76, 100%)
who$_{NOM}$ what$_{ACC}$ advised Daria$_{DAT}$
"Who advised what to Daria?"

b. **Čto** **kto** posovetoval Darii? (ACC>NOM: *n*(5)=75/76, 98.7%)
what$_{ACC}$ who$_{NOM}$ advised Daria
"Who advised what to Daria?"

(57) a. **Kto** **komu** predstavil Petra? (NOM>DAT: *n*(5)=76/76, 100%)
who$_{NOM}$ who$_{DAT}$ introduced Petr$_{ACC}$
"Who introduced Peter to who(m)?"

b. **Komu** **kto** predstavil Petra? (ACC>DAT: *n*(5)=74/76, 97.4%)
who$_{DAT}$ who$_{NOM}$ introduced Petr$_{ACC}$
"Who introduced Peter to who(m)"

(58) a. **Kto** **v** **čem** zainteresovan? (NOM>PREP: *n*(5)=76/76, 100%)
who$_{NOM}$ in what interested
"Who is interested in what?"

b. **V** **čem** **kto** zainteresovan? (PREP>NOM: *n*(5)=72/76, 94.7%)
in what who$_{NOM}$ interested
"Who is interested in what?"

The examples above show that with a very slight preference (there is no significant statistical difference) either order of the wh-phrases is possible. Hence, Superiority does not appear to hold. For Stepanov (1998), Superiority is also the feature that is a diagnostic of the strength of the Q-feature, if Superiority is present then the strong Q-feature is present. Since Superiority appears not to obtain even when there is a Q-feature present in the structure, this Q-feature is weak. In this, Stepanov claims Russian is not unique. In Chinese, for instance, the Q-feature is also weak, and, therefore, Russian patterns with the *in-situ* languages, where overt wh-movement does not exist.

In Russian, however, Superiority "emerges" in certain contexts[14]. This is because all wh-phrases end up in multiple Specifier positions of CP. Following Richards (1997), this should put

[14] Richards (1997) gives a list of properties that differentiate a CP-absorption language from an IP-absorption language. In this sense Russian does not always behave as expected (not similar to Bulgarian); for instance with

Russian into the class of CP-absorption languages – languages where all wh-phrases move to CP overtly as in Bulgarian and covertly as in Chinese. One of the properties of a CP-absorption language is rigid ordering of wh-phrases with respect to each other.

Then, how does the apparent "absence" of Superiority interact with the possibility of a pre-wh phrase in the HOP position? In the previous section, I demonstrated that single wh-phrases can appear in the second position: let us see if this is possible with multiple wh-phrases, and what happens to Superiority in such cases.

2.7 Multiple WH: The "Emergence of Superiority" Program

If the assumption that Russian does not move wh-phrases into a [+wh] position is true, and the process of wh-fronting does not follow the pattern of Bulgarian, we should expect a lack of Superiority in multiple wh-questions. However, there are several contexts in which Superiority does emerge in Russian; both involve HOP – one is when there is a non-wh high overt topic preceding the wh-phrases and the other is with multiple wh-phrases in embedded clauses.

2.7.1 Emergence of Superiority I: High Overt Topic

It is possible to place multiple wh-phrases after a non-wh-topic, this results in multiple wh-questions where wh-phrases do not appear at the left edge of the clause. Thus, questions (56) and (57) above, in which wh-ordering is free, can be asked so that a non-wh word appears on the left edge of the matrix clause; however as the examples below (59),(60) indicate this impacts the possible ordering of the wh-elements:

(59) a. Darii **kto** **čto** posovetoval? (NOM>ACC: *n*(5)=75/76, 98.7%)
Daria$_{DAT}$ who$_{NOM}$ what$_{ACC}$ advised
"Who advised what to Daria?"

b. *Darii **čto** **kto** posovetoval? (*ACC>NOM: *n*(5)=72/76, 94.7%)
Daria$_{DAT}$ what$_{ACC}$ who$_{NOM}$ advised
"Who advised what to Daria?"

---

respect to wh-island violations. On the other hand, Bošković (1997) states that there are environments in which SC exhibits Superiority (i.e. in embedded clauses). This necessarily should lead to a prediction that SC should be insensitive to island conditions (since presumably multiple SpecCPs are available as an escape hatch out of wh-islands). The facts appear to contradict this. Thus the correlation between island-sensitivity and CP-absorption simply may not hold.

(60) a. Petra **kto** **komu** predstavil? (NOM>DAT: *n*(5)=74/76, 97.4%)
Petr$_{ACC}$ who$_{NOM}$ who$_{DAT}$ introduced

b. *Petra **komu** **kto** predstavil? (*ACC>DAT: *n*(5)=74/76, 97.4%)
Petr$_{ACC}$ who$_{DAT}$ who$_{NOM}$ introduced
"Who introduced Peter to who(m)"

Superiority "emerges" in ((59)b) and (60)b) when a non-wh-word occupies a high position at the left edge of the clause, which I claim to be SpecHOP. Why doesn't Superiority hold in multiple wh-constructions when nothing precedes them, and why must it hold in these contexts? When SpecHOP is filled with a non-wh, there is nowhere for a wh-phrases to move to escape the Superiority effects related the CP domain, and the competition over wh features. However HOP allows one WH to move higher; all are equidistant from HOP.

Are there any other environments with multiple wh-phrases where Superiority emerges?

2.7.2 Emergence of Superiority II: Embedded clauses

In embedded clauses, where two (or more) wh-phrases are present, they must appear at the left edge of the embedded clause and Superiority cannot be violated.

(61) a. Maria sprosila, **kto** **čto$_i$** posovetoval t_i Darii?
Maria asked who$_{NOM}$ what$_{ACC}$ advised Daria$_{DAT}$

"Maria asked who advised what to Daria?"
(NOM>ACC: *n*(5)=76/76; 100%)

b. *Maria sprosila, **čto$_i$** **kto** posovetoval t_i Darii
Maria asked what$_{ACC}$ who$_{NOM}$ advised Daria$_{DAT}$

"Maria asked who advised what to Daria?"
(*ACC>NOM: *n*(1)=74/76; 97.4%)

c. Maria sprosila, **kto** **čto** peredal Maximu?
Maria asked who$_{NOM}$ what$_{ACC}$ gave-for Maxim$_{DAT}$

"Maria asked what did Maxim find out from who(m)?"
(NOM>ACC; *n*(5)=75/76, 98.7%)

d. *Maria sprosila, **čto** **kto** peredal Maximu?
Maria asked what$_{ACC}$ who$_{NOM}$ gave-for Maxim$_{DAT}$?

"Maria asked what did Maxim find out from who(m)?"
(*ACC>NOM: *n*(1)=69/76, 90.8%)

(62) a. Ego zainteresovalo, **kto** **komu** predstavil Petra?
he$_{ACC}$ interested who$_{ACC}$ who$_{DAT}$ introduced Petr$_{NOM}$

"He got interested who introduced Petr to who(m)?"
(NOM>DAT: *n*(5)=76/76, 100%)

b. *Ego zainteresovalo, **komu** **kto** predstavil Petra?
he$_{ACC}$ interested who$_{DAT}$ who$_{ACC}$ introduced Petr$_{NOM}$

"He got interested who introduced Petr to who(m)?"
(*DAT>NOM: *n*(1)=69/76, 90.8%)

Thus, we can already see that Superiority emerges in two contexts in multiple wh-constructions: (i) when a non-wh-element is present in left edge position, and (ii) in embedded clauses.

Let us summarize what we have observed thus far. Russian wh-elements must front, and all wh-phrases must front to SpecCP to satisfy [+wh] requirement. I have shown that *in-situ* wh-phrases are not possible in Russian in any contexts except for echo-questions (but their derivation is not discussed here). When we consider Superiority we come across a dichotomy between matrix and embedded clauses: in matrix clauses the Superiority Condition can be

violated, and in this Russian patterns with Serbian/Croatian; however, in embedded clauses, Superiority cannot be violated, and when it comes to embedded clauses, Russian behaves like Bulgarian. Additionally, Russian exhibits another word order in questions, when a non-wh element precedes one or more wh-phrases. When we come across this word order the following observations are in place: with multiple wh-phrases, the Superiority Condition cannot be violated if a non-wh element precedes two wh-phrases. These cases of the emergence of Superiority are expected under a CP-fronting account of Russian multiple wh, given HOP. The unexpected apparent lack of Superiority is simply the effect of either of the fronted whs in SpecCP raising to the single SpecHOP position. This is summarized in Table 3 below.

TABLE 3. MULTIPLE WH OPTIONS

CLAUSE	**Superiority**	**wh-*in-situ***
Matrix	NO	Echo only
WH in 2nd position	YES	
Embedded	YES	Echo only

These differences in Superiority requirements are explained below. However, before I move to the analysis, we have to outline another idiosyncratic fact of Russian multiple wh-movement: fronting of multiple wh-adjunct phrases.

2.8 Multiple WH-adjunct fronting

When two adjuncts are present, multiple fronting of them *in either order* is impossible in Russian.[15] Consider the following Russian examples:

(63) ***Kogda gde** oni vstretjatsja?
when where they meet$_{PL.FUT}$
"When are they going to meet where?"

(64) ***Gde where** oni vstretjatsja?
when where they meet$_{PL.FUT}$
"Where are they going to meet when?"

[15] Interestingly, Citko and Gračanin-Yuksek (2010) claim that multiple fronting of adjuncts is impossible in SC as well.

(65) **Kogda** i **gde** oni vstretjatsja?
where and when they $meet_{PL.FUT}$
"Where are they going to meet and when?"

Evidently, the problem cannot be directly related to violation of Superiority: clearly multiple wh-adjuncts do not exhibit ordering restriction with respect to each other. The problem lies in the way Superiority is derived (51) and the fact that movement paths necessarily cross in such a derivation, which does not create a problem for arguments, but does create a problem for adjuncts.

Rizzi (1990) proposed a distinction in antecedent government of complements and non-complements as a component of the Empty Category Principle (ECP). The notion of Proper Government is no longer possible under current assumptions. However, the fact remains that whatever title and mechanism the modern theory applies, adjunct traces are more restricted than argument traces. The chains that are created as a result of movement are crossing paths. This calls for a formulation of a Multiple WH-Adjunct Ban (66):

(66) **The Multiple WH-Adjunct Ban**
Multiple fronting of two (or more) wh-adjuncts is restricted due to the lack of proper government of traces.

In particular, adjunct traces must be antecedent governed, rendering them subject to Relativized Minimality. In any derivation of a multiple adjunct-wh construction, antecedent Government of the higher adjunct trace is blocked by the intervening lower wh. Thus both of the questions in (63) and (64) crash because they violate the Multiple WH-Adjunct Ban. The lower moved wh-phrase serves as blocker for antecedent government of the trace of the higher wh.[16]

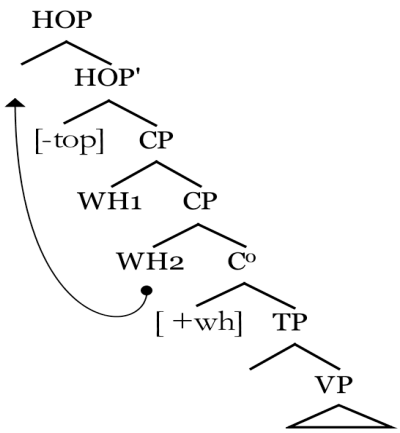
2.9 HOP and Multiple WH

I propose that the paradoxical wh-behavior of Russian wh can be easily accounted for by a single parameter. I have introduced a higher operator phrase HOP in the section on single wh-constructions, and now I show how it works for the multiple wh-constructions. I argue that all wh-phrases in Russian move into SpecCP, behaving as Rudin's [+MFS] languages such as Bulgarian. Note, I claim that each wh-phrase is located in a distinct SpecCP, following Bošković's (2003, 2004). If no further movement takes place or possible, Superiority emerges,

[16] The grammaticality and the derivation of (65) will be discussed in detail in Chapter 3 of this work.

as in Bulgarian. However, the HOP position precedes CP. Either of the two wh-phrases can move to this position (they are equidistant from HOP, each being located in SpecCP). However, if a non-wh XP fills, this position, the WHs are frozen in their Superiority-obeying positions. This is demonstrated in the structure below (67).

(67) Russian Multiple WH Structure ($WH_2 > WH_1$)



The structure in (67) shows that a HOP[17] position in the matrix clause can house either wh-phrase. Therefore, the claim here is that Superiority is present in Russian when HOP is not present (as in embedded clauses), or is not overtly occupied (as in case with a non-wh-element present at the left–edge of the clause). The analysis of Russian wh given in (67) with the High Operator Phrase (HOP) allows me to account for Russian wh-behavior when a non-wh-element is overtly present, the lack of Superiority between the wh-phrases when there is no such phrase, and the unavailability of anything to intervene between the wh-phrases. When an overt non-wh-element is present, Russian reduces to Bulgarian, and hence, should exhibit the same wh-behavior. However, when no overt non-wh-element is present, wh-phrases have a position to move into. This explains how intervening material (such as particles) can appear between wh-phrases in multiple wh-constructions in Russian without breaking their constituency. So when on

[17] While I do believe that this position has Topical properties and can host Topic elements as it will be shown in the forthcoming chapter, I do not want to claim that it is in fact a TopP position. So for the time being I double label it as an operator phrase that can host Top elements (such as unique topics). In fact, that what preceding the wh-phrases non-wh-elements appear to be. In the chapter on particles (chapter 4), I give more evidence for how this position can be utilized for Topics.

overt element occupies the Spec, HOP position, Russian looks like, but does not entirely behave like SC.

In fact, the idea that there exists a pre-CP topic is not novel and has been proposed for Bulgarian by Rudin (1986, 1988). Rudin (1988: 451, ex. 7a) gives the following example in passing (repeated here as (68):

(68) Boris **na** **kogo** **kakvo** kaza [če šte dade]? (Bulgarian)
Boris to whom what said that will $give_{3SING}$
"What did Boris say that (he) would give to whom?"

She states that *Boris* in this question is a Topic, and in the footnote she clarifies that this Topic position is pre-SpecCP, presumably adjoined to CP. This statement about Bulgarian provides additional evidence that Russian is in fact a Bulgarian-type language.[18]

It is not surprising then that several researchers had noticed discourse properties of Russian wh-constructions before (hence, some relevance to topicalization) (Strahov 2000; Scott 2001, Zavitnevich 2005, Dyakonova 2009). Dyakonova, for instance, entertains the idea of D-linking the wh-phrase in the spirit of Pesetsky (2000) but finds this idea not fruitful because she ends up with a matrix/embedded asymmetry that she does not have an elegant way to approach.

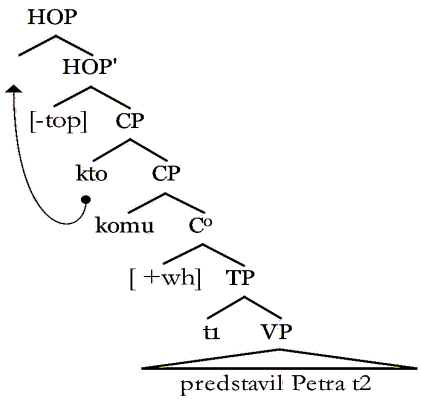
Let me return to the examples of the previous sections and provide their respective derivations:

(69) a. **Kto** **komu** predstavil Petra?
who_{NOM} who_{DAT} introduced $Petr_{ACC}$
wh1 wh2

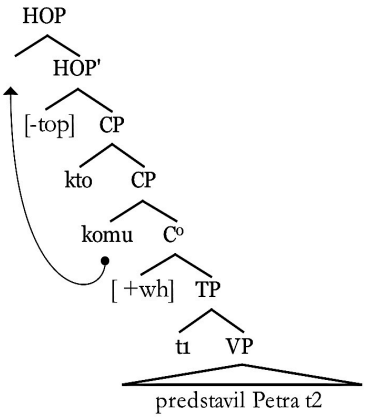
b. **Komu** **kto** predstavil Petra?
who_{DAT} who_{NOM} predstavil $Petr_{ACC}$
wh2 wh1

[18] However the fact that Bulgarian does not have surface violations of superiority in multiple wh-structures indicates merely that this pre-CP position cannot be used for wh-phrases in Bulgarian, though it can in Russian.

(70) Structure of ((69)a)



(71) Structure of ((69)b)

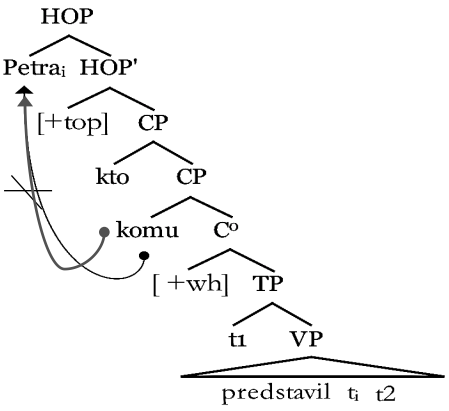


The structures in (70) and (71) show how the wh-phrases can move into the SpecHOP, and hence, get out of Superiority in the main clause. However, when Superiority emerges the path is closed due to the presence of an overt non-wh Topic.

(72) a. Petra **kto** **komu** predstavil?
Petr$_{ACC}$ who$_{NOM}$ who$_{DAT}$ introduced
wh1 wh2

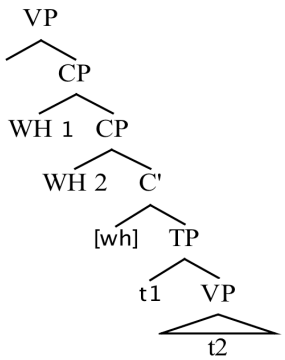
b. *Petra **komu** **kto** predstavil?
Petr$_{ACC}$ who$_{DAT}$ who$_{NOM}$ predstavil
wh2 wh1

(73) Structure of (72) – the Emergence of Superiority (Overt Topic wh in the second position):



Moreover, the fact that HOP is not available in the embedded clause elucidates the embedded facts and shows obedience to Superiority.

(74) Structure of the embedded clause with multiple whs



Russian wh-embedded clauses always look and behave like a Bulgarian- type language due to the lack of HOP position. However, the matrix/embedded, or in other words "root/embedded" asymmetry has been given alternative analyses.

2.9.1 Note on Matrix/Embedded Asymmetry: DenDikken - Giannakidou (2002)

DenDikken and Giannakidou (2002) propose an alternative account of the root/embedded asymmetry with respect to wh-constructions. Their evidence is drawn from the data of Polarity Items (PIs), and aggressively non-D-linked wh-the-hell phrases that are available preceding whs in the root clause, but not in the embedded. Their analysis claims that in the root clause wh-phrases move into SpecFoc position, and not SpecCP position in English. For them, then the structures look as shown in (75) for the matrix clause, and (85b) for the embedded clause:

(75) a. $[_{CP}$ C $[_{TopP}$[a book like this]$_i$ $[_{Top}$ $[_{FocP}$[why]$_j$ [should$_k$ $[_{IP}$ I t_k buy t_i t_j]]]]]]
b. …$[_{CP}$[why]$_j$ [C $[_{TopP}$[a book like this]$_i$ $[_{Top}$ $[_{IP}$ he should buy t_i t_j]]]]]

However, the reason why it is challenging to claim that Russian wh-phrases move into a Focus position has been discussed in detail in the previous sections of this chapter. While their analysis might be attractive for English and Hungarian, it is not warranted for Russian due to the reasons outlined above: the position of the Focus feature on the right rather than on the left of the clause in Russian, and hence, the existence of two foci that occur within the same clause.

The discussion on the interaction of the high position and Superiority returns us to the discussion of why Superiority emerges in some contexts and not others, as well as why some languages obey Superiority while others do not. In the following section I revisit the Superiority condition.

2.10 Superiority Revisited

When two patterns for Superiority exist within one language, it calls for separate analyses for each instance of grammar, which, of course, violates the learnability principle. I claim, however, that Russian can be analyzed uniformly, and not as a contradictory paradigm. In order to confirm the findings of this chapter, let us revisit the ideas behind the Superiority Condition, and discuss how Superiority is derived.

Usually, the lack of Superiority in languages that front multiple wh-phrases is analyzed through movement to IP/FocP/OP (as has been claimed for Serbian/Croatian and Polish.) Bošković (1997), for example, claims that Superiority does not apply because the inherent [foc] movement requirement of [wh] forces movement in each instance, so there is no competition because each wh-element is driven separately. Stepanov and Stateva (2009) state that

Superiority "is simply irrelevant… Each wh-phrase in a multiple wh-question moves for an independent reason". But why should this not also be true for Bulgarian as well where multiple displacement of wh-phrases is forced? Why doesn't movement to IP/FocP/OP obey Superiority? In other words, why is CP absorption subject to Superiority and IP absorption is not? Superiority should hold for *all* multiple wh-fronting languages including Russian and Serbian/Croatian. (As, in fact, we have seen that it does).

For Bulgarian (and now Russian), a theoretical issue remains. Rudin (1988), Richards (1997), Bošković (1997, 2002), Stepanov and Stateva (2009) and others all assume that the wh-phrase that is first in the linear order is the one that moves first in Bulgarian. The second phrase either right adjoins to the first wh-phrase located in SpecCP, as in Rudin, or is *tucked in* to a lower SpecCP, a movement which violates the Extension Condition (Chomsky 1995) and has an *ad hoc* character. This is not a problem for Superiority in languages like English, but in cases of multiple movements of similar elements, where Superiority means that base order is preserved, it is logical to propose an alternative way of deriving the order found that does not violate the Extension Condition. For clarity, let us call the kind of Superiority under discussion *Candidate Superiority*, a descriptive term meant to cover ordering restrictions on multiple required movements of similar elements, such as Bulgarian and Russian wh. What follows, then, does not apply to English Superiority, where only one element is attracted to CP, and the closest is preferred (the usual Minimalist account).

In the spirit of Bošković (2007), I assume that in multiple wh movement, the probing happens upwards, from the wh elements itself (endowed with a strong wh feature). *Candidate Superiority* would then result from the following considerations: derivations are strictly cyclic. The lowest wh element moves first, and targets the edge of the CP domain where the matching feature is hosted. Higher wh elements moves subsequently, targeting an extended position in the same domain, and following the Extension Condition. *Candidate Superiority* thus eliminates tucking in. I adopt multiple specs of CP for the Russian/Bulgarian type languages and multiple specs of IP or FocP for Serbian/Croatian, under the standard Bošković (1997)/ Richards (1998) type analysis. Note that after both elements have landed in the local spec positions, they are now equidistant from any higher probe (such as HOP), an therefore we expect any further single movement to be equally able to target either the higher or lower wh. Hence the apparent lack of

superiority when HOP sits above CP (and, arguably, when a single element rises further in Serbian/Croatian after multiple focus fronting)

While the exact mechanism of upwards probing, and the timing of operations in such a derivation requires further investigation and lies outside the scope of this thesis[19], this approach offers two theoretically attractive results: first, "tucking in" is eliminated; and second, all languages in which multiple elements target a single domain for the same featural reason should show Superiority effects. Thus, Superiority becomes non-parameterized. The latter fits both principles of economy and learnability. The facts also support this idea: The Emergence of Superiority Program presented above, and the emergence of Superiority in similar environments in Serbian/Croatian. Any lack of Superiority effects, then, must be the result of other factors: for instance, the availability of HOP in Russian matrix clauses, into which either of the fronted wh elements can move, masking the effects of Superiority on the CP level.

Now, if the HOP analysis of Russian wh-movement is correct, we should assume that Russian would behave in an analogous manner to Bulgarian with respect to Rudin's (1988) diagnostics concerning parentheticals and wh extractions, so I turn to those now.

2.11 WH-Movement Diagnostics

2.11.1 WH-constituency: Insertion of parenthetical material

In order to show the constituency of wh-phrases in the Spec of CP, or lack thereof, Rudin (1988) shows that parentheticals are not able to intervene between fronted wh-phrases in BG, but can do so in SC.[20]

[19] The primary issue to be resolved involves look-head. When the lowest wh is merged, its eventual target location (CP) is not yet part of the structure. The other wh is already merged well before C. Therefore the computation must keep a record of the order of merge in order to be sure that the lowest element moves first. Alternatively, the first wh could literally move to each local domain edge, followed by the higher one, maintaining the base ordering at each intermediate level until SpecCP is reached. Another alternative could involve Sidewards Movement (Nuñes 2002), in which case the derivation becomes similar to what I propose for coordinated multiple wh constrictions below. However, that approach then leads to potential Coordinate Structure Constraint problems when activating the subsequent move to HOP. However, what matters for our purposes is the notion that (a) the lowest wh moves first and the higher one subsequently, thus allowing the extension Condition to derive base order at the top of the tree, and (b) tucking in can be eliminated from the grammar, an obviously welcome result.

[20] There is a more recent discussion of parentheticals by Lambova (2002) in Bulgarian that in part contradicts Rudin's. Her work to some extent corresponds to the ideas of the current investigation.

(76) **Ko**, po tebi, **šta** kupuje? (SC)
who according-to you what buys

(77) ?***Koj**, spored tebe, **kakvo** kupuva? (Bulgarian)
who according-to you what buys
"Who, according to you, is buying what?"

Rudin argues that the impenetrability of fronted wh-phrases in Bulgarian indicates that they form a constituent. Bošković (2003) notes, however, that (77)

> could also be accounted for under the multiple-specifier analysis of Bulgarian MWF, proposed in Koizumi (1994) and further developed in Richards (1997) and Pesetsky (2000). On this analysis, fronted Wh-phrases in Bulgarian are all located in SpecCP, as in Rudin's analysis. However, they do not form a constituent in that position, contra Rudin (1988). Rather, each wh-phrase is located in a distinct Spec.

Thus Bošković rules out (63) on the basis of a feature clash: a [-wh] element is located in an interrogative [+wh] projection. Either way, in a BG-type language no intervening material can come between the wh-phrases, whereas in a SC-type language it can. Russian should behave the same way, at least in Emergence of Superiority contexts.

To consider this criterion in Russian I have examined the particle ***by*** (a conditional/subjunctive particle). This particle behaves as a second position clitic, attaching to elements preceding it, and can act as a parenthetical in the sense used in Rudin. The example in (78) illustrates the usage of the particle.

(78) a. Ja by posmotrel ètot film
I BY watched this film

b. Ètot film by ja posmotrel
this film BY I watched
"I would watch this film."

c. Maximu by ja podarila takoj podarok
Maxim BY I gave such gift
"I would give Maxim such a gift."

In (78) ***by*** follows clause edge overt topics.

In the examples below I show the distribution of ***by*** with multiple wh-phrases in Russian questions. In (79) the particle appears after both of the wh-words, and in (80) it intervenes between the wh-elements.

(79) **Kto** **kogo** by priglasil na tanec? (WH1 WH2 by: *n*(5)=76/76, 100%)
who who BY invited to dance
"Who would ivite who(m) to dance?"

(80) **Kto** by kogo priglasil na tanec? (WH1 ***by*** WH2: *n*(4,5)=70/76, 92%)
who BY who invited to dance
"Who would invite who to dance?"

The intervening particle is possible also when the order of wh-elements is reversed (ACC> ***by*** >NOM) as shown in (81):

(81) **Kogo** by **kto** priglasil na tanec? (WH2 by WH1: *n*(5)=73/76, 96.1%)
who BY who invited to dance
"Who would invite whom to dance?"

However, some curious facts arise in the following examples. Compare (82) and (83) vs. (84) where two wh-words are separated by are preceded with a non-wh-element in HOP:

(82) Maximu by **kto** **čto** podaril?
$\text{Maxim}_{\text{DAT}}$ by who what present
"Who would give Maxim what as a present?"
(Non-WH by WH WH: *n*(5)=75/76,98.7%)

(83) Maximu **kto** **čto** by podaril?
$\text{Maxim}_{\text{DAT}}$ who what BY present
(Non-WH WH WH by: *n*(4;5)=75/76,98.7%)

(84) *Maximu **kto** by **čto** podaril?
$\text{Maxim}_{\text{DAT}}$ who BY what present
"Who would give Maxim what as a present?
(*Non-WH WH by WH: *n*(1)=74/76,96.1%)

Here we see that the insertion of the particle ***by*** is possible between two wh-phrases, but when a non-wh precedes them, then the intervening material has to appear after both of them or before

both of them following a non-wh-element. This follows from the HOP analysis on the assumption that *by* can appear at the edge of a domain, but not between two specifiers, in the manner described by Rudin and Bošković above.

The HOP structure, therefore, allows me to account for the occurrence of intervening material between wh-phrases in some contexts (no overt topic), and not in others (presence of the overt non-wh-element.) In the former case, one wh element has moved to HOP, something that is impossible when HOP is filled by a non-wh topic.

2.11.2 WH-extractions

Rudin (1988) and Richards (1997) discuss extraction facts for multiple wh-phrases in Bulgarian and Serbian/Croatian. Rudin examines the possibility for movement of multiple wh-phrases out of an embedded clause, which she claims is possible in [+MFS] languages (Bulgarian, Romanian), but not in [-MFS] type (Serbian/Croatian, Polish, and Czech). For Richards the wh-extraction data illuminates island effects indicating the taxonomy on the basis of Subjacency in order to identify CP-absorption (Bulgarian and Chinese), or IP-absorption languages (Serbian/Croatian and Japanese). CP-absorption languages have wh-movement properties: A-bar movement of whs to the SpecCP. The situation with IP-absorption languages is not as straightforward. Richards states that "IP-absorption languages … have somewhat more exotic properties. IP-absorption involves movement to some IP-level projection (or projections), whose position with respect to other IP-level projections is not easily determined." Such movement resembles scrambling found in languages like Japanese (cf. Saito 1992), for instance. Additionally, in some IP-absorption languages, a single wh-word moves obligatorily to Spec CP (Serbian/Croatian), although in Hungarian which is also considered an IP-absorption language it does not.

Still, there is an empirical difference between the language types, and we should consider where Russian falls with regard to this difference. If one claims that Russian is a Bulgarian-type language, as it has been stated in the current analysis, then we should predict Russian (a) to be able to multiply extract wh-phrases out of the embedded clause, and (b) to behave like a CP-absorption language with respect to islands.

For Rudin, the data in (85) indicate that more than one wh-phrase can occupy the SpecCP position. Since there is more than one wh-Spec is available, whs have an escape hatch to the higher clause. The following example comes from Rudin (1988) reproduced here under (85):

(85) **Koj** **kude** misliš [ce e otušil __ __] (Bulgarian)
who where think$_{2ndSG}$ that has gone
"Who do you think is gone where?"

Rudin states that the fact that multiple whs can be extracted in Bulgarin is true virtually for any wh-phrase. Russian behaves similarly to Bulgarian with respect to multiple extractions. Since Russian generally does not allow extractions from *čto* (that)-clauses[21], I am going to consider subjunctive embedded clauses with the subjunctive complementizer *čtoby* ('that/ in order to'), instead of čto (that) parallel to the examples in (85) of Bulgarian shown in (86) and (87) for Russian.

(86) **Kto** **kuda** ty prosiš, [čtoby ušel __ __]? (Russian)
who where you ask that go$_{SUBJ}$
"Who did you ask to go where?"

(87) **Čto** **komu**, ty xotel by, [čtoby Ivan podaril __ __]?
what whom$_{DAT}$ you want would that Ivan gave$_{SUBJ}$
"What would you want Ivan give to whom?"

Therefore, in this sense Russian behaves just like Bulgarian. Note, that S/C shows different behavior (Bošković 1994, 1996): in SC examples such as (85) in Bulgarian and (86) and (87) in Russian are deviant. The fact that multiple wh-extraction is possible in Russian confirms the idea that Russian patterns with Bulgarian-type languages. We have to keep in mind, though, that the extraction data is not clear-cut in Russian, there is a lot of judgment disagreement on the data, especially the part of it that involves wh-islands. The situation with islands is rather challenging. I have discussed islands in section 2.4.1, where I showed that wh-islands do exist in Russian. In Bulgarian, it is believed no such constraint exists (Rudin 1988, Richards 1997). I therefore turn to a discussion of islands now.

[21] The explanation of why extractions out of *čto* (that)-that clauses are bad is beyond the scope of this work; I just use it as a known Russian specific phenomenon.

2.11.2.1 Back to Islands

Rudin shows[22] that Bulgarian allows extraction out of wh-islands while SC does not, based on constructions like (88) and (89):

(88) (Bulgarian)

Vidjah	edna	kniga,	kojato$_i$	se	čudja	**koj**	znae	**koj**	prodava t_i
saw$_{1S}$	a	book	which	REFL	wonder$_{1S}$	who	knows	who	sells

(89) (SC)

*Vidio	sam	knjigu	**koju$_i$**	se	pitam	**ko**	zna	**ko**	prodaje t_i
seen	am	book	which	REFL	wonder$_{1S}$	who	knows	who	sells

"I saw a book which I wonder who knows who sells."

Sentences below demonstrate various island environments in Russian.

(90) *ČTO -island*

a.

Roditeli	dumali,	čto	deti	igrajut	v	sadu	s	ploximi	djadjami
parents	thought	that	children	play	in	yard	with	bad	guys

"The parents thought that the children are playing in the yard with the bad guys."

b.

??**S kem**,	roditeli	dumali,	čto	deti	igrajut	v	sadu	t_i?	
with who	parents	thought	that	children	played	in	yard		

"Who did the parents thought that the children are playing in the yard with?"

Adjunct islands demonstrate sensitivity to extraction of wh-phrases out of adjuncts.

(91) *Adjunct island*

a.

Ty	toropilsja	domoi,	potomu	čto	xotel	zakončit'	rabotu
you	rushed	home	because		wanted	finish	work

"You rushed home because you wanted to finish work".

b.

***Čto$_i$**	ty	toropilsja	domoi,	potomu	čto	xotel	zakončit'	t_i?
what	you	rushed	home	because		wanted	finish	

*"What did you rush home because you wanted to finish?"
(*n(1)=76/76, 100%)

[22] The judgments of the extraction data presented in Rudin (1988) have been disputed by several authors, for instance Lambova (2002), and partly by Richards (1997).

In the case of adjunct islands, the behavior of Russian does not differ from that in English: extraction out of adjunct islands is impossible: (91) is undeniably bad.

Wh-island restrictions have become a diagnostic for the existence of wh-movement in any given language. Sensitivity to islands is the only place where Russian appears to differ from Bulgarian, but in that it is not different from a language with single wh-movement. When it comes to islands, Russian exhibits behavior similar to English.

Bošković (2004) claims that island restrictions might not be a diagnostic of the wh-movement behavior of a language. Stepanov (2007)[23] re-visits Huang's (1982) Constraint on Extraction Domain, where he discusses different types of islands, while Rudin only touches upon the basic facts on island violations in S/C versus island insensitivity in Bulgarian.

The general claim drawn throughout this chapter, however, remains intact: Russian patterns with Bulgarian with respect to wh-behavior, it is a multiple wh-movement language with superiority and an unbreakable cluster of wh elements in (multiple) Specs of CP. I show more of how Russian behaves with respect to wh-islands in the upcoming chapters related to other wh-phenomena, namely sluicing and wh-coordination. Moreover, the evidence provided here for Russian wh-movement enables us to extend the idea that other languages with multiple wh-fronting might be analyzed in the same manner.

2.12 Conclusions

In this chapter, I have considered Russian wh-constructions with one main question in mind: does wh-fronting in Russian constitute true wh-movement for [+wh] reasons? I have examined Russian data and evaluated it against existing analyses of Russian wh behavior. These previous analyses appeared unsatisfactory because each one was able to explain the data only partially, and some make incorrect predictions. Additional investigation of the data collected from a large number of informants allowed me to propose an account that sheds light on the wh-phenomenon in Russian, and possibly to extrapolate this phenomena to other Slavic languages. Müller (1993:150) states that: "in languages like Polish and Russian, only the left-most wh-phrase in multiple questions occupies SpecC – the remaining wh-phrases have undergone the obligatory adjunction to IP, and thus, occupy a non-operator position at S-structure…" I have argued

[23] Sturgeon et al. (2010) discuss subject-islands and suggest that language specific properties such as the existence of a designated structural position for a topic (as in Czech in their experiments) changes the judgments on islands. But in their analysis, they, of course, state that Russian does not possess such a position..

against this statement, showing that Russian exhibits movement to SpecCP for all wh-phrases, and provided evidence for that, namely Superiority violations. I show that all wh-phrases occupy separate Specs of CP. I introduced the High Operator Phrase located higher than CP which can host only one wh-phrase after it has undergone wh movement to SpecCP. Introducing an additional functional category does not violate previous theoretical statements, yet permits us to review the wh-constructions and provide a unique and elegantly simple analysis of the constructions in question. I also show how Superiority might be derived without tucking in and how it applies to wh-behavior in Russian. On this view, Superiority is non-parameterized. This also allows me to rule out multiple wh-fronting of wh-adjuncts by introducing the *Multiple WH-Adjunct Ban* that requires proper government of traces (formerly, ECP).

The goal of the following chapters is to further investigate the nature of Russian wh-movement and the related phenomena and to provide more evidence for the HOP position.

Chapter 3: On Multiple WH Phenomena: Sluicing and Coordination

3.0 Introduction

This chapter discusses two more phenomena: sluicing and coordinated wh-constructions (CMW), both of which are directly related to multiple wh-behavior. While at first sight these phenomena might appear unrelated to each other, accounts of each are typically based on the analysis of non-coordinated, non-sluiced multiple wh constructions. The first part of the chapter is devoted to a discussion of Russian sluicing; the second part addresses the coordination of multiple wh-phrases. Sluicing is a phenomenon related to, but not necessarily defined by, the properties of the wh-behavior in a language (Ross 1969, Levin 1982; Groos & Reimsdjik 1981; Chao 1987; Chung et al. 1995; Lobeck 1995; Romero 1998; Lasnik 2001; and Merchant 2001). It has been shown for some languages that sluicing is directly impacted by the syntactic properties of wh-behavior of these languages (Ross, 1969; Merchant, 1998, 2003; Lipták and Craenenboeck, 2006; Stjepanović, 2003 among others). However, existing analyses of Russian sluicing (Grebenyova 2004, 2006) have suffered from reliance on the analyses of Russian wh-properties the limitations of which have been shown in detail in Chapter 2 (Stepanov's 1998 and Strahov's 2001 analyses). The account offered in this chapter differs from the existing analysis of Russian sluicing, and is based on the findings provided in Chapter 2. Later in this chapter, I argue that my analysis of Russian wh-movement allows for a successful analysis of coordinated multiple wh-phrases (CMW) as well.

(1) The primary conclusions of chapter 2 were:

(a) 'true' wh-movement takes place in Russian (the [+wh] feature is strong), wh-phrases are in the specifier of CP;
(b) Superiority must hold in wh-contexts in Russian; though its effects may be masked in the main clause by the ability of either wh to raise further (to SpecHOP);
(c) There is a High HOP position in matrix clauses only.
(d) Therefore, there is an asymmetry with respect to surface Superiority in matrix vs. embedded clauses: surface Superiority cannot be violated in embedded clauses;

In this chapter, I first consider the sluicing phenomenon in general. Then, I briefly discuss existing analyses of Russian sluicing (Grebenyova, 2004, 2006, 2007) and show how their predictions yield inaccurate results. I specifically discuss asymmetries between matrix and embedded sluicing environments and offer an analysis of multiple sluicing that is different from

those suggested earlier for Russian (Grebenyova 2004, 2007). My account better handles the facts and is consistent with the findings in chapter 2 based on data collected from a statistically significant number of speakers (n=35[1]). It is from these data that an additional puzzle arises, namely coordination between wh-phrases: in some multiple sluicing environments we find a surprising preference to use coordination ("i" *"and"*) in multiple wh-questions. Thus, a comprehensive account of wh-coordination in Russian is called for. I suggest an approach to wh-coordination constructions in the last part of the chapter. I offer an account of multiple coordinated wh-constructions (CMW), which is in accordance with existing analyses of multiple coordinated wh-questions in other languages, however, once again, I show that Russian patterns with languages where multiple wh-fronting is required and Superiority emerges in such derivations. I address existing accounts of wh-coordination in Russian, and while using them as a basis for an account suggested in the current chapter, I point out their weaknesses. As a consequence, I provide an account of multiple wh-sluicing and wh-coordination that is based on the idea that Russian is a wh-movement language. While the current chapter does not add significantly to the discussion of HOP, it gives additional evidence for the wh-phenomenon discussed in detail in the previous chapter.

3.1 Definition of Sluicing

The original term "sluicing" comes from Ross 1969 (followed by Merchant 2001). A description is given below in (2):

(2) Sluicing is a phenomenon of IP-ellipsis, where a wh-phrase moves out of IP, and IP is deleted at PF

Sluicing is a construction where there is an interrogative clause with only a wh-element (wh-elements) pronounced. It occurs in embedded clauses, (3)a and (4), as well as in main clauses, (3)b and (5), (6), in English and Russian respectively.

[1] The method for collecting sluicing and CMW data is identical to the one described in chapter 2. The characteristics of speakers are the same, the number of speakers is different for set of data due to the fact that it was only collected in one location: St.-Petersburg, Russia.

(3) a. John lovessomebody but I don't know who [John loves]. (English)

b. A: John lovessomebody.
B: Who[John loves]

(4) Dima ljubit kogo-to, no ja ne znaju **kogo** [~~Dima ljubit~~] (Russian)
Dima loves somebody, but I not know who [~~Dima loves~~]
"Dima loves somebody, but I don't know who"

(5) A: Dima ljubit kogo-to.
Dima loves somebody.

B: **Kogo** [~~Dima ljubit~~]?
Who [~~Dima loves~~]

(6) A: Ivan čto-to podaril Marii.
Ivan smth. gave Maria

B: **Čto** [~~Ivan podaril Marii~~]?
What? [~~Ivan gave Maria~~]

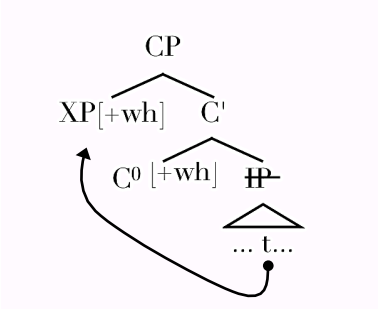
Ross (1969) claimed that in sluicing, the elided part forms a non-constituent. Thus, his solution was to analyze sluicing as regular wh-movement with ellipsis of the remaining part of the wh-phrase (question). Merchant (2001) further detailed Ross's ideas by presenting various sluicing puzzles and providing a comprehensive set of solutions for them. In the current approach to Russian sluicing, I rely on Merchant's findings. The data in (7) and (8) offers more examples of sluicing suggested by Merchant.

(7) a. Jane bought something but I don't know what.
b. Someone called, but I can't remember who.

(8) a. Jennifer was there, but you'd never guess who else.
b. Jack called but I don't know (when/why/who/where from).
c. Sally is out hunting – guess what!
d. A car is parked on the lawn – find out whose.

According to Merchant, these sentences ((7) and (8)) have the structure schematically shown in (9) where the IP-sentential part of the interrogative CP is elided.

(9) IP Ellipsis



The syntax of (9) is based on the syntactic properties of ellipsis (following Lobeck 1995, Saito and Murasugi 1999 among others). Thus, the ellipsis head is an empty category and it is important to understand what licenses the processes of ellipsis. Merchant 2001, Stjepanović 2003, and Takahashi 1993 (among others) assume that sluicing is related to the properties of wh-behavior in a particular language. Craenenbroeck and Lipták (2006) postulate the "Wh/Sluicing Correlation": "the syntactic features that the [e]-feature has to check in a certain language are identical to the strong features a wh-phrase has to check in a regular constituent question in that language" (p.257). Thus, one should assume that if wh-movement is possible, then sluicing is possible. Conversely, if wh-movement is not possible, then, the question arises whether sluicing is possible (Merchant, 1998). Merchant (1998), when discussing sluicing in Japanese, states that the lack of wh-movement creates an environment for what can be called pseudosluicing –"a kind of reduced cleft, where the pivot is a wh-phrase (10)."

(10) *Pseudosluice*:
An elliptical construction that resembles a sluice in having only a wh-XP
as remnant, but has the structure of a cleft, not of a regular embedded question.

a. Somebody kissed me, but I don't know who. [Sluice]
b. Somebody kissed me, but I don't know who it WAS [Pseudosluice]

Pseudosluicing also reportedly requires specific intonation. The argument here is that a language that exhibits wh-movement cannot apply pseudosluicing and the entire IP (TP) can be elided leaving only a wh-phrase behind. In a pseudosluice construction a copula verb usually appears overtly after a wh-phrase: it does not undergo ellipsis as in (10).

Additionally, Takahashi (1994) introduces the notion of multiple sluicing based on the assumption that if multiple wh-movement exists in a language, then multiple sluicing is also

possible when more than one wh-phrase appears at the sluice. Takahashi shows that in a language where multiple wh-movement is not observed (such as English), multiple sluicing cannot occur. Only one wh-phrase undergoes raising to SpecCP in English (11) - (13).

(11) What did John give to whom?

(12) What did John buy where?

(13) Where did John buy what?

Since only one wh-phrase is fronted in these questions, therefore, the prediction is that only one wh-phrase can appear in the sluice as shown in the following examples (14) - (17) (Merchant 2001 for the detailed discussion of the mechanics of these).

(14) John gave something to someone, but I don't know what.

(15) John bought something somewhere, but I am not sure what.

(16) *John bought something somewhere, but I don't remember where what.

(17) *John bought something somewhere, but I don't know what where.

It is only the wh-phrase fronted to SpecCP that can appear in the sluiced part of the sentence while the rest undergoes ellipsis. Notice, that (16) also crashes because of a Superiority violation. Thus, we are faced with several issues: on the one hand, sluicing technically cannot occur if wh-movement is not attested in a language; on the other hand, IP-ellipsis can be licensed by categories other than C^0 counter traditional Ross's view on sluicing (as was claimed for Hungarian, for instance, by Craenenbroeck and Lipták (2006) and for Japanese in Merchant 2001 and Kizu (1997, 1999) (18)).

(18) John-ga dareka-ni at-ta ga watasi-wa sore-ga
John$_{NOM}$ someone$_{DAT}$ meet$_{PAST}$ but I$_{TOP}$ it$_{DAT}$

dare-ni (da) ka sira-nai
who$_{DAT}$ (be) Q know-not

"John met someone, but I don't know who (it is)"

In such cases, however, the resulted ellipsis is not standardly considered sluicing under its original definition, (though it is often still called sluicing). As it stands, if we follow the notion that Russian does not exhibit multiple wh-fronting (suggested by Stepanov 1998, and supported

by Grebenyova) – then, one should expect that Russian does not exhibit multiple sluicing, similar to English as shown in (17), or if sluicing is empirically possible, then, the elision happens at IP(TP)/(FocP) level (whichever one assumes to host wh-phrases).

We have seen, however, that Russian does exhibit multiple wh-fronting. I have claimed that it exhibits multiple wh-movement for reasons discussed in chapter 2. For the theories that claim that multiple wh-fronting occurs for the reasons of focus movement or scrambling, we will have to assume one of the analyses that would accommodate sluicing licensed by a category other than C as in Hungarian, or to resort to an account of sluicing similar to Japanese following ideas of Nishiguachi (1998), Ku or Merchant (1998), who stated that Japanese exhibits pseudosluicing, i.e. clefting followed by ellipsis. By taking the theory presented in chapter 2 as a foundation for the current discussion, we can assume that the regular analysis of sluicing should suffice, an extremelt welcome simplification.

However, the simple application of a Merchant-type analysis is not going to be entirely sufficient for two reasons: first, the lack of Superiority in main clause sluicing constructions (shown below), and secondly, the existence of the apparently optional coordination marker "i" ("and"). The following section is devoted to accommodating the former problem: analysis of Russian sluicing constructions, including the discussion of the existing analyses of Russian sluicing. The latter: existence of coordination, in its turn requires a special discussion in the later sections of the chapter.

3.2 Types of Russian Sluicing

Any type of wh-phrase can appear in the sluiced construction, for example, subjects (19), objects (20), and adverbs (21) respectively.

(19) Ivan čto-to podaril Marii, no nikto ne znaet **čto**.
Ivan smth. gave Maria but no one NEG knows what
"Ivan gave Maria something, but no one knows what".

(20) Kto-to prislal professoru nepriličnoe pis'mo, no nikto ne
someone$_{NOM}$ sent professor$_{DAT}$ obscene letter$_{ACC}$, but no one NEG
dogadyvaetsja **kto**.
guesses who$_{NOM}$
Somebody sent a professor an obscene letter but no one can guess who"

(21) A: Kto-to prislal professoru nepriličnoe pis'mo.
someone sent professor$_{DAT}$ obsene letter
"Someone sent a professor an obscene letter."

B: **Kto**? / Ugadaj, **kto**.
Who? / Guess who.

(22) Maxim kuda-to podeval noski, no ne pomnit kuda.
Maxim somewhere put socks but NEG remember where.
"Maxim put his socks somewhere, but he doesn't remember where".

In chapter 2, I showed that multiple wh-questions are freely available in Russian matrix and embedded clauses. Thus, following Takahashi's 1994 definition of multiple sluicing, whenever two or more wh-phrases appear in the sluice before the elided material, we can consider multiple wh-sluicing constructions. As we would expect, given chapter 2, multiply sluiced wh-phrases can be found in both the embedded (23), (25) and the matrix (24) environments:

(23) Maxim komu-to čto-to rasskazyval, no ja ne uslyšal **komu** **čto**.
Maxim someone smth. told but I not heard who$_{DAT}$ what$_{ACC}$
"Maxim told something to someone, but I didn't hear who he told what".

(24) A: Kto-to kogo-to pervym udaril.
someone smbd. first hit
"Someone hit somebody first."

B: **Kto** **kogo**? (non-echo)
who$_{NOM}$ who$_{ACC}$

(25) Danila kogda-nibud' čto-nibud' izobretet, tol'ko on ne znaet **čto** *(i)[2]
Danila sometime something invent only he not know what$_{ACC}$ (&)
kogda.
when
"Danila will invent something someday, but he doesn't really know when it will happen, and what it will be."

Thus, we can see that the constructions with a single and multiple wh-remnants are available in both matrix and embedded clauses. In the following section, I explore how Russian multiple wh ellipsis constructions fit into the existing theories of sluicing.

3.2.1 Single Sluicing and Its Analysis

The data presented in (26) - (28) show a set of constructions with ellipsis and one wh-phrase, where a corresponding overt referent, also known as "correlate" in Merchant's (2001) terms, is shown:

(26) Maria vybrala sebe čto-to no ja ne zametil **čto** [~~Maria vybrala~~
Maria chose herself$_{DAT}$ smth. but I NEG noticed what [~~Maria chose~~
~~sebe~~.]
~~herself~~$_{DAT}$]
"Maria chose something for herself, but I didn't notice what".

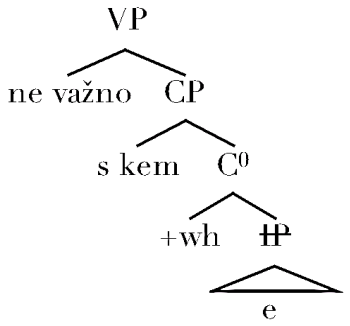
(27) Professor zastavil čitat' kakie-to stat'i, no ja ne pomnju **kakie**
professor ordered to-read some articles but I not remember which
[~~stat'i professor zastavil citat~~']
[~~articles professor ordered to-read~~]
"The professor mde (us) read some articles, but I don't remember which"

[2] Notice, that the sluicing of an wh-argument and an wh-adjunct requires «i» (and) in (25). A discussion of these is forthcoming.

(28) Deti s kem-to podralis' no mne ne važno **s** **kem**.
kids with smbd. fought but me not important with who$_{INS}$
[~~deti podalis'~~].
[~~kids fought~~].
"The kids fought with somebody but I don't care with who."

It has been generally stated that what functional projections license sluicing depends on what positions wh-phrases occupy in sluices (Merchant 2001). The idea that I adopt here is that if wh-phrases occupy SpecCP positions, then, in sluiced constructions they should occupy the same positions. As I showed in Russian, a wh-phrase moves to the SpecCP. Thus, an analysis of IP-deletion offered for the languages, where a wh-phrase is in CP as in (29) showing the sluiced part of (27) should explain single sluicing constructions with a single wh-remnant:

(29) [$_{VP}$ne važno [$_{CP}$ **s** **kem** [$_{IP/TP}$ ~~deti podralis'~~]]
NEG important with who ~~kids fought~~



Then, under the assumption of wh-movement and of what positions wh-phrases occupy in Russian, it should not be problematic to understand the mechanism of single sluiced matrix constructions. However, multiple sluicing constructions where multiple wh-phrases appear at the ellipsis site are somewhat trickier. The problems that arise here are discussed in the following section.

3.2.2 Multiple Wh Sluicing and Analyses

There are two observations that have been made (Grebenyova 2004, 2006; Scott 2004) about Russian multiple sluicing: in the matrix clauses the order of wh-remnants is free; however, in the

embedded clauses the order of the wh-remnants is fixed, in other words the Superiority Condition must hold. Note, that the same asymmetry in respect to Superiority in matrix vs. embedded (non-elided) clauses has been discussed in multiple wh-questions described in the previous chapter.

(30) A: Kto-to komu-to s utra trezvonit.
somebody someone since morning ring-repeatedly
"Somebody is someone from early morning over and over". [NOM>DAT ✔]

B: **Kto** **komu**?
who$_{NOM}$ whom$_{DAT}$ [NOM>DAT ✔]

B': **Komu** **kto**?
whom$_{DAT}$ who [DAT > NOM ✔]

(31) Komu-to kto-to zvonit s utra, no ja ne znaju
someone$_{DAT}$ smbd.$_{NOM}$ called since morning but I not know
kto **komu**
who$_{NOM}$ who$_{DAT}$
"Somebody calls someone in the morning, but I don't know who [calls] whom"
[DAT>NOM ✔: NOM>DAT ✓]

(32) ?*Komu-to kto-to zvonit s utra no ja ne znaju
somebody$_{DAT}$ somone$_{NOM}$ calls since morning but I not know
komu **kto**.
who$_{DAT}$ who$_{NOM}$
[*DAT>NOM ✔: DAT>NOM ✗]

(33) *Kto-to komu-to zvonit s utra no ja ne znaju **komu**
somone$_{NOM}$ someone$_{DAT}$ calls since morning but I NEG know who$_{DAT}$
kto.
who$_{NOM}$
"Somebody calls someone in the morning, but I don't know who [calls] whom."
[*NOM>DAT ✔: DAT>NOM ✗]

The examples in (30) - (33) show elliptical constructions with multiple wh-remnants. Notice that the violation of Superiority in the embedded clause makes sentences (32) and (33) ungrammatical. In example (30), where ellipsis happens in the main clause, either ordering of the wh-remnants is possible.

Russian multiple sluicing constructions were analyzed in detail by Grebenyova (2005, 2006, 2007) who used Stepanov's (1998) analysis of wh-behavior as the basis for her analysis of sluiced constructions. For her, then, the asymmetry in Superiority requirements in the matrix and the embedded clauses are problematic. The main points of her analysis are as follows:

(34) Grebenyova's (2006) Russian sluicing characteristics:

(i) licensing of IP – deletion cannot be accounted for because there is no filled CP to license it (Lobeck 1995 for licensing, Merchant for IP-deletion account[3], Chung, Ladusaw, McCloskey (CLM) 1995 for the deletion of TP);

(ii) here should be another way of accounting for multiple wh-remnant behavior (i.e. superiority effects);

(iii) in other words, since a Merchant-type account does not work for Russian other analysis is necessary.

Grebenyova (2006) provides an alternative account that is based on the assumption that no wh-fronting into the CP domain is available in Russian (following existing accounts of multiple wh in Russian). In her account the emergence of Superiority in the embedded contexts is explained by a *semantic* approach to Superiority in Russian. Superiority emerges only in embedded sluicing, and Grebenyova claims there should be a semantic (rather than a syntactic) explanation for it. In particular, she uses the notion of Semantic Parallelism (35) of Fiengo and May (1994) (also formulated in Rooth 1992 as Direct Parallelism, Fiengo and May 1998, CLM 1995).

(35) *Syntactic Parallelism Condition on IP Ellipsis*:
An IP E can be deleted only if E is morphosyntactically identical to an antecedent IP at LF

Grebenyova (2004) re-formulates this in her own terms:

[3] Chung, Ladusaw, McCloskey (standardly CLM, 1995, 2006) propose that is is not necessarily IP that is deleted but can be another category such as TP dependeding on the internal structure requried by a language. That distinction is ignored for present purposes.

> I suggest that in order to understand this phenomenon, one should consider the Parallelism required in ellipsis. I adopt the definition of Parallelism of Fiengo and May (1994), further developed in Fox and Lasnik (2003), which requires that variables in the elided and antecedent clauses are bound from parallel positions. I assume that in the antecedent clause the variable introduced by the indefinite is bound by existential closure (Kratzer 1998).

Grebenyova needs this device (of parallelism) to explain the sudden appearance of Superiority in sluicing envrionments. Grebenovya derives the facts in (31) - (33) from the properties of wh-behavior in Russian, utilizing the notion of Semantic Parallelism and existing analyses of ellipsis. Such an approach would have been ideal, except that some predictions it makes do not hold. For instance, the account does not work for the matrix clause sluicing constructions, as we have seen in the contrasts outlined in (30) vs. (31) - (33), where Superiority only emerges in the embedded clause, but not in the matrix clause. Thus, it is important to have an explanation for the environments where Superiority is not expected. Following Merchant's view on Japanese sluicing, we know that if C cannot license IP deletion, it does not behave as sluicing and has to entail a different analysis, for instance, pseudosluicing, or another sort of ellipsis as long as it has a licensor. For Grebenyova's analysis to survive, she is obliged to find a way to explain why Superiority holds in Russian constructions in embedded clauses. She refutes another possibility of C^0 licensing ellipsis (in the spirit of Stjepanović's 2003 analysis of S/C): "However, it is difficult to extend this analysis [Stjepanović's] to Russian, merging C^0 overtly cannot result in superiority effects. I would like to explore an alternative account and suggest that the Superiority effects observed under sluicing follow from an independent property of ellipsis, namely, quantifier parallelism."

The notion of Semantic Parallelism comes out of Prince's Parallelism Principle (1981: 226) described as a speaker's assumption "that the hearer will predict, unless there is evidence to the contrary, that (a proper part of) a new (conjoined) construction will be parallel/equivalent in some semantic/pragmatic way(s) to the one just processed." But even this semantic device does not explain the grammaticality of the sentence in ((31)=(37) below) where the ordering of the pronouns in the matrix clause and the ordering of the wh-remnants are reversed (and not parallel to the ones that are just processed), as well as ungrammaticality of either ((32)=(38) below) where the ordering is mirrored (equivalent and parallel in the matrix clause vs. embedded clause), and (33)=(39) where it is reversed and, hence, violating Superiority in the embedded clause compared to (36) which is fine.

(36) A: Kto-to komu-to s utra trezvonit.
somebody someone since morning ring-repeatedly

"Somebody is someone from early morning over and over". [NOM>DAT ✔]

B: **Kto** **komu**?
who$_{NOM}$ whom$_{DAT}$ [NOM>DAT ✔]

B': **Komu** **kto**?
whom$_{DAT}$ who$_{NOM}$ [DAT > NOM ✔]

(37) Komu-to kto-to zvonit s utra, no ja ne znaju .
someone$_{DAT}$ smbd.$_{NOM}$ called since morning but I not know

kto **komu**
who$_{NOM}$ who$_{DAT}$

"Somebody calls someone in the morning, but I don't know who [calls] whom"
[DAT>NOM ✔: NOM>DAT ✔]

(38) ?*Komu-to kto-to zvonit s utra no ja ne znaju
somebody$_{DAT}$ somone$_{NOM}$ calls since morning but I not know

komu **kto**.
who$_{DAT}$ who$_{NOM}$
[*DAT>NOM ✔: DAT>NOM ✗]

(39) *Kto-to komu-to zvonit s utra no ja ne znaju
somone$_{NOM}$ someone$_{DAT}$ calls since morning but I not know

komu **kto**
who$_{ACC}$ who$_{NOM}$

"Somebody calls someone in the morning, but I don't know who [calls] whom."
[*NOM>DAT ✔: DAT>NOM ✗]

Again, in matrix sluicing, either order is possible. In embedded contexts, Superiority determines possible orders, and crucially, not due semantic parallelism. The contrast between (37) and (38) shows that the device to explain the ungrammaticality of some wh-ellipsis constructions on the semantic basis is not sufficient.

However, the question still remains what licenses the deletion: how do the wh-remnants in Russian survive deletion if they are not in SpecCP? It is true that various ellipsis environments

require Parallelism effects (Fiengo & May 1994, Merchant 2001). This is obviously not what seems to be at work here since such an account requires one to deny a strong [+wh] feature in the sluiced setting and assume that some other category licenses of IP/TP -deletion. This suggests two possibilities:

(i) [+wh] and [+foc] features are both capable of licensing TP-deletion

(ii) Foc^0 is the licensor of TP-deletion

Under the view in (ii), not only SpecCP occupants can survive the deletion process. The conclusion then, is that Focus feature can license TP-deletion, i.e. "the projection [which is] immediately below the focus projections is elided in sluicing" (Merchant 2001) (as shown in the linear structure below)

(40) [CP [FocP **wh** [~~TP …t~~]]]

Craenenbroeck and Lipták (2006) use this analysis for sluicing in Hungarian. But once again, such an analysis does not explain why Superiority should hold in embedded sluices and not in matrix ones. The FocusP deletion licensing analysis can be involved is to allow to predict Superiority as it happens in matrix clauses. Additionally, if Foc as a licenser is available in Russian, then we should expect phrases with both filled CP and FocP, and then ellipsis to follow? The following sentences show examples where Focus is overtly present in the 'sluiced' context. The judgments on these and similar sentences are collected from 35 speakers (*n*=35) and are reflected in examples (41) and (42) both of which are ungrammatical with the focus element left non-elided:

(41) *Prepodavateli vydavali studentam diplomy no ja ne zametil **kto**.
professors gave students diplomas but I not noticed who
STEPANU
$STEPAN_{DAT.FOC}$
"The professors gave out diplomas to students, but I didn't notice who [gave a diploma] to STEPAN"

(42) ??Ispolniteli predstavljali raznye pesni na konkurse no ja ne pomnju
singers performed different songs at competition but I not remember

kto "Gorodok"$_{FOC}$
who "GORODOK"

"Singers performed different songs at the competition, but I don't remember who [performed the song] "LITTLE TOWN".

It appears that the focused element cannot remain when deletion is licensed. An alternative analysis of these sentences could come along the lines of gapping. Thus, we could conclude that sluicing is licensed by C analogous to its English counterpart following the analysis of Loebeck (1995) and Merchant (2001) outlined in (9) above: IP-deletion is licensed by C^0. Superiority must hold in multiple sluices, and the analysis based on the principle of Parallelism does not bear on correct predictions.

However, there is another idea that should be considered here, and which follows from Merchant's (1998) view on Japanese sluicing: we know that if C^0 cannot license IP deletion, it does not behave as sluicing and has to entail a different analysis, for example, as pseudosluicing (not a CP-licensed PF deletion), or another sort of ellipsis. Citko (p.c.) suggests that it is possible that Russian (and conceivably similarly Polish) exhibits VP-ellipsis (VPE). For instance, VPE can be a possible analysis for (43):

(43) Kto-to na kogo-to kričal no ja ne slyšal **kto**
someone$_{NOM}$ at smbdy$_{ACC}$ screamed but I not heard who$_{NOM}$

na **kogo** [~~kričal~~].
at whom$_{ACC}$ [~~yelled~~]

"Someone yelled at somebody but I didn't hear who (yelled) at whom".

Nonetheless, a mere VPE explanation is not sufficient here, for if we assume that only the VP is elided, we would end up with incorrect predictions: that an auxiliary will overtly remain after ellipsis takes place. Some immediate evidence (following Johnson 1991) that these constructions cannot be viewed as VP-ellipsis comes from the fact that if it is VPE, then, in cases where an overt copula is present, we should expect the copula to be present at Spell-out. Consider the following examples where the auxiliary remains at the ellipsis site and this results in ungrammaticality:

(44) *Vse ispolniteli budut pet' raznyje pesni no ja ne znaju **kakie**
all performers will sing different songs but I not know which

kto budut/budet.
who will$_{3PL/3SG}$

"All performers will sing different songs, but I don't know who will sing which one".

(45) *Každyj ispolnitel' budet pet' svoju pesnju no my ne uvereny
each performer will sing his$_{REFL}$ song but we not sure

kto kakuju budet
who which will$_{3SG}$

"Each performer will sing his/her song but we are not sure who will sing which song"

(46) Každyj ispolnitel' budet pet' svoju pesnju no my ne uvereny
each performer will sing his$_{REFL}$ song but we not sure

kto kakuju
who which

"Each performer will sing his/her song but we are not sure who will sing which one"

If we assume a case of VP-ellipsis here, we should expect the future copula '*budet'* ('will') to remain overtly. Examples (44) and (45) show otherwise: the presence of the copula yields ungrammatical sentences. Thus, it is unlikely to be a case of VP-ellipsis.

Instead it could have been logical to rely on the analysis of sluicing offered for Serbian/Croatian (S/C) by Stjepanović (2003), who follows Boškovic's (1998b) observation that S/C shows Superiority effects in some contexts: sluicing being one of them. She states that Superiority in S/C emerges in a number of contexts: long-distance questions, embedded questions, and matrix clauses with an overt complementizer – when overt wh-movement must occur, and in such contexts Superiority is operative: below are the examples from Bošković (2002a) of such environments in (47) - (49):

(47) a. [**Ko koga** voli], taj o njemu i govori. (S/C)
who whom loves that-one about him even talks
"Everyone talks about the person they love."

b. ?*[**Koga ko** voli], taj o njemu i govori.
whom who loves that-one about him even talks

(48) a. ?**Ko koga** vrdiš da je istukao?
who whom claim$_{2S}$ that is beaten
"Who do you claim beat whom?"

b. ***Koga ko** t vrdiš da je istukao?
whom who claim$_{2S}$ that is beaten

(49) a. Tom čoveku, **ko** je **šta** poklonio?
that man who aux what bestowed
"On that man, who bestowed what?

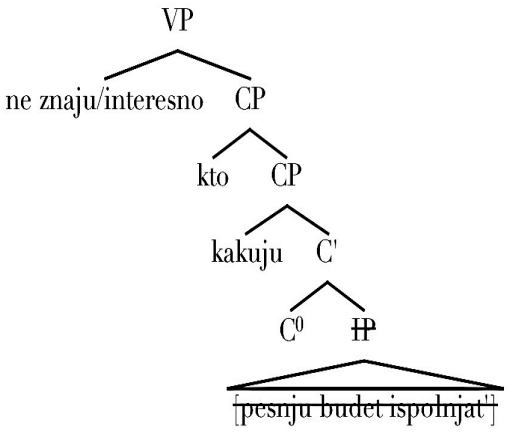
b. ??Tom čoveku, **šta** je **ko** poklonio?
that man what aux who bestowed

Bošković (2001) also notes that the contexts in which S/C must have wh-movement also display selective Superiority effects: the highest wh-phrase prior to movement is first in the linear order, the order of other wh-phrases being free. We have seen that in Russian Superiority emerges in some contexts, and is absent in others, and the differences between Russian and S/C were drawn out. However, sluicing is one construction in which wh-properties of Russian and S/C appear similar. This is expected, since sluicing is an environment in S/C where wh-movement properties surface. Stjepanović (2003) considers sluicing in SC as PF deletion (+E feature) of the constituent after the wh-phrase, which is interpreted at LF. The technical implementation of this proposal is that placement of the [+E] feature triggers deletion at PF (in the spirit of CLM 1995 and Merchant 2001 outlined in (9)). She states that since wh-phrases must be in SpecCP in sluicing constructions, C^0 must be present in overt syntax. A strong [+wh] feature in C^0 triggers Superiority effects. This comes from the sluicing description offered in Ross 1969, Lasnik 1999, Merchant 2001 among others and given in (50):

(50) the interrogative complementizer (i.e. C^0 with features [+wh, +Q]) licenses the deletion of its complement (i.e. TP): wh-movement is to SpecCP followed by the PF deletion of the category below: CP (IP-deletion).

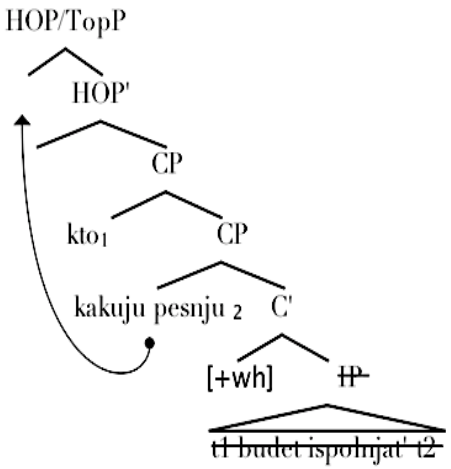
Other analyses seem to be unable to accommodate all the facts of the constructions with wh-remnants in Russian: the asymmetry in Superiority requirements in the matrix and the embedded clauses, the non-requirement for strict semantic Parellelism in the matrix clauses, the unavailability of an overt Focus element (hence, the impossibility of Foc^0-licensing of the deletion), and the unavailability of the overt copula (hence, the improbability of VP-ellipsis). However, if one postulates that wh-movement takes place in Russian (as it has been argued in chapter 2): then, the CP analysis, and the "standard" Merchant-type IP-deletion analysis of sluicing should work for Russian. If such an analysis of wh is correct, one should expect the following to be true: multiple sluicing is available in Russian in both matrix and embedded clauses, and superiority is observed in embedded sluicing. Therefore, the example of Russian sluicing can be analyzed as shown in (51)–as an IP- deletion and C-licensing the sluice:

(51) Embedded Sluicing in Russian: C-licensed IP deletion:



The non-existence of Superiority requirements in the matrix clauses follows from the availability of the HOP position in the matrix clauses. Similarly to the multiple wh – question constructions, the wh-phrases have the escape path into HOP position in the matrix clauses but not in embedded ones. This is shown in (52):

(52) Matrix Sluicing: HOP / C-licensing IP-deletion:



Thus, with the findings of chapter 2 that assert that Russian has multiple wh-movement into the C-domain - the analysis of sluicing becomes a much easier enterprise: it reduces to a Merchant-type sluicing analysis with IP-deletion, and no additional semantic devices are needed.

Nonetheless, the sluicing analysis cannot be entirely complete at that. In addition to the behavior outlined above, Russian contains structures that have provoked questions on how to account for them if the structure above is correct, namely coordinated wh-constructions (CMW). In the next section I turn to wh-constructions in sluicing the analysis of which deem challenging without a satisfactory account of general coordinated wh-constructions in Russian.

3.3 "Wrinkled" Sluicing

As I have shown, analyzing sluicing in Russian reduces to a simpler task assuming an account of [+wh]-behavior in Russian. Adapting the multiple wh-movement approach, where all wh-phrases move into CP, gives a straightforward analysis of sluicing. However, Russian sluicing offers an additional puzzling twist shown below in (53): cases in which two wh-remnants have to be coordinated. The puzzle unveils in two ways: some sluicing constructions allow multiple wh-phrases at the ellipsis site (as in (53): without change in meaning), while others require (as in contrast between (54)a and (54)b), the addition of the conjunction 'i' - "and".

(53) a. ?*Kto-to čto-to čitaet, no ja ne uveren **kto** **čto**.
somebody something reads but I NEG sure who what
"Somebody reads something, but I am not sure who [reads]what"

b. Kto-to čto-to čitaet, no ja ne uveren **kto** i **čto**.
somebody something reads but I NEG sure who and what
"Somebody reads something, but I am not sure who [reads]what"

(54) a. Kto-to kogo-to udaril, no ja ne znaju, **kto** **kogo**.
smbd. $smbd_{ACC}$ hit but I NEG know who who
"Somebody hit someone but I don't know who whom"

b. Kto-to kogo-to udaril, no ja ne znaju, **kto** i **kogo**.
smbd. $smbd_{ACC}$ hit but I NEG know who and who
"Somebody hit someone but I don't know who whom"

(55) a. *Kto-to kogo-to udaril, no ja ne znaju, **kogo** **kto**
somebody $smbd._{ACC}$ hit but I NEG know who who
"Somebody hit someone but I don't know who whom"

b. *Kto-to kogo-to udaril, no ja ne znaju, **kogo** i **kto**
somebody $smbd._{ACC}$ hit but I NEG know who and who
"Somebody hit someone but I don't know who whom"

So far, it appears that the general restrictions on Superiority hold in cases of coordinated wh-phrases in the same way as they do in regular multiple wh sluicing (55).

There are several questions to be asked here about the data in (53), (54), and (55). Is this just an accidental phonological requirement for an exception for an animacy – inanimacy principle considering that a wh-phrase 'čto' ('what') in the Nominative and the Accusative cases are identical? And more globally, how can conjoined sluicing constructions be analyzed? Or, are they related to coordinated wh-behavior in Russian? The answers to these questions come out of the consideration of an independent wh-construction phenomenon: coordinated whs. Coordinated multiple wh-constructions (CMW) pose additional questions. Perhaps, if we have an independent (full clause) account of CMW, we can account for coordinated constructions in

sluicing. In the next section I address CMW, and in the consequent section, I explain the contrast between (53) and (54) – the contrast that arises from the specifics of Russian coordinated wh-constructions in the following sections.

3.4 Coordinated Multiple-WH Constructions (CMW)

Coordinated wh-phrases are not specific to Russian and are attested in many languages: Serbian/Croatian (Browne 1972, Gračanin-Yuksek 2007), Bulgarian, Polish (Citko 2008), Greek (Sinopolou 2009), Vlach (Merchant 2007), and Russian (Kazenin 2002, Gribanova 2009). There are, of course, various differences in the restrictions that these constructions have in each language, and the interpretation that such questions yield in each question. Kazenin (2001) states:

> Although multiple questions in Russian and other Slavic languages generally are relatively well studied, constructions with coordination of Wh-phrases (henceforth, CMW), surprising in many aspects, somehow have not got any considerable attention on behalf of formal syntacticians working at Slavic languages (p.1).

As is well known, coordination is generally possible only for phrases of identical categories that occupy identical syntactic positions, and is impossible for phrases occupying different syntactic positions. This is shown in the contrast between (56), (57) vs. (58), (59):

(56) Orkestranty stučali po tarelkam i čajnikam
orkestra-players hit on plates and teapots
"Orchestra players hit plates and teapots"

(57) Oni slyšali ob ètom ot babuški i (ot) mamy.
they heard about it from grandma and (from) mother
"They heard about it from grandma and mom."

(58) Udaril Vasja (*i) Petju
hit $\text{Vasja}_{\text{NOM}}$ (*and) $\text{Petja}_{\text{ACC}}$
"Vasja hit Petja."

(59) Ob ètom slyšal Vasja (*i) ot Peti.
about this heard Vasja (*and) from Petja
"Vasja heard about it from Petja."

However, coordination of wh-phrases differs in this respect from any other phrasal coordination: wh-phrases can originate from different positions allowing coordination of NPs with different theta roles, or coordination of arguments and adjuncts. Thus, in most cases, they violate the Law of Coordination of Likes (LCL) (Schachter 1977, Williams 1981), which generally prohibits coordination of constituents of different categories[4].

(60) **Kto** i **kogo** udaril?
who$_{NOM}$ and who$_{ACC}$ hit
"Who hit who(m)?"

(61) **Kto** i **ot** **kogo** ob ètom slyšal?
who$_{NOM}$ and from who about this heard
"Who heard this and from whom did s/he hear this?"

The fact that coordinated wh-phrases violate LCL is a truly remarkable property of CMW, they do not have restrictions on compatibility of wh-phrases: a subject wh-phrase can coordinated with any object wh-phrase.

(62) **Kto** i **za** **čem** prixodil?
who and for what came
"Who came and what for (for which thing)?

(63) **Komu** i **o** **čem** ty rasskazyval?
whom and about what you told
"Whom did you tell about what?"

(64) **Kogda** i **čem** zakončilas' Perestrojka?
when and with-what ended Perestrojka
"When did Perestrojka end and how (with what result)?"

[4] The Law of Coordination of Likes is widely described in the literature,; for instance, Progovać (1996) starts from citing Chomsky (1957) "if X and Y are both constituents, but are constituents of different kinds… then we cannot in principle form a sentence by conjunction" (p.35), and continue by using Schacher's (1970:90) formulation of Coordinated Constituent Constraint, where it is stated that "The constituents of a coordinated construction must belong to the same syntactic category and have the same semantic functions". Camacho (2003) refers to the same phenomenon as Wasow's generalization, which requires symmetry between syntactic representations of the phrases in a conjunct.

While the fact that non-alike wh-phrases have been observed for Russain CMW (as well as many other languages), some crucial data of certain restrictions on multiple wh-coordination in Russian have been overlooked; partly because the fact that various wh-phrases exhibit different properties of coordination with respect to Superiority effects, for instance. Hence, now I turn to the typology of CMW in Russian, and the specific properties of each type.

3.4.1 Types of multiple wh-constructions

3.4.1.1 Typology

There exist several approaches to coordinated wh-constructions. The fact is addressed in the literature on other Slavic languages but not on Russian (Skrabalova 2006, Gračanin-Yuksek 2007, Citko 2008 among others). Skrabalova's (2006) classification of multiple wh-questions with and without coordination is given below. She offers the following typology:

(65) Skrabalova's (2006) classification of multiple wh-questions:

(a) Multiple wh-constructions: two or more wh-phrases
(b) Type 1: coordinated wh-constructions where the wh-phrases are (all) **arguments**
(c) Type 2: coordinated wh-constructions where the wh-phrases are (all) **adjuncts**
(d) Type 3: (Mixed type): coordinated wh-phrases where **both argument and adjunct** wh-phrases are present

Multiple wh-constructions (as in (a) in the above typology) have been extensively discussed and analyzed in chapter 2 of the current work. A consideration of the other Types 1, 2, and 3 (Mixed) follows. It is important to discuss each type separately in the analysis of Russian CMW. One of the problems that existed with other analysie of CMW in Russian is that all the data (even though recognized by the virtue of a possibility of various options in coordination): argument type – Type 1, adjunct type – Type 2, and the Mixed Type were given a uniform analysis independently of the type of coordination. I argue thst distinct analyses are required.

3.4.2 CMW: Type 1, Type 2 and the Mixed Type

Gribanova (2009) identifies three properties of CMW constructions in Russian. First, any number of wh-phrases, two or more, can be coordinated. (I am going to consider only the

coordination of just two wh-phrases with the assumption that more than two coordinated wh-phrases would behave in a similar manner.) Second, any type of wh-phrases can be coordinated, and third, there exist no apparent restrictions on the ordering of the wh-phrases. Both the second and the third statement require careful clarification. While it is true that any type of wh-phrases can be coordinated, we are going to see when it comes to two adjunct wh-phrases, that in that case coordination is not only possible, but is necessary to create a grammatical questions. The discussion of why non-coordinated multiple adjunct fronting is impossible was discussed in chapter 2 on multiple wh-fronting (refer to the *Multiple Wh-Adjunct Ban*). And finally, and most importantly, in most contexts of wh-coordination in Russian (both Type 1 and Type 3), Superiority emerges. (There are, of course, no ordering restrictions when two adjuncts are coordinated, as would be expected, sice their underlying positions show no asymmetries.)

Another issue that always goes hand in hand when CMW are being considered is the question of their interpretation: availability of Single-Pair (SP) versus Pair-List (PL) interpretations. Many analyses of CMW are built on the fact that they yield (or less categorically put: 'strongly prefer' (Citko & Gračanin-Yuksek 2010)) only SP answers, and hence, have a specific semantic disposition, and consequently specific syntactic structure different from regular wh-fronting constructions. In short, the difference between multiple wh-constructions without coordination and coordinated multiple wh-constructions is discussed on the semantic level of response interpretations, rather than a syntactic one: Gribanova (2009:134) makes the following point:

> … the syntax of these constructions interacts with their available interpretations: namely, that CP-fronting supports only a pair-list reading, whereas IP-fronting supports both a single – pair and a pair –list reading". For her, each reading corresponds to a different structure. Such argumentation becomes cyclic in its nature. One would need to show independent reasons for a need of a certain structure rather than eliciting a specific meaning. Moreover, a statement that coordinated wh-phrases necessarily yield a SP reading is highly overrated, as I demonstrate below.

Sluicing is often used as a diagnostic for the availability of PL/SP readings. Gribanova (2009: 141) following Grebenyova (2006) asserts that the non-coordinated wh-items[5] that appear in sluicing are infelicitous if the antecedent forces an SP reading.

[5] In section 3.3, it was noted that in certain sluicing constructions coordination of wh-phrases is necessary, and if coordination is present the meaning of the sluice remains the same. The statistical data obtained for the current work next to Grebenyova's examples show the relevant numbers.

(66) Každyj priglasil kogo-to na tanec no ja ne pomnju **kto** **kogo**
everyone invited someone to dance but I NEG remember who_{NOM} who_{ACC}

"Everyone invited someone to dance, but I don't remember who (invited) who."

(67) #Každyj priglasil kogo-to na tanec no ja ne pomnju **kto** i
everyone invited someone to dance but I NEG remember who_{NOM} and

kogo
who_{ACC}

"Everyone invited someone to dance, but I don't remember who (invited) who".
[*n*(5)=33: 94.7% - contrary to the "infelicitous" judgment]

(68) Kto-to priglasil kogo-to na tanec no ja ne pomnju **kto** i
someone invited someone to dance but I NEG remember who_{NOM} and

kogo
who_{ACC}

"Someone invited someone to a dance, but I don't remember who (invited) who".

(69) #Kto-to priglasil kogo-to na tanec no ja ne pomnju **kto**
someone invited someone to dance but I NEG-remember who_{NOM}

kogo
who_{ACC}

"Someone invited someone to a dance, but I don't remember who (invited) who.
[*n*(5)=35: 100% - contrary the infelicitous judgment]

The notation of ("#") is taken from Gribanova (2009). However, as the data show, these sluiced constructions are fine both semantically and syntactically. Let me note that in Serbian/Croatian an analogous example appears ungrammatical and not simply "infelicitous".

(70) *Znam da je neko pozvao nekogo da igra ali (SC)
I-know that aux someone invited someone to dance but

ne znam **ko** i **kogo**
not know who_{NOM} and who_{ACC}

"Someone invited someone to a dance, but I don't remember who (invited) whom."

Gribanova (2009: 136), who attempts an analysis of Russian CMW, has inconclusive data on Russian coordinated wh-constructions, and states:

> The judgments recorded above may vary across speakers, which is not uncommon in the realm of ordering restrictions (one need only check the abundant literature on superiority effects in English to get a sense of how variable judgments are). However, preliminary corpus searches show that orderings like the ones reported here are used frequently.

However, I rely here on speaker data collected in the grammaticality judgment task (n=35)[6] and the percentages of acceptability are recorded next to each example. In this section, I separate and organize the data with some of the target examples coming from the Russian corpora. Such typological organization follows the ideas of Skrabalova (2006), Citko (2009), Gračanin-Yuksek (2007) who state that CMW constructions are not uniform. I also note on the availability of a SP versus a PL reading for each grammatical example. Hence, I address this issue of possible interpretation(s) when considering Russian CMW data. I also indicate the availability of a SP versus a PL reading for each grammatical example. Hence, I address this issue of possible interpretation(s) when considering Russian CMW data in the following sub-sections.

3.4.2.1 Type 1: Wh-argument & Wh-argument

Type 1 constructions consist of two coordinated wh-arguments. Please note that the glosses of these sentences are difficult in English for English lacks such constructions.

(71) ***Čto** i **kto** govorit?
what$_{ACC}$ and who$_{NOM}$ says
"Who says and what?" ("What is said and who says it?")
[*ACC>NOM: n(1)=33: 94.3%; n(2)=2: 5.7%)[7]

(72) **Kto** i **čto** govorit?
who$_{NOM}$ and what$_{ACC}$ says
"Who says what?"
[NOM>ACC: n(5)=32: 92.4%] [SP/PL✔]

---

[6] While these speaker data are not perfectly uniform, they are consistent from speaker to speaker with a low percentage of variability among the utterances.

[7] Data notation: n-number of speakers who gave the judgment score of (1-5) (maximum n=35). The mode of the scores given is reflected in parenthesis, from (1) to (5) possible score. For example, "n(1)=33" means that 33 speakers out of 35 assigned a score of "1" (not acceptable at all) to a given example.

(73) **Kto** i **kogo** obmanyvaet?
who$_{NOM}$ and who$_{ACC}$ cheats

"Who cheats whom?"
[NOM>ACC: *n*(5)=34: 97.1%] [SP/PL: ✔]

(74) ***Kogo** i **kto** obmanyvaet?"
whom$_{ACC}$ and who$_{NOM}$ cheats

"Who cheat s whom?"
[*ACC>NOM: *n*(1) = 35: 100%]

(75) **Kto** i **v čem** ubeždaet Petra?
who$_{NOM}$ and in-what$_{PREP}$ convinces Petr$_{ACC}$

"Who convinces Peter and of what?" lit. "Who convinces Peter to do what?"
[NOM>PREP: *n*(5) = 34: 97.1%] [SP/PL: ✔]

(76) ***V čem** i **kto** ubeždaet Petra?
in-what and who$_{NOM}$ convinces Petr$_{GEN}$

"Who convinces Petr and of what?"
[*PREP>NOM: *n*(1) = 32: 92.4%]

Notice the contrast in grammaticality in examples where Superiority is preserved as in (72), (72), and (75) are fine, but where it is violated: as in (71), (74) and (76) are ungrammatical[8].

All of these examples show the following distribution: coordinated wh-phrases with a coordinator "*i"* ("and") occur pre-verbally: Superiority between a Nominative and the Accusative, or Nominative and oblique case arguments is observed. When Superiority is violated, the questions become ungrammatical. Such questions "always"[9] have both Single-Pair (SP) and Pair-List (PL) interpretations available, contra Gribanova (2009).

---

[8] There exists an additional strategy to create CMW in Russian, which is shown in (i) and (ii) below:
(i) **Kto** ubeždaet Petra i **v cem**?
who$_{NOM}$ convinces Peter and in what$_{PREP}$
(ii) ***V čem** ubeźdaet Petra i **kto**?
in-what$_{PREP}$ convinces Petr and who$_{NOM}$
lit. who convinces Peter in what?
While the Superiority facts are the same here, I ssum thee constructions have a distinct derivation, and leave them out of the discussion.

[9] The word "always" is in quotes to reflect the fact that there are, or course, cases where either only SP or only PL interpretation is available due to very specific contexts. For instance, in question "Who killed whom? (Who killed and who it was that was killed?" (to reflect the CMW context) (from Pesetsky's 2009 lecture), the answer depends on the context: only single pair is available when only two people are contextually involved (a duel between Burr

3.4.2.2 Type 2: Wh-adjunct & Wh-adjunct

Type 2 constructions consist of two coordinated adjunct wh-phrases. The most curious fact about this type is that when two wh-phrases are adjuncts multiple wh-fronting no direct derivation is available but multiple fronted coordinated wh-adjuncts are possible. (A similar phenomenon is described by Citko & Gračanin-Yuksek 2010 in Croatian).

(77) **Kogda** i **gde** vy vystupaete?
when and where you perform

"When do you perform and where?"
[WH-ADJ & WH-ADJ: *n*(5) = 35: 100%] [SP/PL: ✓]

(78) ***Kogda** **gde** you vystupaete?
when where you perform

"When do you perform where?"
[*WH-ADJ WH-ADJ: *n*(1) = 35: 100%]

(79) *?**Gde** **kogda** vy vystupaete?
where when you perform?

"Where do you perform when?"
[WH-ADJ & WH-ADJ: *n*(1) = 33: 94.7%; *n*(2)=2:5.7%]

(80) **Gde** i **kogda** vy vystupaete?
where and when you perform

"Where do you perform and when?"
[WH-ADJ & WH-ADJ: *n*(5) = 35: 100%] [SP/PL: ✓]

(81) **Gde** i **počemu** vy vystupaete?
where and why you perform

"Where do you perform and why?"
[WH-ADJ & WH-ADJ: *n*(5) = 35: 100%] [SP/PL: ✓]

(82) ?***Gde** **počemu** vy vystupaete?
where why you perform

"Where do you perform why?"
[*WH-ADJ WH-ADJ: *n*(1) = 33: 94.7%; *n*(2)=2: 5.7%]

---

and Hamilton in 1804), or only pair-list reading is available if many people are present as in a battle (as in the battle of Waterloo in 1815).

Any discussion of Superiority, of course, is not relevant in this case, since the initial ordering of wh-adjuncts is free (the explanation is below). However, coordination is mandatory and both SP and PL are possible when two adjunct wh-phrases are coordinated.

3.4.2.3 Type 3. Mixed Type: Wh-argument & Wh-adjunct

In this type of constructions the wh-argument is conjoined with a wh-adjunct.

(83) **Kto gde** budet spat'?
who where will sleep
"Who will sleep where?"
[WH-ARG WH-ADJ: *n*(5) = 35: 100%] [SP/PL: ✓]

(84) **Gde kto** budet spat'?
where who will sleep
"Who will sleep where?"
[WH-ARG WH-ADJ: *n*(5) = 35: 100%] [SP/PL: ✓]

(85) **Kto** i **gde** budet spat'?
who and where will sleep
"Who will sleep and where?"
[WH-ARG & WH-ADJ: *n*(5) = 34: 97.4%] [SP/PL: ✓]

(86) ***Gde** i **kto** budet spat'?
where and who will sleep
"Who will sleep where?"
[*WH-ADJ & WH-ARG: *n*(2) = 30: 85.7%] [SP/PL: ✓]

In the coordinated Mixed type constructions Superiority once again emerges (85) vs. (86). Both PL and SP interpretations are available. The fact that for all of the types both interpretations are available contradicts the existing interpretation for these constructions in other languages[10]. The observations about these constructions can be summarized in the following table.

[10] Tomaszewicz (2010) states that CMW in Polish have PL interpretations in special cases. I take this claim further and state that an interpretation of CMW in Russian is not special and PL readings are available across the board.

TABLE 1. MULTIPLE WH TYPOLOGY

	Type 1 Arg&Arg	**Type 2 Adj&Adj**	**Mixed Type 3 Arg &Adj**	**Multiple wh**
Coordination Mandatory	no	yes	no	no (not available for wh-adjuncts)
Superiority	yes	no	yes/no	yes/no
SP Interpretations	yes	yes	yes	yes
PL Interpretations	yes	yes	yes	yes

This summarizes all of the above observations and allows for the analysis that follows in the next sections.

3.5 Analysis of CMW

A number of analyses exist for CMW for different languages: Browne's 1972 for Serbo-Croatian, Gračanin – Yuksek's 2007 for Czech, Citko's 2008 for Polish, Merchant's 2008 for Vlach, Kazenin's 2002 and Gribanova's 2009 for Russian. All of these authors consider CMW as an instance of coordination of two XPs at the left periphery of the clause or clauses. Existing accounts of Coord-wh pattern into two groups: monoclausal approaches (i) (Kazenin 2002, Lipták 2001, Skrabalova 2006, Gribanova 2009), and (ii) biclausal approaches in terms of ellipsis (Bánréti 1992, Giannakidou & Merchant 1998, Camacho 2003, Tomaszewicz 2011), or sharing (Citko & Gračanin-Yuksek 2010).

(87) Two kinds of CMW analyses:

(i) [CP [&P wh1 **and** wh2] [TP *t*1 ... *t*2]]

(ii) [&P [CP wh1 [TP *t*1 ...]] **and** [wh2 [TP ... *t*2]]]

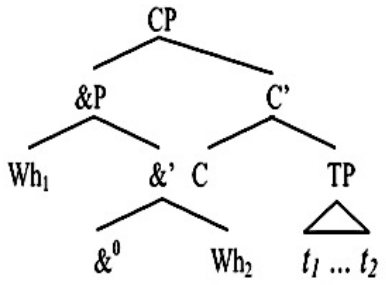
All authors agree on the point that there is no unique uniform structure for CMW and that structures vary not only cross-linguistically, but also within a single a language. Giannakidou and Merchant (1998) coin the term "reverse sluicing" as the analysis for coordination in English,

in which they assume wh-phrases in two clauses. In other words, there is an empty IP immediately after a wh-phrase that results from sluicing or IP ellipsis. Following Chung et al. (1995), they argue that this empty position undergoes IP-copy or "IP recycling" in the interpretive component of LF (the relevant example is below):

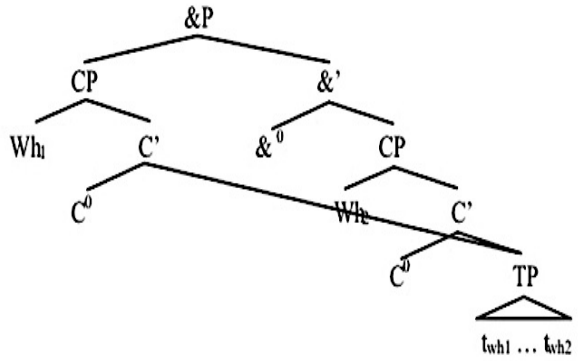
(88) The journalists want to know if [~~IP the suspect will make a statement t1~~] **and when**1
wh & wh
[IP suspect will make a statement t*i*]

Other bi-clausal structures are proposed by Citko & Gračanin-Yuksek (2010). Citko and Gračanin-Yuksek (CGY, 2010) claim that there are three types of CMW: one mono-clausal, and two bi-clausal, and all three can be present within a language (as in Romanian), while sometimes just one is sufficient to account for CMW (as in Bulgarian). Essentially CGY propose three separate structures for CMW (two of them are shown below). In (89), I reproduce a mono-clausal structure, in (90) - a bi-clausal structure with bulk sharing. In (91), I reproduce a structure of backwards (reverse) sluicing coordination.

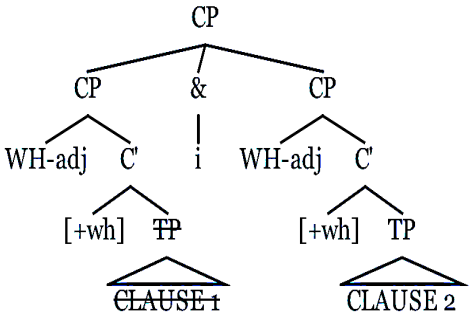
(89) Mono-clausal CMW (=(7)a, CGY 2010 p. 5)



(90) Bi-clausal CMW with bulk-sharing (=(7)b, CGY 2010, p.5)



(91) Backwards Sluicing (bi-clausal) (Bánréti 1992, Giannakidou & Merchant 1998, Camacho 2003)



CGY (2010) suggest that these structures are able to account for coordinated wh-behavior in various languages: English, Bulgarian, and SC among others. They also state that the properties of non-coordinated wh-behavior do not necessarily correspond to the properties of coordinated multiple wh-constructions. I will consider ((89)=(7)a) and (90)=(7)b)) whether this approach can apply to Russian. I also suggest that such structures might be unnecessary for Russian.

Following the idea that several structures of CMW can possibly be present within one language, I propose to consider each CMW Type separately not only typologically but also for the purposes of analysis. My analysis is based on various aspects of all of the CMW accounts: I take aspects of Kazenin's (2001) and Gribanova's (2009) analyses of Russian, CGY (2010)'s multi-language structures, and consider Merchant and Giannakidou's (1998) ideas of backwards (reverse) sluicing in coordination. Thus, there are several possibilities to be considered.

3.5.1 Backwards (Reverse) Sluicing: Kazenin 2001

Kazenin (2001) refutes the backward sluicing analysis for "some" Russian coordinated wh-constructions and proves that CMW in Russian is mono-clausal. His rejection of the backwards sluicing analysis comes from the fact that if one allows for a such a process to take place, one will end up with constructions that are impossible in Russian: an uninterpreted cataphoric antecedent, as shown in (92)-(94) resulting in a non-identical deleted and pronounced clause:

(92) [**Kogo** Petja izbil] i [za čto Petja ego$_i$/??pro$_i$ izbil]? (Kazenin, 2001)
whom Peter beat and for what Peter him$_{ACC}$ beat
"Who did Peter beat up and what for [did Peter beat him up]?"

Under sluicing where we can elide 'Petja izbil' ("Peter beat") on either clause, we will expect this to become (93) (if the phrase is elided in the first clause):

(93) [**Kogo** Petja izbil] i [za čto Petja egoi/??pro*i* izbil]? (Kazenin, 2001)
whom Peter beat and for what Peter him$_{ACC}$ beat

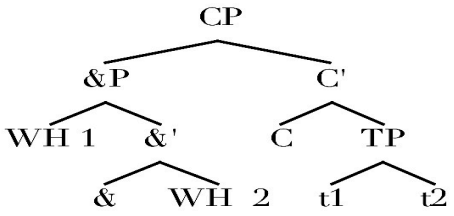
resulting in an ungrammatical (94):

(94) *[Kogo Δ] i [za čto **ego$_i$/??pro$_i$** Petja izbil]?

However, this problem only pertains when *arguments* are involved in the derivation. With two adjuncts, a backwards sluicing analysis remains possible. That is, backwards sluicing is generally available, but independently ruled out for arguments in the manner Kazenin indicates, that is for Type1 and Type 3, which accounts for the lack of free ordering in these Types.

Since a backwards sluicing analysis does not work work for Types 1 and 3 in Kazenin's (2001) view, (and the same problems have been noticed but resolved for English coordinated wh-adjuncts by Giannakidou and Merchant (1998)), he offered an alternative followed by Gribanova in (2009) (and essentially applied for Hungarian by Lipták (2009)) that uses a coordinated &P structure where wh-phrases are themselves coordinated as shown in (95):

(95) Mono-clausal &P for CMW (Kazenin 2001, Gribanova 2009)



The structure in (95) is the only derivation available for Type 1 and Type 3 CMW constructions, and these are the cases in which Superiority emerges. Why does (95) lead to Superiority effects? Notice that the structure in (95) is identical to the one in (89) proposed by CGY. Gribanova (2009), who uses this structure to derive CMW in Russian, does not provide the details of a derivation that would account for the emerging Superiority effects or the availability of PL readings.

3.5.2 Mono-clausal Coordination and Sideward Movement: CGY 2010

CGY provide a possible analysis that involves the mechanism of sideward movement (Nuňes 2001, 2004), as proposed in Zhang (2007, 2009). Sideward movement is different from standard "upward" movement in that it is the kind of movement that takes place across two distinct tree structures before they are combined to form a single structure. On such an analysis, each wh-phrase in a CMW question moves sidewards first, to become part of the coordinated phrase ConjP, which is subsequently merged with the original structure (the exact mechanics is outlined in (103).

CGY suggest three kinds of diagnostics to determine whether a CMW has a mono- or a bi-clausal structure: the parallelism in Superiority effects between CMWs and multiple wh-fronting without coordination, the grammaticality of mixed CMWs with obligatorily transitive verbs, and the possibility of conjoining two arguments. On the basis of these diagnostics CGY propose a mono-clausal structure with a ConjP in (89) as the analysis of Bulgarian CMW presenting us with the following Bulgarian data: consider (96) and (97) that show no difference in Superiority requirements between multiple wh-movement and CMW.

(96) a. **Koj** **koga** ste si hodi v Bulgaria? (Bulgarian)
who when will REFL go in Bulgaria [NOM>ADJ no coordination]
"Who is going to Bulgaria when?"

b. ***Koga** **koj** ste si hodi v Bulgaria? [*ADJ>NOM no coordination]
when who will REFL go in Bulgaria
*"When is who going to Bulgaria?"

(97) a. **Koj** i **koga** ste si hodi v Bulgaria? [NOM&ADJ]
who and when will REFL go in Bulgaria
"Who is going to Bulgaria and when?"

b. ***Koga** i **koj** ste si hodi v Bulgaria? [*ADJ&NOM]
when and who will REFL go in Bulgaria
*"When is who going to Bulgaria?"

The Superiority requirement is the same in Bulgarian multiple wh-constructions and CMW: it must hold. In some environments, though, both orders are allowed as in ((98) - (100)).

(98) a. **Kakvo** **koga** jade Ivan? (Bulgarian)
what$_{ACC}$ when ate Ivan
"What did Ivan eat when?"

b. **Koga** **kakvo** jade Ivan?
when what$_{ACC}$ ate Ivan
"What did Ivan eat when?"

(99) a. **Kakvo** i **koga** jade Ivan?
what$_{ACC}$ and when ate Ivan
"What did Ivan eat when?"

b. **Koga** i **kakvo** jade Ivan?
when and what$_{ACC}$ ate Ivan
"When did Ivan eat what?"

(100) **Koj** i **kakvo** e kupil?
who and what AUX bought
"Who bought what?"

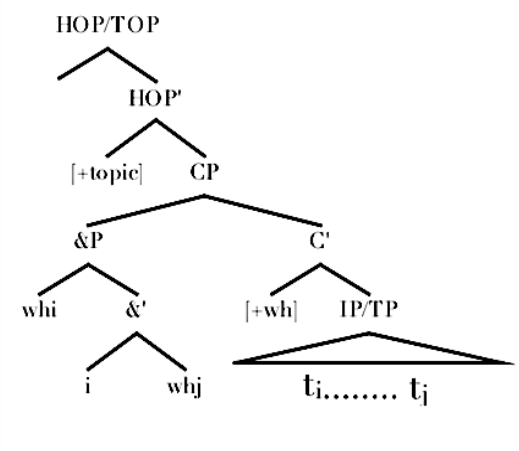
The conclusions that CGY draw from this set of Bulgarian data are that Bulgarian follows or lacks Superiority in CMWs exactly in the same environments as it would in multiple wh-fronting questions ((96) and (97), (100); that the mixed type of coordination is allowed with transitive verbs (98), (99), and that two arguments are allowed to be coordinated (100). They state that the parallelism of the Superiority requirements in multiple wh-constructions and CMW is relevant because "it provides strong evidence that CMW in Bulgarian are derived by a strategy that is at work in MWHs [*multiple wh-constructions*], namely, multiple wh-movement."(p.26). These three diagnostics allows them to use the proposed mono-clausal structure to account for CMW in Bulgarian.

3.5.3 Type 1 Analysis: Russian

If one looks at the facts of wh-coordination of arguments (Type 1), it becomes clear that they are identical to Bulgarian. When two arguments are coordinated, we observe the emergence of Superiority. Since I have claimed that in Russian true multiple wh-movement takes place and the Superiority Condition applies to Russian wh-phrase frontning in exactly the same manner as it does in multiple wh-constructions in Bulgarian, this behavior is not unexpected. The only reason multiple non-coordinated wh-fronting questions look as if Superiority does not hold in Russian is due to the availability of HOP located higher than CP. If the structure in (95) is adopted to analyze CMW in Russian then the emergence of Superiority in CMW reduces to the nature of ConjP[11]: whatever restrictions on Superiority in multiple wh-fronting exist in a language they continue to exist in CMW cases. Even the presence of HOP does not allow the escape hatch of the coordinated wh-phrases out of the coordinated phrase (ConjP) in light of the Coordinated Structure constraint. If the entire ConjP raises, Superiority is still observed. Crucially, wh_2 cannot reach HOP from within ConjP. The structure for Russian Type 1 is presented in (101) and the derivation of (95) shows a step-by-step derivation of ConjP in (103) and where I discuss the reason for it exhibiting Superiority:

[11] ConjP = &P (the difference is merely notational)

(101) Type 1 CMW Structure for Russian (HOP is not available for wh-reach out of &P)



According to Gribanova (2009), such a structure also captures the lack of PL readings, by assuming a strict structural locality condition on Quantifier Absorption at LF, an operation that takes adjacent unary quantifiers and returns a single quantifier (following Higginbotham and May 1981), under the assumption that in order to undergo Quantifier Absorption quantifiers must be structurally adjacent. The structural adjacency comes from the definition in (102):

(102) Structural adjacency: α and ß are structurally adjacent iff:
  a. α c-commands ß
  b. α c-commands no head that c-commands ß

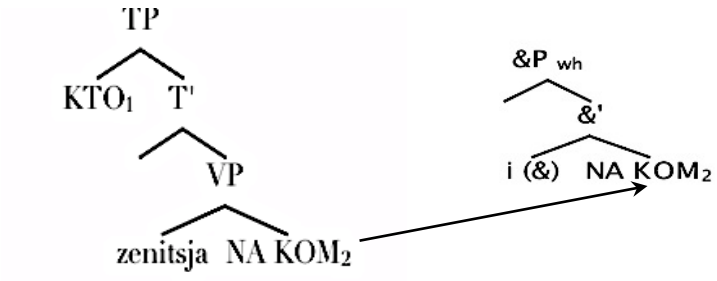
However, it is baffling that all the existing analyses of CMW in Russian are based on an erroneous assumption about CMW: that they do not yield PL readings. Quantifier Absorption is an additional mechanism that has to be superficially applied on the structure in order to limit it to just SP readings, otherwise the structure itself does not require this additional mechanism. Hence, since we have seen the PL readings are readily available, the structure in (101) (following the analysis in (89) and (95)) accounts for the facts of the Type 1 CMW.

Gribanova (2009) who proposes a similar mono-clausal structure (as in (95)) does not discuss, however, a crucial issue: how do the wh-arguments end up in the Coordination phrase - ConjP? CGY (2010) assume a sidewards movement where movement takes place in two separate trees before they are merged together following Nuňes (2001). In such movement each wh-phrase moves to &P (ConjP) structure from the original theta/case position of the original

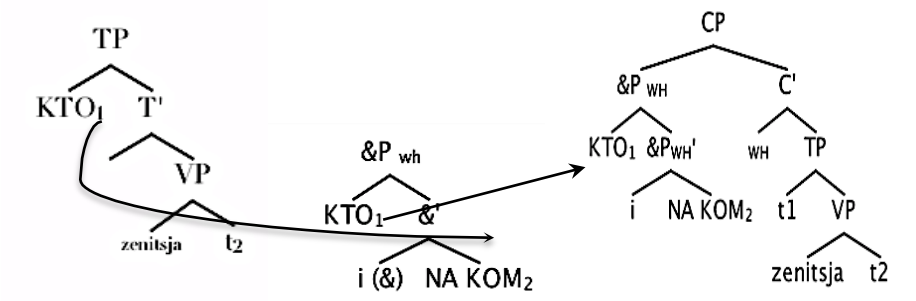
tree first before this ConjP structure is merged with the original tree. The process is triggered in a bottom-up fashion to satisfy the requirement for coordination as soon as the "matching" phrase enters the derivation. In a language where the building of a multiple wh-question and a coordinated multiple wh-questions are parallel, the mechanism should be parallel as well, in a sense that a single C head triggers movement (Gribanova 2009). A step-by-step building of such a structure is spelled out below in (103):

(103) **&P derivation and merge**

a. wh2 [VP to &P]



b. wh1 c. &P to CP



In addition to postulating sidewards and bottom-up movement, one needs a clear idea why the Law for Coordination of likes can be also violated when it comes to coordinating wh-phrases. CGY (2010: fn. ii) state "Thus, while the Law of Coordination of Likes can be obviated only through sideward movement […], this possibility is subject to independent constraints." The discussion of this is necessary to rule out the ungrammaticality of (104), for example.

(104) *Cheerfully and the watermelon John ate. (CGY, 2010)

It is not clear how just the mechanism of a "sideward" movement can obviate the requirement of Likes. Zhang (2009) notes that 'in order to check the uninterpretable feature of the attractor, the moved element may not contain any checking features with conflict values'. Wh-phrases possess theta-roles and case – both satisfied in the original derivation (before sideward movement takes place), but other than that they are feature unspecific except for the [+wh] feature itself. Thus, one can postulate a feature-specific restriction on ConjP phrase, and this should take care of the problem with respect to wh-argument coordination (and it should hold for coordination in general); the feature [wh] should mark a ConjP created separately (in a sideward space) resulting in $ConjP_{wh}$. Such a $ConjP_{wh}$ can only be moved into the Specifier of CP, which has the same feature. Thus, the case and theta roles of wh-arguments are checked in their original positions, and then the sideward movement into $ConjP_{wh}$ applies one argument at a time. Under such movement it can be expected that ordering of whs would come out of their position in the original tree before any movement and merging as outlined in a step-by-step derivation in (103). This structure and derivation of the coordinated wh-arguments should predict the occurrence of wh-arguments in coordinated phrase independently of their original cases (NOM, ACC, oblique) and of their theta roles, since these are checked in the original position. This structure also explains the parallelism in Superiority effects between multiple fronted wh-questions, and coordinated ones. The HOP phrase cannot be at play here (not a reachable spot for coordinated wh-arguments out of ConjP). As it has been shown by Kazenin the bi-clausal sluicing analysis is impossible in the case of Type 1 coordination because the elided IP-material (in Giannakidou & Merchant's terms) is non-symmetrical between the clauses under the assumption that clauses must be identical for elision.

3.5.4 Type 2 analysis

The same mono-clausal structure should be possible but might be difficult to postulate for the analysis of the Type 2 CMW with two adjuncts. It is not clear how such a structure can be derived from the original positions of adjuncts in the original tree[12]. Bošković (1998) notes that

[12] Additionally, while the ConjP analysis captures the facts of Type 2 coordination, it does not give an explanation for why adjunct wh-phrase fronting is only possible through coordination. The facts of impossibility of regular multiple wh-fronting of adjuncts was discussed in Chapter 2 (see the Multiple Wh-adjunct Ban). In Croatian, analogous to Russian sequences of wh-phrases that are allowed in multiple wh-fronting are different from those that are allowed in CMW, thus restricting non-coordinated wh-adjuncts from occurring.

as in Russian, in SC two adjuncts cannot both be fronted in regular questions but are perfectly fine in a CMW, as illustrated in (105):

(105) a. ***Gde** **kada** Ivan nastupa? (SC)
where when Ivan performs
"Where does Ivan perform when?"

b. **Gde** i **kada** Ivan nastupa? (SC)
where and when Ivan performs
"Where does Ivan perform when?"

If well-formedness of (105)b depended on the well-formedness of (105)a, then (105)b should not be possible to derive, contrary to the evidence. We observe the same phenomenon in wh-adjunct sequences in Russian: while multiple fronting of two wh-adjunct phrases is impossible, the coordinated structure of wh-adjuncts is welcome (106)a and (106)b repeating examples (79) and (80):

(106) a. *?**Gde** **kogda** vy vystupaete? (Russian)
where when you perform?
"Where do you perform when?"

b. **Gde** i **kogda** vy vystupaete?
where and when you perform
"Where do you perform and when?"

CGY state that either of the bi-clausal structures in (90) or (91) could capture these SC facts. I do agree that a bi-clausal analysis of this Type of CMW is possible. However, I suggest backwards sluicing can be the analysis for the coordination of Type 2 with two wh-adjuncts only. Such an analysis would reflect the fact that wh-phrases do raise to CP as they are supposed to in Russian. On the other hand, when it comes to coordinating adjuncts, the problem outlined by Kazenin of the cataphoric dependency becomes irrelevant since two identical clauses are available for ellipsis as shown in (107), where either clause can be elided. The Kazenin problem disappears, in fact, exactly in the case of Type 2 (two adjunct wh-phrases). That is because only in Type 2 the elided constituents are truly identical. In the case with wh-adjuncts the ellipsis happens to one of the identical parts.

(107) a. [$_{CP}$ **Gde**$_1$ [$_{VP}$ ~~vy vystupaete~~ t_1]] i [$_{CP}$ **kogda**$_2$ [$_{VP}$ vy vystupaete t_2]]?
where you perform and when you perform

b. [$_{CP}$ **Gde**$_1$ [$_{VP}$ vy vystupaete t_1]] i [$_{CP}$ **kogda**$_2$ [$_{VP}$ ~~vy vystupaete~~ t_2]]?

The structure in (91) reflects the derivation in (107). It is a potential analysis for Type 2 in several respects. First, it captures the fact that the coordination of wh-adjunct phrases can happen independently of multiple wh-fronting restrictions. Second, it also captures the fact that Superiority is irrelevant in these cases since adjuncts are generally inserted in the structure late and can appear in any order. It does not involve either of the bulk-sharing constructions proposed by CGY that need a lot of motivation for movement such as Multidominance accounts. Finally, it allows us to capture an additional set of coordination data (as in (107) for free where a coordinated wh-adjunct linearly appears clause-finally: here, of course, it is obvious that it appears clause initially and the rest of the clause undergoes ellipsis. Even though Gribanova claims that her analysis in general should work for "wh-final" coordination, it is not immediately conceivable how the structure in (95) can capture these facts. Thus, considering each type of CMW separately has undeniable advantages. On the other hand, the structure in (95) cannot truly be ruled out for adjunct coordination.

Thus far, I have provided accounts for Type 1 and Type 2 CMW where I captured the facts of emergence of Superiority. One of the approaches would be to analyze Type 3 together with Type 1, and that proves logical following either of the following accounts Gribanova's, Giannakidou-Merchant, or GCY ((91)=(7)b) accounts.

3.5.5 Type 3- Mixed Type Analysis

In the Mixed Type a wh-argument and a wh-adjunct are coordinated. The Superiority emerges in the contexts where a subject wh-argument is coordinated with a wh-adjunct only when wh-argument is a subject.

(108) a. **Kto** i **kogda** pocelujet Mariju?
who and when $kiss_{FUT}$ $Maria_{ACC}$

b. ***Kogda** i **kto** pocelujet Mariju?
when and who $kiss_{FUT}$ $Maria_{ACC}$
"Who will kiss Maria and when?"

At first sight there are two problems that arise with applying mono-clausal ConjP account to the mixed type of wh-coordination: First, Superiority effects are not parallel to the ones in Type 1 coordination: Superiority can sometimes be violated when a wh-argument and a wh-adjunct are coordinated (but only when a wh-argument is not Nominative), notice the existence in contrast between (108)a and (108)b vs. a lack thereof (109)a and (109)b.

(109) a. **Kogda** i **komu** Ivan prines dokumenty?
when and who_{DAT} $Ivan_{ACC}$ brought documents
"Who did Ivan bring the documents for and when?"
[ADJ>NOM: *n*(4)=30:85.7%; *n*(5)=14.3%]

b. **Komu** i **kogda** Ivan prines dokumenty?
who and when Ivan brought documents
"Who did Ivan bring the documents for and when?"
[NOM>ADJ: *n*(5)=33: 94.2%]

Applying a backwards (reverse) sluicing analysis is problematic for two reasons. First, if Superiority emerges in some contexts, then how such an analysis can account or predict the ordering of the wh-phrases. Secondly, the same problem with asymmetrical IP-ellipsis arises as discussed in Type 1 and shown by Kazenin (2001). When using a mono-clausal ConjP analysis, the problems outlined above can all be avoided. As it has been shown such an analysis should predict the emergence of Superiority effects (compare to the facts with non-coordinated wh-constructions where Superiority can be violated but after raising to CP utilizing HOP). The postulation of the additional feature onto the ConjP such as [wh] allows us to solve the problem with 'unlike' wh-phrases. Expanding Lebeaux's (1988) ideas, Chomsky (1993:37) analyzes the distinction between complements and adjuncts: "The extension property for substitution entails that complements can only be introduced cyclically, hence before wh-extraction, while adjuncts

can be introduced non-cyclically, hence adjoined to the wh-phrase after raising to [Spec,CP]". Such distinction explains several possibilities on how adjunct whs can enter a derivation outlined in (103) with one addition: adjuncts can enter the derivation late, and hence, will take the open left-over position in ConjP. Analogous to the derivation of wh-arguments, the wh-argument in the Mixed Type receives its case and/or theta role in its original position.

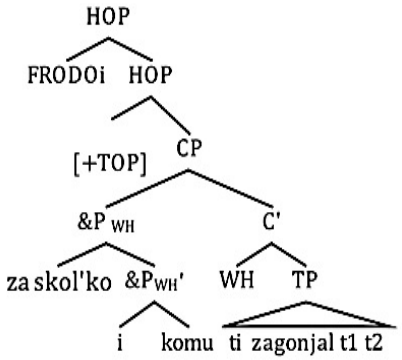
(110) Mixed Type Russian CMW (matrix clause)

a .	Frodo	**komu**	i	**za**	**skol'ko**	Kol'co	zagonjal?
	$\text{Frodo}_{\text{NOM}}$	who_{DAT}	and	for	how-much	Ring	tried-to-sell
b.	Frodo	**za**	**skol'ko**	i	**komu**	Kol'co	zagonjal?
	$\text{Frodo}_{\text{NOM}}$	for	how-much	and	who_{DAT}	Ring	tried-to-sell

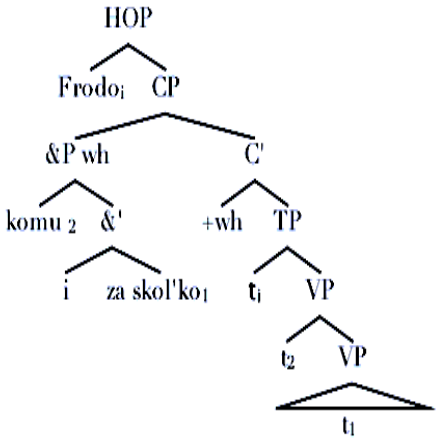
"Who was Frodo trying to sell the Ring and for how much?"

Questions as in (110) can be derived in the following manner: notice that HOP is occupied by "Frodo" here, but it is not what allows the ordering of wh-phrases as expected, but the derivation of the ConjP phrase merged into the original tree as shown above with the assumption of the possibility of late adjunct insertion.

(111) Mixed Type coordination structure



(112) Type 3 Structure with reverse wh-positions in the original derivation



The derivation of &P (ConjP) is the same as presented in (101), but the non-wh-element can reach out into HOP.

Thus, the three types of Russian CMW are analyzed with two possible structures: $ConjP_{wh}$ and backwards sluicing only in the case of Type 2 that by itself presents a somewhat interesting case. These two analyses reflect the main facts of coordinated wh-structures: Superiority effects and the necessity for coordination of adjunct wh-phrases. They predict the possibility of coordination of wh-phrases of different categories as well, unlike regular coordination that poses the restriction on coordination of "unlike" categories.

3.5.6 CMW Account Summary

In this section, I have given an account of CMW in Russian. The consideration of multiple wh coordination should not be seen as a uniform phenomenon, but as three different instances of coordination analyzed independently. Such a view allows me to postulate separate structures accounting for each of the Types and capture the facts that are outlined in Table 1: Superiority effects, presence of both SP and PL readings, and the differences in sequences of wh-constructions in regular multiple wh-fronting and CMW. It also allowed me to preserve the fact that Russian is a wh-movement language; and hence, some of the facts that other analyses did not capture: parallelism of the Superiority effects and the manner of wh-adjunct coordination, using the restrictions that both multiple wh-fronting posits and the restriction on coordinated wh-phrase distribution. While a grammar that allows different mechanisms to account for the same

phenomenon seems non-economical, the fact that these phenomena exhibit different properties calls for a proposal of separate accounts. Also, no new processes are proposed to handle the facts. Of course, more work in this area and advances in the syntactic field might suggest a more successful approach to CMW in Russian.

3.6 Back to Sluicing (not backwards): "Ironing out the 'wrinkles'"

The entire consideration of CMW in Russian grew out of the puzzle of coordination of wh-phrases at the sluicing site (as outlined in section 3.3). If one accepts the analysis of CMW in Russian for wh-constructions of various types, then the problematic sluicing facts discussed earlier fall out naturally. We should assume that in sluicing, when CMW are present, one of the structures argued for above would appear at the site of ellipsis (with the process of movement into ConjP space intact). Thus, when two coordinated wh-argument phrases are present, a ConjP structure at CP would be possible, and the ellipsis of IP/TP would be licensed by C^0. In the case of coordinated wh-adjuncts we should assume "double" sluicing might take place under the possible coordination as ellipsis analysis discussed above: C^0-licensed sluicing, and then, the additional backward sluicing to end up with wh-adjunct remnants, the only absolutely identical remnants. In such cases we should assume that the clause deletion of the first clause would be licensed by the sluicing processes (C_1 as in (113)) and that the "backward" sluicing" would only have the second clause (C_2 as in (113)) to "take care of".

(113) Gde-to kogda-to [ljudi peli pesni prosto tak], no nikto
somewhere some-time people sang songs simply so but nobody

ne pomnit [$_{CP1}$ **gde** [~~$_{TP}$ ludi peli pesni prosto tak~~] i
not remember where [people sang songs simply so] and

[$_{CP2}$ **kogda** [~~$_{TP}$ludi peli pesni prosto tak~~]
when [people sang songs simply so].

"Somewhere sometime (ago) people sang songs just because, but nobody remembers **where** [people sang songs just because] and **why** [people sang songs just because]".

Since HOP cannot participate in the process of coordinated wh-phrases for wh-phrases themselves, we would predict that Superiority would always hold where it is naturally expected (the Emergence of Superiority cases) by the virtue of entering the derivation and checking off necessary features (i.e. θ/case) before movement into ConjP if such procedure is necessary. The Superiority restrictions, then, are borne out of the general restrictions on Superiority in multiple

wh-fronting, or based on the fact of an adjunct appearing in the derivation. Thus, the accounts given for the CMW in Russian allow us to account for the coordinated sluicing constructions in Russian that presented a puzzle of contradictions at first sight.

3.7 Conclusion

In this chapter I have discussed Russian sluicing, specifically multiple sluicing, and provided an analysis of it. The phenomenon has been analyzed before, but some of the predictions made by earlier analyses fell short of covering all the variety of data offered in the language. The findings of chapter 2, claiming that true Bulgarian-style multiple-wh movement happens in Russian, allow us to reduce an analysis of sluicing to the standard analysis (Ross 1969, Merchant 2001). However, the multiple sluicing data offer an additional puzzle – coordinated multiple wh-phrases (CMW) in the sluice. While CMW has received active attention in recent years, the issue of coordinated sluicing was not addressed. The approach to coordinated multiple sluicing comes from the approaches to CMW in any given language. Russian as some other languages offers three types of multiple coordinated constructions: coordinated wh-arguments, coordinated wh-adjuncts, and the Mixed Type: coordinated wh-adjuncts and wh-arguments. The analysis of CMW is not trivial. Several approaches have been considered in the current chapter.

The final analysis suggests that these three different types should be analyzed separately: with the proposed earlier ConjP coordination at CP with the revised mechanism of how sideward movement takes place: an issue not clearly outlined in the analyses of Russian CMW; and with the utilization of the backwards sluicing approach for the coordination of multiple wh-adjuncts, where only coordinated multiple wh-fronting is possible, in order to capture this exact fact. The coordinated wh-phrases analysis proposed by Gribanova (2009) for Russian successfully works only with respect to coordination of two wh-arguments: Type 1 CMW. Her analysis, however, misses a few generalizations, such as the typology of coordination, and the presence of PL interpretation. Additionally, where wh-coordination is not parallel to wh-fronting possibilities as in the case with adjuncts, I have suggested using the second account: backwards sluicing, which allows for a bi-clausal structure to be present, and captures the facts true of this type of coordination. In other words, the consideration of different strategies of coordination separately allows us to account for a range of possibilities in Russian multiple coordinated wh-constructions, including sluicing, while sluicing itself is reduced to the Merchant style IP-

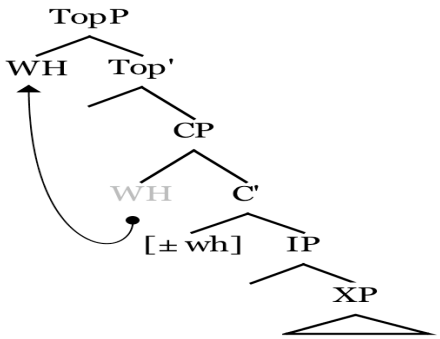
deletion process if wh-behavior is analyzed as wh-movement into CP with the additional functional category which is discussed in detail in chapter 2.

Chapter 4. Particle –TO: a HOP particle

4.0 Introduction

Below is the structure that I proposed as a layout of the left-periphery of the Russian main clause. The structure assumes a unique high HOP[±topic] position. This position is different from the TopP* suggested in Rizzi (1997), for instance, in that it is not iterative. Thus, its uniqueness is literal.

(1) HOP over CP



There were several pieces of evidence provided for this position. I showed asymmetries between matrix and embedded clauses: the unique HOP is not available in the embedded clause. This chapter provides additional syntactic evidence for the existence of such a HOP.

When any unique position is postulated, it is always attractive to have empirical evidence that the head of such position is filled. Of course, in many cases, the head is filled with a [+feature] of the respective phrase, i.e the interrogative head carries a Q feature, and the TP head carries a tense feature. However, in this particular case, it might be the case that unique HOP can be filled overtly. I refer here to the particle –TO that tends to appear clause initially attached to Topicalized elements. There is a question whether –TO is in fact in the head of HOP or is attached to the element it "modifies". Hence, there are a couple of questions that arise can be formulated in the following way:

Questions about –TO:

(i) which elements does the particle –TO occur on, and which ones can it not occur with, in other words what kinds of XPs can be –TO marked?

(ii) –TO appear in HOP[+topic] by movement or is it base-generated?
   a. if –TO appears in HOP[+topic] by movement, it presumably moves with the element it postposes
   b. if –TO is base generated, it presumably occupies the head of HOP[+topic] position, and hence, acts as a trigger for movement.

(iii) what are the restrictions on its distribution: i.e. occurrence of –TO *in-situ*, and does it follow island restrictions since either option (a) or (b) of (ii) assume movement?

(iv) does it appear in the embedded clauses? (Based on the findings of chapter 2, it does not).

The objective of this chapter is to provide answers to these questions. The chapter has the following structure: first, it examines the distribution of –TO. Then, it discusses the possibilities of how –TO occupies a certain position(s) in an utterance and for what reasons, what triggers its occurrence in these positions, as well as restrictions on its distribution. The restrictions are viewed in the environment of movement out of islands: what kinds of extractions are possible and which are not. This allows us to finalize the analysis of the place of this particle in the syntax of Russian.

4.1 Semantico – Syntactic Distribution of –TO

Discourse particles in different languages share certain properties. Bayer and Obenauer (2011), in an overview of German discourse particles, suggest the following list of properties that is commonly shared by discourse (cf. Biberauer, Holmberg & Roberts 2009, Sheehan 2009, Bailey 2010):

- Properties of Discourse Particles:

(i) Discourse particles are often adverbial in nature but show clearly distinct properties.

(ii) Unlike adverbs, discourse particles are usually immobile. They can neither be moved to designated landing sites (such as SpecCP) nor extraposed.

(iii) Discourse particles normally cannot bear stress.

(iv) Discourse particles mostly have counterparts to which they are historically related. They are the result of grammaticalization.

(v) Discourse particles are in their typical occurrences monosyllabic. German has *bloß, denn, doch, halt, ja, nur, schon, wohl*, but also "exceptions" such as *vielleicht, aber, eben, etwa*. The latter are historically younger and may not have been fully reanalyzed yet.

The same is true for Russian: *-TO*, *ŽE*, *VED'*, *-KA*, vs. *tol'ko* (only*), prjamo/prjam* (an emphatic particle from the adverb "straight/straight away"), for instance.

(vi) Discourse particles are modificational and as such appear to be "optional"[1].

(vii) Due to their sensitivity to sentence types and their impact on illocutionary force, discourse particles are generally confined to root-clauses. There are, of course, exceptions, and this will be mentioned below.

In this chapter, I show that the Russian particle "-TO" has undergone grammaticalization. It is fairly apparent from the diachronic studies and the studies of contemporary Russian dialects that this particle has a historical relation to the demonstrative pronoun "TO" ('that'[2]) (Nikolaeva, 1985). I show how it has grammaticalized into several separate homonymous particles as a part of the indefinite pronouns discussed below and the discourse particle in question. It is immobile in that it cannot be moved, but has a specific place in the structure confined in the matrix clause. In Russian, this particle[3] looks like an unstressed clitic: it attaches to the word (or a number of words) it modifies and is monosyllabic and unstressed. It is pronounced together with the item it attaches to.[4] I claim a constituent that appears with -TO to be a Topic. I have provided a discussion of what it means to be a Topic in Chapter 1. It is rather inconsequential what definition of Topics we assume here: several definitions of Topics have been suggested. Here is just a brief overview of the basics, taken from Molnár (2002):

[1] Bayer and Obenauer (2011) suggest that one should be careful in claiming that discourse particles are optional, there are certain exceptions to the case where they are in fact mandatory (at least in the case of German) (see their discussion of this issue. Bailey et al. (2009) make a similar claim.

[2] "that" here means *–over there*, as an opposite to this, similar to the French *voilà.* As in: *Prinesi mne TO* (*Bring me that*).

[3] There are other homophonous particles in Russian as well, for instance the one that marks the indefinite pronoun as in *kto-**to*** – 'somebody'. I believe their distribution to differ. While this usage is relevant here and will be discussed below, other irrelevant usages of the homophonous particles remain beyond the scope of this investigation.

[4] In spelling the attachment is indicated by a hyphen. In the rest of the chapter, I continue using the hyphen before the particle to maintain this rule in order to indicate the place of attachment

> As to topic, three definitions dominate the linguistic landscape. Firstly those which define topic as the notion of aboutness (Reinhart 1982) or as an "address pointer", also called "link" by Vallduvì (1992). According to another influential view, the topic should be regarded as the notion of frame: "the topic sets a spatial, temporal, or individual framework within which the main predication holds" (Chafe 1976: 51). A third definition of topic is given with recourse to old information: the topic is either identified as given information or in weaker versions, the "givenness condition" is an important part of topicality (p.100).

The understanding of essence of Topic is relevant here merely because when particle –TO attaches to an item in an utterance, it is undoubtedly serves the function of Topic or a Contrastive Topic, and marks constituents as Topicalized[5] (these ideas are expressed by Marshall 2002, McCoy 2001, 2002, among others.)

4.2 Particle –TO Distribution

In this section I consider the distribution of the particle –TO with arguments: I demonstrate where in a sentence it occurs, which elements can be –TO-picalized, and show whether there are any restrictions on its occurrence with subjects, direct, indirect and oblique objects.

4.2.1 Arguments [NP]+TO

4.2.1.1 Subject NP

The example below demonstrates the occurrence of –TO attached to an argument in the subject position.

(2) Maxim ljubit pivo.
Maxim likes beer
"Maxim likes beer."

(3) Maxim -to ljubit pivo.
Max_{NOM} TO likes beer
"As for Maxim, he likes beer."

[5] It is important to note that anything that is marked with –TO can be elicited from previous discourse, and that does not necessarily mean 'immediate' discourse' as it is commonly assumed for D-linking; it might refer to the common knowledge[5] of the interlocutors. In section 4.2 I present data of –TO distribution.

(4) Max -to čto ljubit?
Max TO what likes
"As for Max, what does he like?"

(5) Max gde rabotaet?
Max where works
"As for Max, where does he work?"

The analysis for examples such as (6) without TO was discussed in chapter 2.

(6) Max -to gde rabotaet?
Max TO where works

This question only means that out of the set of people the person is asking about s/he is only interested in the place of work of Maxim.

(7) [Maria] -to ne vyxodit iz doma uže god.
Maria TO not comes out house already year
"As for Maria, she hasn't left the house for a year already".

(8) a. Uže god Maria ne vyxodit iz doma.
Already year Maria not leaves out house

b. *Uže god [Maria] -to ne vyxodit iz doma.
Already year Maria TO not leaves out house
"As for Maria, she hasn't left the house for a year".

Notice, that in (8) a temporal adjunct can appear before Maria, if Maria is not topicalized, however, when it is –TO marked, the occurrence of the adjunct in the high position is impossible.

(9) a. [Maria] -to skol'ko vremeni ne vyxodit iz doma?
Maria TO how-much time not comes out house
"How long hasn't Maria left her house for (tell me about Maria)?"

b. Skol'ko vremeni [Maria] ne vyxodit iz doma?
how-much time Maria not comes out house
"How long hasn't Maria left her house for?"

c. *Skol'ko vremeni [Maria] -to ne vyxodit iz doma?
how-much time Maria TO not comes out house
"As for Maria, how long hasn't she left the house for?"

Examples ((9)a-c) demonstrate that TO-marked NP subjects require to be fronted, and when they appear in questions with wh-words, they occur before a wh-element.

On the other hand, the fronting of other elements that are not – TO-marked when a TO-marked element is present is ungrammatical, as in example (10) where (a) shows the neutral word order, and (b) –TO marking on the subject.

(10) a. Max ljubit pivo.
Max likes beer
"Max likes beer"

b. *Pivo Max -to ljubit t_i.
beer Max TO likes
"As for Max, beer he likes."
*"As for beer, Max likes it"

Notice, that shifting –TO to the object NP would repair the grammaticality, but would also change the meaning of the sentence as shown in (c):

c. Pivo -to Max ljubit t_i.
$beer_{ACC}$ TO Max likes
"As for beer, Max likes it."

((10)c) can easily be an answer to the question about Maxim's alcoholic preferences for beer, but not a question about Maxim (as ungrammaticality of one of the (10)b) glosses shows): as an answer to the question: "Does Max like beer?". Here, of course, 'pivo' ("beer") is the object of

the verb 'ljubit'("like"), and not the subject of the sentence. This is shown in (11) as a variant of (10), where –TO is *in-situ* and is ungrammatical:

(11) *Max lujbit pivo -to.
Max likes beer TO
"As for beer, Max likes it."

4.2.1.2 Object NP

In this subsection, I show how –TO marks direct, indirect, and oblique objects.

When a direct object NP is marked with –TO, it has to appear at the left edge of the clause in order for the sentence to be grammatical (as in (12)). If the –TO marked object is left *in-situ* the sentence is ungrammatical (12). If the –TO- marked object does not appear at the very edge of the clause, the sentence is still ungrammatical (12). The occurrence of two -TO-marked elements results in ungrammaticality, even though the –TO marked object is at the very left edge as in (12).

(12) a. Oni dolgo vypolnjali zadanie.
they longtime worked$_{IMP}$ assignment
"They were working on the assignment for a long time."

b. [Zadanie] -to oni dolgo vypolnjali.
assignment TO they longtime worked$_{IMP}$
"As for the assignment, they were working on it for a long time."

c. *Oni dolgo vypolnjali zadanie -to.
they longtime worked$_{IMP}$ assignment TO

d. *Oni zadanie -to dolgo vypolnjali.
they assignment TO longtime work$_{IMP}$

e. *Zadanie -to oni -to dolgo vypolnjali.
assignment TO they TO longtime worked$_{IMP}$

(12) shows that the typical post-verbal direct object of (12), when TO-marked, appears in initial position (12). Other orders ((12)c-d) are ungrammatical. (12) shows that two elements cannot be TO-marked in the same sentence. In sentences without an overt subjects, the behavior of – TO-marked objects is the same as above: the -TO-marked element should appear at the front of the clause, otherwise, the sentence is ungrammatical as in shown in contrasts below (13) - (16).

(13) Otpusk -to kogda berete t_i?
Vacation$_{ACC}$ TO when take
"When do you plan to take your vacation?"

(14) *?Kogda berete otpusk -to?
when take vacation$_{ACC}$ TO
"As for vacation, when do you plan to take it?"

(15) Bilety -to na ètu prem'eru ešče vozmožno kupit'?
Tickets$_{ACC}$ TO for this premier still possible buy$_{INF}$
"Are there still tickets for this premiere?"

(16) *Na ètu prem'eru eščo vozmožno kupit' bilety -to?
for this premier still possible buy$_{INF}$ tickets$_{ACC}$ TO
"Are there still tickets for this premiere?"

When an oblique object NP is marked with –TO it also must appear at the left edge of the clause. Observe the contrast between (18) and (19), and (22) and (23) respectively.

(17) Obez'jany bojatsja snega.
monkeys afraid snow$_{GEN}$
"Monkeys are afraid of the snow."

(18) Snega -to obez'jany bojatsja.
Snow$_{GEN}$ TO monkeys afraid
"As for snow, monkeys are afraid of it."

(19) *Obez'jany bojatsja snega -to.
monkeys afraid snow$_{GEN}$ TO
"As for snow, monkeys are afraid of it."

(20) *Snega -to obez'jany -to bojatsja.
snow$_{GEN}$ TO monkeys$_{N}$ TO afraid
"As for snow, monkeys are afraid of it."

(21) Vse volšebniki vosxiščajutsja ved'moj iz-za ee masterstva.
all wizards admire witch$_{INSTR}$ for her mastery
"All wizards admire the witch for her mastery."

(22) *Vse volšebniki vosxiščajutsja ved'moi -to iz-za ee masterstva.
all wizards admire witch$_{INSTR}$ TO for her mastery
"As for the witch, all wizards admire her for her mastery."

(23) Ved'moj –to vse volšebniki vosxiščajutsja iz-za ee masterstva.
witch$_{INSTR}$ TO all wizards admire for her mastery
"As for this witch, all wizards admire her for her mastery."

(24) *Vse volšebniki ved'moi -to vosxiščajutsja iz-za ee masterstva.
all wizards witch$_{INSTR}$ TO admire for her mastery
"As for the witch, all wizards admire her for her mastery."

(25) *Ved'moj –to vse volšebniki -to vosxiščajutsja iz-za ee masterstva.
witch$_{INSTR}$ TO all wizards TO admire for her mastery
"As for this witch, all wizards admire her for her mastery."

As with direct objects, when oblique objects are –TO marked, the sentence is ungrammatical when a –TO-marked objects is fronted but not to the very left edge of the sentence (as in (24)); and doubling of the –TO particle is impossible as in (25).

So far, we have observed that the particle –TO occurs at the very left edge of the sentence and attaches to the element it marks ("-TO marking"), and it is unique, there is only one – TO allowed per clause, and its occurrence *in-situ* with NPs in subject and object positions is restricted. In the following section, I will take a closer look in its distribution with adjuncts.

4.2.2 Adjunct + TO

4.2.2.1 PP + TO

In this section I show the data of –TO marked PPs.[6] So far we have seen this particle occurring on various NPs. Should we expect that this particle would as freely occur with prepositional

[6] Many examples given here with –TO require a specific dialog to precede them to be grammatical, in other words they require a certain context for –TO to appear on the elements. The usage of –TO is in fact, much more productive

phrases (PPs)? Compare an example containing a PP (a), and PP-TO (b), as well as PP-TO *in-situ* (c) constructions. Notice, that (c) example is ungrammatical, as expected.

(26) a. On ne ustaet mne povtorjat' o Rasputine.
he Neg tires me repeat$_{INF}$ about Rasputin
"He doesn't get tired of talking to me about Rasputin."

b. [O Rasputine] -to on mne ne ustaet povtorjat'.
about Rasputin TO he me not tired repeat$_{INF}$
"As for Rasputin, he doesn't get tired of talking about him to me."

c. *On ne ustaet mne povtorjat' [o Rasputine] -to.
he Neg tires me repeat$_{INF}$ about Rasputin TO
"He doesn't get tired of talking to me about Rasputin."

The particle –TO can occur with locative PPs as well, but again only on the left edge of the clause.

(27) a. Ja včera ne byl v magazine.
I yesterday Neg was in store
"I wasn't in the store yesterday."

b. [V magazine] -to ja včera ne byl.
in store$_{PPREP}$ TO I yesterday Neg was
"As for the store, I wasn't there yesterday."
(context: that is why my fridge is completely empty.)

c. *Ja včera ne byl v magazine -to.
I yesterday Neg was in store TO
"I wasn't in the store yesterday."

---

in conversational Russian, and in Southern dialects. In Southern dialects of Russian bordering with Ukraine or on contemporary Ukrainian territories (Crimea), it is commonly substituted by "-taki", and such utterances might even sound awkward for a Russian ear. That does not guarantee that particle –taki has the same distribution, but at first sight it appears very similar in many respects.

(28) a. My budem pit' pivo v parke.
we will drink beer in park
"We will drink beer in the park."

b. [V parke] -to my budem pit' pivo.
in park TO we will drink beer
"As for the park, we'll drink beer there."

c. *My budem pit' pivo [v parke] -to.
we will drink beer in park TO
"As for the park, we'll drink beer there."

– TO can also occur with modificational and instrumental prepositional phrases, as shown in the examples below.

(29) ?[S trudom] -to my ego ugovorili provesti u nas ostatok večera.
with difficulty TO we him convinced spend at us remaining evening
"It is with difficulty that we convinced him to spend the rest of the evening with us."

(30) *My ego ugovorili provesti u nas ostatok večera [s trudom] -to.
we him convinced spend at us remaining evening with difficult TO

(31) *My ego ugovorili [s trudom] -to provesti u nas ostatok večera.
we him convinced with difficulty TO spend at us remaining evening
"It is with difficulty that we convinced him to spend the rest of the evening with us."

(32) [Ne bez truda]-to my ugovorili ego nam pomoč'.
Neg without work TO we convinced him us help
"It is with hard effort that we convinced him to help us."

(33) [S mečom] -to každyj durak budet geroem.
with sword TO every fool be_{FUT} $hero_{INST}$
"It is with a sword that every idiot can be a hero."

(34) *Každyj durak [s mečom]-to budet geroem.
every fool with sword TO be_{fut} $hero_{INST}$

(35) *?Každyj durak budet geroem [s mečom] -to.
every fool be$_{FUT}$ hero$_{INST}$ with sword TO

(36) [Bez nožnic] -to kak prikažeš' èto razrezat'?
without scissors TO how order this cut
"How would you like me to cut it without any scissors?"

(37) Oni bez truda rasslyšali zvuki tankov v tišine.
they without work heard sounds tanks in silence
"They heard the sounds of the tanks in the silence without difficulty."

(38) [V tišine] -to oni bez truda rasslyšali zvuki tankov.
in silence TO they without work heard sounds tanks$_{GEN}$
"In the silence that they heard the sounds of the tanks without difficulty."

(39) *Oni rasslyšali [v tišine] -to zvuki tankov bez truda.
they heard in silence TO sounds tanks without work

(40) *Oni bez truda rasslyšali zvuki tankov [v tišine] -to.
they without work heard sounds tanks in silence TO
"They heard the sounds of the tanks in the silence without difficulty."

4.2.2.2 Adverbs and Temporal PPs

In Russian adverbs can occupy several positions (as shown in ((41)a-c)); however, when an adverb is marked with – TO, its occurrence on the left-edge is mandatory.

(41) a. *My xodim v sportivnyi zal [po utram] -to.
we go to sport hall in mornings TO
"In the mornings, we go to the gym."

b. [Po utram] -to my xodim v sportivnyi zal.
in mornings TO we go to sport hall
In the mornings, we go the gym."

c. *My xodim [po utram] -to v sportivnyi zal.
we go in mornings TO to sport hall
"In the mornings, we go the gym."

(42) a. Zavtra -to im nezačem prixodit'.
Tomorrow TO them nothing-for come$_{INF}$
"As for tomorrow, there's no reason to come."

b. *Im nezačem prixodit' zavtra -to
them nothing-for come$_{INF.}$ tomorrow TO

c. *Im zavtra -to nezačem prixodit'.
them tomorrow TO nothing-for come$_{INF}$
"There's no reason for them to come tomorrow."

(43) a. [V voskresen'e] -to oni guljat' ne xodjat.
on Sunday TO they walk$_{INF}$ not go
"As for Sundays, they don't go for walks."

b. *Oni guljat' ne xodjat [v voskresen'e] -to.
they walk$_{INF}$ not go on Sunday TO
"As for Sundays, the don't go for walks."

The data in this section show that the particle –TO marks various adverbial expressions including those expressed by prepositional phrases: locative, modicative, temporal, causal, and instrumental. All cases in which the particle –TO occurs anywhere but on the left edge of the sentence are ungrammatical. In Russian adverbs can occupy several positions; however, when an adverb is marked with – TO, its occurrence on the left-edge is mandatory. This is because of the association of the particle with the left-edge HOP category.

4.2.2.3 [NP [PP]] +TO = [NP-to PP] or [NP PP]-to?

Now let us consider what happens when there is a PP inside an object NP as it is shown in (44) and (45) respectively. Example (a) shows a construction that is not –TO-marked, and not topicalized, (b) shows the attachment of –TO *–in-situ* to the entire constituent, and (c) the attachment of –TO *in-situ* to the N-head of the NP. Both (b) and (c) are ungrammatical.

(44) a. Studenty dolgo vypolnjali [$_{NP}$zadanie [$_{PP}$po matematike]].
students longtime worked$_{IMP}$ assignment in math
"Students worked on a math assignment for long time."

b. *Studenty dolgo vypolnjali [$_{NP}$zadanie [$_{PP}$po matematike]] -to.
students longtime worked assignment in math TO
"As for the math assignment, students were working on it for a long time."

c. *Studenty dolgo vypolnjali [$_{NP}$zadanie -to [$_{PP}$po matematike]].
students longtime worked$_{IMP}$ assignment TO in math

d. *[$_{NP}$Zadanie -to [$_{PP}$po matematike]] studenty vypolnjali dolgo.
assignment TO in math students worked$_{IMP}$ longtime
"As for the math assignment, students worked on it for a long time."

e. [$_{NP}$Zadanie [$_{PP}$ po matematike]] -to studenty vypolnjali dolgo.
assignment in math TO students worked$_{IMP}$ longtime
"As for the math assignment, students worked on it for a long time."

Notice that, movement of the entire constituent allows the attachment of –TO only to the entire phrase (44), and not to the N-head only (44). Since we have postulated the uniqueness of – TO occurrence, we would also expect to rule out the [NP-TO PP-TO] combination. This is shown in (f) below.

f. *[$_{NP}$Zadanie -to [$_{PP}$po matematike]] -to studenty vypolnjali dolgo.
assignment TO in math TO students worked$_{IMP}$ longtime
"As for the math problem, the students worked on it for a long time."

Of course, such a distribution would make semantic sense as well: -TO marking on both elements would result in a topicalization tautology.

Example (45) below gives an additional example of a PP within an NP constituent: in this case a PP represents an inalienable possession "*dlinnye volosy*" ('long hair').

(45) a. Lena poljubila [$_{NP}$junošu [$_{PP}$s dlinnymi volosami]] s pervogo
Lena fell-in-love guy$_{ACC}$ with long hair with first
vzgljada.
sight
"Lena fell in love with a guy with long hair at first sight."

b. *Lena poljubila [$_{NP}$junošu [$_{PP}$s dlinnymi volosami]] -to s pervogo
Lena fell-in-love guy$_{ACC}$ with long hair TO with first
vzgljada.
sight
"As for the guy with long hair, Lena fell in love with him at first sight."

c. *[$_{NP}$Junošu -to[$_{PP}$s dlinnymi volosami]] Lena poljubila s pervogo
guy$_{ACC}$ TO with long hair Lena fell-in-love with first
vzgljada.
sight

d. [$_{NP}$Junošu [$_{PP}$s dlinnymi volosami]]-to Lena poljubila s pervogo
guy$_{ACC}$ with long hair TO Lena fell-in-love with first
vzgljada.
sight
"As for the guy with long hair, Lena fell in love with him at first sight."

Both examples (44) and (45) show that when there's a PP inside NP, the entire NP – the full constituent - must be what is -TO-marked, otherwise the sentence is ungrammatical. Of course, one more logical possible occurrence of –TO in this case might be on the adjective as in (e) below.

e. *[7][$_{NP}$Junošu[$_{PP}$ s dlinnymi-to volosami]] Lena poljubila s pervogo
guy$_{ACC}$ with long TO hair Lena fell-in-love with first
vzgljada.
sight
"As for the long-haired guy, Lena fell in love with him at first sight."

Even though (e) is ungrammatical, notice the difference in glosses of (e) and (b-d). –TO-marking shifts the meaning of the sentence, but occurrence of the –TO-marked element away from the edge of clause results in ungrammaticality.

[7] (45) is possible only under the contrastive reading and will be discussed in the later part of the chapter, but is not possible as is in the context of discussion provided here.

From the observations above, we can conclude that –TO attaches to various XPs: NP-subjects, direct and indirect objects, different prepositional phrases, and direct objects with prepositional phrases inside them, if and only if they act as a constituent and are moved to the left edge of the sentence.

4.2.2.4 AP + TO

In this section, I consider the distribution of TO with adjective phrases. Consider the following examples with predicate adjective phrases first:

(46) a. On nikogda ne byl dovol'nym.
he never not was $content_{INSTR}$
"He was never content."

b. *?On nikogda ne byl dovol'nym -to.
he never not was $content_{INSTR}$ TO
"As for funny, he never had that quality."

c. Dovol'nym -to on nikogda ne byl $t_{i.}$.
content TO he never not was
"As for funny, he never had that quality."

d. *Dovol'nym -to on-to nikogda ne byl $t_{i.}$
content TO he TO never not was
"As for funny, he never had that quality."

As before, when the particle TO attaches to an AP, the adjective phrase must appear on the left edge of the sentence (as the contrast between (46) and (46) shows), and doubling of the particle is impossible (as shown in (46).

Example (47) demonstrates attachment of the particle –TO to an attributive adjective phrase.

(47) a. Maria polučila [prekrasnuju rol']
Maria got [wonderful part]$_{ACC}$
"Maria got a wonderful part (in a play)."

b. [Prekrasnuju rol'] -to Maria polučila.
[wonderful part]$_{ACC}$ TO Maria got
"As for that wonderful role, Maria got it."

c. *[Prekrasnuju-to rol'] Maria polučila t_i.
[wonderful TO part] Maria got
"As for wonderful, Maria got that role."

(47) shows the neutral word order with no –TO –marking. (47) shown – TO-marking of the entire NO containing the attributive adjective, where the –TO-particle occurs after the entire constituent. (47) shows –TO marking on the adjective when the adjective and the noun are split by the particle, and that split of the constituent results in ungrammaticality.[8] When it comes to –TO marking on predicative and attributive adjective phrases, the –TO marked AP must appear at the edge of the clause. Double occurrence of the particle is impossible as shown in (46). The particle should appear after the entire constituent, the split of the phrase results in ungrammaticality as shown in (47).

To summarize the data above we can state that -TO attaches to any XP constituent (subjects, objects, prepositional phrases, and adjectives). When –TO-marking occurs, the –TO marked element appears at the left edge of the sentence with –TO following it, and adding the meaning of relation to discourse' or 'belonging to the set' to the sentence, i.e. 'as for'.

4.2.3 VP+TO

4.2.3.1 Infinitive Verbs -TO

How does –TO-marking work with VPs? Infinitive verbs in Russian can front, and therefore, we can expect that –TO can attach to infinitival verbs (V_{INFIN} /VP_{INFIN}). Example (48) shows –TO-marked V_{INF} *in-situ*, and is ungrammatical, and (48) shows V_{INF}-TO at the left edge of the clause.

[8] The sentence in (47) can be grammatical, if the opposing clause is present as demonstrated in the example below.
(i) [Prekrasnuju -to rol'] Maria polučila, a deneg ne zarabotala.
wonderful TO part Maria got but money NEG earned

(48) a. Ivan ljubit rasskazyvat'.
Ivan likes tell$_{INFIN}$
"Ivan likes to tell (stories)."

b. *Ivan ljubit rasskazyvat' -to.
Ivan likes tell$_{INFIN}$ TO
"As for story-telling, Ivan likes it."

c. Rasskazyvat' -to Ivan ljubit.
tell$_{INF/IN}$ TO Ivan likes
"As for story-telling, Ivan likes it."

Example (49) shows VP_{INFIN} with the –TO particle. Notice, that the splitting of the VP constituent as in (49) is ungrammatical.

(49) a. Danik obožaet [$_{VP}$smotret' fil'my].
Danik loves watch$_{INF}$ films

b. *[Smotret' -to fil'my] Danik obožaet $t_{i.}$
watch TO films Danik loves
"As for watching films, Danik loves it."

c. [Smotret' fil'my] -to Danik obožaet $t_{i.}$.
watch films TO Danik loves
"As for watching films, Danik loves it."

Again, in the case when -TO attaches to $V_{INFIN,}$ or to [$_{VP}$ V_{INFIN} NP], the –TO marked constituent must appear at the left–edge of the clause, and the constituent cannot be split by the particle.

In summary, in section 4.2 we have seen that –TO attaches to any constituent XP: NP subjects and objects, various PPs; APs, and infinitival VPs. –TO marking results in topcalization of the entire constituent, the split of the constituent results in ungrammaticality, and a –TO-marked constituent must appear at the very left edge of the matrix clause. Doubling of the particle within one clause is impossible, its occurrence in the clause is unique. Thus, we can formulate the following –TO Generalization:

(50) ***–TO Generalization***

–TO is generated in the head of HOP and attracts to the Spec of HOP a topicalized constituent

Bayer and Obenauer (2011) claim that there is "controversy about the X-bar status of discourse particles, cf. Meibauer (1994)." Some researchers view them as adverbs and therefore, they are of some impoverished kind (see Cardinaletti 2007, Coniglio 2007 for discussion). Bayer (1996) and Bayer and Obenauer (2011) argue for their head status.

If the particle occupies a head position then how does a relevant constituent (the one that a particle would modify) get there? I assume that the topicalized constituents discussed here are not base-generated and cannot be referred to as left dislocated hanging topics (HTLD), and hence movement should take place. Note that in hanging topical constructions or base generated constructions, the element the Topic –marked element requires a resumptive pronoun, particularly within islands, whereas when a Topic-marked element is moved the resumptive pronoun is impossible since there is trace left as a result of movement. Consider the following examples from English.

(51) a. Syntax, Deidra likes. (English)
b. *Syntax, Deidra likes it.
c. *As for syntax, Deidra likes.
d. As for syntax, Deidra likes it.

(52) a. *Sintaksis Dina ego ljubit. (Russian)
syntax Dina $it_{SG\text{-}ACC}$ likes

b. Sintaksis Dina ljubit.
syntax Dina likes

c. Sintaksis-to Dina ljubit.
syntax TO Dina likes

There is a contrast between (a) and (b), and (c) and (d) respectively with respect to how the sentences are derived. In the following section, I consider the properties of movement to the [+topic] position, and claim that they are similar to (a) topicalization rather than to (c) topicalization.

4.3 Evidence for Movement; restrictions on movement

The discussion above has outlined several facts about this particle: it serves certain discourse necessities of the language - topicalization. It is used in a colloquial register to emphasize the "givenness" of a certain item of the utterance. The position of the particle with elements on the

left-edge of the clause with most XPs, makes it likely to assume that –TO is located in the HOP domain, and that it moves there. We can entertain two possibilities here, one is that –TO moves to Spec of HOP together with the element it is attached to, and the second one is that –TO is located in the head of HOP and the constituent moves there (formulated as question (i)). How do we exactly divorce these two approaches?

We have seen that –TO attaches to an entire constituent, otherwise the sentence is ungrammatical, thus, we can state that TO does in fact occupy the head of the HOP position, and the constituent ends up there via movement. If this is true, then such movement should observe movement restrictions. Let us consider whether –TO-marked elements are sensitive to island conditions. Example (53) shows that movement out of a wh-relative clause in Russian causes ungammaticality:

(53) *Čto ty znaeš' ved'mu, kotoraja ljubit?
what you know witch$_{ACC}$ which likes
"*What do you know a witch that likes?"

As one would expect, the –TO marked object shown in (52) that occurs at the left edge of the matrix clause is impossible when it is moved out of a relative clause (compare (53) to (51)).

(54) Moroženoe -to ved'ma ljubit.
ice-cream$_{ACC}$ TO witch likes
"As for ice-cream, the witch likes it."

(55) *Moroženoe -to, ja znaju ved'mu, kotoraja ljubit t_i.
ice-cream TO I know witch which likes
"As for ice-cream, I know a witch that likes [it.]"

We have established that a –TO-marked constituent left *in-situ* results in ungrammaticality; therefore, there is no way to get out of an island to the left periphery (as in (56)).

(56) *Ja znaju ved'mu, kotoraja ljubit moroženoe -to.
I know witch which likes ice-cream TO
"As for the ice-cream, I know which witch likes it."

When extraction is impossible by itself, then – TO-marking is impossible; the element cannot reach the position in the high left-periphery it is targeting. Example (55) shows an adjunct island, and (56) shows the –TO marked object moved out of the adjunct island, both are expectedly bad. This is because even without the –TO-marking the extraction is impossible (as in (57)). Hence, the object cannot move to the HOP out of the island.

(57) *Kogo$_i$ Victor ubežal, potomu čto ljubit t_i?
who$_{ACC}$ Victor ran-off because loves
"*Who did Victor run off because he loves?"

(58) Victor ubežal, potomu čto ljubit Mašu.
Victor ran-off because loves Masha$_{ACC}$
"Victor ran off because he loves Masha".

(59) *Mašu$_i$ Victor ubežal, potomu čto ljubit t_i.
Masha$_{ACC}$ Victor ran-off because loves
"Victor ran off because he loves Masha."

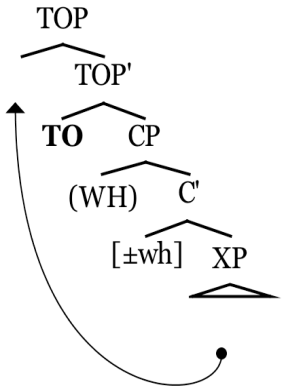
(60) *Mašu$_i$ -to Victor ubežal, potomu čto ljubit t_i.
Masha$_{ACC}$ TO Victor ran-off because loves
"Victor ran off because he loves Masha."

(61) *Kem$_i$, ty znaeš studentov, kotorye vosxiščajutsja t_i?
whom$_{INST}$ you know students that admire
"*Who do you know students that admire?"

(62) *Džonom$_i$ -to, ty znaeš studentov, kotorye vosxiščajutsja t_i.
John$_{INST}$ TO you know students that admire
"*As for John, you know many students that admire [him]."

Constructions with -TO are thus sensitive to islands, as demonstrated in adjunct and relative islands respectively. I propose that constituents that are –TO marked move to the Spec of HOP, attracted by the [+TOP] features –TO, which is located in the head of HOP and triggers the movement. I put forward the following structure for –TO in the matrix clause shown in (63)

(63) –TO Structure



We have also seen that while –TO acts as a trigger for movement, the movement fails when there are other structural restrictions imposed on it. Thus, we have to revisit the –TO generalization.

(64) ***TO-generalization Revisited:***

> – TO is generated in the head of HOP and attracts a topicalized constituent that can independently move to the Spec of HOP.

The structure in (63) also captures the fact that in cases where –TO-marked XP occurs in combination with wh-phrases, XP-to appears before the wh, given the examples of non-initial WH we saw in Chapter 2.

4.3.1 Structures of XP-TO

The assumption here is that –TO resides in the head of the HOP phrase that is marked with a [+topic] feature. For the ease of structures I just call it TopP here, however, this position is identical to HOP[+topic].

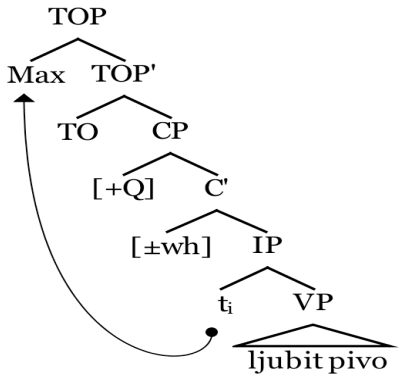
The structure in (63) and the –TO generalization allows us to give a structural analysis to the data used in the earlier sections of the chapter. Here I give structure examples for subject NP-TO, PP-TO, and even a [clause]-TO constructions.

4.3.1.1 NP-TO

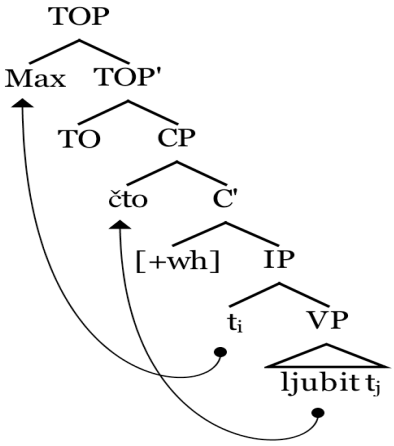
(65) a. Max -to ljubit pivo.
Max TO likes beer

b. Max -to čto ljubit?
Max TO what likes

(c) Tree-structure of (65)



(d) Tree structure of (65)



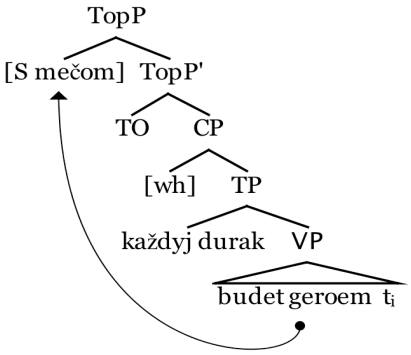
In (65) Max is topicalized, so it is moved to the HOP[+topic], -TO is located in the head of Top'.
In (65) the –TO-marked subject "Max" appears before 'what' (čto).

4.3.1.2 PP-TO

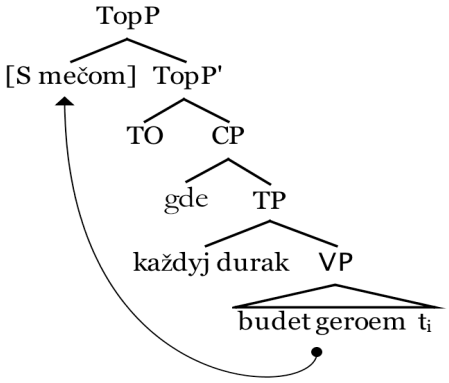
(66) a. S mečom -to každyj durak budet geroem.
with sword TO every fool be_{FUT} hero
"With a sword every idiot can be a hero?"

b. S mečom -to gde každyj durak budet geroem.
with sword TO where every fool be_{FUT} hero
"With a sword, where would every idiot be a hero?"

(c) Tree –structure of (66)



(d) Tree structure of (66)



4.3.1.3 [clause]-TO

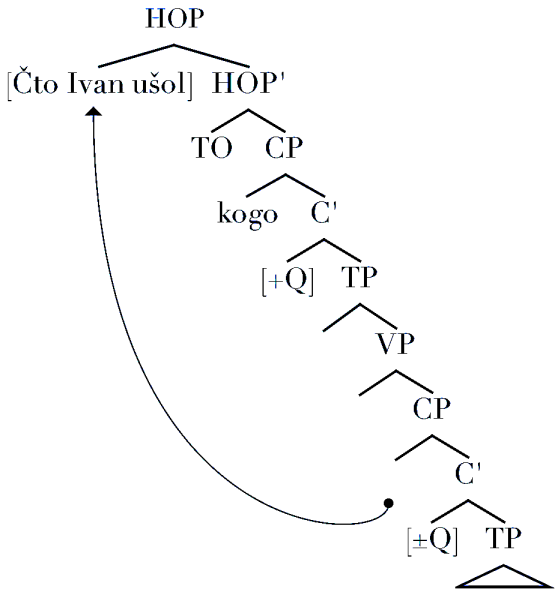
Everything we have seen thus far predicts that TO-marking on an entire moved clause should be possible. This is shown in (67). Structure in (69) shows that as expected a WH question follows the TO marked phrase. Other orders are impossible, as in (68).

(67) a. [$_{CP}$Čto Ivan ušel] -to, vse uznali sliškom pozdno.
that Ivan left TO all found out too late
"As for the fact that Ivan left, everyone found out too late."

b. [$_{CP}$Čto Ivan ušel] -to, kogo volnuet?
that Ivan left TO who worries
"As for the fact that Ivan left, who cares?"

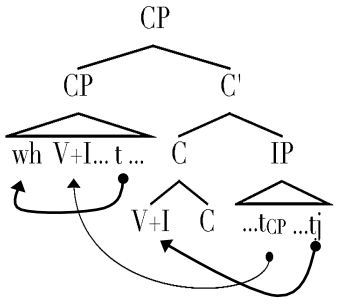
(68) *Kogo volnuet, [čto Ivan ušel] -to
who worries that Ivan left TO
"Who cares that Ivan left?"

(69) Structure of (67)



The structure in (69) shows how the entire clause can move into the HOP[+topic]. The structure in (69) can also be analyzed as a clausal pied-piping. Arregi (2003) suggested such movement for Basque. The structure is shown below in (70):

(70) Clausal pied-piping (Arregi 2003):



However, there appear to be certain differences on restrictions here: in Basque the clause movement happens to CP, in Russian it can only move to HOP[+topic] since it is shown here that *čto* ('that') is a [-wh] complementizer of the embedded clause, and hence, it cannot appear in CP.

This leads to the discussion of what is distribution of TO in the embedded clause. The next section discusses this issue in detail.

4.4 -TO in Embedded Clauses

In this section, I consider the distribution of the particle in embedded clauses. In the earlier chapters, I have claimed that the unique HOP[+topic] is only available in matrix clauses. In the first part of the chapter, I have shown the behavior of–TO marked elements in matrix clauses. Therefore, we predict that the occurrence of –TO marked elements in embedded clauses is ruled out for two reasons: first, *in-situ* –TO-marking is impossible (72), and secondly, there is no available position in the embedded clause for the –TO marked element to move into – no left-edge clause position of HOP in the embedded clause (73).

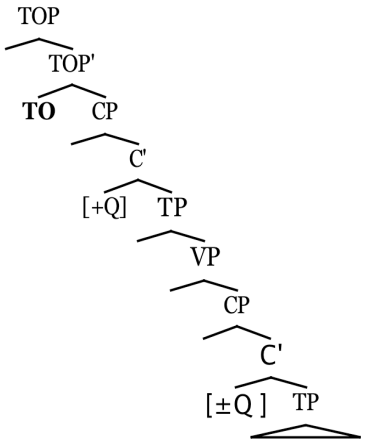
(71) Mne interesno, **kto** poljubit Ivana?
me$_{DAT}$ interesting who love$_{FUT}$ Ivan$_{ACC}$
"I am interested, who will fall in love with Ivan."

(72) *Mne interesno, **kto** poljubit Ivana -to?
me$_{DAT}$ interesting who love$_{FUT}$ Ivan$_{ACC}$ TO
"I am interested, who will fall in love with Ivan."

(73) *Mne interesno, Ivana -to **kto** poljubit?
me$_{DAT}$ interesting Ivan$_{ACC}$ TO who love$_{FUT}$
"I am interested, who will fall in love with Ivan."

Since high HOP is not available in the embedded clause, the sentence (73) crashes due to the fact that there is nowhere for the –TO marked constituent to move, and no head to host the particle. Thus, pre-wh in embedded clauses is out as predicted by the lack of lower HOP. The structure of the embedded clause vs. the matrix clause is reflected in (74):

(74) TO-clause with embedding



Thus,–TO –marking in embedded clauses is unavailable in the same way as it is available in matrix clauses; however, when long distance extraction out of the embedded clause into the Spec HOP of the matrix clause is possible – then –TO-marking is available, as long as nothing is blocking such extraction. Thus, we should expect that no argument of the embedded clause can be –TO-marked as in (76), since it cannot appear *in-situ,* and it cannot move to the edge of the embedded clause (as in (77)).

(75) Ženja uverena, [čto Ivan sdast èkzameny].
Zhenya sure that Ivan $pass_{FUT.}$ exams
"Zhenya is sure that Ivan will pass the exams."

(76) *Ženja uverena, čto Ivan sdast èkzameny -to.
Zhenya sure that Ivan $pass_{FUT.}$ exams TO
"As for exams, Zhenya is sure that Ivan will pass them".

(77) *Ženja uverena, èkzameny -to (čto) Ivan sdast
Zhenya sure exams TO (that) Ivan $pass_{FUT}$
"Zhenya is sure that as for exams, Ivan will pass them."

However, if the movement of the element that is targeted for –TO-marking into the matrix clause is possible, then the –TO attachment to a subordinate argument is possible. This is, of course, true in cases where no island is blocking the movement and long-distance extraction is available. This is shown in (78).

(78) Èkzamenyi -to, Ženja uverena, (čto)[9] Ivan sdast ti
exams TO Zhenya sure (that) Ivan passFUT.
"As for exams, Zhenya is sure that Ivan will pass them".

In (78) the –TO marked element appears in the matrix clause HOP[+topic], and thus, the sentence is possible. This is expected. However, there is also another possibility for –TO –marking in embedded clauses: the –TO marked element occupies the position below CP. This is shown in (79) and (80) below: the –TO marked element appears at the position right below the edge of the clause, and below the wh-element in these cases.

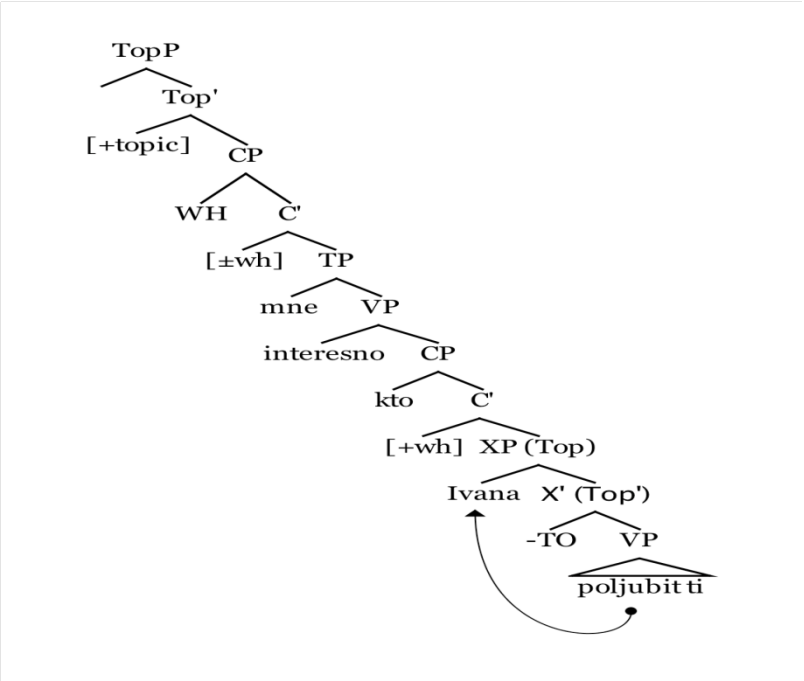
(79) ?Mne interesno kto Ivana -to poljubit.
meDAT interesting who Ivan TO loveFUT
As for Ivan, I wonder who will love him."

(80) ?Ja ne uveren, otkuda ved'ma-to priletela.
I not sure where-from witch TO flew
"As for the witch, I am not sure where she flew from."

In the embedded clause the –TO marked element can appear directly *following* a wh-phrase. Since the unique HOP position described above is not available in the embedded clause, what hosts the –TO-marked element? Such a matrix/embedded dichotomy is not unusual with respect to the Topic position and was claimed for English by den Dikken and Giannakidou (2002)[10]. They suggest a high topic position in the matrix clause, and a lower position in the embedded clause. According to them, when a dichotomy arises, another position can host an element in the embedded clause. Thus, we can assume that some sort of a partial movement to a higher, but not the highest clause edge position takes place. The –TO marked element appears below any wh-elements, which demonstrates that they do not reach CP domain. I will call this construction "Partial TO-marking" (recall also Dyakonova's 2009 TopP > topP): the structure is given in (81)

[9] In (78) either presence or absence of the complementizer "čto" (that) does not impact the grammaticality judgments.
[10] Some details of den Dikken and Giannakidou (2002)'s analysis is discussed in the previous chapter. Their idea of explaining the dichotomy can be applied here.

(81) Partial TO-marking in Embedded Clauses



Therefore, while –TO –marking in embedded clauses is not possible in the same way as it is in the matrix clause, it is available when long distance movement to the main clause is not restricted, and it is also possible within the clause to the lower than CP position. The nature of TO occurrences in the embedded clause requires further investigation.

4.5 Other relevant usage of –TO: Non-initial, Exclamative and Contrastive, Wh-TO

In this section, I present another set of constructions with – TO that do not fit the picture drawn in this chapter so far, namely TO-marking on tensed verbs, adjectives, and wh-phrases.

4.5.1 WH+TO

It is logical to assume that if every constituent can be marked with –TO then wh-phrases should be possible to mark with TO as well. This is true: wh-words can be marked with –TO. However, here we run into a problem. It can be formulated in the following way, wh-word +TO creates an indefinite pronoun. That is why it is common in the literature to differentiate between these "different" TOs. There are at least three: one – *to* that creates an indefinite pronoun; second TO that acts as a definite pronoun, and the third –TO which acts as a discourse particle. The

consensus is that they are just homophonous. I have a somewhat different explanation for the last two. Below I show a table of all the wh-phrases the result of wh+-*to* merge.

TABLE 1. RUSSIAN WH-WORDS AND INDEFINITE PRONOUNS

Case	*Question words*	*Indefinite pronouns with -TO*	*Indefinite pronouns with –nibud'*	*Gloss of the Pronouns*
Nom	kto, čto	kto-, čto - to	kto-, čto - nibud'	someone/something$_{NOM}$
Acc	kogo, čto	kogo-, čto - to	kogo-, čto - nibud'	someone/something$_{ACC}$
Gen	kogo, čego	kogo-, čego - to	kogo-, čego - nibud'	someone/something$_{GEN}$
Dat	komu, čemu	komu-, čemu- to	komu-,čemu- nibud'	someone/something$_{DAT}$
Instr	kem, čem	kem-, čem - to	kem-, čem - nibud'	someone/something$_{INSTR}$
Prep	kom, čem	kom-, čem- to	kom-, čem- nibud'	someone/something$_{PREP}$
ADV	kak	kak -to	kak –nibud'	somehow
	gde	gde -to	gde – nibud'	somewhere (locational)
	kuda	kuda-to	kuda-nibud'	somewhere (directional)
	kogda	kogda -to	kogda – nibud'	sometime
	počemu	počemu-to	počemu - nibud'	for some reason

Notice, that there exist two types of indefinite pronouns. One can observe a difference in their distribution with and without –TO (wh-word-*to* vs. wh-*nibud'*). They also do carry different meanings. Here, I postulate that a wh-word+*nibud'* is an unspecific indefinite pronoun, and wh-word+*TO* results in a specific indefinite pronoun. The historical development of this particle in a language briefly outlined above plays a role in creating of this distinction. It is to say that when a pronoun marked with –TO there exists a known (indefinite, but not the indefinite in number) set of items that such a pronoun can refer to. It is not surprising, then, that a wh-phrase in an information question cannot be marked with –TO. Imagine that a wh-phrase moves into HOP where –TO is present. This would result in clash of features when it comes to HOP. The wh-phrase occurrence in HOP is only possible when HOP is marked with [-topic] feature: that would

allow a wh-word to move there; however, when –TO is present overtly, HOP is marked with a [+topic] feature. Moreover, when stacking of particles occurs the combination wh-*nibud' –TO* is possible (82), whereas wh-*to-TO* (83) is not.

(82) Kogo-nibud'-to ty točno vstretiš na ètom koncerte.
someone TO you surely meet on this concert
"You will definitely meet someone at this concert (someone worth meeting)"

(83) *Kogo-to -to ty točno vstretiš na ètom koncerte.
someone TO you surely meet on this concert
"You will definitely meet someone at this concert (the people that you know)"

This contrast is not surprising, in (82) the particle –TO does not make an indefinite pronoun specific, it just makes a mark that this pronoun is available from previous discourse. As in the context of an answer to the following question (84):

(84) Context for (82):

Ty dumaeš, ja vstreču kogo-nibud' na koncerte?
you think I $meet_{FUT}$ somebody on concert
"Do you think I will meet somebody at the concert?"

The stacking of this particle with other particles is possible. Bayer (1996) states that when stacking of the particles occur, the ordering of the particles is usually rigid. While other particles (most likely Focus ones, more on these can be found in the works of Nikolaeva 1985, Bonnot 1986, 1988, 1990) can be stacked in no particular order, -TO probably by its nature of enclitization always appear first in that row as in an example (85) below:

(85) To -to že ved' i ono!
that TO $_{PRT}$ $_{PRT}$ $and_{EXCL.}$ $it_{NEUT.SING, NOM}$
"This is what it is!"

Additionally, a –TO marked element can co-occur with a question particle *li*. The discussion of the distribution of *li* is not in the scope of this work (for details see Rudnitskaya 1998 analysis). In brief, the crucial issue here is that *li* is a question particle that marks a yes/no question and it cannot co-occur with wh-phrases (unlike Bulgarian, for instance). That is how an information

question differs from a yes/no question in Russian. When *li* occurs in a question, one might expect a verbal inversion, which is discussed in Bailyn (1995). Notice, that the occurrence of *li* particle in the embedded clause is mandatory (87), while it can be omitted in the matrix clause (88). When this particle does not occur in the clause, no inversion is necessary. Consider the following examples:

(86) *Komu li Maria pozvonit?
who$_{DAT}$ Q Maria call$_{FUT}$
"Who is Maria going to call?"

(87) Pozvonit li Maria segodnja?
call$_{FUT}$ Q Maria today
"Will Maria call today?"

(88) Maria pozvonit segodnja?
Maria call$_{FUT}$ today
"Will Maria call today?"

(89) Mne ljubopytno, pozvonit li Maria segodnja?
Me$_{DAT}$ curious call$_{FUT}$ Q Maria today
"I am curious whether Maria is going to call today?"

(90) *Mne ljubopytno, pozvonit Maria segodnja?
Me$_{DAT}$ curious call$_{FUT}$ Maria today
"Will Maria call today?"

Notice, that the difference between the requirement of the presence of *li* (Q) in the embedded clause where it is mandatory as we can see from the contrast between (89) and (90), and the root clause where the occurrence of *li* is optional. While this fact is interesting in itself, I do not to attempt to explain it here. If one assumes the occurrence of *li* question particle in the C', then the following constructions should be of interest here.

(91) Maria -to pozvonit li segodnja?
Maria TO call$_{FUT}$ Q today?
"Maria, will she call today?"

(92) *Mne ljubopytno Maria -to pozvonit li segodnja?
me$_{DAT}$ curious Maria TO call$_{FUT}$ Q today
"I am curious whether Maria is going to call today?"

Notice that while a –TO marked Topic element is possible in the matrix clause (91) it is not possible in the embedded clause where CP is interpreted as a question (92).

Thus, even though, the combination of a wh-phrase +TO result in an indefinite pronoun, and hence, block the information question from occurring for a reason of a feature being set [+topic], the general distribution between a Topic particle and a CP remains the same. Moreover, the fact that a wh-element when combined with –TO creates an indefinite pronoun in Russian does not rule out the general possibility of combining wh-phrases with Topic markers. Examples of such combinations exist. Grohmann (1997) notes "Interestingly, Miyagawa (1987) investigates a possible interaction of the topicalization marker *wa* and wh-movement in Japanese. He concludes that in certain contexts the Wh-element may indeed be marked with *wa.*" Miyagawa stated that one of the properties of the particle *wa* is "set-anaphoric", in other words it determines a set. The properties of such a set is also discussed in detail. This particle can occur according to him with various wh-words assuming that this meaning is preserved. One of the examples is provided below in (93):

(93) (Miyagawa, 1987: 189 (12))

dare wa sanji ni itte dare wa yoji ni itte
who TP 3 o'clock at go$_{GER}$ who TP 4 o'clock at go$_{GER}$
dare wa goji ni itta no?
who TP 5 o'clock at went QP
"Who went at 3 o'clock, who went at 4 o'clock, who went at 5?"

In the following section I consider cases that are somewhat problematic for –TO distribution. I have stated that –TO marks virtually any XP, but have avoided the discussion of the tensed verbs. Here, I show how – TO marks tensed verbs.

4.5.2 VP_{TENSE} + TO

When it comes to distribution of –TO with tensed verbs (VP_{TENSE}) the picture changes from everything that we have seen so far: instead of appearing at the left edge of the clause, the –TO marked VP_{TENSE} appears on the right of the clause.[11,12]. Consider the following examples (94):

(94) a. Začem oni pozvonili -to?
why they called TO
"What was the reason for their call?"

b. *Začem oni pozvonili -to tak rano?
why they called TO so early
"Why did they call (even thought of calling) so early?"

c. ?Začem oni tak rano pozvonili -to?
why they so early call TO
"Why did they call (even thought of calling) at such an early hour?"

---

[11] There is another distribution of –TO doubling on verbs that is restricted to one type of constructions (these constructions are discussed in detail in McCoy (2001) and Abels (2001))

(i) Pozvonili – to oni pozvonili, no ničego konkretnogo ne soobščili.
called TO they called but nothing concrete NEG informed
"As for calling, they called, but they didn't say anything informative.

As in the situation with adjectives, these are contrastive, and also utilize the reduplication of the sort to create the construction. These probably need separate discussion.

(ii) Staraetsja –to on staraetsja, no vse ravno polučaet dvojki.
$try_{REFL,3rd}$ TO. he $try_{REFL3rd}$ but all same gets twos
"As for trying, he tries, but still receives Ds."

Compare (ii) to (iii a, b), and (iv)

(iii) a. On staraetsja, no vse ravno polučaet dvojki.
he $try_{REFL3rd}$ but all same gets twos

b. *On staraetsja-to, no vse ravno polučaet dvojki.
he $try_{REFL3rd}$ TO. but all same gets twos

(iv) a. *Pozvonili – to oni, pozvonili no ničego konkretnogo ne soobščili.
called TO they call but nothing concrete NEG informed

b. Pozvonili oni no ničego konkretnogo ne soobščili.
called they but nothing concrete NEG informed

[12] It is also important to keep in mind that historically the particle –TO was, of course, a pronominal particle, and it agreed in number and in gender with the noun it attached to. This agreement is still used in some dialects of Southern Russian. Hence, its agreement with verbs can be complicated, and in the process of historical transition.

(94) shows that a tensed verb with –TO-marking only works when the verb that is marked with –TO occurs on the very right of the clause. Compare the contrasts between (94) vs. (94) and (95) vs. (95) with an adjunct and a direct object following the –TO marked verb or preceding it respectively.

(95) a. *Začem oni pozvonili -to Ivanu?
why they call TO $Ivan_{DAT}$
"Why did the call (even thought of calling) Ivan?"

b. ?Začem oni Ivanu pozvonili -to?
why they $Ivan_{DAT}$ call TO
"Why did the call (even thought of calling) Ivan?"

All the sentences in (94) are questions, however (94) and (95) are marginally available if at all. The declarative sentences with – TO marked on VP_{TENSE} are not available as demonstrated in examples (96) - (99).

(96) *Oni pozvonili-to v Moskvu.
they called TO in Moscow
"They called Moscow."

(97) *Oni v Moskvu pozvonili -to
they in Mosow called TO
"They called Moscow."

(98) *Pozvonili -to oni v Moskvu.
called TO they in Moscow
"They called Moscow."

(99) *[Pozvonili v Moskvu] -to oni.
called in Moscow TO they
"They called Moscow."

The unavailability of the construction does not depend on any of the restrictions discussed above: it is not the position of the –TO marked element and not the entirety of the constituent that makes it ungrammatical. Such constructions are simply not available. Thus, on one hand, we can state

that the same –TO-marking that we have observed on various XPs is not available on $VP_{TENSE.}$ On the other hand, there are sentences like the one in (94), and additional examples indicated below in (100) - (101) that have particle –TO appearing after V_{TENSE}.

(100) Ty začem pozvonil-to?
you why called TO
"Why is it that you called?" (Compare to "Why did you call?")

(101) Ty gde živeš'-to?
you where live TO
"Where is it that you live?" (Compare to: "Where do you live?")

In the embedded clause the situation is different, however, -TO particle cannot occur to the right of the embedded tense verb as shown in the following examples:

(102) ?Mne interesno, kogo Ivan priglasil -to na prazdnik?
me_{DAT} interesting who_{ACC} Ivan invite TO on party
"I am interested who Ivan invited to the party?"

(103) *Ja xoču uznat', kogda oni pošli -to v kino?
I want know when they went TO in cinema
"I want to know when they went to the movies."

Particle –TO is not welcome as a modifier of a tensed verb in the embedded clause. In general, despite the examples given above in (79) and (80), the distribution of –TO in the embedded clause is fairly muddled. However, one thing is clear, it does not have the same distribution and/or meaning in the embedded clause as it does in the matrix clause.

There are several possible explanations for this behavior. First, let us keep in mind that this behavior does not violate one part of the –TO Generalization: it still attaches to an entire constituent. The problem arises in that VP_{TENSE} do not move, and on the surface –TO marked V_{TENSE} occurs at the right edge. There are several possible approaches here. One is that the high –TO binds the –TO below and the lower copy is pronounced. However, this approach does not explain the ungrammaticality of declarative sentences. Another approach is that –TO clausal pied-piping discussed above (Arregi 2003) takes place, where the entire clause is moved to the relevant position in the derivation. – TO attaches to the entire sentence, thus a sentence acts as a

clause constituent. This approach seems to be on the right track considering the special meaning that such questions gain from the –TO marking present in them. This would also explain why it is unavailable in declarative clauses. This option can work for the explanation of (100) and (101) as well as for the cases with V_{INF} where -TO also occurs at the very right edge of the sentence rather than on the left. Such cases are discussed in the subsection below.

4.5.3 V_{INFIN} +TO = Exclamative

The particle –TO that occurs on the right edge of the questions changes the meaning of the sentence, which is indicated by the glosses. In this they differ from regular information questions.

(104) Skol'ko možno bolet' -to?
How much can be-sick TO
"How long can one be POSSIBLY sick for?" (as an exclamation)

(105) a. Čto vy xotite rasskazat' -to?
what you want $tell_{INFIN}$ TO
"What is it that you want to tell us?" (vs. "What do you want to tell us?")

b. Čto vy xotite rasskazat'?
what you want $tell_{INFIN}$
"What do you want tell us?"

The difference in meaning in (a) and (b) is obvious here.

(106) *?Oni ne xoteli ničego rasskazyvat' -to.
they not wanted nothing $tell_{INFIN}$ TO
"As for telling, they didn't want to tell anything"

(107) Rasskazyvat' -to oni ničego ne xoteli.
$tell_{INF}$ TO they nothing neg wanted
"As for telling, they didn't want to tell anything"

vs.

(108) Počemu oni ne xoteli ničego rasskazyvat' -to?
why they not wanted nothing $tell_{INFIN}$ TO
"Why is it that they didn't want to say anything?"

Examples (104) through (108) show different (im)possibilities of –TO-marking: (106) is expectedly ungrammatical – the –TO marked element is *in-situ*, while (105) is possible because V_{INF}, which is –TO marked occurs in the high position on the left edge. The grammaticality of (104), (105) and (107) arises from a completely different fact: these are questions with exclamation on a par with 'wh-the hell' constructions, since (108) can easily be interpreted as "Why the hell didn't they want to say anything?" Thus, the entire clause is being –TO-marked. Thus, the entire clause receives emphasis – I call this usage – TO–exclamative. This would explain linear occurrence of the particle on the right edge of the clause with both V_{INFIN} and $V_{TENSE,}$ which is in fact should be interpreted as [$_{Q}$ CLAUSE]-TO. These, of course, only refer to question clauses.

Another productive usage of the particle is Contrast. In fact several researches commented on its contrastive features (McCoy 2002, Nikolaeva 1985 among others).

4.5.4 -TO of Contrast

-TO is very productive in a number of contrastive environments shown in (109) - (111) below, where the element of the first part of the sentence is opposed to the parallel element in the second part of the sentence. In each case the element in the first clause can be –TO-marked.

(109) a. Dissertaciju –to on zaščitil, a rabotu najti ne smog.
dissertation TO he defended but job find NEG could
"As for dissertation, he defended it, but he couldn't find a job."

b. *Dissertaciju –to on zaščitil, a rabotu-to najti ne smog.
dissertation TO he defended but job TO find NEG could
"As for dissertation, he defended it, but he couldn't find a job."

c. ?Dissertaciju on zaščitil, a rabotu-to najti ne smog.
dissertation he defended but job TO find NEG could
"He defended the dissertation, but as for a job, he couldn't find any."

(110) a. Svetu -to on ljubit, a Rozu ne boitsja.
Sveta$_{ACC}$ TO he loves but Rosa$_{ACC}$ NEG afraid
"As for Sveta, he loves her, but Rosa he isn't afraid of."

b. *Svetu -to on ljubit, a Rozu-to ne boitsja.
Svetu TO he loves but Rosa TO NEG afraid

c. ?Svetu on ljubit, a Rozu-to ne boitsja.
Svetu he loves but Rosa TO NEG afraid
"He loves Sveta, but as for Rosa, he isn't afraid of her."

(111) a. Zoja -to krasavica, a Lena umnica.
Zoya TO beauty but Lena clever
"As for Zoya, she's a beauty, but Lena's clever".

b. *Zoja-to krasavica, a Lena-to umnica.
Zoya TO beauty but Lena TO clever
"As for Zoya, she's a beauty, but Lena's clever".

In such contrastive contexts the occurrence of –TO marked elements in both opposing clauses results in ungrammaticality as in (109), (110), and (111). When –TO is used contrastively the requirement of its appearance on the full constituent does not exist anymore. This creates an additional puzzle, and gives additional data for the theory of left branch extraction (LBE).

4.5.4.1 Adjectives +TO Revisited

The example in (112) is similar to the example with attributive adjective phrase from the earlier section. However, here, I consider another possibility of –TO marking that result not in mere topicalization, but additionally, in contrast.

(112) a. On zadal studentam interesnuju zadaču.
he gave students$_{DAT}$ [interesting problem]$_{ACC}$

b. Interesnuju on zadal studentam zadaču.
interesting he$_{NOM}$ gave students$_{DAT}$ [t problem]$_{ACC}$
"He gave the students an interesting problem."

Example (112) demonstrates the possibility of the left-branch extraction without –TO-marking, which is not uncommon for Russian, where a part of a constituent is left-dislocated. Simple –TO

marking of the adjective in such case is not grammatical as has been stated before and is repeated in (113):

(113) *Interesnuju -to on zadal studentam zadaču.
interesting TO he_{NOM} gave $students_A$ [t $problem]_{ACC}$
"He gave students an interesting problem."

However, if the –TO marked element has a contrastive counterpart in the opposing clause, the sentence becomes grammatical: in other words, in a *contrastive* situation, the usage of –TO attached to an adjective is possible. Imagine a situation in which Bob is buying a car. And Bob is hesitating between a green and a blue one, and finally, buys a blue car: you report on this purchase (114):

(114) Sinjuju -to mašinu Bob kupil, a zelenuju – net.
Blue TO car Bob bought but green not
"As for the blue car, Bob bought it, but he didn't buy a green one."

Analogously, (113) can be repaired in the following fashion:

(115) Interesnuju -to on zadal studentam zadaču, a trudnuju – ne stal.
interesting TO he_N gave $students_A$ problem but difficult not $become_{PAST}$
[~~zadavat'~~]
~~assign~~$_{INF}$
"As for the interesting problem, he assigned it to the students, but a difficult one decided not to assign."

The Contrastive usage of –TO also explains the grammaticality of(116)=(47)[13] repeated here:

(116) [Prekrasnuju -to rol'] Maria polučila, a deneg ne zarabotala.
wonderful TO part Maria got but money not earned
"As for the wonderful role, Maria got it, but she didn't make any money."

This contrastive usage poses further problem as to where these contrasted –TO marked elements land, since the -TO generalization is violated in these cases: the constituents can be split as we can see in (115) and (116). The answer to this question would rely on the adopted theory of LBE.

[13] Note, that (47b) main clause is by itself ungrammatical without a contrastive clause as illustrated in (116).

Several such theories exist (Bošković 2005, 2011 for Slavic generally; Wiland 2010 for Polish; Pereltsvaig 2007 for Russian). Depending on which view of constituency is taken in the case of LBE, this phenomenon of contrastive – TO-marking can be explained.

In section 4.5, I have discussed of –TO occurrences that do not fit the picture outlined throughout the rest of the chapter. I have discussed how –TO can appear on the right edge of the sentences and how –TO can add the meaning of Contrast when the opposing element and/or clause is present. In the case of contrasting, -TO still must appear on the left –edge of the matrix clause, and no doubling of the particle (i.e. on both of the opposites) is possible. In the case where –TO appears on the right edge, -TO marks the entire question clause with exclamation and adds to it the meaning similar to "wh-the –hell" constructions.

4.6 Chapter Summary

In this chapter, I considered the particle -TO that has been given attention in the semantic literature, but does not have syntactic analysis. Here, I proposed that the particle occupies the head of the HOP phrase of the matrix clause. Due to the fact that only one and unique HOP is available, only one –TO per clause is available. The elements that are –TO-marked must move to the left edge of to appear before the particle. The elements that are -TO-marked must create a constituent. I showed that it attaches to any element, any XP, to satisfy the discourse requirement to mark Topicality. The analysis provided in the chapter shows that the particle is sensitive to restrictions on movement: island conditions. On the other hand, it can be extracted out of the embedded clause in the environments where movement is not blocked. Its distribution with verbal elements is somewhat different for verbs do not raise in Russian; however, as it has been shown in the previous chapter, HOP can bind elements, and hence, this particle can appear attached to an element on the right edge of the clause, and the clausal pied-piping can be utilized. Its distribution in embedded clauses is restricted due to the fact that the unique HOP position does not exist in the embedded clauses, and the root/embedded dichotomy is observed here once again. When it does appear in the embedded clauses, its position is below CP, and hence, is different from the unique HOP position claimed for the main clause in Russian. This particle apparently occupies a unique, literally and figuratively, position in Russian syntax.

Chapter 5. Syntactic and Semantic Nature of HOP

5.0 Introduction

Throughout the previous chapters I presented an analysis of wh-behavior and accounted for various word orderings by introducing the high operator phrase (HOP). The introduction of this category led us to reconsider the Superiority Condition: we entertained an important theoretical point of how Superiority should unfold, and that Superiority effects are not parameterized. So far, HOP is a category that is used to extend the structure of CP-domain. Such a proposal, while apparently successful in the account put forward here, might at first sight appear syntactically unorthodox because of its dichotic feature [±topic] (elements that this position can host), and the labeling of the phrase as an *operator*. Thus, a more detailed consideration of the syntactic and semantic nature of HOP is needed. For instance, on several occasions throughout this work, it has been stated that HOP is a matrix-clause-only phenomenon, and clausal asymmetry data were presented, but no additional evidence beyond the asymmetrical distribution of –TO marked phrases was given. In chapter 2, the proposal was that HOP can host both [+wh]-elements, and overt high topics. In chapter 4, the latter part has been expanded to show that particle –TO, a topic marker, marks high overt topicalized elements moved into HOP bearing a [+topic] feature. Thus, the following theoretical questions remain:

Remaining Theoretical Questions:

(i) what are the consequences of labeling HOP a high *operator* phrase?

(ii) what is the semantic nature of HOP: what does it mean for a wh-element to appear in HOP, a position that hosts topics?

(iii) what is the syntactic nature of HOP? What is the evidence that HOP is exclusively a matrix[49] (root) clause phenomenon?

[49] Throughout this work I have been dividing clauses into matrix and embedded. Another common terminological approach is to divide clauses into root/embedded or root/subordinate categories. All of the names are in essence the same. For the purposes of this chapter, I will have to use root/embedded notation due to the fact that historically the only matrix clause phenomenon has received a name of "root transformations" (Emonds 1969).

And consequently, if postulation of HOP is valid for Russian, and does in fact provide a clean analysis of Russian word order variation, then the following question should be in place:

(iv) is HOP used cross-linguistically, specifically in other Slavic languages?

Answering the questions above would, perhaps, require an extensive theoretical discussion: each answer leading to a separate part of another thesis on the matter. In this chapter, I provide preliminary discussion of each in its turn, and the directions in which the answers might take us. In order to address questions (i) and (iii), we are going to run "root transformation" diagnostics (Edmonds 1969, Hooper and Thompson 1973) to determine whether HOP is in fact just a matrix clause phenomenon (MCP) – a more recent name for RT. We will also briefly investigate the relation between HOP and quantifiers to determine its *operator* status. Answering (ii) and (iv) will unfold from these discussions, for answers to these questions are interrelated. Some issues are rather well-established. For instance, the fact that *operator* phrases should host quantifiers and wh-phrases. Thus, we start the discussion of the syntactic nature of HOP by discussing its Operator status.

5.1 Syntactic nature of HOP

5.1.1 Operator status of HOP

Labeling a phrase is never a trivial matter, especially a phrase that can host items with different features[50]. Syntactic category labels describe what an element is, while syntactic functions describe what an element does. By labeling a category as an *operator* phrase, one assumes that it should host items that possess operator properties or function as operators. Traditionally, wh-words, quantifiers and topic elements are called operators. In German, arguably, any of these can be hosted in SpecCP. Therefore, there does not exist a one-to-one correspondence between the "label" of the German CP phrase, and the elements that can appear in its Specifier. Therefore, one might assume that a phrase labeled as an operator, would not necessarily have all the properties of the Operator Phrase, but would act upon the properties of the phrases it hosts. The motivation behind labeling a high phrase over CP an operator is because it can host operators: wh-words, topics, on the other one hand, and cannot host non-operators. The

[50] Important discussion on labeling can be found in Thráinsson (1996) and Collins (2002)

difference between HOP and CP, then, is that CPs are built on the feature composition of the head C', while HOP is built upon the requirement of HOP Operator host in Specifier.

In classical Government and Binding (GB) theory, an operator is usually understood to be a wh-word or a quantifier in an A'-position. The examples below consist of different types of operators.

(1) $\text{Who}_{i\ OP}$ said he_i killed John?

(2) Everyone_{OP} likes someone.

Operators can also be non-overt (or null):

(3) John_i is easy [OP_i PRO to please t_i]

According to Koopman and Sportiche (1982) all operators are subject to the Bijection Principle (Sportiche 1985:467)

(4) *Bijection Principle* (Sportiche 1985:467)
Operators locally bind one and only one A- position.

The High Operator Phrase (HOP) can host such operators. HOP can host both wh-elements and non wh-elements such as topics; this is essential for the account proposed here. This is similar to Citko (1998), who proposes that wh-words in Polish occupy an operator position. Thus, I state that the High Operator Phrase must possess some kind of a [± topic] feature to host both topicalized and non-topicalized elements. I return to the discussion how this is possible in section 5.1.2.

The proposal that a dichotic [± topic] feature marks the head of the Operator phrase is also not new. English CP can host wh-operators, and hence can carry a [+wh] feature, but can also host non wh-elements (i.e. that) and hence carry a [-wh] feature. From the syntactic literature we also know that SpecCP, a common site for wh-words, is available to host Topics in German (Frascarelli, Hinterhölzl, 2007[51]). Thus, SpecHOP can host topicalized material when marked with a [+topic] feature, and wh-phrases when it is not.

Let us first see if generalizations pertinent to operator phrases hold in this case: we should assume that bare quantifiers are possible in SpecHOP position, and that we should also observe weak crossover effects (WCO) as one would in the case of an operator present. In order to

[51] The possibility of SpecCP base-generation of topics or movement is entertained.

investigate WCO, we should establish whether bare quantifiers can occur in the SpecHOP position.

Examples of bare quantifiers in English and Russian are given in (5) and (6) respectively:

(5) John likes everyone.

(6) Ivan ljubit vsex (**?***každogo)
Ivan loves everyone$_{PL\text{-}ACC}$ /everyone$_{SG\text{-}ACC}$
"Ivan loves everyone."

Note that the expected equivalent of 'everyone' in Russian is '*každyj'*. However, it is not simple to create examples with '*každyj'* where it acts as a bare quantifier because of its morphological adjective-like nature: it is assumed that there is a null-NP that all Long Form adjectives as '*každyj'* modify (Babby 1975, Bailyn 1994) nouns, such as '*čelovek'* ("person"), '*student'* (student), 'mal'čik' (boy), for example. Therefore, even constructing examples for verification with a quantifier moved to SpecHOP is somewhat challenging. Usually extracting just '*každyj'* is difficult, its plural equivalent '*vse'* ("all") works somewhat better, for it can stand alone in many cases.

(7) a. Ivan (vsex) ljubit (vsex).
Ivan everyone $_{PL}$ loves everyone$_{PL}$
"Ivan loves everyone."

b. Vsex Ivan ljubit.
everyone$_{PL\text{-}ACC}$ Ivan$_{NOM}$ loves

In (7) '*vsex'* ('everyone') can occur pre-verbally and post-verbally without a change in meaning (more discussion of this appears in Dyakonova (2009) and is not immediately relevant to the discussion here). Example (8) demonstrates, however, that *vsex* can occur in the high left position, presumably in SpecHOP. The picture would become even clearer, if we use a combination of *vsex* and a wh-word:

(8) Vsex **kto** ljubit?
everyone$_{PL}$ who loves
"Who loves everyone?"

(9) Vsem **o** **čem** Ivan rasskazal?
everyone $_{PL\text{-}DAT}$ about what Ivan told
"What did Ivan tell everyone?"

Notice that questions in (8) and (9) are impossible in embedded clauses:

(10) *Mne interesno, vsex **kto** ljubit?
me$_{DAT}$ interesting everyone$_{PL}$ who loves
"I wonder who loves everyone?"

(11) *Mne interesno, vsem **o** **čem** Ivan rasskazal?
me$_{DAT}$ interesting everyone$_{PL\text{-}DAT}$ about what Ivan told
"I wonder what did Ivan tell everyone?"

Under the hypothesis that HOP is not available in embedded clauses, the unavailability of (10) and (11) is predicted: there is nowhere for the quantifier to front to overtly in the embedded clause. It is important to note that topicalizing a quantifier is not accepted across the board in Russian[52], but the contrast between its availability in the fronted position over a wh-word in the matrix and the embedded clause is striking.

Now, let us turn to weak crossover (WCO) effects. WCO restricts binding and coreference possibilities between a pronoun and a variable in the case of a crossover of the antecedent of the variable over the pronoun (for a detailed discussion of Russian quantifier behavior: WCO effects and QR, refer to Antonyuk – Yudina 2006, forthcoming). General examples are shown in (12) and (13) where (14) and (15) show Russian WCO violations.

(12) *Who$_i$ does his$_i$ mother love <who> ? (wh-mvt causes WCO)

(13) *His$_i$ mother loves everyone$_i$. (QR causes WCO)

(14) ***Kogo** ego mama ljubit? (wh-mvt causes WCO)
who$_{ACC}$ [his mother]$_{NOM}$ loves
"Who does his mother love?"

[52] The exact nature of this variability is, indeed, a question for further research in this area.

(15) *Ee$_i$ mama pričesala [každuju devočku]$_i$ (QR causes WCO)
her mother combed every girl

"Her mother combed every girl."

Now, let us consider how WCO acts with respect to HOP position. To do so, I compare the -TO construction to a Focus version of the pure fronting HOP construction, to see whether the Topic/Focus status of the moved element is relevant, which it appears to be, as predicted by Rizzi (2004). First, note that simply using the SpecHOP position does not appear to cause a WCO violation. This is shown in (16):

(16) a. ?Ženju$_i$ **gde** ee$_i$ sestra videla?
Jane where her sister saw

"Where was Jane seen by her sister?"

b. ?Ivana$_i$ **gde** ego$_i$ žena zastala s ljubovnice$_j$?
Ivan$_{ACC}$ where his wife caught with lover

"Where was Ivan caught by his wife with a lover?"

The same holds for the TO-construction:

(17) a. Ženju$_i$-to ee$_i$ sestra videla.
Jane TO her sister saw

"Jane was seen by her sister."

b. Ivana$_i$-to ego$_i$ žena zastala s ljubovnicej.
Ivan TO his wife caught with lover

However, when the fronted element is strongly Focused, the WCO effect appears:

(18) a. *ŽENJU$_i$ **gde** ee$_i$ sestra videla včera?
Jane$_{FOC}$ where her sister saw yesterday
"Where was JANE seen by her sister?"

b. *IVANA$_i$ **gde** ego$_i$ žena zastala s ljubovnicej?
IVAN where his wife caught with lover
"Where was IVAN caught by his wife with a lover?"

These data show that with surface word orders that look identical, we observe WCO in Russian in some cases (18) but not in others (16) - (17). Let us look at each one separately. Examples ((16)a, b) are subject to a crossover violation, however, we observe only a very slight deviation in judgments: the possibility of co-indexation exists. Both "*Ženja*" and "*Ivan*" occupy SpecHOP here. When the topicalization is associated with -TO, as in ((17)a, b) (which I show is a topic particle housed in HOP in Chapter 4), the examples become fully grammatical. Note that the contrast between (16) and (17) is minor, unlike the contrast between ((16)a, b) and ((18)a, b) where examples in (18) are completely unacceptable. The difference in interpretation is that "*Ženja*" and "*Ivan*" are focused in ((18)a) and ((18)b) respectively. This set of data shows that while weak crossover effects are present with respect to HOP, the restrictions on crossover are in accordance with the general tendency of WCO in relation to Topics and Focus. This is fully consistent with expectations if we follow Rizzi's (1997) evidence that topic movement does not induce weak crossover effects while Focus does. Rizzi (1997:290) gives the following Italian examples where quantifiers and topicalization are incompatible (19), but are fine in a focus position (20).

(19) Gianni$_i$, sua$_i$ madre lo$_i$ ha sempre apprezzato (Rizzi 1997: ex. 17)
"Gianni, his mother always appreciated him"

(20) ??GIANNI$_i$, sua$_i$ madre ha sempre apprezzato t$_i$ (non Piero) (Rizzi 1997: ex. 18)
"GIANNI, his mother always appreciated, not Piero."

Thus, we can state that the operator status of HOP is not unfounded when it comes to its interaction with quantifiers. In the next section, I consider HOP as a host for topics and wh-phrases, and what the latter means.

5.1.2 HOP as a host for topics and wh-phrases

Jaeger (2004) using Rudin's (1985) data discusses topics and clitic doubling in Bulgarian. He makes several important observations that are directly relevant to the discussion of the validity of HOP as it is stated here both as a host for overt topics and wh-phrases, and as main clause phenomenon (MCP). Let us start by introducing the most important observations about Bulgarian made by Jaeger (2004): overt topics preceed wh-words as shown in (21):

(21) Ivan **na kogo** dade knigite? (Jaeger 2004: Rudin 1985: 92)
Ivan to whom gave books
"Who did Ivan give books to?"

Additionally, Billings and Rudin (1998:6) give an example where Superiority in Bulgarian is violated (as in (22)). The possibility of Superiority violation is the fact debated in the literature. However, this example does not contradict either the account proposed throughout this work, and the view on Candidate Superiority put forward in Chapter 2. However, there should be a valid reason for such a violation. The idea here is that Superiority can be violated for two reasons: strong D-(linking) and specificity. Let us consider these ideas for the following examples:

(22) a. **Koj na kogo** mu xaresva?
who whom LOC pleases

b. **Na kogo koj** mu xaresva?
whom who LOC pleases
"Who likes whom?"

Jaeger (2004) following Billings and Rudin's (1998) ideas, assigns the question in (22)b special pragmatic and prosodic properties along the lines of D-(linking) (Pesetsky 1987). Such a question encodes special discourse information found relevant to the previous discourse. He formulates several hypotheses for Bulgarian: "WH-Topic Fronting Hypothesis" and "Topics First! Hypothesis". Therefore, the order of wh-phrases can be changed in Bulgarian when certain discourse contexts are present: either D-linking (in Pesetsky's terms,) or discourse saliency of the Speech act (according to Miyagawa 2012).

(23) **WH-Topic Fronting Hypothesis**

Clitic-doubling in Bulgarian wh-question indicates that clitic doubled wh-phrase is the topic of the question. Thus, fronting of clitic-doubled wh-phrases is due to the same feature (topicality) that causes topic-fronting in non-question clauses.

Thus, he claims that wh-phrases appear in a topic position. The relevant question is, of course, what does it mean for a wh-phrase to appear in a topic position, when wh-phrases are considered *inherently focused.* Jaeger associates this with the D-linking phenomenon (Pesetsky 1987). "All fronted *wh*-phrases are foci and sometimes a *wh*-phrase can also be the topic of the question (cf. Gundel 1988:210; Leafgren 1997:127; Steedman 2000:659)" – he states. It was also suggested for Sanskrit by Hale 1987 (presented below), and for Spanish by Zubizarreta (1998) who states that "Topics typically precede wh-phrases in root clauses".

Jaeger's second hypothesis is formulated as:

(24) **Topics First! Hypothesis**

Like order constraints on the left-periphery of non-interrogatives, wh-phrase ordering (including so-called Superiority effects) in Bulgarian wh-question is (partly) determined by topicality.

When Topics appear in the left-periphery, it is only "the main" topic in Erteshik-Shir's terms (which is also nicely noted by Dyakonova (2009) that such a "special" Topic exists). Frascarelli & Hinterhölzl (2003) distinguish three types of topics, two of which are immediately relevant to our concerns:

(a) The ABOUTNESS TOPIC, which occupies the highest Topic position in the left periphery. Frascarelli & Hinterhölzl say: 'it is cognitively speaking important for such Topics to occur Tat the beginning of the sentence.' (cf. Lambrecht 1994: 194). ABOUTNESS Topics which are located in a *higher position with respect to WH/Focus* constituents.

(b) The FAMILIARITY TOPIC occupies the lowest TopP projection. FAMILIAR Topics are located lower than WH/Focus constituents and they can be realized in either peripheries.

Since FAMILIARITY TOPIC and FOCUS do not have a strict syntactic position in Russian, it is difficult to ascribe a rigid function to the hosting category. A similar analysis was proposed by Hale (1987) for Vedic Sanskrit and discussed in Kiparsky (1995). Hale identified two types of

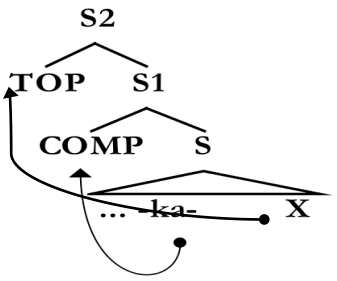
XP-fronting: 'weak' topicalization, which signals that the XP is used emphatically, and wh-movement. These two types of XP-movement target different positions at the left periphery. Hale assumes that the topic construction involves movement to a higher position, while wh-movement to a wh-landing site. Hale (1987) provides the following examples from Sanskrit:

(25) rátham kó nír avartayat
chariot$_{ACC}$ who down rolled
"Who rolled out the chariot?" (RV10.135.5; from Kiparsky 1995:154, Hale 1987)

(26) índrah kím asya sakhyé cakāra
Indra$_{TOP}$ what did do friendship$_{LOC}$
"What did Indra do in his friendship?" (RV 6.27. 1b from Hale 1987:42)

These examples show that similar constructions to the ones discussed have been analyzed similarly.

(27) Analysis of Sanskrit: Hale 1987: 41



Thus, the proposal put forward here is not entirely new, and does not contradict existing assumptions about the relationship between topicalization and wh-movement.

The next issue to consider is "Topics First! Hypothesis" (Jaeger 2004). If such hypothesis is put forward and HOP is, in fact, a category that hosts left-edge topics, we need to consider its uniqueness to appear in matrix clauses.

5.1.2 HOP as a Main Clause Phenomenon (MCP)

In chapter 1, we saw that in the Cartographic tradition, all clauses are expected to be structurally the same. While the proposed HOP category expands the CP-domain, it is not a Cartographic solution. I demonstrated that Russian exhibits asymmetries in the the allowed categories in

matrix and embedded clauses. I argued that HOP is available in matrix clauses, but is not available in embedded clauses. The easiest explanation for such an asymmetry would be to state that VP does not select HOP in the embedded clause. While this explanation holds, it is important to look into details of how clauses are constructed.

The unavailability of Topic material in the left-periphery in the embedded clause compared to the matrix clause is known from various languages, including Yiddish and German, for instance (Diesing, 2005:4):

> ….main clauses in Yiddish have a different structure from main clauses in asymmetric V2 languages (such as German), in which all V2 results from movement to CP. Furthermore, it can be shown that within Yiddish, matrix questions and embedded questions have a different structure. Evidence for this comes in part from the fact that they display different word orders.

Since Emonds (1969) it has been assumed that matrix clauses have a special status, and certain transformations can only take place on the left periphery of the clause. He called such phenomena "root transformations" (RT). Emonds proposed that RTs were subject to a syntactic restriction: they could apply only to root sentences - that is, S-nodes which themselves were either not dominated by any other nodes or dominated only by other S-nodes. Thus, the definitions of the *root clause* and a *root transformation* are given in (28) and (29)

(28) *Root sentence:* "a root will mean either the highest S in a tree, an S immediately dominated by the highest S or the reported S in direct discourse." (Emonds 1969: 6)

(29) *Root transformation*: "a RT is one in which any constituents moved, inserted or copied are immediately dominated by a root in the derived structure." (Emonds 1969: 7)

Hooper and Thompson (1973) (henceforth, H&T) provide a long list of RTs, including but not limited to VP preposing, complement preposing, adverb dislocation, tag question formation, negative formation verbal inversion, directional adverb preposing, right dislocation, left disclocation, and topicalization. Albrecht, Haegeman, and Nye (2012: 2) state that "the proper characterization of root transformations, or MCP as they are now widely referred to, and of the contexts in which they can occur, has been debated." Emonds (1969) claims that MCP are restricted to "assertive" structures. This seems to capture the distribution of the phenomenon in general; however, multiple researchers (H&T, Haegeman, Goldsmith 2004, Heycock 2006) agree that the exact understanding of what "assertive" structure entails remains unclear. H&T provide an alternative stating that "it is not appropriate to emphasize elements of a sentence whose

proposition is already known whose truth is presupposed, and whose content is relegated in the background" (H&T: 495). Thus, topicalization appears to be a good candidate for MCP.

While most of the other MCP items on the H&T's list are not available in Russian, topicalization and left dislocation are. Left dislocation is usually analyzed as clefting. Hooper and Thompson claim that while topicalization is a root transformation, clefting is not. Haegeman (2009) confirms that clefting is not a matrix clause phenomenon (MCP). H&T, analyzing exclusively English structures, state that Topicalization is an NP-NP-VP sequence, which cannot be generated by the phrase structure rules of English. Clefting is NP-V-VP-S $_{\text{RELATIVE}}$ sequence, which the phrase structure rules can easily generate. Thus they conclude, "Topicalization is an RT, and Clefting is not; Clefting is a structure preserving transformation in Emonds's terms" (Hooper & Thompson 1973:472, with examples in (30) and (31) respectively).

(30) Topicalization: This book you should read.

(31) Clefting: It's this book that you should read.

Thus, before going into more details about contexts in which MCP occur that might be relevant to current work, let us consider some Russian data that fits H&T's discussion of the RT phenomenon as a diagnostic. Consider the following examples:

(32) Ivan uslyšal ot druzej, [$_{CP}$ čto Maša ljubit Tolstogo].
Ivan heard from friends that Masha loves Tolstoy
Ivan heard from friends that Masha loves Tolstoy.

(33) [$_{CP}$Čto Maša ljubit Tolstogo], Ivan uslyšal ot druzej.
that Masha likes Tolstoy Ivan heard from friends

In (32) we see a sentence with a matrix clause and the embedded clause. When the embedded clause is moved over the matrix clause, the sentence remains ungrammatical (33). However, embedding such a construction as shown in (34) is impossible.

(34) *Vse govorjat, čto, [$_{CP}$čtoMaša ljubit Tolstogo], Ivan uslyšal ot druzej.
everyone says that that Masha loves Tolstoy Ivan heard from friends
"*Everyone says that that Masha loves Tolstoy Ivan heard from friends"

The sentence in (35) is a modification of (32) into an embedded question using the construction of a single WH discussed with an over topic occupying a HOP position in chapter 2. Example (36) shows that clause inversion is possible in such case, but embedding is not as shown in (37):

(35) Ivan **gde** uslyšal, čto Maša ljubit Tolstogo?
Ivan where heard that Masha loves
"Where did Ivan hear that Masha loves Tolstoy?"

(36) Čto Maša ljubit Tolstogo, Ivan **gde** uslyšal?
that Masha loves Tolstoy Ivan where heard
"That Masha loves Tolstoy, where did Ivan hear[this]?"

(37) *Vse govorjat čto čto Maša ljubit Tolstogo, Ivan **gde** uslyšal?
everyone says that that Masha loves Tolstoy Ivan where heard
"Everyone says that that Masha loves Tolstoy, where did Ivan hear?"

These examples show that the word order where wh-word appears after an overt topic is a root phenomenon. The examples here reiterating the examples used in Chapter 2 in the discussion of single wh.

Thus, we have established that HOP consisting structure in these examples appears in matrix clause, and its embedding leads to ungrammaticality. If we assume that at least some types of MCP is available in Russian as outlined in examples above, we can proceed to the discussion of what MCP entail. One of the relevant types of MCP is topicalization.

Miyagawa (2012) contributes to the discussion of MCP by analyzing data that sheds light on MCP, and uses only strictly unembeddable contexts (unlike H&T who discuss both RT and RT that are embeddable). For Miyagawa, root phenomenon information-structure related effects such as focus and topic are irrelevant. Root phenomenon in his view encodes the relationship between speaker and hearer. He claims that more "genuine" root phenomena "depend on availability of additional structural layer above the CP, which serves to anchor the utterance to speaker and to the discourse context" (Aelbrecht, Haegeman, Nye 2012: 9). To implement this proposal Miyagawa uses speech act projection (SaP) dominating CP. Therefore, we can see that very salient discourse items carrying strong speech act topics should appear in the position dominating CP. In this sense this is exactly what HOP is for: it is an "anchor of the utterance" to

context. Its uniqueness distinguishes it from any other discourse interpretative categories.

In addition, other linguistic phenomenon is said to pattern with Miyagawa's root phenomena. Several researchers (Munaro 2010, Haegeman, Hill 2010; Poletto and Zannutini 2010 among others) analyze edge clause particles as an instanitations of a structural layer dominating CP, which encodes 'anchoring of clause to the discourse'. The discussion of Chapter 4 of –TO is directly linked to these ideas: special discourse particles (-TO in Russian) emphasize the clausal link, and only appears at the clause edge.

From this discussion of MCP, we can see direct similarities between the structure postulated here as a matrix clause only category: HOP that hosts topicalized material in its Specifier, and a left edge clausal topic particle in its head. The entire category creates a domain outside CP in the matrix clause only and shares discourse encoding ideas. I discuss relevant semantic properties of items appearing in HOP below. Such a proposal is not new, and is attested to in other languages. Fine graining of MCP is still debated as it has been mentioned before. One aspect of the discussion is clear: topics can appear in the high category dominating CP. I have discussed the approaches to wh-topic fronting earlier, even though the occurrence of wh-phrase in a topic position is considered incompatible with MCP under most views.

Since it has been claimed that wh-phrases that occur in HOP might be considered D-linked, as well as non-wh-topics that appear in HOP are special "strong" or "aboutness" topics, it would be logical to suggest direction in identifying the possible semantic properties of this category.

5.2 Semantic Role of HOP

The idea that HOP hosts both overt topics appearing at the left–periphery and D-linked wh-phrases is discussed in the previous section. It is surprising that a wh-phrase can appear in a topic-like position (as was noticed by Jaeger for Bulgarian). It appears that such an idea is not new, and was suggested for other languages. The extension of the CP-domain with some kind of a syntactic shell category (or CP dominating category) falls out of the idea that this high category reflects discourse-related instantiations and encodes discourse information that is relevant to the speaker and the hearer (something that is not easily captured syntactically). In Russian, when a wh-phrase occupies the HOP position in a question, the question is discourse grounded. There is a difference in interpretation between (38) and (39):

(38) Kto kogo ubil?
who$_{NOM}$ who$_{ACC}$ killed
"Who killed whom?"

(39) Kogo kto ubil?
who$_{NOM}$ who$_{ACC}$ killed
"Who killed whom?"

The question in (38) has a generic interpretation of multiple pairs: for every *x* who was a person that killed *y*. In (39), however, the interpretation differs: which *y* is there that *x* killed, and might even have a set of *x* and *y* presupposed. In the following sentence (40), when HOP is occupied by an overt topic, no preexisting set interpretation is available, for neither of wh-phrases can move into HOP, and therefore, does not have any discourse grounding.

(40) Ivanu kto kogo predstavil?
Ivan$_{DAT}$ who$_{NOM}$ who$_{ACC}$ introduced
"Who introduced who to Ivan?"

The context for (40) is that there is a set of *x* and *y* where *x* introduced *y* to Ivan. The only discourse-linked element here is *Ivan*.

Therefore, the argument here is two-fold. The item appearing in HOP[+topic] satisfies two requirements: discourse saliency, and specificity. I discuss what it means for HOP to host "specific" items in the following section.

The items that HOP hosts can carry the feature of Specifity. Russian does not have an overt lexical item or a category to indicate Specifity. Specificity in Russian is retrieved from contextual, intonational, and antecedent cues. However, it is fair to state that HOP allows an overt realization of the feature of Specificity. Thus, the item that occupies the HOP position is specific. It has been noted in chapter 4 that the particle –TO that marks topicalization (and contrast) does not co-occur with indefinite pronouns due to a phonological redundancy. The examples of such a redundancy are given below in (41) and (42):

(41) *Kogo-to **to** on ubil.
someone$_{ACC}$ TO he killed
"There is someone that he killed."

(42) *Kogda-to **to** oni uznajut pravdu.
sometime TO they learn$_{FUT}$ truth
"Some day they will know the truth".

But it is not only a phonological redundancy that prevents such utterances; it is also a clash of features of the indefinite pronoun and the marking of Specificity. Indefinite pronouns do not have an antecedent in discourse, and therefore, cannot be specified. The intuition is that HOP might carry the feature of discourse-grounding through Specificity. The previous discussion on what elements appear in this position sheds some light on how the semantic analysis of this position should unfold; however, at this point the idea is embryonic since it requires additional data collection on the matter.

5.3 HOP cross-linguistically

It has been shown that HOP serves well for the analysis of Russian. The intuitions behind its postulation are not entirely novel based on the previous discussion of Bulgarian (Rudin 1985, Jaeger 2004). Therefore, one should question whether a category such as HOP appears in other languages. The discussion of high Bulgarian topics supports the idea that HOP can be a plausible analysis of the constructions evaluated by Jaeger (2004) on the one hand satisfying his "Topics First! Hypothesis". On the other hand, HOP might also be a useful device in the analysis of Serbian/Croatian, especially considering the contexts in which S/C exhibits Superiority. It is also not unlikely that such a high category can be applied to the analysis of Hungarian data as well (cf. Lipták 2001; Lipták, A.K. 2001). Its dichotic feature and operator status allows different items to be hosted there. Thus, it is not unlikely that HOP can be used for the analysis cross-linguistically. We can create a tentative language typology that would reflect the interaction of the existence of HOP in a language together with "strong" [+wh] feature, as well as the language's requirement for fronting of wh-phrases. The typology distinguishes between the languages with a strong [+wh] feature that has multiple wh-fronting (such as Bulgarian), with a strong [+wh] feature that has only one wh-phrase moving into CP (such as English) (CP-absorption languages in Richards' terms); and [-wh] feature with multiple wh fronting (such as

Serbian/Croatian) (an IP-absorption language), and wh-*in-situ* languages. Then, it proposes the existence of HOP category or lack thereof based on what we already know about these languages. Table 1 reflects the intuitions what positions what languages would occupy in such a typology.

TABLE 1. HOP/WH/FRONTING LANGUAGE TYPOLOGY

	+HOP		**-HOP**	
	+WH CP-absorption	**-WH IP-asborption**	**+WH CP-absorption**	**-WH IP-absorption**
+Multiple WH Fronting	Russian Bulgarian (?)	Serbian/Croatian Hungarian	Bulgarian (?)	Polish (?)
-Multiple WH Fronting	French	Japanese	English (-V2) German (+V2)	Chinese

Such a typology would assume that Bulgarian and Russian can, in fact, occupy the same place, and that Bulgarian uses HOP in discourse related contexts, where a wh-phrase or an overt topic preceding wh-phrases can occur (following Jaeger's ideas discussed above). Additional data are required to confirm the place where these languages would belong to in such a typology. However, this issue requires additional data and prediction investigations.

5.4 Chapter Conclusions

This chapter investigated some of the consequences of claims that have been made throughout this dissertation: HOP as matrix clause phenomenon, the occurrence of wh-phrases in a position that hosts topics, and issues with labeling the phrase "Operator Phrase". Several facts have been established. First, labeling HOP as an operator does not contradict its behavior with respect to quantifiers, and it can host elements that are considered operators: such as wh-phrases, so long as they are D-linked.

The fact that HOP only appears in matrix clauses has been discussed. The debate on MCP is ongoing: therefore, we can only rely on the facts that seem to be more or less well established

in a sense of what MCP constitute. However, despite Cartographic claims, matrix and embedded clauses differ in the list of categories in a derivation. And so the matrix-only presence of HOP contributes to the debate about Cartography on the side of the skeptics.

Lastly, I briefly speculated on the possibility that HOP appears in other languages, but not in all. The idea does not seem unfounded, especially after consideration of Bulgarian data following the general claim of chapter 2 that Russian patterns with Bulgarian. The possibility of utilizing HOP for Serbian/Croatian should be also investigated further for S/C exhibits Superiority in some contexts, and HOP can resolve inconsistencies in obedience to Superiority Condition

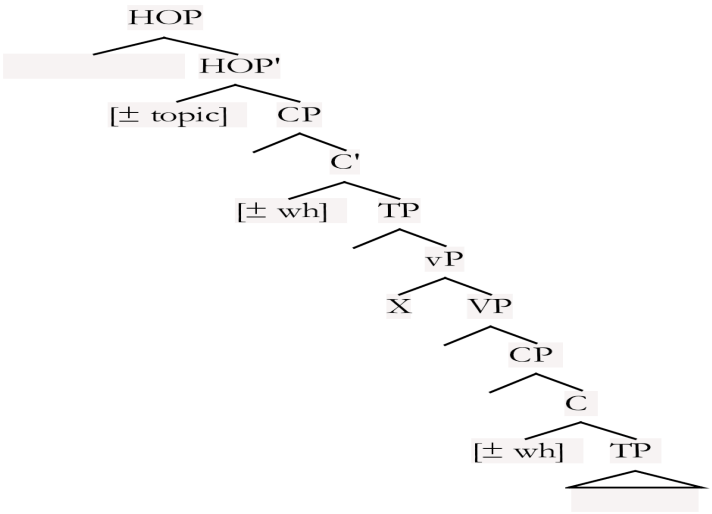
Thus, this chapter outlined the possible place of HOP cross-linguistically and the consequences of the postulation of such a category, as well as directions for further syntactic investigations in several domains.

Chapter 6. Conclusions and Directions

The current dissertation accomplishes two main goals: on the one hand, it provides an analysis of wh-behavior in Russian, and accommodates the structure proposed for this analysis into a general account of Russian word order, on the other hand, taking into consideration the notions of Information Structure. Accounting for Information Structure is traditionally done through understating how a structure reflects discourse related categories: various accounts exist. Their interplay is discussed in Chapter 1.

The account put forward here provides the geometry of the 'left-periphery' of Russian sentence structure, and is able to account for a number of constructions: multiple wh questions; sluicing, including multiple sluicing; multiple wh-coordinated questions; and topicalized constructions with and without the usage of a discourse particle. While most of these Russian constructions have been analyzed before (Stepanov 1998; Strahov 2001; Grebenyova 2005-2007; Gribanova 2009; McCoy 2001, 2002; to name a few), no account appears to encompass all of the phenomena. One of the advantages of the research here is that it simplifies the approach by claiming that all of these constructions can be unified by addressing the underlying structure of the Russian left-periphery. It claims that there is a functional category located in the high left field (over CP) in Russian that is able to host a number of items that are featurally diverse. The structure repeated below is the keystone for this account: it represents a geometrical arrangement of the Russian main clause left-periphery. The structure assumes a unique high HOP position (see 1). This position is different from the TopP* suggested in Rizzi (1997), for instance, in that it is not iterative. Thus, its uniqueness is literal. The evidence for its uniqueness is provided throughout this work, and with more detail in Chapter 5.

(1) HOP [±topic] over CP



One of the important properties of the position postulated is that it is only available in matrix clauses. I explicitly show resultant asymmetries between matrix and embedded clauses: the unique HOP is not available in the embedded clause. The unavailability of this position is based on selection properties: it cannot be selected by V. Its uniqueness is discussed as a specific property of a category that allows just one element to appear there.

Proposing an extra functional category in the domain of the left-periphery necessarily leads to considering parallels with the cartographic research program. I show that while possibly being a cartographic extension of CP, the view taken here is not cartographic in its nature. I consider an existing account of Russian word order presented within the cartographic framework (Dyakonova 2009), and show how it is designed so as to be successful in predicting any outcome and conclude that it is therefore unfalsifiable and fails. Hence, although I proposed a cartographic-looking structure, I argue that it is not cartographic acting, evidence for which I provide (Chapter 1), while still allowing me to account for a number of Russian word-order phenomena with a *minimum* of means (assuming that main theoretical tenets hold). This is an interesting discussion in itself: the sufficiency of Cartography as a research program, while the limitations on optionality of the Minimalist program, make accounts of IS challenging. Both require an alternative.

In this sense, HopP provides the correct architectural mapping of Russian syntactic structures with wh-movement, wh-related, and discourse-related utterances. These are discussed in detail in Chapters 2, 3, and 4.

By analyzing wh-constructions in Russian, I re-consider the existing Slavic wh-typology. The typology is based on the parameterized property of a language to follow or violate the Superiority Condition. Traditionally, Russian has been assigned into the group of languages that do not obey Superiority. Moreover, the fact of the existence of wh-movement in Russian at all has been doubted. The main claim I make here about multiple wh-behavior is that if one assumes wh-movement in Russian (and reasons are provided to support such a conclusion in Chapter 2), then Superiority in Russian must hold: it emerges in very specific and predicted environments. The reason it is commonly not seen in surface word orders is due to the availability of HOP. Each environment of the Emergence of Superiority is discussed: HOP is blocked by an overt non-wh-element, movement to HOP is blocked by the requirements of the phrase where the movement originates (ConjP, for instance), and HOP is not available, as in embedded clauses.

Several additional puzzles unfold here and a solution or a path to one are suggested. The first puzzle is unavailability of multiple wh-fronting of adjuncts. If one claims that 'true' wh-movement exists in Russian, then one should not expect that two adjuncts cannot front in a language (the same is true for Serbian/Croatian, but not for Bulgarian, for example). This is an important finding. The properties of proper government and late insertion of adjuncts (wh-adjuncts) are discussed in detail in Chapter 2 and 3. It is the proper government failure that disallows multiple wh-adjunct fronting. In order to solve this problem, I introduced the *Multiple WH-Adjunct Ban* (Chapter 2 (67)) that requires proper government of traces (formerly, ECP) when two or more adjuncts are present. The overall results of the Russian wh-behavior investigation can be summarized as follows:

(a) 'true' multiple wh-movement takes place in Russian ([+wh] feature is strong), wh-phrases multiply fill the specifier of CP;
(b) multiple wh-adjunct fronting is restricted in Russian;
(c) Superiority must hold in wh-contexts in Russian; though its effects may be masked in the main clause by the ability of either wh to raise further (to SpecHOP). This is an environment of the Emergence of Superiority.

(d) HOP position is available in matrix clauses only (the selection requirements are discussed).
(e) there is an asymmetry with respect to surface Superiority in matrix vs. embedded clauses: surface Superiority cannot be violated in embedded clauses, due to the lack of HOP

The re-analysis of Russian wh leads to the discussion of the phenomenon that is commonly related to wh-behavior in a given language: sluicing. Here, I specifically consider multiple sluicing (Takahashi 1993) and provide an account of it. The account becomes a theoretically easier enterprise if true Bulgarian-style multiple-wh movement into Specs of CP is claimed for Russian. This allows us to reduce an analysis of sluicing to the standard IP-deletion ellipsis analysis (Ross 1969, Merchant 2001). However, the multiple sluicing data offer yet another puzzle – coordinated multiple wh-phrases (CMW) in the sluice. While CMW has received active attention in recent years, the issue of coordinated sluicing was not addressed. On the one hand, the approach to coordinated multiple sluicing should logically be based on an approach to CMW in a language. On the other hand, the correspondence between multiple wh-fronting and multiple wh-coordination is questionable (Citko and Gračanin-Yuksek 2010 for Slavic vs. Gribanova 2009 for Russian). It is important to clarify in which ways CMW are related and in which ways this process differs from multiple wh-fronting.

First, I stress that it is crucial to recognize that Russian (on a par with other languages) exhibits three types of multiple coordinated wh-behavior: coordination of two (or more) wh-arguments (Type 1), coordination of two (or more) wh-adjuncts (Type 2), and coordination of wh-adjuncts and wh-arguments (the Mixed Type). Secondly, I put forward several ideas on how each type can be analyzed.

The CMW approach here suggests that each type should be accounted for separately: with the proposal from earlier works of ConjP-coordination at CP. The mechanism of how wh-phrases end up in ConjP has to be revised, however. Following Nuñes 2001, I fill in details of the Sideward Movement mechanism that takes place: this is an issue not clearly outlined in previous analyses of Russian CMW. I also consider the possibility the utilization of the backwards sluicing approach (Giannakidou & Merchant 1998) for the coordination of multiple wh-adjuncts, the fronting of which is only possible through coordination (recall that multiple wh-adjunct fronting is banned). I also discuss availability of various readings of CMW: pair-list (PL) vs. single-pair (SP). I claim that the availability of readings in multiple whs and CMW are identical

despite previous statements. Thus, I show that the CMW analysis based primarily on the availability of readings (PL/SP) recently proposed for Russian by Gribanova (2009) works successfully only with respect to coordination of two wh-arguments: Type 1 CMW. At least with the evidence and mechanisms that she provides. Her analysis, however, misses a few important generalizations, such as the typology of coordination, and the presence of PL interpretation.

Applying the mechanism of Sideward Movement with explanation on how the derivation is achieved sheds light on CMW in Russian. Plus, I show that backwards sluicing is an option for the coordination of adjunct wh-phrases (argued against by Kazenin 2002) due to the fact that the elided parts are identical. Since ConjP is postulated here, Superiority effects are expected to hold analogously to multiple wh-fronting facts: they emerge where one would expect it, and are violated where Superiority is not expected (two adjuncts). HOP is unreachable for wh-phrases out of ConjP (constraint on extraction from a coordinate structure).

Therefore, a careful investigation of CMW allows to provide an account for them at the site of ellipsis. While sluicing itself is reduced to the Merchant style IP-deletion process, the coordination takes place at the ellipsis site following the mechanisms discussed for CMW before ellipsis takes place. These mechanisms are discussed in Chapter 3 in length.

In the final part of this work, I move away from strictly wh-phenomena and consider another type of constructions that provide more evidence that some kind of HOP-category must exist in Russian to accommodate it. I consider constructions with particle –TO-marking. This particle has been given rather broad attention in the semantics and discourse literature, but does not have a satisfactory syntactic analysis. This particle has historically grammaticalized and there are several homonymous remnants functioning differently. This makes its analysis more challenging. Here, I proposed that the particle occupies the head of the high HOP position of the matrix clause. When –TO is present in the head of HOP, HOP spreads [+topic] feature (Rizzi 1990). Topicalized elements move to SpecHOP to receive –TO marking. Due to the fact that only one unique HOP phrase is available, only one –TO per clause is available. Elements that are –TO-marked must move to the very left edge of the main clause to appear before the particle. I show that elements that are -TO-marked must create a constituent. I illustrate that -TO can attach to virtually any XP, to satisfy the discourse requirement to mark Topicality. The analysis provided in the chapter shows that the particle is sensitive to restrictions on movement: island conditions. On the other hand, it can be extracted out of the embedded clause in the

environments where movement is not blocked. Its distribution with verbal elements is problematic for verbs do not raise in Russian. I demonstrate that HOP can bind elements, and by assuming clausal pied-piping (in the spirit of Arregi 2003) in certain contexts this particle can appear attached to an element on the very right edge of the clause. Its distribution in embedded clauses is restricted as expected since unique HOP position is not available in embedded clauses: root/embedded asymmetry is observed here once again. The Russian language has a number of particles, but -TO occupies a unique, literally and figuratively, position in Russian syntax.

As a result of this work several important theoretical issues have surfaced. (i) The first one is related to what Superiority is and what the exact mechanisms behind Superiority are? (ii) The second one is more applied and has to do with how and why adjuncts are different from arguments. (iii) And last but not least, whether the account put forward here can be applied to other languages: how universal is the proposal here?

All three of these questions require more theoretical and empirical investigations. The response to the last question should ideally be affirmative. The structure proposed here should theoretically be applicable to other languages such as Serbian/Croatian, and perhaps, the asymmetry between matrix and embedded clauses in this language lies exactly in the availability of such a position. It has been tested (as a part of this work) that Serbian/Croatian unlike Bulgarian does not allow high topicalized elements. Recall that this position has a [-topic] feature, and perhaps, never hosts topics in S/C; however, it can provide an escape path out of Superiority for wh-phrases.

The second issue of wh-adjunct behavior is puzzling. The question why some languages would have a multiple wh-adjunct ban (Russian, S/C), while others would not (Bulgarian) is not an easy one to tackle. Notice, that multiple wh-adjunct fronting is also an environment where multiple wh-fronting (if coordinated) can occur in English. This particular issue requires a thorough cross-linguistic investigation.

The ideas of the universality of the account that claims Emergence of Superiority in certain contexts cannot be complete without addressing question (i) about Superiority. Research in this area should concentrate along the following lines: In Minimalism, Superiority reduces to closest (or "minimally linked"). Richards (1997) derives the lack of Superiority through IP-absorption "assuming that Superiority constrains A-bar movement but not A- movement" while Bruening (2001) assumes that there is a "generalized [optional] P-feature of C. If a language makes use of

this feature, movement takes place in a way that obeys Superiority [but not] if the language does not make use of it." Both claims weaken UG unnecessarily. Then, one might claim that Superiority is unparameterized (despite apparent violations), and economy conditions remain exceptionless. This would be an important advance in syntactic theory.

References

Antonyuk-Yudina, S. 2006. The cope of Quantifier Phrases in Russian: A QR analysis. *Linguistics in the Big Apple*

Abels, K. 2001. The Predicate Cleft Construction in Russian. *Formal Approaches to Slavic Linguistics 9*. pp.1-19. Ann Arbor: Michigan Slavic Publications

Abels, K. 2004. Right Node Raising: Ellipsis or ATB movement? *North East Linguistics Society, 34:* 44-59

Abels, K. 2009. Some implications of improper movement for cartography. In *Alternatives to Cartography*, J. Craenenbroeck (ed.). pp. 325-360. Berlin, Germany: de Gruyter Mouton

Aelbrecht L, Haegeman, L. and Nye, R. 2012. *Main Clause Phenomena: New horizons*. Cambridge University Press

Arregi, K. 2003. Clitic Left Dislocation is Contrastive Topicalization. In *Proceedings of the 26th Annual Penn Linguistics Colloquium*, Penn Working Papers in Linguistics 9.1:31-44. Philadelphia: Penn Linguistics Club

Babby, L. 1975. A Transformational Analysis of Transitive -sja Verbs in Russian. *Lingua 35*: 297-332

Bailey, L. R. 2010. Sentential word order and the syntax of question particles. *Newcastle Working Papers in Linguistics 16*: 23-43

Bailyn, J.F. 1995. A Configurational Approach to Russian "Free" Word Order. Doctoral Dissertation, Cornell University

Bailyn, J. F. 2003. A (purely) Derivational Approach to Russian Scrambling. In B. Partee et. al. (eds), *Formal Approaches to Slavic Linguistics 11*. University of Michigan Press

Bailyn, J. F. 2004. Generalized Inversion. *Natural Language and Linguistic Theory 22*:1-49

Bailyn, J. F., 2011. *The Syntax of Russian*. Cambridge University Press

Bailyn, J. F., 2011 Alternatives to cartography: Review. *Language 87* (3)

Bánréti, Z. 1992. A mellérendelés. [Coordination]. In Kiefer, F (ed.) *Strukturális magyar nyelvtan. Mondattan. [The structure of Hungarian Syntax]*, pp.715-796. Budapest: Akadémiai Kiadó

Bayer, J. , Obenauer H-G., 2011. Discourse particles, clause structure, and question types. *The Linguistic Review 28*: 449-491

Belletti, A. 2001. Inversion as Focalization. In A. Hulk and J.Y.Pollock (eds.), *Subject Inversion in Romance and the Theory of Universal Grammar*, pp. 60-90. New York: Oxford University Press

Belletti, A. (ed.) 2004. Structures and beyond: The cartography of syntactic structures, vol. 3. Oxford: Oxford University Press

Biberauer, T, Holmberg, A. & Roberts, I. 2009. Linearization and the Architecture of Grammar: A view from the Final-over-Final Constraint. In V. Moscati & E. Servidio (eds.) *Proceedings of the Incontro di Grammatica Generativa XXXV*

Bitextin, A. B. 1994. Chastitsy -TO, ZHE, VED' i vvodnye konstruktsii tipa KAK IZVESTNO kak sredstva ukazanija na izvestnost' propozitsional'nogo soderzhanija predlozhenija slushajuschemu. Doctoral Dissertation, Moskovskij gosudarstvennyj universitet (Moscow State University), Moscow, Russia

Boeckx, C., Grohmann K. (eds.). 2003. *Multiple wh-fronting*. John Benjamins: Amsterdam

Bonnot, C. 1986. Emplois de la particule -to. In *IVe coloque de linguistique russe*, pp. 21-30. Toulous: Service des Publications U.T.M

Bonnot, C., 1988. To particule de rappel et de thématisation. *Les particules énonciatives en russe contemporain*(2). Paris: Institut D'Études Slaves

Bonnot, C. 1990. La particule -to et la polémique chachée en russe moderne: A propos du statut énuonciatif du thème. *Revue Des Études Slaves*, LXII(1-2): 67-75

Bonnot, C. 1999. Pour une définition formelle et fonctionnelle de la notion de thème (sur l'exemple du russe moderne). In Cl. Guimier (éd.) *La thématisation dans les langues, Actes du colloque de Caen*, pp. 15-31. 9-11 octobre 1997, Bern, Berlin, Frankfurt/M., New York, Paris, Wien, Peter Lang

Bošković, Ž. 1994. ECP, Spec-Head Agreement, and Multiple Wh-Movement in Overt Syntax, In M. Lindseth, S. Franks (eds.) *Proceedings of Formal Approaches to Slavic Linguistics II*, pp. 119-143. Ann Arbor, Michigan Slavic Publications Bloomington, Indiana.

Bošković, Ž. 1996. Fronting Wh-phrases in Serbo-Croatian. In M. Lindseth and S. Franks (eds.) *Proceedings of Formal Approaches to Slavic Linguistics V*, pp. 86-107. Ann Arbor, Michigan Slavic Publications

Bošković, Ž. 1997a. On Certain Violations of The Superiority Condition, AgrO, and Economy of Derivation. *Journal of Linguistics*, *33*: 227-254

Bošković, Ž. 1997b. Superiority effects and multiple wh-fronting in Serbo-Croatian. *Lingua* 1-2:1-20

Bošković, Ž. 1998a. On the interpretation of multiple questions. In J. Fodor, S.J. Keyser & A. Brand (eds.) *A Celebration: Essays for Noam Chomsky"s 70th birthday*

Bošković, Ž.1998b. Wh-phrases and Wh-Movement in Slavic. *Comparative Slavic Morphosyntax*

Bošković, Ž. 1999. On Multiple Feature Checking: Multiple WH Fronting and Multiple Head Movement. In S.D. Epstein and N. Hornstein (eds.) *Working Minimalism*, pp. 159-188. MIT Press, Cambridge, MA

Bošković, Ž. 2001. Sometimes in [Spec,CP], sometimes in-situ. In R. Martins, D.Michaels, J. Uriagereka (eds.) *Step by step. Essays on minimalism in honour of Howard Lasnik,* pp. 53-87. Cambridge Mass. MIT Press

Bošković, Ž. 2002. Multiple wh-fronting. *Linguistic Inquiry* 33:351-383

Boskovic, Ž. 2003. On Wh-islands and obligatory Wh-movement contexts in South Slavic C. Boeckx and K. Grohmann (eds.) *Multiple wh-fronting*, pp. 27-50. Amsterdam: John Benjamins

Bošković, Ž. 2004. Topicalization, focalization, lexical insertion, and scrambling. *Linguistic Inquiry 35*:613-638

Boskovic, Ž. 2005, On the locality of left branch extraction and the structure of NP. *Studia Linguistica 59*: 1-45

Boškovic, Ž. 2011. Last Resort with Move and Agree in Derivations and Representations. In C. Boeckx (ed.) *The Handbook of Linguistic Minimalism*

Bošković, Ž., Takahashi, D.1998. Scrambling and Last Resort. *Linguistic Inquiry 24:*1-23

Bouchard, D., 2009. A solution to the conceptual problem of cartography. In J. Craenenbroeck (ed.). *Alternatives to Cartography*, pp. 245-274. Berlin, Germany: de Gruyter Mouton

Brody, M. 1990. Some remarks on the Focus field in Hungarian. *UCL Working Papers in Linguistics 2*: 201-226. UCL

Browne, W. 1972. Conjoined questions and the limitation on English surface structure. *Linguistic Inquiry 3*, 223–226

Bruening, Benjamin. 2001. QR obeys superiority: Frozen scope and ACD. *Linguistic Inquiry 32* (2): 233-273.

Camacho, J. 2003. The Structure of Coordination: Conjunction and Agreement Phenomena in Spanish and Other Languages. *Studies in Natural Language and Linguistic Theory, 57*. Dordrecht: Kluwer

Cardinaletti, A. 2007. Für eine syntaktische Analyse von Modalpartikeln. In E. M. Thüne, F. Ortu (eds.) *Gesprochene Sprache und Partikeln*, pp. 89-101. Frankfurt, Peter Lang

Chafe, W.L.1976. Givenness, contrastiveness, definiteness, subjects, topics and point of view. In Charles N. Li (ed.) *Subject and Topic*, pp. 27-55. New York, Academic Press

Chao, W., 1987. On Ellipsis. Doctoral Dissertation. University of Massachusetts, Amherst

Chomsky, N. 1957. *Syntactic Structures*. The Hague, Mouton

Chomsky, N.1973. Conditions on transformations. S.R. In Anderson & P. Kiparsky (eds): *A Festschrift für Morris Halle*, pp. 232-286. New York. Rinehart &Winston

Chomsky, N. 1993. A Minimalist Program for Syntactic Theory. In K. Hale & S. J. Keyser (eds.) *The View from Building 20*, pp. 1-52. MIT Press, Cambridge, MA

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

Chomsky, N. 2001. Derivation By Phase. In Michael Kenstowicz (ed.) *Ken Hale: A Life in Language*, pp. 1-52. Cambridge, MA: MIT Press

Chomsky, N. 2004. Beyond Explanatory Adequacy. In A. Belletti(ed.), *Structures and beyond: The cartography of syntactic structures,* 3, pp. 104–131. Oxford: Oxford University Press

Chung, S., Ladusaw, W. A., McCloskey, J. 1995. Sluicing and logical form. *Natural Language Semantics 3:* 239-282

Cinque, G. 2005. Deriving Greenberg's universal 20 and its exceptions. *Linguistic Inquiry 36*: 315–32

Cinque, G., Rizzi, L. 2008. The Cartography of Syntactic Structures. *Studies in Linguistics 2*, CISCL 2008 Working Papers.

Cinque, G. 2004. "Restructuring" and Functional Structure. In A. Belletti (ed.) *Structures and Beyond. The Cartography of Syntactic Structures*, 3, OUP, New York

Citko, B. 1997. On Multiple Wh-movement in Slavic. In Z. Boškovič, S. Franks and W. Snyder (eds.) *Proceedings of Formal Approaches to Slavic Linguistics VI*, pp. 97-113. Ann Arbor, Michigan Slavic Publications

Citko, B. 2008. An argument against assimilating appositive relatives to coordinate structures. *Linguistic Inquiry, 39*: 633-655

Citko, B., Gračanin-Yuksek M. 2010. Towards a New Typology of Coordinated Wh-Questions. Ms.

Citko, B., Grohmann, K. 2001. The non-uniqueness of multiple wh-fronting. German=Bulgarian. In Franks, S. & M. M. Yadroff (eds.) *Formal Approahces to Slavic Linguistics*, pp. 117-136. Ann Arbour, Michigan. Michigan Slavic Publications

Collins, C. 2002. Eliminating labels. In S. Epstein, D. Seely (eds.) *Derivation and Explanation in the Minimalist Program*, pp. 42-64. Malden, MA: Blackwell

Comrie, B. 1988. Topics, Grammaticalized Topics, and Subjects. In S. Axmaker, Shelley, A. Jaisser, H. Singmaster (eds.) *General Session and Parasession on Grammaticalization,* Berkeley Linguistics Society, Berkeley, CA

Coniglio, M. 2007. "Deutsche Modalpartikeln: ein Vorschlag zu ihrer syntaktischen Analyse". In E. M. Thüne, F. Ortu (eds.) *Gesprochene Sprache und Partikeln*, pp. 103-113. Frankfurt, Peter Lang

Coniglio, M. 2006. German Modal Particles in the Functional Structure of IP. *Working Papers in Linguistics 16*:57-95. University of Venice

Craenenbroeck, J., Liptàk, A. 2008. On the Interaction between Verb Movement and Ellipsis: New Evidence from Hungarian. In C. B. Chang and H. J. Haynie (eds). *Proceedings of the 26th West Coast Conference on Formal Linguistics*, pp.138-146. Cascadilla Proceedings Project, Somerville, MA

Den Dikken, M., Giannakidou A. 2002. From hell to polarity: aggressively non-D-linked wh-phrases as polarity items. *Linguistic Inquiry 33*:31-61

Diesing, M. 2005. The Upper Functional Domain in Yiddish. In W. Abraham (ed.), *Focus on Germanic Typology, Studia Typologica 6*, pp. 195-209. Berlin: Akademieverlag

Domínguez, L. 2004. *Mapping Focus: The syntax and Prosody of Focus in Spanish.* Boston: Boston University

Dyakonova, M. 2009. *A phase-based approach to Russian free word order*. Doctoral Dissertation, LOT Publications

Emonds, J. 1969. *"Root and Structure-preserving Transformations."* Ph.D. diss., Massachusetts Institute of Technology.

Erteschik-Shir, N. 1997. *The Dynamics of Focus Structure.* Cambridge: Cambridge University Press

Erteschik-Shir, N. 2007. *Information Structure: the Syntax-Discourse Interface.* New York: Oxford University Press

Fiengo, R., May R.1994. Indices and Identity. Oxford, MA: MIT Press

Fox, D. and H. Lasnik. 2003. Successive cyclic movement and island repair: The difference between Sluicing and VP Ellipsis. *Linguistic Inquiry 34:*143-154

Frascarelli, M & Hinterhölzl, R. 2003. "Types of Topics in German and Italian." *Workshop on Information Structure and the Architecture of Grammar: A Typological Perspective, February 1-2, 2004*

Frascarelli, M., Hinterhölzl, R., 2007. Types of topics in German and Italian In S. Winkler and K.Schwabe (eds.) *On Information Structure, Meaning and* Form, pp. 87-116. Amsterdam: John Benjamins

Gallego, A., 2009. Phases and variation : exploring the second factor of the faculty of language. In J. Craenenbroeck (ed.) *Alternatives to cartography*, pp. 109-152. Berlin: Mouton de Gruyter

Gelderen van, V. 2003. Scrambling Unscrambled. Doctoral Dissertation, University of Leiden

Giannakidou, A., Merchant, J. 1998. Reverse sluicing in English and Greek. The *Linguistic Review 15*

Gračanin-Yuksek, M. 2007. About Sharing. Doctoral dissertation, MIT

Grebenyova, L. 2004. Interpretation of Slavic Multiple Wh-Questions. In *Proceedings of the Annual Workshop on Formal Approaches to Slavic Linguistics XII*

Grebenyova, L. 2005. Sluicing and Left-branch Extraction out of Islands. In *Proceedings of the West Coast Conference on Formal Linguistics 24*

Grebenyova, L. 2006. Sluicing Puzzles in Russian. In *Proceedings of the Annual Workshop on Formal Approaches to Slavic Linguistics 14*

Grebenyova, L. 2007. Sluicing in Slavic. *Journal of Slavic Linguistics 15*(1): 49-80

Grenoble, L. A. 1998. *Deixis and information packaging in Russian discourse.* Philadelphia: John Benjamins

Grewendorf, G. 2001. Multiple Wh-Fronting. *Linguistic Inquiry 32*: 87–122

Gribanova, V. 2009. Structural adjacency and the typology of interrogative interpretations. *Linguistic Inquiry 40 (1).* 133-154

Grohmann, K. 1997. German Superiority. *Groninger Arbeiten zur germanistischen Linguistik 40*: 97-107

Groos, A., van Reimsdijk, H. 1981. The Matching Effects in Free Relatives: A Parameter of Core Grammar. In A. Belletti, L.Brandi , L. Rizzi (eds.) *Theory of Markedness in Generative Grammar*, pp. 171-216. Pisa: Scuola Normale Superiore

Hale, M. 1987. Studies in the comparative syntax of the oldest Indo-Iranian languages. Doctoral Dissertation. Harvard University, Cambridge, MA

Heycock, C. 2002. "*Embedded Root Phenomena*." Ms. University of Edinburgh.

Higginbotham, J., May. R. 1981. Questions, quantifiers, and crossing. *The Linguistic Review 1*: 41–79

Hooper, J., and Thompson S. 1973. "On the Applicability of Root Transformations." *Linguistic Inquiry 4*: 465-97

Horvath, J. 1995a. Structural Focus, Structural Case, and the Notion of Feature-Assignment. In K. É. Kiss, ed., *Discourse Configurational Languages*, Oxford Studies in Comparative Syntax, pp. 28-64. New York/Oxford: Oxford University Press

Horvath, J. 1995b. Partial Wh-Movement and Wh "Scope-Markers". In I. Kenesei, ed., *Approaches to Hungarian Vol.5*, pp. 89-122. Szeged, Hungary: JATE Press

Huang, C -T J. 1982. Logical Relations in Chinese and the Theory of Grammar Doctoral Dissertation, MIT, Cambridge, MA

Jackendoff, R. 1977. *X-Bar Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press

Jaeger, T. F. 2004. Topicality and superiority in Bulgarian wh-questions. In *Proceedings of FASL XII: The Ottawa Meeting 2003*, pp. 207–228. Ann Arbor: Michigan Slavic Publications

Johnson, K. 1991. Object Positions. *Natural Language and Linguistic Theory 9*: 577-636

Junghanns, U., Zybatow, G. 1997. Syntax and Information Structure of Russian Clauses. In Browne, E.W. et al. (eds), *Proceedings of FASL* XXXIX, pp. 289-319. Ann Arbor: Michigan Slavic Publications

Kazenin, K. 2001/2002. On coordination of wh-phrases in Russian. Ms., University of Tuebingen

King, T.H. 1995. *Configuring topic and focus in Russian*. Stanford, CA: CSLI Publications.

Kiparsky, P. 1995. Indo-European Origins of Germanic Syntax. In Ian Roberts and Adrian Battye (eds.), *Clause Structure and Language Change*, p. 140-167. Oxford University Press

Kiss, K.E 1995. *Discourse Configurational Languages: Introduction*. In K. E. Kiss (ed.), Discourse Configurational Languages, pp. 3-27 Oxford: Oxford University Press

Kiss, K.E. 1998. Identificational Focus versus Information Focus. *Language 74*: 245-273

Kiss, K.E. 2002. EPP in a Topic-Prominent Language. In P. Svenonius (ed.), *Subjects, Expletives, and the EPP*. New York: Oxford University Press, 107- 124

Kiss, K. E. 2007. Topic and Focus: Two Structural Positions Associated with Logical Functions in the Left Periphery of the Hungarian Sentence. In Féry, C., G. Fanselow, and M. Krifka (eds.): *The Notions of Information Structure* (Interdisciplinary Studies on Information Structure 6): 69–81

Kitahara, H. 1997. *Elementary Operations and Optimal Derivations*. Cambridge, MA MIT Press

Koizumi, M. 1994. Nominative objects: The role of TP in Japanese. In Masatoshi Koizumi, Hiroyuki Ura (eds.) *Formal Approaches to Japanese Linguistics 1*, pp. 211-230. MIT Working Papers in Linguistics 24. MIT, Cambridge, MA

Koopman, H., Sportiche, D. 1982. Variables and the Bijection Principle. In J. Kaye, H. Koopman and D. Sportiche (eds.) *Projet sur les langues kru: Premier rapport*, pp.176-202. Université du Quebec a Montreal

Kratzer, A. 1994. The representation of focus. In A. von Stechow & D. Wunderlich (eds.) *Handbook of Semantics*, pp. 825-834. Berlin: Walter de Gruyter

Kratzer, A. 1998. Scope or Pseudoscope? Are there Wide-scope Indefinites? In S. Rothstein (ed.) *Events in Grammar,* pp. 163-196. Dordrecht: Kluwer

Krifka, M. 1992. A compositional semantics for multiple focus constructions. In J. Jacobs (ed.) *Informationsstruktur und Grammatik*, pp. 17-53. Opladen: Westdeutscher Verlag

Kuno, S., Robinson, J. 1972. Multiple Wh Questions. *Linguistic Inquiry 3*(4): 463-487

Lambova, M. 2002. On A'-movements in Bulgarian and their interaction. *The Linguistic Review 18*: 327-374

Lambrecht, K. 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representation of Discourse Referents*. Cambridge: Cambridge University Press

Lasnik, H. 1999. Pseudogapping puzzles. In S. Lappin and E. Benmamoun (eds.) *Fragments: Studies in Ellipsis and Gapping*, pp. 141-174. Oxford: Oxford University Press

Lasnik, H. 2001. When can you save a structure by destroying it? In M. Kim and U. Strauss (eds*.) Proceedings of the North East Linguistic Society 31* (2) GLSA, pp. 301-320

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

Levin, L. S. 1982. Sluicing: A lexical interpretation procedure. In J. Bresnan (ed.) *The mental representation of grammatical relations*, pp. 590-654. Cambridge, MA: MIT Press

Liakin, D. 2005. On Wh-Movement and Wh-in-situ in Russian. In S. Franks, F. Gladney, and M. Tasseva-Kurktchieva (eds.), *Proceedings FASL XXXIII*: *the Columbia Meeting 2004*, pp. 218–229. Ann Arbor, MI: Michigan Slavic Publications

Lipták, A.K. 2001. The Syntax of Wh-Items in Hungarian. Doctoral Dissertation, University of Leiden

Lipták, A. K. 2003. Conjoined questions in Hungarian. In Cedric Boeckx and Kleanthes Grohmann (eds.) *Multiple wh-fronting*, pp. 141–160. Amsterdam: John Benjamins

Lipták, A.K. 2009. On the correlative nature of Hungarian left peripheral relatives. In Shear et al., (eds.) *Dislocated elements in discourse,* pp. 398-430. London: Routledge

Lipták, A.K. 2011. The structure of the topic field in Hungarian. In: Beninca, P & Munaro, N (eds.) *Mapping the left periphery*, pp. 163-200. Oxford University Press

Lipták, A. K., van Craenenbroeck, J. 2006. The cross-linguistic syntax of sluicing: evidence. *Syntax, 9* (3): 248-274

Lobeck, Anne. 1995. *Ellipsis: Functional heads, licensing, and identification*. New York, NY: Oxford University Press

Marshall, T. A., 2002. Connotations and functions of Russian discourse markers ved', zhe and -to. Doctoral Dissertation, Cornell University

McCoy, S. G. 2001. Colloquial Russian particles -TO, ZHE, and VED' as set-generating ("kontrastive") markers: A unifying analysis. Doctoral dissertation, Boston University

McCoy, S. G., 2002. Semantic and discourse properties of colloquial Russian construction of the form "X-to X, a..." *Glossos(3)*

Meibauer, J. 1994. Modaler Kontrast und konzeptuelle Verschiebung: Studien zur Syntax und Semantik deutscher Modalpartikeln. In Linguistische Arbeiten, 314. Tübingen, Max Niemeyer Verlag

Merchant, J. 1998. 'Pseudosluicing': Elliptical clefts in Japanese and English. In Alexiadou et al. (eds.). *ZAS Working Papers in Linguistics 10*. Zentrum für Allgemeine Sprachwissenschaft: Berlin

Merchant, J. 2001a. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford University Press: Oxford

Merchant, J. 2001b. Islands and LF-movement in Greek sluicing. *Journal of Greek Linguistics 1*(1):39-62

Merchant, J. 2005. Revisiting syntactic identity conditions. Handout, Workshop on Ellipsis, University of California, Berkeley, October 8, 2005

Merchant, J. 2006. Sluicing. In M. Everaert and H. van Riemsdijk (eds.), *The Syntax Companion*, pp. 269-289. Blackwell: London

Merchant, J. 2007. Spurious coordination in Vlach multiple wh-fronting. MALC Handout

Miyagawa, S. 1987. Restructuring in Japanese. In T. Imai & M. Saito (eds*.) Issues in Japanese Linguistics*, pp. 273-300. Foris Publications

Miyagawa, S. 2012. Agreements that occur mainly in the main clause. In *L. Aelbrecht, L. Haegeman, and R. Nye* (eds.) *Main Clause Phenomena: New horizons.* Cambridge University Press

Molnár, V. 2001. Contrast from a contrastive perspective. In I. Kruijff-Korbayová & M. Steedman (eds.), *ESSLLI 2001 Workshop on Information Structure, Discourse Structure and Discourse Semantics*

Molnár, V. 2006. On Different Kinds of Contrast. In V. Molnár and S. Winkler (eds.) *The Architecture of Focus*, pp. 197-234. The Hague: Mouton de Gruyter

Müller, G., Sternefeld, W. 1993. Improper Movement and Unambiguous Binding. *Linguistic Inquiry 24:* 461–507

Munaro, N. 2004. ."Computational puzzles of conditional clause preposing". In R. Delmonte & A. DiSciullo, (eds.), *UG and External Systems*, John Benjamins: Amsterdam/Philadelphia

Neeleman, A., Titov, E. de Koot, H., Vermeulen, R. 2009. A syntactic typology of topic, focus, and contrast. In J. Craenenbroeck (ed.). *Alternatives to cartography*, pp. 15-52. Berlin: Mouton de Gruyter

Nikolaeva, T. M. 1985. Funktsii chastits v vyskazyvanii na materiale slavianskikh jazykov. Moskva: Nauka

Nishiguchi, S. 2007 `Multiple Sluicing' in Japanese and the Functional Nature of wh-Phrases. *Journal of East Asian Linguistics, 7(2*): 121—152

Nuňes, J. 2001. Sideward movement. *Linguistic Inquiry 32*: 303–344

Nuňes, J. 2004. *Linearization of chains and sideward movement.* Cambridge, MA: MIT Press

Olsvay, Cs. 2004. The Hungarian Verbal Complex: An Alternative Approach. In K. É Kiss & H. van Riemsdijk (eds.) *Verb clusters: A study of Hungarian, German and Dutch*, pp. 290–333. Amsterdam: John Benjamins

Pereltsvaig, A. 2004. Topic and Focus as Linear Notions: Evidence from Italian and Russian. *Lingua 114*: 324-344

Pereltsvaig, A. 2007. On the Universality of DP: A View from Russian. *Studia Linguistica 61*(1): 59-94

Pesetsky, D. 1987. Wh-in-situ: Movement and Unselective Binding. In E.Reuland and A. Meulen (eds.), *The Representation of (In)definiteness*. Cambridge

Pesetsky, D. 2000. *Phrasal movement and its kin*. MIT Press

Pesetsky, D., Torrego, E. 2007. The Syntax of Valuation and the Interpretability of Features. In S. Karimi, V. Samiian and W. Wilkins (eds.), *Phrasal and Clausal Architecture: Syntactic Derivation and Interpretation*, pp. 262-294. Amsterdam: John Benjamins

Prince, E. 1981. Toward a taxonomy of given-new information. In P. Cole (ed.) *Radical Pragmatics*, pp. 223-56. NY: Academic Press

Prince, E. 1998. On the limits of syntax, with reference to Left-Dislocation and Topicalization. In P. Culicover and L. McNally (eds). *Syntax and semantics. Vol. 29*. The limits of syntax, pp. 281-302 NY: Academic Press

Progovać, L. 1996. Clitics in Serbian/Croatian: Comp as the Second Position. In A. Halpern and A. Zwicky (eds.) *Approaching Second: Second Position Clitics and Related Phenomena*, pp. 411-428. CSLI Publications, Stanford University

Reinhart, T. 1982. Pragmatics and Linguistics: an Analysis of Sentence Topics. *Philosophica 27:* 53-94

Richards, N. 1997. What Moves Where When in Which Language? Doctoral Dissertation. MIT

Richards, N. 2001. *On movement in language*. Oxford: Oxford University Press.

Riemsdijk, H.van, 1998. Head Movement and Adjacency. *Natural Language and Linguistic Theory 16*: 633–678

Rizzi, L. 1990. *Relativized minimality*. Cambridge, MA: MIT Press

Rizzi, L. 1997. The Fine Structure of the Left Periphery. In L. Haegeman (ed.) *Elements of Grammar*, pp. 281-337. Kluwer, Dordrecht

Rizzi, L. 2004. Locality and left periphery. In A. Belletti (ed.), *Structures and beyond: The cartography of syntactic structures,* 3. Oxford University Press, New York

Rizzi, L. 2005. On some properties of subjects and topics. In Brugé, L., Giusti, G., Munaro, N., Schweikert, W. and Turano, G. (eds.). *Proceedings of the XXX Incontro di Grammatica Generativa*. Venezia, Cafoscarina

Rodionova, E. 2001. Word Order and Information Structure in Russian Syntax. MA Theses. University of North Dakota

Romero, M. 1998. Sluicing and Focusing a Wh-Phrase. In E. Benedicto, M. Romero and S. Tomioka (eds.)*, Proceedings of the Workshop on Focus*, pp. 191-208. Amherst, MA: GLSA

Rooth, M. 1992. A Theory of Focus Interpretation. *Natural Language Semantics 1:* 75–116

Ross, J.R. 1969. Guess who? In R.I. Binnick, A. Davison, G.M. Green, and J.L. Morgan (eds.) *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society*, pp. 252-286. Chicago: Chicago Linguistic Society

Rudin, C. 1986. *Aspects of Bulgarian Syntax: Complementizers and Wh-Constructions*. Columbus, OH: Slavica.

Rudin, C. 1988. On multiple questions and multiple wh-fronting. *Natural Language & Linguistic Theory 6:* 445–501

Rudnitskaya, E. 2000. The Derivation of Yes-No Li Questions in Russian: Syntax and/or Phonology? In T. H. King & I. A. Sekerina (eds.) *Proceedings of FASL at Philadelphia*. pp. 347-362. Ann Arbor: Michigan Slavic Publications

Saito, M. 1992. Long distance scrambling in Japanese. *Journal of East Asian Linguistics 1*(1): 69-118

Saito, M., Murasugi, K. 1999. Subject predication within IP and DP. In K. Johnson & I. Roberts (eds.) *Beyond Principles and Parameters: Essays in Memory of Osvaldo Jaeggli*, pp. 167–188. Dordrecht: Kluwer

Schachter, P. 1973. Focus and relativization. *Language 49:* 19–46

Schahter, P. 1977. Constraints on coordination. *Language 53:* 86-103

Scott, T. 2001. Multiple Wh in Russian. Ms. Stony Brook University

Scott, T. 2003. CP-Topic Parameter and multiple WH. *Presented at the Linguistics Society of America Annual Meeting 79.*

Scott, T. 2010. Spurious coordination in Russian multiple WH. Paper presented at the 2010 LSA Annual Meeting, Baltimore, MD

Sekerina, I. 1997. The Syntax and Processing of Scrambling Constructions in Russian. Doctoral Dissertation, CUNY

Selkirk, E. O. 1995. Sentence prosody: Intonation, stress and phrasing. In J. Goldsmith (ed.) *Handbook of phonological theory*, pp. 550–569. Cambridge, MA: Blackwell

Sheehan, M., 2009. The Final-over-Final Constraint as an effect of Complement Stranding. *Newcastle Working Papers in Linguistics 15*: 104-125

Sinopoulou, O. 2009. Wh&wh questions in Greek. *Paper presented at the 19th International Symposium on Theoretical & Applied Linguistics, Aristotle University of Thessaloniki, April 3–5*

Skrabalova, H. 2006. Parataxe apparente et coordination des interrogatifs en tchèque. In I. Bril, G. Rebuschi (eds) *Coordination et subordination : typologie et modélisation. Faits de Langue 28*. Paris: Ophrys

Slioussar, N. 2007. Grammar and Information Structure. A study with reference to Russian. Utrecht: LOT Publications

Sportiche, D. 1988. A theory of floated quantifiers and its corollaries for constituent structure. *Linguistic Inquiry 19:* 425-450

Sprouse J., Almeida, D. 2010. A quantitative defense of linguistic methodology. Ms. http://ling.auf.net/lingBuzz/001075

Stepanov, A. 1998. On Wh-Movement in Russian. In P. N. Tamanji & K. Kusumoto, (eds.) *Proceedings of the 28th Meeting of the North East Linguistic Society*. pp. 453-467. Amherst, MA: GLSA, University of Massachusetts

Stepanov, A., Stateva, P. 2009. "When QR disobeys Superiority". *Linguistic Inquiry 40*

Stjepanović, S. 1995. Short Distance Movement of Wh-Phrases in Serbo-Croatian Matrix Clauses. Ms. University of Connecticut, Storrs

Stjepanović, S. 2003. Multiple wh-fronting in Serbo-Croatian matrix questions and the matrix sluicing construction. In C. Boekx, K. K. Grohmann (eds.) *Multiple Wh-fronting*, pp. 255-284. Amsterdam: John Benjamins

Strahov, N. 2001. A Scrambling Analysis of Russian Wh-Questions. In S. Franks et. al. eds.) *Proceedings of FASL IX*, pp, 293-311. Ann Arbor, MI: Michigan Slavic Publications

Strawson, P. 1964. Identifying Reference and Truth-Values. *Teoria, 30:* 96–118

Sturgeon A., Polinsky M., Kravtchenko E., Gallo C. G., Medová L., Koula V. 2010. Subject Islands in Slavic: The Syntactic Position Matters. University UC Santa Cruz, Syntax Lab March 5, 2010

Švedova, N. Yu. (ed.) 1980. *Russkaja grammatika*. Moscow. Russia

Takahashi, D. 1993. Movement of wh-phrase in Japanese. Natural Language and *Linguistic Theory, 11:* 655-678

Takahashi, D. 1994. Sluicing in Japanese. *Journal of East Asian Linguistics 3:* 265–300

Thrainsson, H. 1996. On the (non)-universality of functional projections. In W. Abraham, S.D. Epstein, H. Thrainsson & C. J.-W. Zwart (eds.) *Minimal ideas*, pp. 253-81. Amsterdam: Benjamins

Tomaszewicz, B. 2010. Against spurious coordination in multiple wh-questions. *Paper presented at West Coast Conference on Formal Linguistics 28*, University of Southern California, Los Angeles, CA

Vassilyeva, A. N. 1972. *Particles in colloquial Russian*. Moscow: Progress Publishers

Vallduvi, E. 1990. The informational component. Doctoral Dissertation, University of Pennsylvania

Vallduví, E. 1992. *The Informational Component*. NY, London: Garland Publishing Inc.

Van Valin, R. 1999. A typology of the interaction of focus structure and syntax. In E. Raxilina and J. Testelec (eds). *Typology and Linguistic theory: from description to explanation* pp. 511- 524. Moscow: Languages of Russian Culture

Wagner, M. 2009. Focus, Topic, and Word Order: A Compositional View. In J. Craenenbroeck (ed.) *Alternatives to cartography*, pp. 53-86. Berlin: Mouton de Gruyter

Wiland, B. 2010. Prefix stacking, syncretism, and the syntactic hierarchy." *Workshop on Slavic Prefixes and Prepositions, Formal Description of Slavic Languages (FDSL 8.5*), Brno, November 2010

Williams, E. 1981. Transformationless grammar. *Linguistic Inquiry 12*: 645– 653

Williams, E. 2009. There is no alternative to cartography. In J. Craenenbroeck (ed.) *Alternatives to cartography*, pp. 361-374. Berlin: Mouton de Gruyter

Yadroff, M. 1991. The syntactic properties of adjunction in Russian. Ms., Indiana University

Yanko, T. E. 2001. Kommunikativnyje strategii russkoj reči. Jazyki slavjanskoj kul'tury. Moscow, Russia

Yokoyama, O. 1986. *Discourse and Word Order*. Amsterdam, Philadelphia: John Benjamins

Zavitnevich, O. 2005. On Wh- Movement and the Nature of Wh-Phrases: Case Re-Examined. *Skase Journal of Theoretical Linguistics, 2 (3):* 75-100

Zemskaya, E. A. 1973. Sovremennyj russkij jazyk. Prosveščenie. Moscow, Russia

Zemskaya E.A. 2004. Russkaja razgovornaja reč: lingvističeskij analiz i problem obučenija. Moscow: Nauka

Zhang, N. 2007. The Syntactic Derivations of Two Paired Dependency Constructions. *Lingua 117 (12):* 2134-2158

Zhang, N. 2009. *Coordination in Syntax*. Cambridge, UK: Cambridge University Press

Zolotova G. A. 2000. *Kommunikativnaja grammatika russkogo jazyka*. MGU, Moscow, Russia

Zubizarreta, M.L. 1998. *Prosody, Focus, and Word Order.* Linguistic inquiry monographs 33. Cambridge, Mass.: MIT Press

Zubizaretta, M. L. 2001. "Preverbal Subjects in Romance Interrogatives." In Aafke Hulk and Jean-Yves Pollock, eds., *Subject Inversion in Romance and the Theory of Universal Grammar*. Oxford University Press., New York, pp. 183-204

Zwart, C. J. W. 2005. Verb second as a function of Merge. In M. den Dikken and C.M. Tortora (eds.) *The function of function words and functional categories*, pp. 11-40. Amsterdam: John Benjamins